ISSN 2308-3840 (online)

World Journal of *Meta-Analysis*

World J Meta-Anal 2015 February 26; 3(1): 1-81

Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

A peer-reviewed, online, open-access journal of meta-analysis

Editorial Board

2013-2018

The *World Journal of Meta-Analysis* Editorial Board consists of 400 members, representing a team of worldwide experts in clinical meta-analysis research. They are from 41 countries, including Argentina (2), Australia (3), Austria (1), Belgium (5), Brazil (11), Canada (16), Chile (3), China (121), Croatia (1), Egypt (1), Finland (4), France (2), Germany (9), Greece (11), Hungary (2), India (12), Iran (2), Ireland (1), Israel (2), Italy (41), Japan (5), Lithuania (1), Malaysia (1), Netherlands (8), New Zealand (1), Norway (1), Peru (1), Poland (4), Portugal (6), Romania (1), Saudi Arabia (4), Singapore (3), South Africa (1), South Korea (7), Spain (8), Sri Lanka (2), Sudan (1), Switzerland (2), Thailand (4), Turkey (3), United Kingdom (22), and United States (64).

Alessandro Cucchetti, *Bologna*
Gianfranco Damiani, *Rome*
Fabrizio D'Ascenzo, *Turin*
Massimo Del Fabbro, *Milano*
Valeria Fadda, *Florence*
Alessandro Fancellu, *Sassari*
Giuseppe Ferrante, *Milan*
Virginia Festa, *Rome*
Francesco Fiorica, *Ferrara*
Guglielmo Giraldi, *Rome*
Jenny Guidi, *Bologna*
Lorenzo Loffredo, *Rome*
Andrea Messori, *Firenze*
Eliano Pio Navarese, *Bydgoszcz*
Stefano Omboni, *Solbiate Arno*
Alvisa Palese, *Udine*
Stefano Palomba, *Reggio Emilia Stefano*
Carlo Perricone, *Rome*
Mario Petretta, *Naples*
Alessandro Pezzini, *Brescia*
Gianluca Pontone, *Milan*
Paolo Emilio Puddu, *Rome*
Andrea Rognoni, *Novara*
Giuseppe Scalabrino, *Milan*
Città della Salute e della Scienza, *Turin*
Fabrizio Sgolastra, *L'Aquila*
Maria Lucia Specchia, *Rome*
Stefano Trastulli, *Terni*
Fabio Tine, *Palermo*
Nereo Vettoretto, *Chiari*
Alberto Vianello, *Perugia*
Luigi Zorcolo, *Cagliari*

Japan

Nguyen Tien Huy, *Nagasaki*
Hiroharu Kamioka, *Tokyo*
Koji Kawakami, *Kyoto*
Keitaro Matsuo, *Nagoya*
Kazushi Okamoto, *Nagoya*

Lithuania

Edmundas Kadusevicius, *Kaunas*

Malaysia

SP Pani, *Ipoh*

Netherlands

Michel van den Bekerom, *Amsterdam*
Dimitra Dodou, *Delft*
D Haverkamp, *Amsterdam*
Vassilios Koussoulas, *Drachten*
BJ Polder, *Nijmegen*
Theo Stijnen, *Leiden*
RNM Weijers, *Amsterdam*
Joost de Winter, *Delft*

New Zealand

Shaofeng Li, *Auckland*

Norway

Eivind Berge, *Oslo*

Peru

Rafael Bolaños Díaz, *Lima*

Poland

Maciej Banach, *Lodz*
Krzysztof Jonderko, *Sosnowiec*
Jolanta Lissowska, *Warsaw*
Maciej Plaszewski, *Warsaw*

Portugal

Daniel Caldeira, *Lisboa*
J Costa, *Lisbon*
Ana Miguel, *Coimbra*
Manuel Morgado, *Covilh*
Bárbara Peleteiro, *Porto*
Rui Torres, *Paredes*

Romania

Fratila Ovidiu, *Oradea*

Saudi Arabia

Hazem M Al-Mandeel, *Riyadh*
Ezzeldin M Ibrahim, *Jeddah*
Mutahir A Tunio, *Riyadh*
Hayfaa A Wahabi, *Riyadh*

Singapore

Nikos LD Chatzisarantis, *Singapore*
Edwin Choon Wyn Lim, *Singapore*
Roger Ho, *Singapore*

South Africa

Alaine Umubyeyi Nyaruhirira, *Pretoria*

South Korea

Jung-Hee Kim, *Cheonan*
Hyangsook Lee, *Seoul*
Myeong Soo Lee, *Daejeon*
Chi-Un Pae, *Bucheon*
Yong Hyun Park, *Yangsan*
Jae Hong Seo, *Seoul*
Yong Sang Song, *Seoul*

Spain

Pablo Avanzas, *Oviedo*
Joan Cid, *Barcelona*
Joaquin de Haro, *Madrid*
Joan Guardia-Olmos, *Barcelona*
Nabil Halaihlel, *Zaragoza*
JA Monge-Argilés, *Alicante*
Raul Moreno, *Madrid*
Inés Velasco, *Aracena*

Sri Lanka

Ranil Jayawardena, *Colombo*
Priyanga Ranasinghe, *Colombo*

Switzerland

Jay P Singh, *Zurich*
Giorgio Treglia, *Bellinzona*

Sudan

Samir MH Shaheen, *Khartoum*

Thailand

Chuenjid Kongkaew, *Phitsanulok*
Manop Pithukpakorn, *Bangkok*
Piyamitr Sritara, *Bangkok*
Surasak Saokaew, *Phayao*

Turkey

Nese Demirturk, *Afyonkarahisar*
Nilüfer Ozabaci, *Eskisehir*
Ilke Sipahi, *Istanbul*

United Kingdom

Omar M Aboumarzouk, *Wales*
Abeer Al-Namankany, *London*
Ernest A Azzopardi, *Cardif*
Umberto Benedetto, *Cambridge*
Joanne Brooke, *London*
Noriko Cable, *London*
David Chan, *Cardiff*
YC Cheong, *Southampton*
Andrew Currie, *Harrow*
G Nabi, *Dundee*

Lesley A Anderson, *Belfast*
Valentina Gallo, *London*
Gianpiero Gravante, *Leicester*
Peter N Lee, *Surrey*
Igho Onakpoya, *Oxford*
Ashish Pradhan, *Huntingdon*
Evridiki Patelarou, *London*
Jian-Qing Shi, *Newcastle*
Surendra P Singh, *Wolverhampton*
Natalie Taylor, *Leeds*
Yousef Shahin, *Hull*
Zheng Ye, *Cambridge*

United States

L Joseph Su, *Rockville*
Olusola Adesope, *Pullman*
Mike Allen, *Milwaukee*
Bhupinder Anand, *Bellaire*
Stephen Aronoff, *Philadelphia*
KoKo Aung, *San Antonio*
William L Baker, *Farmington*
Moritz C Wyler von Ballmoos, *Milwaukee*
Matthew L Bechtold, *Columbia*
Atul Bhardwaj, *Hershey*
Somjot S Brar, *Los Angeles*
Hui Cai, *Nashville*
Subhash Chandra, *Towson*
Wen-Pin Chang, *Omaha*
Yong Chen, *North Wales*
Myunghan Choi, *Phoenix*
John Coverdale, *Houston*
Prakash C Deedwania, *Fresno*
Eugene Demidenko, *Hanover*
Hong-Wen Deng, *New Orleans*
Eric M Deshaies, *Syracuse*
Tao Fan, *Whitehouse Station*
Shinga Feresu, *Bloomington*
Janvier Gasana, *Miami*
Kaveh Hajifathalian, *Boston*
Mohammad Obaidul Hoque, *Baltimore*
Larissa R Brunner Huber, *Charlotte*
Imran H Iftikhar, *Columbia*
Vijayvel Jayaprakash, *Buffalo*
Xuezhi Jiang, *Weat Reading*
Shuo Jiao, *Seattle*
Evelyn Johnson, *Boise*
Le Kang, *Silver Spring*
SR Kapadia, *Cleveland*
Lior Katz, *Houston*
Daniel M Laskin, *Richmond*
Yu Liang, *Foster City*
Paul Ellis Marik, *Norfolk*
Lynne V McFarland, *Seattle*
Marcovalerio Melis, *New York*
Brian J Miller, *Augusta*
Pavlos Msaouel, *New York*
Joshua E Muscat, *Hershey*
Chee Yuan Ng, *Loma Linda*
Nghi C Nguyen, *Saint Louis*
Brandi S Niemeier, *Whitewater*
Thomas D Parsons, *Denton*
Nidal Abi Rafeh, *New Orleans*
Praveen Roy, *Marshfield*
Ali Salavati, *Philadelphia*
Ankur Sethi, *Chicago*
Tatyana A Shamliyan, *Minneapoli*
Qian Shi, *Rochester*
Zhongjie Shi, *Philadelphia*
Param Puneet Singh, *Chicago*
KV Slavin, *Chicago*
Ali El Solh, *Buffalo*
Jielin Sun, *Winston-Salem*
Richard G Trohman, *Chicago*
Laurah Turner, *Cincinnati*
Sheila Wilhelm, *Detroit*
Alex K Wong, *Los Angeles*
Xiaohui Xu, *Gainesville*
Lu Yin, *Nashville*

WJMA World Journal of Meta-Analysis

Contents Bimonthly Volume 3 Number 1 February 26, 2015

EDITORIAL

META-ANALYSIS

World Journal of Meta-Analysis
Volume 3 Number 1 February 26, 2015

Contents

NAME OF JOURNAL
World Journal of Meta-Analysis

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Bimonthly

EDITOR-IN-CHIEF
Giuseppe Biondi-Zoccai, MD, Assistant Professor, Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
World Journal of Meta-Analysis
Room 903, Building D, Ocean International Center, No. 62 Dongsihuan Zhonglu, Chaoyang District, Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
8226 Regency Drive,
Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
February 26, 2015

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.1

World J Meta-Anal 2015 February 26; 3(1): 1-3
ISSN 2308-3840 (online)

EDITORIAL

Why meta-analyses are important for complementary and alternative medicine research

Holger Cramer

Holger Cramer, Department of Internal and Integrative Medicine, Kliniken Essen-Mitte, Faculty of Medicine, University of Duisburg-Essen, 45276 Essen, Germany
Author contributions: Cramer H solely contributed to this editorial.
Supported by Rut- and Klaus-Bahlsen Foundation, Germany.
Conflict-of-interest: The supportive foundations had no influence on the content of this editorial.

Correspondence to: Dr. Holger Cramer, PhD, MSc, Department of Internal and Integrative Medicine, Kliniken Essen-Mitte, Faculty of Medicine, University of Duisburg-Essen, Am Deimelsberg 34a, 45276 Essen, Germany. h.cramer@kliniken-essen-mitte.de
Telephone: +49-201-17425015
Fax: +49-201-17425000
Received: October 20, 2014
Peer-review started: October 20, 2014
First decision: November 27, 2014
Revised: December 3, 2014
Accepted: December 29, 2014
Article in press: December 31, 2014
Published online: February 26, 2015

Abstract

Complementary and alternative medicine (CAM) is defined as a group of interventions that are not generally considered part of conventional medicine. This definition already implies that CAM interventions are often not systematically studied; and the research evidence from single trials on CAM is often limited by small sample sizes, unclear methodology, and inadequate statistics. As a result, both, significant and insignificant results are often hard to interpret based on single trials. Summarizing the evidence from single CAM trials, qualitative systematic reviews still have to deal with the same problems as individual trials as they can only rely on the original reports. Thus, effects of CAM interventions are often underestimated or overestimated based on single trials or qualitative systematic reviews. While meta-analyses still are limited by the methodological shortcomings of the included studies, a well-conducted meta-analysis can deal with two common problems of CAM trials: inadequate statistics that rely on within-group comparisons and small underpowered sample sizes. Although large and high quality trials are urgently needed for most CAM interventions, funding often is limited. Until higher quality research is available, meta-analyses provide a useful tool to investigate the actual level of evidence of currently published CAM trials. This editorial presents examples of meta-analyses in the field of CAM and discusses how they contribute to the consolidation of evidence.

Key words: Complementary therapies; Meta-analysis; Review; Randomized controlled trial; Bias

Core tip: The research evidence from single trials on complementary and alternative medicine (CAM) is often limited by small sample sizes, unclear methodology, and inadequate statistics. Qualitative systematic reviews still have to deal with the same problems as individual trials as they can only rely on the original reports. While meta-analyses still are limited by the methodological shortcomings of the included, they can deal with two common problems of CAM trials: inadequate statistics that rely on within-group comparisons and small underpowered sample sizes.

Cramer H. Why meta-analyses are important for complementary and alternative medicine research. *World J Meta-Anal* 2015; 3(1):

1-3 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/1.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.1

TEXT

The National Center for Complementary and Alternative Medicine of the National Institute of Health defines complementary and alternative medicine (CAM) as "a group of diverse medical and health care systems, practices, and products that are not generally considered part of conventional medicine"[1]. A growing number of randomized controlled trials (RCTs) aimed to investigate the effectiveness of CAM therapies in varied medical conditions. However, while these trials are urgently needed to consolidate evidence for interventions that have been - by definition - rarely studied systematically in the past, the research evidence from single trials on CAM is often limited by small sample sizes, unclear methodology, and inadequate statistics. Both, significant and insignificant results are often hard to interpret based on single trials. On the one hand, while randomized trials are generally conducted to compare effects of two or more different treatments on a specific condition, especially CAM trials often solely rely on within-group comparisons that do not take into account unspecific effects. Thereby, the value of having a control group is lost and it is often impossible to estimate the real specific effects of the intervention; overestimating its actual specific efficacy. On the other hand, small trials that often have to deal with marked baseline differences between groups are often underpowered to detect specific effects of the intervention. Based on those problem, it has even been encouraged to abandon RCT designs in CAM research altogether. However, most problems of CAM RCTs can be adequately addressed by proper methodology use. In recent years, a continuously growing number of systematic reviews have been conducted in order to summarize evidence from single CAM trials. While this tendency is definitely useful to consolidate evidence, qualitative reviews still have to deal with the same problems as individual trials as they can only rely on the original - often heavily biased - reports. This is where meta-analyses should come into play. While meta-analyses still are limited by the methodological shortcomings of the included trials and badly conducted meta-analyses can even worsen the situation, a well-conducted meta-analysis can deal with two common problems of CAM trials: inadequate statistics that rely on within-group comparisons and small underpowered sample sizes. By ignoring the original statistical method of the published analyses and by quantitatively pooling the results of several trials, a between-group analysis based on a larger sample can be created that compensates for at least some of the shortcomings of the original trials. Although a trained statistician would surely be able to assess the real effect sizes from the published data without relying on a published meta-analysis, CAM trials are often used to guide clinical decision making; and a clinician often will not be able to re-evaluate the statistics of a published trial. While it does not reduce the risk of bias of the original studies, a meta-analysis can address the problem of inadequate statistics and improve the power of the analysis. Take, *e.g.*, yoga. Yoga has now become a popular means to improve health and well-being and several studies have investigated yoga's effectiveness in varied medical conditions. In 2011, a systematic review on yoga for low back pain included a total of seven RCTs published until March 2011[2]. Five of the RCTs found significant effects of the yoga interventions while the other two did not. The systematic review concluded that yoga might be able to alleviate low back pain but that any definitive claims should be treated with caution[2]. A second systematic review on the same condition included trials that were published until January 2012[3]. While the first review refrained from meta-analyzing the data due to heterogeneity of the included trials, the second review was able to include a meta-analysis. Based on ten RCTs, this meta-analysis found strong evidence for short-term effectiveness and moderate evidence for long-term effectiveness of yoga for chronic low back pain in the most important patient-centered outcomes[3]. As a consequence, this review concluded that yoga can be recommended as an additional therapy to chronic low back pain patients. While the obvious divergences in conclusions might also be accounted to the increased number of included RCTs in the second review, they are likely to be at least partly based on the inclusion of a meta-analysis whose findings can go beyond just balancing the results of individual trials against each other. On the other hand, meta-analyses can also help to revise falsely overoptimistic conclusions of single trials. A 2012 systematic review on the effects of yoga for schizophrenia included three RCTS that were published until October 2011[4]. Despite the low number of eligible trials, the overall positive findings of those led to the conclusion that yoga could be helpful in reducing general psychopathology, positive, and negative symptoms in patients with schizophrenia. However, as these conclusions were based on the results that were reported in the original articles, and one out of three RCTs reported only within-group pre-post comparisons rather than between-group comparisons, these results were not robust against reporting bias. Accordingly, based on the same trials, a recent meta-analysis failed to find any effects of yoga on schizophrenia psychopathology[5]; resulting in the counterintuitive finding that a meta-analysis of three RCTs that all reported positive effects resulted in insignificant group differences.

A problem of meta-analyses - especially for but not limited to CAM research - is heterogeneity between trials; specifically clinical heterogeneity (differences in, *e.g.*, interventions that might be labeled with the same term) which often results in statistical heterogeneity (differences in the interventions' effects). While studies in a meta-analysis will inevitably differ from each other, substantial statistical heterogeneity can reduce the precision of effect estimates. Thus, authors of CAM-related meta-

analyses should be aware of the heterogeneity of CAM interventions and define the focus of their meta-analysis as precisely as possible.

Small underpowered and poorly conducted trials are by no means only a problem of CAM research. However, as external funding for CAM trials is limited to non-existent in most countries, large well-conducted trials are especially difficult to conduct in this research area. While meta-analyses cannot compensate for low-quality original research, they provide a useful tool to investigate the actual level of evidence of currently published CAM trials until higher-quality research evidence is available.

REFERENCES

1 National Center for Complementary and Alternative Medicine (NCCAM) [Internet]. Bethesda, MD: National Center for Complementary and Alternative Medicine. Available from: URL: http: //nccam.nih.gov/

2 **Posadzki P**, Ernst E. Yoga for low back pain: a systematic review of randomized clinical trials. *Clin Rheumatol* 2011; **30**: 1257-1262 [PMID: 21590293 DOI: 10.1007/s10067-011-1764-8]

3 **Cramer H**, Lauche R, Haller H, Dobos G. A systematic review and meta-analysis of yoga for low back pain. *Clin J Pain* 2013; **29**: 450-460 [PMID: 23246998 DOI: 10.1097/AJP.0b013e31825e1492]

4 **Vancampfort D**, Vansteelandt K, Scheewe T, Probst M, Knapen J, De Herdt A, De Hert M. Yoga in schizophrenia: a systematic review of randomised controlled trials. *Acta Psychiatr Scand* 2012; **126**: 12-20 [PMID: 22486714 DOI: 10.1111/j.1600-0447.2012.01865.x]

5 **Cramer H**, Lauche R, Klose P, Langhorst J, Dobos G. Yoga for schizophrenia: a systematic review and meta-analysis. *BMC Psychiatry* 2013; **13**: 32 [PMID: 23327116 DOI: 10.1186/1471-244X-13-32]

P- Reviewer: Allen M, Floriani I, Grundmann O, Xiong XJ
S- Editor: Song XX **L- Editor**: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.4

World J Meta-Anal 2015 February 26; 3(1): 4-10
ISSN 2308-3840 (online)

META-ANALYSIS

Prophylactic tracheal intubation for upper GI bleeding: A meta-analysis

Ashraf A Almashhrawi, Rubayat Rahman, Samuel T Jersak, Akwi W Asombang, Alisha M Hinds, Hazem T Hammad, Douglas L Nguyen, Matthew L Bechtold

Ashraf A Almashhrawi, Rubayat Rahman, Samuel T Jersak, Akwi W Asombang, Alisha M Hinds, Hazem T Hammad, Matthew L Bechtold, Division of Gastroenterology and Hepatology, University of Missouri, Columbia, MO 65212, United States

Douglas L Nguyen, Division of Gastroenterology and Hepatology, University of California, Irvine, CA 92697, United States

Author contributions: Almashhrawi AA, Nguyen DL and Bechtold ML were responsible for the conception and design of the study; Almashhrawi AA, Rahman R, Jersak ST and Hinds AM collected the data and organized data extraction sheets; Hammad HT, Nguyen DL and Bechtold ML statistically analyzed the data; Almashhrawi AA, Rahman R and Jersak ST drafted the manuscript with critical revision being performed by Asombang AW, Hammad HT, Nguyen DL and Bechtold ML.

Conflict-of-interest: The authors have no conflicts of interest for this manuscript.

Data sharing: No additional data are available.

Correspondence to: Matthew L Bechtold, MD, FACP, FASGE, FACG, Division of Gastroenterology and Hepatology, University of Missouri, Five Hospital Drive, Columbia, MO 65212, United States. bechtoldm@health.missouri.edu
Telephone: +1-573-8821013
Fax: +1-573-8844595

Received: October 28, 2014
Peer-review started: October 28, 2014
First decision: November 14, 2014
Revised: December 22, 2014
Accepted: January 9, 2015
Article in press: January 12, 2015
Published online: February 26, 2015

Abstract

AIM: To evaluate usefulness of prophylactically intubating upper gastrointestinal bleeding (UGIB) patients.

METHODS: UGIB results in a significant number of hospital admissions annually with endoscopy being the key intervention. In these patients, risks are associated with the bleeding and the procedure, including pulmonary aspiration. However, very little literature is available assessing the use of prophylactic endotracheal intubation on aspiration in these patients. A comprehensive search was performed in May 2014 in Scopus, CINAHL, Cochrane databases, PubMed/Medline, Embase, and published abstracts from national gastroenterology meetings in the United States (2004-2014). Included studies examined UGIB patients and compared prophylactic intubation to no intubation before endoscopy. Meta-analysis was conducted using RevMan 5.2 by Mantel-Haenszel and DerSimonian and Laird models with results presented as odds ratio for aspiration, pneumonia (within 48 h), and mortality. Funnel plots were utilized for publication bias and I^2 measure of inconsistency for heterogeneity assessments.

RESULTS: Initial search identified 571 articles. Of these articles, 10 relevant peer-reviewed articles in English and two relevant abstracts were selected to review by two independent authors (Almashhrawi AA and Bechtold ML). Of these studies, eight were excluded: Five did not have a control arm, one was a letter the editor, one was a survey study, and one was focused on prevention of UGIB. Therefore, four studies (N = 367) were included. Of the UGIB patients prophylactically intubated before endoscopy, pneumonia (within 48 h) was identified in 20 of 134 (14.9%) patients as compared to 5 of 95 (5.3%) patients that were not intubated prophylactically (P = 0.02). Despite observed trends, no significant

differences were found for mortality (P = 0.18) or aspiration (P = 0.11).

CONCLUSION: Pneumonia within 48 h is more likely in UGIB patients who received prophylactic endotracheal intubation prior to endoscopy.

Key words: Prophylactic endotracheal intubation; Upper gastrointestinal bleeding; Endoscopy; Complication; Pneumonia; Aspiration

Core tip: Patients with upper gastrointestinal bleeding (UGIB) require endoscopic treatment with variable outcomes of aspiration, pneumonia, non-endoscopic interventions, and mortality. It is suggested that endotracheal intubation prior to endoscopy might reduce aspiration, pneumonia, and mortality. Few studies have evaluated this issue. We performed a meta-analysis of observational studies examining endotracheal intubation *vs* no intubation in UGIB patients. We found that patients intubated had higher incidence of pneumonia within 48 h. There was no significant increase in aspiration and mortality in the intubated group. This meta-analysis demonstrates the need for randomized controlled trials to assess the issue.

Almashhrawi AA, Rahman R, Jersak ST, Asombang AW, Hinds AM, Hammad HT, Nguyen DL, Bechtold ML. Prophylactic tracheal intubation for upper GI bleeding: A meta-analysis. *World J Meta-Anal* 2015; 3(1): 4-10 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/4.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.4

INTRODUCTION

Upper gastrointestinal bleeding (UGIB) is still significant in the United States[1]. Health-resources utilization in those with UGIB is significantly higher than those without UGIB[2-5]. Although UGIB hospitalizations have decreased in the last decade, likely because the use of acid suppression therapy[6,7], mortality has not decreased and UGIB continues to be a significant cause of hospital admissions[8-13].

Many strategies have been implemented to reduce the morbidity, mortality, and cost associated with UGIB, including scoring systems, appropriate resuscitation, and improvements in endoscopic and non-endoscopic therapies[14-19]. In an attempt to reduce aspiration and aspiration pneumonia in patients presenting with UGIB, prophylactic tracheal intubation prior to performing endoscopy has been used, but is there any evidence to support such a practice. Tracheal intubation might prevent aspiration in selected cases but outcomes could be related to how experienced medical personnel performing the intubation is and how sick the patient is, *i.e.*, with altered mental status or massive bleeding[20-22]. However, controversy does exist, even at our own institution, of the utility of prophylactic intubation in patients with UGIB. The largest reason for this controversy is that limited observational studies have addressed the utilization of tracheal intubation in the setting of UGIB[23-27]. These studies evaluated outcomes, including aspiration, mortality, aspiration pneumonia, and hospital length of stay. As our knowledge to answer the question of the utility of tracheal intubation in the setting of UGIB is still lacking, we conducted a meta-analysis to further evaluate such limited data.

MATERIALS AND METHODS

Search of literature

A complete search of Scopus, CINAHL, Cochrane databases, PubMed/Medline, and Embase was completed in May 2014. Search terms were used individually or in various combinations and included "endotracheal intubation", "tracheal intubation", "upper gastrointestinal bleeding", "upper gastrointestinal hemorrhage", "variceal hemorrhage", "non-variceal hemorrhage", "esophagogastroduodenoscopy", "peptic ulceration", "duodenal ulceration", and "gastric ulceration". Peer-reviewed studies on UGIB patients that compared prophylactic to no prophylactic intubation were selected and reviewed. References of relevant papers were searched as well for possible additional articles that were not identified in the original search. Search also included published abstracts in the major digestive disease conferences in the United States in the last 10 years. Three investigators reviewed all studies selected for inclusion criteria. Studies in children or in languages other than English were excluded from this meta-analysis.

Data extraction

All included studies were reviewed with two investigators (AA, MB). At least two of three primary outcomes were evaluated in all included studies. If a study had missing data on these subjects or clarification was needed, attempts were made to contact the authors to obtain the necessary information. Data from the studies chosen were extracted by two investigators individually and were settled by mutual agreement.

Statistical analysis

This meta-analysis followed principles of the MOOSE guidelines[28]. Meta-analysis was performed comparing the results of UGIB patients by calculating pooled estimates presented as odds ratio (OR) of outcomes with Mantel-Haenszel (if no heterogeneity) or DerSimonian and Laird models (if heterogeneity). Heterogeneity analyzed by calculating I^2 measure of inconsistency (significant if $P < 0.10$ or $I^2 > 50\%$). A sensitivity analysis was done if heterogeneity was statistically significant. RevMan 5.2 (Copenhagen: The Nordic Cochrane Centre, The

Table 1 Details of the studies

Ref.	Study type	Country	Time	No. of patients	Group	*n*	Age (mean/ median years)	Gender (% male)	Population (inclusion criteria)
Koch *et al*[25]	Retrospective	United States	1995-2002	62	PI	42	49	74	Bleeding varices
					No PI	20	48	65	No radiographic or clinical respiratory issues
Rehman *et al*[24]	Retrospective	United States	2002-2006	98	PI	49	62[b]	61	Cirrhosis
					No PI	49	68[b]	82	Hematemesis Shock
[a]Perisetti *et al*[26]	Retrospective	United States	2000-2013	138	PI	69	61[b]	NA	Endoscopy with intubation
					No PI	69	66[b]	NA	Matched controls
[a]Tang *et al*[27]	Retrospective	United States	2008-2013	69	PI	43	53	69.8	Endoscopy in suspected variceal bleeding
					No PI	26	55	61.5	

[a]Abstract; [b]Median. PI: Prophylactic intubation; NA: Data not available.

Table 2 Quality assessment of studies included in meta-analysis

Ref.	Study design	Selection bias	Confounders	Blinding	Data collection methods	Withdrawals and dropouts	Intervention integrity	Analyses	Quality assessment
Koch *et al*[25]	Retrospective Moderate	Moderate	Moderate	Weak	Strong	NA	Moderate	Moderate	Moderate
Rehman *et al*[24]	Retrospective Moderate	Moderate	Strong	Weak	Strong	NA	Strong	Moderate	Moderate
Perisetti *et al*[26]	Retrospective Moderate	Moderate	Weak	Weak	Strong	NA	Moderate	Moderate	Weak
Tang *et al*[27]	Retrospective Moderate	Moderate	Strong	Weak	Strong	NA	Moderate	Moderate	Moderate

NA: Data not available.

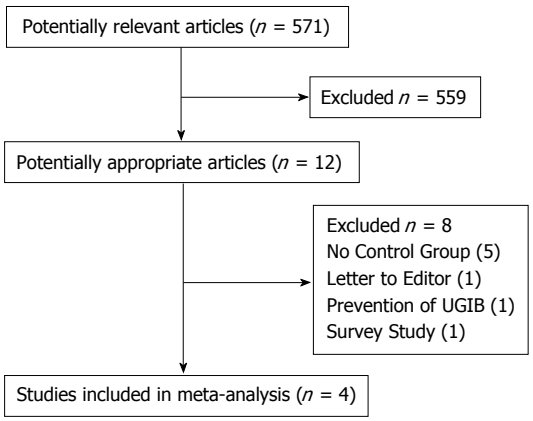


Figure 1 Details of article search. UGIB: Upper gastrointestinal bleeding.

Cochrane Collaboration, 2012) used for statistical analysis. Funnel plots were visually inspected for publication bias assessment.

Study quality assessment

The Effective Public Health Practice Project model was used to assess study quality[29]. This scale is based upon strong, moderate, or weak rankings for analysis, interventional integrity, withdrawal/dropout descriptions, data collection, blinding, confounders, design, and potential selection bias. Study quality is determined by how many weak ratings in each category (no weak ratings is strong, one weak is moderate, and ⩾ 2 weak is weak).

Biostatistics

The corresponding author (Bechtold ML) is a biostatistician and has reviewed and approved all statistical data in the manuscript. Four of the authors (Hinds AM, Hammad HT, Nguyen DL, Bechtold ML) are extensively trained in the statistics used in meta-analysis.

RESULTS

Search of literature

Initially, 571 articles were discovered in the electronic databases (Figure 1). Ten relevant peer-reviewed articles in English and two relevant abstracts were selected for review by two independent authors (Almashhrawi AA and Bechtold ML). Of these studies, eight were excluded: Five did not have a control arm, one was a letter the editor, one was a survey study, and one was focused on prevention of UGIB. Therefore, four studies were identified as meeting inclusion criteria[24-27]. All the four studies included (N = 367) were retrospective cohorts. The studies were conducted throughout the United States and were published 2007 to 2014. All included studies examined the impact of prophylactic endotracheal intubation on UGIB outcomes (Table 1). The study quality was adequate based upon the Effective Public Health Practice Project model (Table 2).

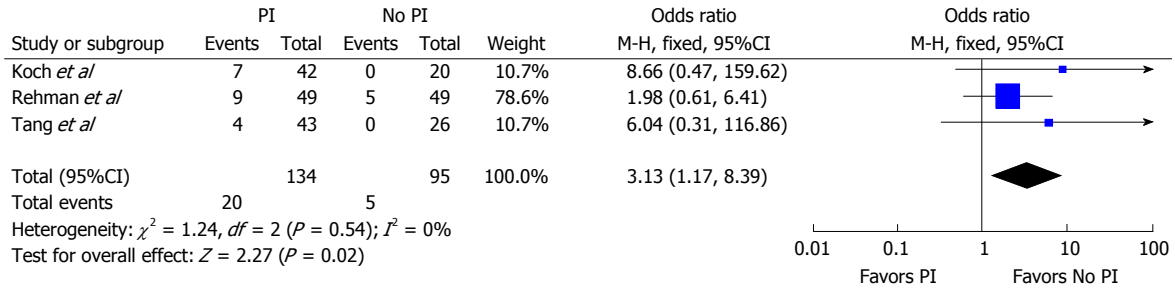


Figure 2 Forest plot demonstrating comparison of prophylactic intubation *vs* no intubation for patients with upper gastrointestinal bleeding for pneumonia within 48 h. PI: Prophylactic intubation.

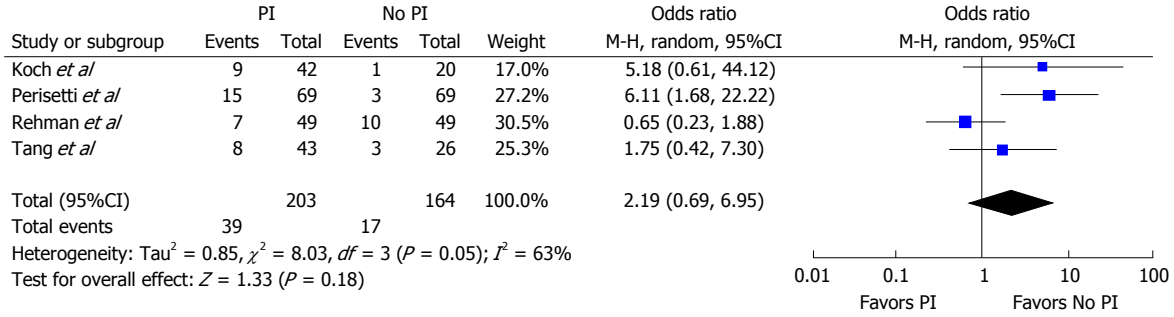


Figure 3 Forest plot demonstrating comparison of prophylactic intubation *vs* no intubation for patients with upper gastrointestinal bleeding for mortality. PI: Prophylactic intubation.

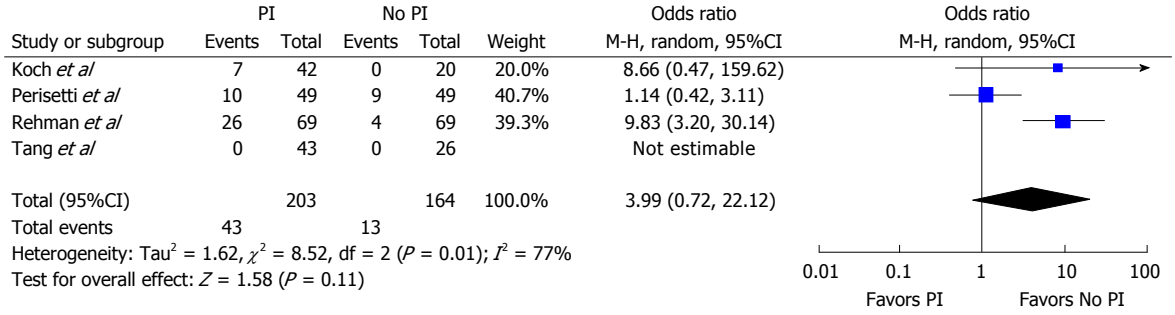


Figure 4 Forest plot demonstrating comparison of prophylactic intubation *vs* no intubation for patients with upper gastrointestinal bleeding for aspiration. PI: Prophylactic intubation.

Pneumonia within 48 h

Pneumonia within 48 h was examined in three studies (N = 229)[24,25,27]. With prophylactic intubation, 20 of 134 (14.9%) patients with UGIB developed pneumonia. For those not being intubated, 5 of 95 (5.3%) patients with UGIB developed pneumonia within 48 h. Those UGIB patients who underwent prophylactic intubation had higher amount of pneumonia than those not prophylactically intubated with odds ratio of 3.13 (95%CI: 1.17-8.39; P = 0.02) with no statistically significant heterogeneity (I^2 = 0%, P = 0.54) (Figure 2).

Mortality

Mortality was examined in four studies (N = 367)[24-27]. Mortality was noted in 39 of 203 (19.2%) patients with UGIB prophylactically intubated and 17 of 164 (10.4%) patients with UGIB not prophylactically intubated. No statistically significant higher mortality was noted for those patients prophylactically intubated (OR = 2.19; 95%CI: 0.69-6.95; P = 0.18) with statistically significant heterogeneity observed (I^2 = 63%, P = 0.05) (Figure 3). Given significant heterogeneity, a sensitivity analysis was performed by excluding the Rehman *et al*[24] study which demonstrated a statistically significant higher mortality for those patients with prophylactically intubated as compared to those not intubated without significant heterogeneity with OR = 3.72 (95%CI: 1.55-8.92; P < 0.01).

Aspiration

Aspiration was analyzed in four studies (N = 367)[24-27]. Aspiration was noted in 43 of 203 (21.2%) patients with UGIB prophylactically intubated and 13 of 164 (7.9%) patients with UGIB not intubated. Statistically non-significant higher aspiration was noted in patients with UGIB prophylactically intubated (OR = 3.99; 95%CI: 0.72-22.12; P = 0.11) with statistically significant heterogeneity (I^2 = 77%, P = 0.01) (Figure 4). Given significant heterogeneity, a sensitivity analysis was per-

formed by excluding the Perisetti *et al*[26] study which demonstrated a statistically significant more episodes of aspiration for those patients with prophylactically intubated as compared to those not intubated without significant heterogeneity (OR = 9.67; 95%CI: 3.40-27.52; $P < 0.01$; $I^2 = 0\%$, $P = 0.94$).

Publication bias

Publication bias was not observed in any outcomes in this meta-analysis based upon funnel plots.

DISCUSSION

In an effort to provide airway protection and reduce aspiration complications, providers may elect to perform tracheal intubation for patients presenting with UGIB. Unfortunately, there are no published guidelines to direct the use of endotracheal intubation in this group of patients, partly because of the lack of evidence-based recommendations. Emergent tracheal intubation is clearly indicated as a measure to protect airways in specific clinical presentations such as patients with altered mental status or those hemodynamically unstable. On the other hand, complications can arise directly from emergent tracheal intubations and the benefits of tracheal intubation should be weighed against the risks in each case individually. Schwartz *et al*[30] found that emergency intubation results in esophageal intubation in 8%, new pulmonary infiltrates identified post-intubation in 4%, and 3% died within 30 min of intubations, although those who died were those hemodynamically unstable before intubation. Only few studies evaluated this important subject and all were of retrospective design and varied in results[24-27].

Koch *et al*[25] evaluated the outcomes of 62 patients with 69 episodes of variceal bleeding who were either prophylactically intubated or not intubated prior to endoscopy and discovered significantly more aspiration in those who were prophylactically intubated. However, no differences were noted for mortality or length of stay[25]. Rehman *et al*[24] utilized 49 matched controls to 49 patients with UGIB and shock, cirrhosis, or hematemesis. Although cardiopulmonary complications are common in this population, no difference was discovered between the prophylactic intubation and no intubation in matched controls for mortality, length of stay, pneumonia, or aspiration[24]. Similarly, an abstract by Tang *et al*[27] in 69 patients with suspected variceal hemorrhage showed no significant differences between prophylactic intubation *vs* no prophylactic intubation for mortality, pneumonia, and length of stay. In contrast, an abstract by Perisetti *et al*[26] demonstrated that prophylactic intubation in patients with UGIB resulted in significantly more aspiration, length of stay, and mortality during hospitalization. Therefore, results has varied among the retrospective studies in regards to important outcomes such as aspiration, pneumonia, and mortality.

Due to this variability, we conducted this meta-analysis to evaluate the available evidence from four published retrospective studies that compared outcomes in UGIB patients who were prophylactically intubated and those who were not prophylactically intubated.

All studies evaluated mortality and aspiration[24-27], while only three studies evaluated pneumonia within 48 h as an outcome[24,25,27]. Our results show that there was a significant higher amount of pneumonia within 48 h in patients with UGIB who received endotracheal intubations prophylactically in comparison with those who were not intubated. In regards to aspiration and mortality, trends were noted toward worse outcomes in those patients who were prophylactically intubated but no statistically significant differences were noted. However, given significant heterogeneity, the sensitivity analyses demonstrated statistically significant worse outcomes for mortality and aspiration in those patients undergoing prophylactic intubation.

Strengths of our study are as follows. First, this is the first meta-analysis that evaluates outcomes difference between prophylactic intubation and no intubation in UGIB patients. Second, a large extensive search for relevant studies was conducted using several electronic search engines and three major gastroenterology and endoscopy conferences proceedings and abstracts for the period from 2004-2014. Third, each study included and evaluated at least two of the three outcomes studied in our meta-analysis. Four, the study populations were from different geographic areas in the United States and different time periods over 10 years, making it relevant to a large population. Fifth, no heterogeneity was identified in the pneumonia outcome. Finally, no publication bias was noted by the funnel plot. On the other hand, limitations were also apparent. First, a small number of studies were included. However, these studies are the only studies to-date on the subject. Second, all studies were observational with no randomized controlled trials on this issue which requires attention when forming conclusions from this meta-analysis and taken into consideration. Finally, significant heterogeneity was identified in two of the three outcomes (mortality and aspiration). Therefore, the DerSimonian and Laird model was utilized, limiting heterogeneity impact. Also, sensitivity analyses were performed which demonstrated statistically significant higher mortality and more aspiration in those patients undergoing prophylactic intubation. However, given the limited number of studies, subgroup analysis for sources of heterogeneity (such as location, timing, abstract exclusion) was not performed.

In conclusion, this meta-analysis demonstrates that patients with UGIB who received prophylactic endotracheal intubation have higher odds of having pneumonia within 48 h. Trends showing higher odds of mortality and aspiration in those prophylactically intubated were noted but no statistically significant differences were seen in comparison to those not intubated. Although these results must be interpreted with caution in light of the small number of studies in this meta-analysis leading to one or two studies having significant weight on the results, this

study addresses prophylactic intubation in UGIB patients prior to endoscopy. Based upon these results, prophylactic tracheal intubation is not beneficial in patients with UGIB and should not be recommended.

COMMENTS

Background

Endoscopic treatment is the main treatment for upper gastrointestinal bleeding (UGIB) and preventing complications during endoscopy is important. Endotracheal intubation might be used to protect airways and prevent aspiration, pneumonia, and reduce mortality. This study shows no evidence to support this practice generally and appropriateness of endotracheal intubation should be determined for each case individually.

Research frontiers

The authors performed the first meta-analysis comparing prophylactic endotracheal intubation to no intubation for UGIB to evaluate for pneumonia (within 48 h), aspiration, and mortality.

Innovations and breakthroughs

This is the first meta-analysis comparing prophylactic endotracheal intubation to no intubation for UGIB. The authors found that there is no evidence to support universal use of prophylactic endotracheal intubation prior to endoscopy. On the contrary, significantly more episodes of pneumonia occurred with the intubated group, and trends for worse aspiration and mortality were seen as well in the intubated group although not statistically significant.

Applications

Endotracheal intubation should be determined on an individual case-by-case approach when considered prior to endoscopy for UGIB treatment. Further studies, preferably randomized controlled trials, are likely needed to fully assess the practice of prophylactic intubation in UGIB patients prior to endoscopy.

Terminology

Odds ratio: Statistical term for the odds an event did or did not occur. Mean difference: Statistical term of difference between the means for a given variable. Heterogeneity: Test for uniformity in composition of studies included. Publication bias: Phenomenon where positive studies are more likely to be published than negative studies, leading to possible misrepresentation of data in meta-analysis.

Peer review

This is a very early systematic review and meta-analysis investigating the impact of prophylactic tracheal intubation on iatrogenic pneumonia, all-cause mortality and aspiration arising from complications due to endoscopy for upper GI bleeding.

REFERENCES

1 **Lewis JD**, Bilker WB, Brensinger C, Farrar JT, Strom BL. Hospitalization and mortality rates from peptic ulcer disease and GI bleeding in the 1990s: relationship to sales of nonsteroidal anti-inflammatory drugs and acid suppression medications. *Am J Gastroenterol* 2002; **97**: 2540-2549 [PMID: 12385436 DOI: 10.1111/j.1572-0241.2002.06037.x]

2 **Zaman A**, Goldberg RJ, Pettit KG, Kaniecki DJ, Benner K, Zacker C, DiCesare J, Helfand M. Cost of treating an episode of variceal bleeding in a VA setting. *Am J Gastroenterol* 2000; **95**: 1323-1330 [PMID: 10811347 DOI: 10.1111/j.1572-0241.2000.02020.x]

3 **Viviane A**, Alan BN. Estimates of costs of hospital stay for variceal and nonvariceal upper gastrointestinal bleeding in the United States. *Value Health* 2008; **11**: 1-3 [PMID: 18237354 DOI: 10.1111/j.1524-4733.2007.00208.x]

4 **Cryer BL**, Wilcox CM, Henk HJ, Zlateva G, Chen L, Zarotsky V. The economics of upper gastrointestinal bleeding in a US managed-care setting: a retrospective, claims-based analysis. *J Med Econ* 2010; **13**: 70-77 [PMID: 20047365 DOI: 10.3111/13696990903526676]

5 **Gleeson F**, Clarke E, Lennon J, MacMathuna R, Crowe J. Outcome of accident and emergency room triaged patients with low risk non-variceal upper gastrointestinal haemorrhage. *Ir Med J* 2006; **99**: 114-117 [PMID: 16972584]

6 **Laine L**, Yang H, Chang SC, Datto C. Trends for incidence of hospitalization and death due to GI complications in the United States from 2001 to 2009. *Am J Gastroenterol* 2012; **107**: 1190-1195; quiz 1196 [PMID: 22688850 DOI: 10.1038/ajg.2012.168]

7 **Zhao Y**, Encinosa W. Hospitalizations for Gastrointestinal Bleeding in 1998 and 2006: Statistical Brief #65. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs [Internet]. Rockville (MD): Agency for Health Care Policy and Research (US), 2006 Feb-2008 Dec

8 **Hreinsson JP**, Kalaitzakis E, Gudmundsson S, Björnsson ES. Upper gastrointestinal bleeding: incidence, etiology and outcomes in a population-based setting. *Scand J Gastroenterol* 2013; **48**: 439-447 [PMID: 23356751 DOI: 10.3109/00365521.2012.763174]

9 **Botianu A**, Matei D, Tantau M, Acalovschi M. Mortality and need of surgical treatment in acute upper gastrointestinal bleeding: a one year study in a tertiary center with a 24 hours / day-7 days / week endoscopy call. Has anything changed? *Chirurgia* (Bucur) 2013; **108**: 312-318 [PMID: 23790778]

10 **Lanas A**. Editorial: Upper GI bleeding-associated mortality: challenges to improving a resistant outcome. *Am J Gastroenterol* 2010; **105**: 90-92 [PMID: 20054306 DOI: 10.1038/ajg.2009.517]

11 **Charatcharoenwitthaya P**, Pausawasdi N, Laosanguaneak N, Bubthamala J, Tanwandee T, Leelakusolvong S. Characteristics and outcomes of acute upper gastrointestinal bleeding after therapeutic endoscopy in the elderly. *World J Gastroenterol* 2011; **17**: 3724-3732 [PMID: 21990954 DOI: 10.3748/wjg.v17.i32.3724]

12 **Marmo R**, Koch M, Cipolletta L, Bianco MA, Grossi E, Rotondano G. Predicting mortality in patients with in-hospital nonvariceal upper GI bleeding: a prospective, multicenter database study. *Gastrointest Endosc* 2014; **79**: 741-749.e1 [PMID: 24219820 DOI: 10.1016/j.gie.2013.10.009]

13 **Sostres C**, Lanas A. Epidemiology and demographics of upper gastrointestinal bleeding: prevalence, incidence, and mortality. *Gastrointest Endosc Clin N Am* 2011; **21**: 567-581 [PMID: 21944411 DOI: 10.1016/j.giec.2011.07.004]

14 **Balaban DV**, Strâmbu V, Florea BG, Cazan AR, Brătucu M, Jinga M. Predictors for in-hospital mortality and need for clinical intervention in upper GI bleeding: a 5-year observational study. *Chirurgia* (Bucur) 2014; **109**: 48-54 [PMID: 24524470]

15 **Tammaro L**, Buda A, Di Paolo MC, Zullo A, Hassan C, Riccio E, Vassallo R, Caserta L, Anderloni A, Natali A. A simplified clinical risk score predicts the need for early endoscopy in non-variceal upper gastrointestinal bleeding. *Dig Liver Dis* 2014; **46**: 783-787 [PMID: 24953205 DOI: 10.1016/j.dld.2014.05.006]

16 **Wang CH**, Chen YW, Young YR, Yang CJ, Chen IC. A prospective comparison of 3 scoring systems in upper gastrointestinal bleeding. *Am J Emerg Med* 2013; **31**: 775-778 [PMID: 23465874 DOI: 10.1016/j.ajem.2013.01.007]

17 **Clarke MG**, Bunting D, Smart NJ, Lowes J, Mitchell SJ. The surgical management of acute upper gastrointestinal bleeding: a 12-year experience. *Int J Surg* 2010; **8**: 377-380 [PMID: 20538082 DOI: 10.1016/j.ijsu.2010.05.008]

18 **Beggs AD**, Dilworth MP, Powell SL, Atherton H, Griffiths EA. A systematic review of transarterial embolization versus emergency surgery in treatment of major nonvariceal upper gastrointestinal bleeding. *Clin Exp Gastroenterol* 2014; **7**: 93-104 [PMID: 24790465 DOI: 10.2147/CEG.S56725]

19 **Targownik LE**, Murthy S, Keyvani L, Leeson S. The role of rapid endoscopy for high-risk patients with acute nonvariceal upper gastrointestinal bleeding. *Can J Gastroenterol* 2007; **21**: 425-429 [PMID: 17637943]

20 **Thibodeau LG**, Verdile VP, Bartfield JM. Incidence of aspiration after urgent intubation. *Am J Emerg Med* 1997; **15**: 562-565 [PMID: 9337361]

21 **Waye JD**. Intubation and sedation in patients who have

emergency upper GI endoscopy for GI bleeding. *Gastrointest Endosc* 2000; **51**: 768-771 [PMID: 10840326]

22 **Prather AD**, Smith TR, Poletto DM, Tavora F, Chung JH, Nallamshetty L, Hazelton TR, Rojas CA. Aspiration-related lung diseases. *J Thorac Imaging* 2014; **29**: 304-309 [PMID: 24911122 DOI: 10.1097/RTI.0000000000000092]

23 **Rudolph SJ**, Landsverk BK, Freeman ML. Endotracheal intubation for airway protection during endoscopy for severe upper GI hemorrhage. *Gastrointest Endosc* 2003; **57**: 58-61 [PMID: 12518132 DOI: 10.1067/mge.2003.46]

24 **Rehman A**, Iscimen R, Yilmaz M, Khan H, Belsher J, Gomez JF, Hanson AC, Afessa B, Baron TH, Gajic O. Prophylactic endotracheal intubation in critically ill patients undergoing endoscopy for upper GI hemorrhage. *Gastrointest Endosc* 2009; **69**: e55-e59 [PMID: 19481643 DOI: 10.1016/j.gie.2009.03.002]

25 **Koch DG**, Arguedas MR, Fallon MB. Risk of aspiration pneumonia in suspected variceal hemorrhage: the value of prophylactic endotracheal intubation prior to endoscopy. *Dig Dis Sci* 2007; **52**: 2225-2228 [PMID: 17385037 DOI: 10.1007/s10620-006-9616-0]

26 **Perisetti A**, Khan H, Sahmoun A, Newman W MR. Role of prophylactic pre-esophagogastroduodenoscopy (EGD) endotracheal intubation (ETI) in upper gastrointestinal bleed (UGIB). A retrospective study. *Am J Gastroenterol* 2013; **108**: S15-S16 (Abstract)

27 **Tang Y**, Wang Y, Wang WW. Elective endotracheal intubation prior to emergent EGD in patients with suspected variceal hemorrhage: An evaluation of outcome and complications. *Gastrointest Endosc* 2014; **79**: AB515-AB516 (Abstract)

28 **Stroup DF**, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008-2012 [PMID: 10789670]

29 **Armijo-Olivo S**, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract* 2012; **18**: 12-18 [PMID: 20698919 DOI: 10.1111/j.1365-2753.2010.01516.x]

30 **Schwartz DE**, Matthay MA, Cohen NH. Death and other complications of emergency airway management in critically ill adults. A prospective investigation of 297 tracheal intubations. *Anesthesiology* 1995; **82**: 367-376 [PMID: 7856895]

P- Reviewer: Goral V, Jonaitis L, Mentes O, O'Rorke M
S- Editor: Song XX **L- Editor**: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.11

World J Meta-Anal 2015 February 26; 3(1): 11-19
ISSN 2308-3840 (online)

Risk of infectious diseases and cutaneous tumours in solid organ recipients: A meta-analysis of literature

Paola Savoia, Giovanni Cavaliere, Paolo Fava

Paola Savoia, Giovanni Cavaliere, Paolo Fava, Department of Medical Science, University of Turin, 10126 Torino, Italy
Author contributions: Savoia P and Cavaliere G contributed equally to this work; Savoia P and Cavaliere G designed the research; Savoia P, Cavaliere G and Fava P performed the research, analysed the data and wrote the paper.

Correspondence to: Paola Savoia, MD, Department of Medical Science, University of Turin, Via Cherasco 23, 10126 Torino, Italy. paola.savoia@unito.it
Telephone: +39-11-6335849
Fax: + 39-11-674034
Received: July 22, 2014
Peer-review started: July 23, 2014
First decision: August 28, 2014
Revised: September 8, 2014
Accepted: November 17, 2014
Article in press: November 19, 2014
Published online: February 26, 2015

Abstract

AIM: To compare the risk of cutaneous infections and tumours in kidney transplant recipients with data recently published about this topic.

METHODS: In the present work, we evaluated the incidence of bacterial, fungal and viral cutaneous infectious diseases and the development of skin cancers in a cohort of 436 patients who undergone a renal transplantation. The median age at transplantation of our patients was 50 years and the median duration of the immunosuppression was of 7.2 years. Data obtained from our cohort were compared with those obtained by a systematic review of the literature of the last 20 years about the same topic.

RESULTS: Infectious diseases were the most frequent dermatological disorders that were diagnosed after transplantation, affecting about the 16.5% of patients. Herpes virus reactivation occurs in about the 35% of patients and is more common within 6 mo from transplantation, whereas when the immunosuppression is reduced, skin infections are mainly represented by Human Papilloma Virus infections and localized mycosis, such as pityriasis versicolor and superficial candidiasis. Bacterial infections were relatively rare and occur mainly in the first months after transplantation. The cumulative risk to develop skin cancer enhance significantly over the time, as consequence of long-term immunosuppressive regiments. Endogenous and exogenous risk factors, as well as the schedule of immunosuppression can play a role and justify the different incidence of skin cancer in the various series.

CONCLUSION: Skin infections and cancer, commonly diagnosed in transplanted patients, impact on survival and life-quality, justifying the realization of follow-up programs for the early diagnosis and treatment.

Key words: Skin infectious disease; Cutaneous tumours; Transplantations; Risk; Solid organ recipients

Core tip: Patients who underwent solid organ transplantation frequently suffer from skin infections and malignancies, due to the effects of long-term immunosuppressive therapy. Here, we compare our data about the risk to develop infectious disease and non-melanoma skin cancer in solid organ transplantation recipients, together with a meta-analysis of data recently reported

by literature about this topic.

Savoia P, Cavaliere G, Fava P. Risk of infectious diseases and cutaneous tumours in solid organ recipients: A meta-analysis of literature. *World J Meta-Anal* 2015; 3(1): 11-19 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/11.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.11

INTRODUCTION

It has been shown that patients receiving solid organ transplants have an increased risk of developing cutaneous infectious disease and skin tumours, as consequence of the long-term immunosuppressive treatment[1-3]. Infective complications are an important cause of morbidity and mortality, and the introduction of potent immunosuppressive agents like tacrolimus or mycophenolate mophetil may results in an increased risk of bacteric, fungine or viral infections[4]. The risk to develop skin cancer increase over the time and the various immunosuppressive regimens has different oncogenetic potential: the risk to develop a skin cancer is mainly related with azathioprine and cyclosporine[5], whereas mammalian target of rapamycin inhibitors have been associated with a lower incidence of *de novo* skin cancer[6]. However, the risk could also depend by other endogenous and exogenous conditions: history of sunburns and ultraviolet (UV) exposure, life habits, skin phototype, concomitant infections and specific genetic signatures can have a major impact in the onset and progression of these specific tumours[6-8]. In particular, different oncogenic and non-oncogenic Human papilloma virus (HPV) strains are frequently isolated from both normal skin and cutaneous tumours in transplant recipients, but their carcinogenetic role should be definitively established.

In this study, we report data about infective skin diseases and cutaneous tumours in a group of 436 renal transplant recipients followed-up at our centre. We also provide an overview and meta-analysis of data published in the recent literature about this topic.

MATERIALS AND METHODS

Data about 436 renal transplant recipients with a dermatological follow-up at our centre were recorded. The 61.3% of these were males and the 38.7% females; median age at transplantation was 50 years and the median duration of immunosuppression was 7.2 years. For each patient, we evaluated the presence of any infectious dermatological disease and the development of skin cancers.

Moreover, a review of the English language literature of the last 20 years was performed using the MEDLINE database, using the key words "infectious skin diseases", "cutaneous tumours" and "solid organ recipients". We included only peer reviewed series with more than 50 cases; single case reports and series with less than 50 cases were excluded as well as articles published on journals without peer review system. No restrictions on the basis of ethnicity were applied.

Fifty-two papers were considered for the analysis.

Statistical analysis was performed with SPSS software (SPSS, Chicago, IL) and with Kaplan-Mayer curves.

RESULTS

Infective disease

Viral, bacterial and fungal diseases were frequently reported in almost all the solid organ recipients cohort published in literature. In our cases, infectious diseases were the most frequent dermatological disorders who were diagnosed after transplantation, affecting the 16.7% of patients.

Viral infections

Herpes simplex virus (HSV) infections are relatively frequent in organ transplant recipients. Infections with reactivated HSV occur with an incidence of up to 35% primarily in the first three weeks following transplantation[9]. Marrow transplant patients are most at risk, but also solid organ transplant recipients show an higher incidence of HSV infections than immunocompetent people especially when preventive antiviral treatment was not performed[10]. There are very different incidence rates of HSV infections in literature depending on the type of immunosuppressive treatment, the geographical area considered and the mean time from transplantation (Table 1). However, there are no remarkable differences when considering the type of organ transplanted. We found a prevalence of HSV recurrent infections (2.4%) similar to those reported by Bakr *et al*[11] and Belloni-Fortina *et al*[12].

Human herpesvirus 6 and 7 (HHV-6 and HHV-7), ubiquitous in humans, cause exanthema subitum in childhood and remain in a latent form in the body after primary infection. Two to three weeks following transplantation up to 30% of all transplant recipients have a reactivation of HHV-6 even if most infections remain asymptomatic[13].

Primary or recurrent varicella zoster virus (VZV) infections can occur in 1%-30% of solid organ transplant recipients with a mean time of onset from transplantation of 9-23 mo and a peak after 6 mo[14]. As it can be seen in Table 1 some authors report lower incidence rates of VZV infections probably because herpetic eruptions develop more commonly during the first year after transplantation[15] and the mean time since transplantation was lower than 1 year[11,12]. Cito megalo virus (CMV), another member of Herpesvirus, is found in 50%-75% of solid organ transplant recipients. CMV rarely causes cutaneous infections but can facilitate other opportunistic skin infections by modulating cell-mediated immunity[14].

Viral warts and condiloma acuminata are clinical expression of HPV infection. Viral warts are frequent in long-term immunosuppressed patients with prevalence rates

Table 1 Incidence of viral infections in solid organ recipients *n* (%)

Ref.	No. of cases/ population	Age (yr)	M/F	Median follow-up time	Therapy	HSV infections	VZV infections	HPV infections
Greenberg *et al*[53]	68/Kidney	NA	NA	NA	NA	10 (14.7)	NA	NA
Hogewoning *et al*[15]	134/Kidney	32.6 ± 10.3	80/54	NA	NA	9 (6.7)	24 (17.9)	NA
Bakr *et al*[11]	302/Kidney	35.9 ± 11.3	216/86	0.25-23 yr	Miscellaneous	9 (3)	3 (1)	33 (10.9)
Savoia *et al*[65]	286/Kidney	NA	273/163	9.3 (0.1-39.8) yr	66.7% combination therapy (tacrolimus in 90%)	11 (2.5)	NA	45 (10.3)
Belloni-Fortina *et al*[12]	161/Liver	47.4 ± 11	116/45	NA	NA	3 (2)	3 (2)	30 (19)
Lima *et al*[22]	53/Kidney	44	28/25	52.5% of patients over 5 yr	44% prednisone 31% mycophenolate	3 (5.7)	4 (7.5)	14 (26.4)

For each literature study has been indicated number of included patients, type of solid organ transplantation, median age, sex, median follow-up time after solid organ transplantation, type of immunosuppressive treatment and the percentage of patients affected by herpes simplex virus (HSV), varicella zoster virus (VZV) and human papilloma virus (HPV) infections. NA: Data non available in the considered study; M: Male; F: Female.

ranging from 35% and 85% 5 years after transplantation[15,16]. We found a prevalence of 12.2% similar to those reported in different studies conducted on other kidney and liver transplant patients[11,12].

Viral warts usually develop on sun-exposed areas, especially in fairer skin-type patients. They are usually multiple and display fewer tendencies for spontaneous regression than in immunocompetent individuals. Their extension may be so widespread to constitute general verrucosis. The types of human HPV found in organ transplant recipients may be different from that seen in the general population. In a study, nine of 10 HPV detected in organ transplant recipients were gamma-PV and one belonged to the genus beta-PV[17]. Other authors report that the most frequent HPV types are HPV-5 and HPV-8, *i.e.*, the same types that can be easily found in epidermodysplasia verruciforme (EV)[18].

Bacterial infections

In the first month by transplantation there is an high frequency of generally trivial nosocomial diseases. The frequent wound infections that can be seen in this period are increasingly been caused by antibiotic resistant strains [vancomycin-resistant enterococci and methicillin-resistant Staphilococcus Aureus (MRSA)][11,19].

In immunosuppressed individuals Staphylococcus aureus infections manifest frequently as pyoderma. However, subcutaneous abscesses, erysipelas, and impetigo may develop in the long term[14].

Interestingly a prospective study on 604 heart transplant patients report an high prevalence of infections in the first year from transplantation with a majority of bacterial infections[20], probably as a consequence of higher dosage of immunosuppressive treatment in order to avoid acute rejections. However, when considering only skin infections, the prevalence was similar to those reported by Perera *et al*[21] and Lima *et al*[22]. When confronted to other reports (Table 2), bacterial infections were relatively rare in our experience and occurred only in 1.4% of the patients.

Necrotizing fascitis (NF) is a devastating infectious disease with 0.04 cases per 1000 person-years in the general population. The mortality rate is 25% to 30% and the most common pathogen in type Ⅱ NF is Streptococcus pyogenes[23]. The characteristics of NF in renal transplant patients are poorly understood due to the rarity of NF in this population. To date, there have only been described 12 cases[24]. When comparing with NF in immunocompetent individuals, fungal etiology appears more common but, surprisingly, the overall mortality rate is lower (16.7% *vs* 25%-30%). Age and use of mycophenolate are associated with an increased risk of death[24].

Nocardiosis is a rare opportunistic infection caused by aerobic Actinomycetes Nocardia and can be associated with severe complications in kidney transplant recipients. Studies showed, in the last 2 decades, that the incidence of Nocardia infection in kidney transplant recipients was approximately 0.4%-1.3%[25]. To date, more than 70 cases of Nocardiosis in renal transplant recipients have been described. Nocardiosis appears after a mean time of 34.1 mo from transplantation and is more frequent in patients with a prior history of acute rejection and in treatment with cyclosporine. Lung, brain, skin, and subcutaneous tissue were the most frequently involved organs[26]. The mortality rate varies between 16.67%[26] and 25%[27].

Although the incidence of tuberculosis in renal transplant recipients is 5 times higher than in the general population tuberculosis is still rare in organ transplant recipients with reported rates of 0.35%-15% depending on the geographical area considered[19]. Among infected transplant recipients, 63% have a pulmonary involvement, 25% have systemic dissemination and 12% have an exclusively extrapulmonary involvement[28].

Skin involvement is generally a sign of disseminated tuberculosis and imposes the research of a visceral involvement. Only 18 cases of cutaneous miliary tuberculosis in patients older than 15 have been described in literature from 1889 and 1991[29-31].

Atypical mycobacterioses are rarer than M. tuberculosis infections and are seen in 0.16%-2.8% of solid organ transplant recipients[32]. Among them, some sporadic cases of infections by *M. Abscessus* and *M. Marinum* are reported in literature[33,34].

Table 2 Incidence of bacterial infections in solid organ recipients *n* (%)

Ref.	No. of cases/ population	Age (yr)	M/F	Median follow-up time	Therapy	Bacterial infections
Bakr *et al*[11]	302/Kidney	35.9 ± 11.3	216/86	0.25-23 yr	Miscellaneous	47 (16) Mainly folliculitis and impetigo
Hogewoning *et al*[15]	134/Kidney	32.6 ± 10.3	80/54	NA	NA	28 (20.9) Mainly folliculitis, ectyma and erysipelas (3)
Alangaden *et al*[66]	127/Liver	47 ± 12	79/62	NA	79% prednisone 72% tacrolimus 28% sirolimus	17 (13) Mainly wound infections and skin and soft tissue infections
Perera *et al*[21]	100/Liver	42.5	NA	5.5 (0.75-16) yr	35% cyclosporine, azathioprine and prednisone 48% prednisone and tacrolimus 17% tacrolimus	5 (5) Mainly folliculitis and 1 case of erythrasma
Sánchez-Lázaro *et al*[20]	604/Heart	51	506/98	First year after transplantation	NA	36 (5.9)
Savoia *et al*[65]	286/Kidney	NA	273/163	9.3 (0.1-39.8) yr	66.7% combination therapy comprising tacrolimus in 90%	6 (1.4) All cases of erysipelas
Lima *et al*[22]	53/Kidney	44	28/25	52.5% of patients over 5 yr	44% prednisone 31% mycophenolate	3 (5.7) 2 cases of furuncle and 1 cellulitis

For each literature study has been indicated number of included patients, type of solid organ transplantation, median age, sex, median follow-up time after solid organ transplantation, type of immunosuppressive treatment and the percentage of patients affected by bacterial infections; when available, type of bacterial infection has been specified. NA: Data non available in the considered study; M: Male; F: Female.

Table 3 Incidence of fungal infections in solid organ transplant recipients *n* (%)

Ref.	No. of cases/ population	Age (yr)	M/F	Median follow-up time	Therapy	*Candida spp*	*Malassezia furfur*	*Dermatophytes*
Virgili *et al*[36]	73/Kidney	22-68	44/29	0.25-26 yr	50.7% association of prednisone, cyclosporine and azathioprine	4 (5.4)	20 (27.4)	7 (9.6)
Güleç *et al*[37]	102/Kidney	31.9 ± 10.3	68/34	4.5 ± 4.55 yr	38.2% association of prednisone, mycophenolate and cyclosporine	31 (30.4)	37 (36.3)	10 (9.8)
Perera *et al*[21]	100/Liver	42.5	NA	5.5 (0.75-16) yr	48% association of prednisone and tacrolimus	19%	4%	11%
Lima *et al*[22]	53/Kidney	44	28/25	NA	83% prednisone 58.5% mycophenolate 50.9% cyclosporine	14 (22.6)	9 (17)	8 (15)

For each literature study has been indicated number of included patients, type of solid organ transplantation, median age, sex, median follow-up time after solid organ transplantation, type of immunosuppressive treatment and the percentage of patients affected by fungal infections. NA: Data non available in the considered study; M: Male; F: Female.

Fungal Infections

Among superficial fungal infection, candidiasis of the mouth and intertriginous skin areas is frequent in the early post-transplant time[35], probably as a consequence of the higher dosage of immunosuppressant treatment in this period (Tables 3 and 4).

Pityriasis versicolor (PV) is such as frequent as superficial candidiasis. Some authors reported prevalence rates of this infection caused by *Malassezia* furfur higher than 30% in cohorts of renal transplant patients[36,37]. Whereas, there are very few literature reports about the prevalence of PV in other solid organ transplant recipients. Perera *et al*[21] report a prevalence ratio of PV of 4% in a group of liver transplant recipient. In our cohort, mycosis, mainly represented by onychomicosis, tinea cruris and genital candidiasis, were observed in the 1.8% of cases.

Deep fungal infections comprise two distinct group of conditions, the subcutaneous and systemic mycoses. Subcutaneous mycoses are caused by fungi that have been introduced directly into the skin through a penetrating injury[36,38]. Systemic dissemination is rare in the immunocompetent patients but could be more frequent in immunosuppressed subjects. Sporotrichosis, mycetoma and chromoblastomycosis are the most frequent subcutaneous infections observed in this group of patients.

Systemic mycoses are fungal infections whose initial portal entry into the body is usually a deep site (*e.g.*, lung and gastrointestinal tract). Skin is usually affected as consequence of systemic dissemination but it may be the primary site in the immunocompromised patients that usually develop systemic candidiasis, aspergillosis, histoplasmosis and cryptococcosis[39-42]. Incidence rates of

Table 4 Incidence of fungal infections in solid organ transplant recipients *n* (%)

Ref.	No. of cases/ population	Age (yr)	M/F	Median follow-up time	Therapy	*Sistemic* infections	*Candida spp*	*Aspergillus spp*
Collins *et al*[45]	158/Liver	46	84/74	NA	Cyclosporine, azathioprine and prednisone	34 (21.5)	28 (17.7)	5 (3.2)
Briegel *et al*[44]	141/Liver	47 ± 12	79/62	NA	Prednisone, cyclosporine and azathioprine	25 (17.7)	10 (7)	11 (7.8)
Kanj *et al*[46]	73/Heart-Lung	NA	NA	NA	NA	37 (50.6)	19 (26)	18 (24.6)
Abbott *et al*[47]	33479/Kidney	43	20154/13325	NA	72.2% with cyclosporine, 65.2% with mycophenolate	595 (1.7)	445 (1.3)	80 (0.2)
Singh *et al*[39]	130/Liver	NA	NA	NA	tacrolimus	11 (14)	6 (5)	4 (3)
Alangaden *et al*[66]	127/Kidney	47.1 ± 12.5	76/51	NA	72% tacrolimus	5 (3.9)	5 (3.9)	NA
Pugliese *et al*[67]	278/Miscellaneous	NA	NA	5.5 ± 5.9 yr	Various	46 (16.5)	45 (16.2)	1 (0.3)
Tessari *et al*[43]	3293/Miscellaneous	NA	2384/909	NA	NA	22 (0.7)	NA	NA

For each literature study has been indicated number of included patients, type of solid organ transplantation, median age, sex, median follow-up time after solid organ transplantation, type of immunosuppressive treatment and the percentage of patients affected by systemic fungal infections. NA: Data non available in the considered study; M: Male; F: Female.

deep fungal infections in solid organ transplant recipients varies from 0.5% to 30%[39,40]. These studies, however, do not clearly distinguish between primary deep skin mycoses and systemic infections. In an Italian series of 3293 consecutive organ transplant recipients with a mean follow-up time since transplantation of 2.5 ± 2 years, only 22 cases of deep mycoses were detected with a prevalence ratio of 0.7%. Six patients had subsequent systemic involvement and three died of systemic dissemination[43]. In a US study conducted on 130 liver transplant patients the authors found 6 cases of systemic candidiasis and 4 of aspergillosis[19]. Other older series conducted on liver transplant patients and exclusively based on detection of systemic mycosis found higher rates of *Candida* and *Aspergillus* infections[44,45]. This higher incidence could probably derive by an higher use of cyclosporine and azathioprine as tacrolimus and sirolimus weren't still commonly used until the late 90 s.

When considering lung transplant recipients, invasive fungal infections occur in 15% to 35% of the patients with *Aspergillus* species accounting for nearly half of them[46-48]. The reported prevalence of *Candida* infections is similar[46].

On the other, hand kidney transplant patients seem to be less frequently affected by invasive fungal infections as reported in some United States series[38,47]. This incidence could be affected by a lower dose of immunosuppressive treatment and a higher use of tacrolimus instead of cyclosporine and azathioprine when confronted with lung and liver transplant recipient.

Cutaneous tumours

Data about the risk to develop non-melanoma skin cancer (NMSC) and the clinical characteristics of the various published series are resumed in Table 5.

The percentage of NMSCs diagnosed after a solid organ transplantation varied from 25% to 35% in the larger series published by literature[3,49,50]. The Basal cell carcinoma (BCC)/Squamous cell carcinoma (SCC) ratio was from 1:1.2 to 1:7[3,49-51]. Fekecs *et al*[52] report a significantly lower percentage of solid organ recipients affected by NMSC (9.5%), but in this study there is a bias due to the relatively short follow-up period.

The 23.5% of our patients developed a NMSCs in the post-transplant period, with a BCC/SCC ratio of 2.45:1. This percentage was similar to those reported in our previous work, conducted on smaller series[53]. Fifty-four per cent of BCCs and 81% of SCCs develop on sun-exposed areas. Patients who developed skin cancers were preferentially males ($P = 0.0017$) and were characterized by a significantly higher age at transplantation ($P < 0.001$) and by a significantly longer duration of immunosuppressive regimen ($P < 0.0001$), according with data reported by others authors[3,50,54]. Also elderly patients[51] showed a higher risk to develop cutaneous tumours. In our experience, exogenous risk factors significantly linked to NMSC risk were outdoor job ($P = 0.0413$), as well as demonstrated in others series[52,53], and incorrect use of sunscreen ($P = 0.0252$). We failed to demonstrate a significant association between lower phototypes and risk of NMSC, as demonstrated by several literature series[3,50,51,53].

In the majority of published studies, cyclosporine and/or azatyoprine-based immunosuppressive regiments showed a significant correlation with the risk of developing skin cancer[3,49,51,52]. On the contrary, we could not identify a specific immunosuppressive drug as a distinctive factor for the development of NMSC.

DISCUSSION

Organ transplantation ensures a prolonged life expectancy and a better quality of life for patients affected by chronic renal, liver, lung or hearth failure. However, long-term immunosuppressive therapy causes important inhibitory effects on immune defence mechanism, leading to frequent skin infections and malignancies that are an important cause of morbidity and mortality for solid organ transplant recipients[1-3].

The schedule of immunosuppressive drugs influences

Table 5 Risk to develop non-melanoma skin cancer and clinical characteristics of the various published series

Ref.	No. of cases/ population	NMSC	BCC/SCC ratio	Median age at transplantation	Median follow-up time	Risk factors associated with NMSC
España *et al*[54]	92/Heart	15.2%	1:1.5	NA	NA	Immunosuppression UV exposure Skin type
Ong *et al*[68]	455/Heart Australia	31%	3:1	NA	NA	Caucasian origin Age at transplantation Duration of follow up
Hiesse *et al*[5]	1710/Kidney France	7.5%-8.2%	NA	35.5 yr	9 yr	Cyclosporin
Moloney *et al*[8]	1755/Kidney Irland	27.7%	1:2	40 yr	5.35 yr	Age at transplantation Duration of immunosuppression
Mackenzie *et al*[49]	384/Kidney New Zeland	25%	1:1.2	41.5 yr	5.3 yr (0.01-33.4)	Age at transplantation Male sex Cyclosporine/Azathioprine
Sandoval *et al*[63]	91/Kidney Chile	16%	1:1-9	NA	7.3 yr (1 mo-29 yr)	Duration of immunosuppression
Fekecs *et al*[52]	116/Kidney, pancreas Hungary	9.5%	1:4	49.3 yr	NA	Painful sunburns Occupational UV exposure Cyclosporine

For each literature study has been indicated number of included patients, type of solid organ transplantation, median age at transplantation, median follow-up time after solid organ transplantation, percentage of patients affected by NMSC and BCC/SCC ratio. Risk factors significantly associated to the development of non-melanoma skin cancer in each study are indicated in the right column. NA: Data non available in the considered study; UV: Ultraviolet; NMSC: Non-melanoma skin cancer; BCC: Basal cell carcinoma; SCC: Squamous cell carcinoma.

the type and the timing of skin disease. The main problems in the first moths are usually represented by wound infections and HSV reactivations, whereas opportunistic infections and herpes zoster develop mainly within 6 mo from transplantation. Thereafter, as immunosuppression is reduced, the more frequently observed skin infections are represented by mycoses and HPV infections[55]. On the contrary, the risk to develop skin cancer increase over the time: the cumulative incidence of skin cancers enhance significantly with the duration of graft, increasing from 5% after 1 year to 43% after 10 years, as demonstrated in several European series[51,52]. In the kidney recipients from our series the median time to onset of skin tumours was 9.9 years from the transplantation.

Moreover, tacrolimus and micophenolate mophetyl are mainly related to the risk to develop skin infections, whereas the higher carcinogenetic risk has been described for azathioprine and cyclosporine[3,6,49,51]. The oncogenic power of cyclosporine in solid organ recipients was confirmed by a large retrospective study[5] that demonstrated a risk of skin cancer significantly higher in the group of CyA-treated patients in comparison with the historical group of patients treated with azatioprine-steroids regimens. Moreover, it has recently been demonstrated that azathioprine induces chronic oxidative stress by forming reactive oxygen species (ROS) causing mutagenic damage of the DNA, that could led to development of NMSC in organ transplantation recipients.

In literature[55-57], the frequency of HPV infections in transplant recipient varies from 6% to 92%, depending on the type and the duration of the immunosuppressive protocol. We observed viral warts in 10.3% of patients from our series, a percentage superimposable to that of 8.2% recently reported in another Italian study[58], probably due to the similarity in the immunosuppressive treatment schedules. Despite some investigations demonstrated that persistent HPV infections can induce malignant transformation of squamous epithelial cells by inactivation of p53, and clinical and histological analyses show progression of viral warts *via* dysplastic lesions up to invasive squamous cell carcinomas, the pathogenic role of HPV in skin tumorigenesis is still in part unclear[59]. With the use of PCR methods, a prevalence of HPV in 69%-88% of squamous cell carcinoma in transplant recipients was found, in particular high-risk HPV types like HPV-16 and epidermodysplasia verruciformis associated HPV types. The prevalence in organ transplant recipients is significantly higher in comparison to immunocompetent patients (about 50%)[60]. On the other side, there were no significant differences of HPV prevalence in basal cell carcinoma between immunocompromised and immunocompetent individuals[1].

Herpes zoster was diagnosed in 2.1% of our patients; this percentage is relatively low in comparison with data reported by other authors[55]. However, no significant differences from other series were found when data were stratified on the basis of different age groups. In fact, Herpes zoster affects essentially patients over 60 years, whereas median age of our population was 50 years. In transplanted patients, HSV and HVZ usually provoke limited infections but can also generate diffuse, hemorrhagic, ulcerated and widespread skin lesions more frequently than in immunocompetent individuals[61]. Also visceral implication are not rare.

When confronted to other reports, bacterial infections were relatively rare in our experience. This could be

considered a consequence of an higher mean time from transplantation in our cohort, as it has been seen that bacterial infections develop more frequently in the first month from transplantation. Moreover we didn't consider folliculitis because they were all of minor entity and we believed that they were more associated to chronic use of steroid rather than to bacterial infections.

A wide variation (7%-75%) in the frequency of superficial fungal infections is reported in several studies; literature data suggest that cutaneous fungal infections in renal transplant recipients are more common in tropical and sub-tropical countries[37]. However, different authors report similar prevalences of dermatophytosis in immunosuppressed and immunocompetent people. Probably that could derive by the necessity of the coexistence of an environmental exposure to pathogenic fungi together with the administration of immunosuppressive agents[37]. Also in our experience, the incidence of superficial fungal infections was low, and only 3 cases of onichomycosis (1.1%) were identified. Systemic fungal infections occur in the 5%-20% of solid organ recipients, mainly caused by *Candida* or *Aspergillus*[55].

The problem about increased risk of skin cancer in solid organ transplant recipients is well known in literature. In particular, it has been estimated a 10-fold increased risk for BCC and a 50-100-fold for SCC. In our experience, the percentage of patients who developed NMSC was 24.8%. This percentage and the BCC/SCC ratio were similar to those reported in recent studies conducted in Italy[7,58] and Spain[50,62] (22% and 25.2%, respectively), probably due to the similarity in skin phenotype, exogenous risk factors exposure and in the immunosuppressive treatment schedule[49,63]. On the contrary, the prevalence of skin cancers in a group of Australian kidney transplant recipients was significantly higher (35%), supporting the importance of latitude and sun exposure on tumour development[3]. Moreover, differences in the median age at transplantation in the various series could partially justify the variability in the percentage of patients that develop a NMSC. Higher age at transplantation is in fact a factor strictly related to the risk of skin cancer in the majority of published series[3,49,50,64]. The length of follow-up could also represent a bias in the different series; the majority of the authors state in fact that the risk to develop cutaneous tumours increase over the time, as the consequence of the longer immunosuppression period[8].

In conclusion, solid organ transplant recipients today have a prolonged life expectancy and a better quality of life. However, cutaneous infections and NMSCs can heavily impact on the quality of life and prognosis of these patients. For this reason it is necessary to perform periodical accurate dermatological controls in order to promptly identify any suspicious lesions. Individual follow-up programs should be realized on the basis of specific risk factor analysis, to optimize the cost-benefit ratio.

COMMENTS

Background

Cutaneous disorders are frequent in chronic renal failure. The majority of these dermatological disorders disappear after kidney transplantation; however, infectious diseases and cutaneous malignancies occur frequently in organ transplant recipients, mainly as a consequence of the long-term immunosuppressive treatment. Infectious skin diseases were frequently diagnosed after transplantation, affecting about the 16.5% of patients whereas dermatological screening identify cutaneous tumours in about 35% of KTR patients. The relative risk of developing skin cancer is 20 to 40 fold increased, in comparison with the general population.

Research frontiers

Type and duration of the immunosuppressive treatment are currently considered as the major factors related to the development of infective and malignant skin lesions in patients receiving solid organ transplants. However other endogenous and exogenous risk factors can justify the different prevalence ratios reported in several literature studies.

Innovations and breakthroughs

Comparing data from the English language literature of the last 20 years with the results from our cohort of 436 kidney transplant recipients, the authors highlight the characteristics and risk factors for the different skin diseases occurring in transplant recipients.

Applications

Development of an integrated risk stratification protocol for skin diseases in transplant recipients with the aim of optimizing cost-benefit ratio of their treatment.

Peer review

The authors have performed a good study, the manuscript is interesting.

REFERENCES

1 **Stockfleth E**, Ulrich C, Meyer T, Arndt R, Christophers E. Skin diseases following organ transplantation--risk factors and new therapeutic approaches. *Transplant Proc* 2001; **33**: 1848-1853 [PMID: 11267539 DOI: 10.1016/S0041-1345(00)02743-3]

2 **Durando B**, Reichel J. The relative effects of different systemic immunosuppressives on skin cancer development in organ transplant patients. *Dermatol Ther* 2005; **18**: 1-11 [PMID: 15842607 DOI: 10.1111/j.1529-8019.2005.05007.x]

3 **Gallagher MP**, Kelly PJ, Jardine M, Perkovic V, Cass A, Craig JC, Eris J, Webster AC. Long-term cancer risk of immunosuppressive regimens after kidney transplantation. *J Am Soc Nephrol* 2010; **21**: 852-858 [PMID: 20431040 DOI: 10.1681/ASN.2009101043]

4 **Hwang EA**, Kang MJ, Han SY, Park SB, Kim HC. Viral infection following kidney transplantation: long-term follow-up in a single center. *Transplant Proc* 2004; **36**: 2118-2119 [PMID: 15518767 DOI: 10.1016/j.transproceed.2004.08.008]

5 **Hiesse C**, Rieu P, Kriaa F, Larue JR, Goupy C, Neyrat N, Charpentier B. Malignancy after renal transplantation: analysis of incidence and risk factors in 1700 patients followed during a 25-year period. *Transplant Proc* 1997; **29**: 831-833 [PMID: 9123545 DOI: 10.1016/S0041-1345(96)00153-4]

6 **Wisgerhof HC**, Edelbroek JR, de Fijter JW, Haasnoot GW, Claas FH, Willemze R, Bavinck JN. Subsequent squamous- and basal-cell carcinomas in kidney-transplant recipients after the first skin cancer: cumulative incidence and risk factors. *Transplantation* 2010; **89**: 1231-1238 [PMID: 20410852 DOI: 10.1097/TP.0b013e3181d84cdc]

7 **Naldi L**, Fortina AB, Lovati S, Barba A, Gotti E, Tessari G, Schena D, Diociaiuti A, Nanni G, La Parola IL, Masini C, Piaserico S, Peserico A, Cainelli T, Remuzzi G. Risk of nonmelanoma skin cancer in Italian organ transplant recipients. A registry-based study. *Transplantation* 2000; **70**: 1479-1484 [PMID: 11118094 DOI: 10.1097/00007890-20001127

0-00015]
8 **Moloney FJ**, Comber H, O'Lorcain P, O'Kelly P, Conlon PJ, Murphy GM. A population-based study of skin cancer incidence and prevalence in renal transplant recipients. *Br J Dermatol* 2006; **154**: 498-504 [PMID: 16445782]
9 **Schmied E**, Dufour JF, Euvrard S. Nontumoral dermatologic problems after liver transplantation. *Liver Transpl* 2004; **10**: 331-339 [PMID: 15004757 DOI: 10.1002/lt.20089]
10 **Slifkin M**, Doron S, Snydman DR. Viral prophylaxis in organ transplant patients. *Drugs* 2004; **64**: 2763-2792 [PMID: 15563248 DOI: 10.2165/00003495-200464240-00004]
11 **Bakr NI**, El-Sawy E, Hamdy AF, Bakr MA. Skin infections in Egyptian renal transplant recipients. *Transpl Infect Dis* 2011; **13**: 131-135 [PMID: 20849434 DOI: 10.1111/j.1399-3062.2010.00568.x]
12 **Belloni-Fortina A**, Piaserico S, Bordignon M, Gambato M, Senzolo M, Russo FP, Peserico A, De Matteis G, Perissinotto E, Cillo U, Vitale A, Alaibac M, Burra P. Skin cancer and other cutaneous disorders in liver transplant recipients. *Acta Derm Venereol* 2012; **92**: 411-415 [PMID: 22377797 DOI: 10.2340/00015555-1316]
13 **Le J**, Gantt S. Human herpesvirus 6, 7 and 8 in solid organ transplantation. *Am J Transplant* 2013; **13** Suppl 4: 128-137 [PMID: 23465006 DOI: 10.1111/ajt.12106]
14 **Ulrich C**, Christophers E, Sterry W, Meyer T, Stockfleth E. [Skin diseases in organ transplant patients]. *Hautarzt* 2002; **53**: 524-533 [PMID: 12221466 DOI: 10.1007/s00105-002-0358-4]
15 **Hogewoning AA**, Goettsch W, van Loveren H, de Fijter JW, Vermeer BJ, Bouwes Bavinck JN. Skin infections in renal transplant recipients. *Clin Transplant* 2001; **15**: 32-38 [PMID: 11168313 DOI: 10.1034/j.1399-0012.2001.150106.x]
16 **Krüger-Corcoran D**, Stockfleth E, Jürgensen JS, Maltusch A, Nindl I, Sterry W, Lange-Asschenfeldt B, Ulrich C. [Human papillomavirus-associated warts in organ transplant recipients. Incidence, risk factors, management]. *Hautarzt* 2010; **61**: 220-229 [PMID: 20165825 DOI: 10.1007/s00105-009-1860-8]
17 **Köhler A**, Gottschling M, Manning K, Lehmann MD, Schulz E, Krüger-Corcoran D, Stockfleth E, Nindl I. Genomic characterization of ten novel cutaneous human papillomaviruses from keratotic lesions of immunosuppressed patients. *J Gen Virol* 2011; **92**: 1585-1594 [PMID: 21471318 DOI: 10.1099/vir.0.030593-0]
18 **Majewski S**, Jablonska S. Do epidermodysplasia verruciformis human papillomaviruses contribute to malignant and benign epidermal proliferations? *Arch Dermatol* 2002; **138**: 649-654 [PMID: 12020228 DOI: 10.1001/archderm.138.5.649]
19 **Singh N**, Gayowski T, Wagener MM, Doyle H, Marino IR. Invasive fungal infections in liver transplant recipients receiving tacrolimus as the primary immunosuppressive agent. *Clin Infect Dis* 1997; **24**: 179-184 [PMID: 9114144]
20 **Sánchez-Lázaro IJ**, Almenar-Bonet L, Martínez-Dolz L, Buendía-Fuentes F, Agüero J, Navarro-Manchón J, Raso-Raso R, Salvador-Sanz A. Post-heart transplant tumors: chronology and impact on survival. *Transplant Proc* 2010; **42**: 3201-3203 [PMID: 20970651 DOI: 10.1016/j.transproceed.2010.05.052]
21 **Perera GK**, Child FJ, Heaton N, O'Grady J, Higgins EM. Skin lesions in adult liver transplant recipients: a study of 100 consecutive patients. *Br J Dermatol* 2006; **154**: 868-872 [PMID: 16634888 DOI: 10.1111/j.1365-2133.2006.07154.x]
22 **Lima AM**, Rocha SP, Reis Filho EG, Eid DR, Reis CM. Study of dermatoses in kidney transplant patients. *An Bras Dermatol* 2013; **88**: 361-367 [PMID: 23793196 DOI: 10.1590/abd1806-4841.20131859]
23 **Sarani B**, Strong M, Pascual J, Schwab CW. Necrotizing fasciitis: current concepts and review of the literature. *J Am Coll Surg* 2009; **208**: 279-288 [PMID: 19228540 DOI: 10.1016/j.jamcollsurg.2008.10.032]
24 **Tsai SF**. Necrotizing fasciitis in patients who underwent renal transplantation. *Transplant Proc* 2013; **45**: 2807-2810 [PMID: 23972528 DOI: 10.1016/j.transproceed.2013.02.142]
25 **Queipo-Zaragozá JA**, Broseta-Rico E, Alapont-Alacreu JM, Santos-Durantez M, Sánchez-Plumed J, Jiménez-Cruz JF. Nocardial infection in immunosuppressed kidney transplant recipients. *Scand J Urol Nephrol* 2004; **38**: 168-173 [PMID: 15204409 DOI: 10.1080/00365590410025353]
26 **Yu X**, Han F, Wu J, He Q, Peng W, Wang Y, Huang H, Li H, Wang R, Chen J. Nocardia infection in kidney transplant recipients: case report and analysis of 66 published cases. *Transpl Infect Dis* 2011; **13**: 385-391 [PMID: 21824241 DOI: 10.1111/j.1399-3062.2011.00607.x]
27 **Husain S**, McCurry K, Dauber J, Singh N, Kusne S. Nocardia infection in lung transplant recipients. *J Heart Lung Transplant* 2002; **21**: 354-359 [PMID: 11897524 DOI: 10.1016/S1053-2498(01)00394-1]
28 **Aguado JM**, Herrero JA, Gavaldá J, Torre-Cisneros J, Blanes M, Rufí G, Moreno A, Gurguí M, Hayek M, Lumbreras C, Cantarell C. Clinical presentation and outcome of tuberculosis in kidney, liver, and heart transplant recipients in Spain. Spanish Transplantation Infection Study Group, GESITRA. *Transplantation* 1997; **63**: 1278-1286 [PMID: 9158022 DOI: 10.1097/00007890-199705150-00015]
29 **Stein A**, Purgus R, Drancourt M, Olmer M. Photo quiz. Diagnosis: cutaneous miliary tuberculosis. *Clin Infect Dis* 1999; **29**: 1126-1127; quiz 1307-1308 [PMID: 10524951]
30 **del Giudice P**, Bernard E, Perrin C, Bernardin G, Fouché R, Boissy C, Durant J, Dellamonica P. Unusual cutaneous manifestations of miliary tuberculosis. *Clin Infect Dis* 2000; **30**: 201-204 [PMID: 10619756 DOI: 10.1086/313587]
31 **Park KW**, Kim US, Shin JW, Yoo CG, Oh MD, Choe K. Disseminated erythematous papules in a renal transplant recipient: a case of disseminated tuberculosis. *Scand J Infect Dis* 2002; **34**: 775-777 [PMID: 12477335 DOI: 10.1080/00365540210147930]
32 **Pandian TK**, Deziel PJ, Otley CC, Eid AJ, Razonable RR. Mycobacterium marinum infections in transplant recipients: case report and review of the literature. *Transpl Infect Dis* 2008; **10**: 358-363 [PMID: 18482202 DOI: 10.1111/j.1399-3062.2008.00317.x]
33 **Morales P**, Gil A, Santos M. Mycobacterium abscessus infection in transplant recipients. *Transplant Proc* 2010; **42**: 3058-3060 [PMID: 20970610 DOI: 10.1016/j.transproceed.2010.08.004]
34 **Doucette K**, Fishman JA. Nontuberculous mycobacterial infection in hematopoietic stem cell and solid organ transplant recipients. *Clin Infect Dis* 2004; **38**: 1428-1439 [PMID: 15156482 DOI: 10.1086/420746]
35 **Ribeiro PM**, Bacal F, Koga-Ito CY, Junqueira JC, Jorge AO. Presence of Candida spp. in the oral cavity of heart transplantation patients. *J Appl Oral Sci* 2011; **19**: 6-10 [PMID: 21437462]
36 **Virgili A**, Zampino MR, La Malfa V, Strumia R, Bedani PL. Prevalence of superficial dermatomycoses in 73 renal transplant recipients. *Dermatology* 1999; **199**: 31-34 [PMID: 10449954]
37 **Güleç AT**, Demirbilek M, Seçkin D, Can F, Saray Y, Sarifakioglu E, Haberal M. Superficial fungal infections in 102 renal transplant recipients: a case-control study. *J Am Acad Dermatol* 2003; **49**: 187-192 [PMID: 12894063 DOI: 10.1067/S0190-9622(03)00861-2]
38 **Tessari G**, Cagalli A, Girolomoni G. Opportunistic deep cutaneous mycoses in solid organ transplant recipients. *G Ital Dermatol Venereol* 2014; **149**: 417-422 [PMID: 25068229]
39 **Singh N**. Fungal infections in the recipients of solid organ transplantation. *Infect Dis Clin North Am* 2003; **17**: 113-134, viii [PMID: 12751263 DOI: 10.1016/S0891-5520(02)00067-3]
40 **Gabardi S**, Kubiak DW, Chandraker AK, Tullius SG. Invasive fungal infections and antifungal therapies in solid organ transplant recipients. *Transpl Int* 2007; **20**: 993-1015 [PMID:

17617181 DOI: 10.1111/j.1432-2277.2007.00511.x]
41 Fungal infections. *Am J Transplant* 2004; **4** Suppl 10: 110-134 [PMID: 15504225]
42 **Snydman DR**. Epidemiology of infections after solid-organ transplantation. *Clin Infect Dis* 2001; **33** Suppl 1: S5-S8 [PMID: 11389515 DOI: 10.1086/320897]
43 **Tessari G**, Naldi L, Piaserico S, Boschiero L, Nacchia F, Forni A, Rugiu C, Faggian G, Dall'olio E, Fortina AB, Alaibac M, Sassi F, Gotti E, Fiocchi R, Fagioli S, Girolomoni G. Incidence and clinical predictors of primary opportunistic deep cutaneous mycoses in solid organ transplant recipients: a multicenter cohort study. *Clin Transplant* 2010; **24**: 328-333 [PMID: 19712084 DOI: 10.1111/j.1399-0012.2009.01071.x]
44 **Briegel J**, Forst H, Spill B, Haas A, Grabein B, Haller M, Kilger E, Jauch KW, Maag K, Ruckdeschel G. Risk factors for systemic fungal infections in liver transplant recipients. *Eur J Clin Microbiol Infect Dis* 1995; **14**: 375-382 [PMID: 7556225 DOI: 10.1007/BF02114892]
45 **Collins LA**, Samore MH, Roberts MS, Luzzati R, Jenkins RL, Lewis WD, Karchmer AW. Risk factors for invasive fungal infections complicating orthotopic liver transplantation. *J Infect Dis* 1994; **170**: 644-652 [PMID: 8077723]
46 **Kanj SS**, Welty-Wolf K, Madden J, Tapson V, Baz MA, Davis RD, Perfect JR. Fungal infections in lung and heart-lung transplant recipients. Report of 9 cases and review of the literature. *Medicine* (Baltimore) 1996; **75**: 142-156 [PMID: 8965683 DOI: 10.1097/00005792-199605000-00004]
47 **Abbott KC**, Hypolite I, Poropatich RK, Hshieh P, Cruess D, Hawkes CA, Agodoa LY, Keller RA. Hospitalizations for fungal infections after renal transplantation in the United States. *Transpl Infect Dis* 2001; **3**: 203-211 [PMID: 11844152 DOI: 10.1034/j.1399-3062.2001.30404.x]
48 **Paradowski LJ**. Saprophytic fungal infections and lung transplantation--revisited. *J Heart Lung Transplant* 1997; **16**: 524-531 [PMID: 9171271]
49 **Mackenzie KA**, Wells JE, Lynn KL, Simcock JW, Robinson BA, Roake JA, Currie MJ. First and subsequent nonmelanoma skin cancers: incidence and predictors in a population of New Zealand renal transplant recipients. *Nephrol Dial Transplant* 2010; **25**: 300-306 [PMID: 19783601 DOI: 10.1093/ndt/gfp482]
50 **Bernat García J**, Morales Suárez-Varela M, Vilata JJ, Marquina A, Pallardó L, Crespo J. Risk factors for non-melanoma skin cancer in kidney transplant patients in a Spanish population in the Mediterranean region. *Acta Derm Venereol* 2013; **93**: 422-427 [PMID: 23303600 DOI: 10.2340/00015555-1525]
51 **Keller B**, Braathen LR, Marti HP, Hunger RE. Skin cancers in renal transplant recipients: a description of the renal transplant cohort in Bern. *Swiss Med Wkly* 2010; **140**: w13036 [PMID: 20652847 DOI: 10.4414/smw.2010.13036]
52 **Fekecs T**, Kádár Z, Battyáni Z, Kalmár-Nagy K, Szakály P, Horváth OP, Wéber G, Ferencz A. Incidence of nonmelanoma skin cancer after human organ transplantation: single-center experience in Hungary. *Transplant Proc* 2010; **42**: 2333-2335 [PMID: 20692474 DOI: 10.1016/j.transproceed.2010.05.021]
53 **Greenberg MS**, Friedman H, Cohen SG, Oh SH, Laster L, Starr S. A comparative study of herpes simplex infections in renal transplant and leukemic patients. *J Infect Dis* 1987; **156**: 280-287 [PMID: 3036965]
54 **España A**, Redondo P, Fernández AL, Zabala M, Herreros J, Llorens R, Quintanilla E. Skin cancer in heart transplant recipients. *J Am Acad Dermatol* 1995; **32**: 458-465 [PMID: 7868716 DOI: 10.1046/j.1523-1747.2000.0202a-3.x]
55 **Ulrich C**, Hackethal M, Meyer T, Geusau A, Nindl I, Ulrich M, Forschner T, Sterry W, Stockfleth E. Skin infections in organ transplant recipients. *J Dtsch Dermatol Ges* 2008; **6**: 98-105 [PMID: 17995969 DOI: 10.1111/j.1610-0387.2007.06431.x]
56 **Seçkin D**, Güleç TO, Demirağ A, Bilgin N. Renal transplantation and skin diseases. *Transplant Proc* 1998; **30**: 802-804 [PMID: 9595105 DOI: 10.1016/S0041-1345(98)00055-4]
57 **Lally A**, Casabonne D, Imko-Walczuk B, Newton R, Wojnarowska F. Prevalence of benign cutaneous disease among Oxford renal transplant recipients. *J Eur Acad Dermatol Venereol* 2011; **25**: 462-470 [PMID: 20738465 DOI: 10.1111/j.1468-3083.2010.03814.x]
58 **Formicone F**, Fargnoli MC, Pisani F, Rascente M, Famulari A, Peris K. Cutaneous manifestations in Italian kidney transplant recipients. *Transplant Proc* 2005; **37**: 2527-2528 [PMID: 16182734 DOI: 10.1016/j.transproceed.2005.06.067]
59 **Pfister H**. Chapter 8: Human papillomavirus and skin cancer. *J Natl Cancer Inst Monogr* 2003; (**31**): 52-56 [PMID: 12807946]
60 **Purdie KJ**, Surentheran T, Sterling JC, Bell L, McGregor JM, Proby CM, Harwood CA, Breuer J. Human papillomavirus gene expression in cutaneous squamous cell carcinomas from immunosuppressed and immunocompetent individuals. *J Invest Dermatol* 2005; **125**: 98-107 [PMID: 15982309 DOI: 10.1111/j.0022-202X.2005.23635.x]
61 **Shiley K**, Blumberg E. Herpes viruses in transplant recipients: HSV, VZV, human herpes viruses, and EBV. *Hematol Oncol Clin North Am* 2011; **25**: 171-191 [PMID: 21236397 DOI: 10.1016/j.hoc.2010.11.012]
62 **Fuente MJ**, Sabat M, Roca J, Lauzurica R, Fernández-Figueras MT, Ferrándiz C. A prospective study of the incidence of skin cancer and its risk factors in a Spanish Mediterranean population of kidney transplant recipients. *Br J Dermatol* 2003; **149**: 1221-1226 [PMID: 14674900 DOI: 10.1111/j.1365-2133.2003.05740.x]
63 **Sandoval M**, Ortiz M, Díaz C, Majerson D, Molgó M. Cutaneous manifestations in renal transplant recipients of Santiago, Chile. *Transplant Proc* 2009; **41**: 3752-3754 [PMID: 19917380 DOI: 10.1016/j.transproceed.2009.05.041]
64 **Ducroux E**, Boillot O, Ocampo MA, Decullier E, Roux A, Dumortier J, Kanitakis J, Jullien D, Euvrard S. Skin cancers after liver transplantation: retrospective single-center study on 371 recipients. *Transplantation* 2014; **98**: 335-340 [PMID: 24621534 DOI: 10.1097/TP.0000000000000051]
65 **Savoia P**, Stroppiana E, Cavaliere G, Osella-Abate S, Mezza E, Segoloni GP, Bernengo MG. Skin cancers and other cutaneous diseases in renal transplant recipients: a single Italian center observational study. *Eur J Dermatol* 2011; **21**: 242-247 [PMID: 21382788 DOI: 10.1684/ejd.2011.1272]
66 **Alangaden GJ**, Thyagarajan R, Gruber SA, Morawski K, Garnick J, El-Amm JM, West MS, Sillix DH, Chandrasekar PH, Haririan A. Infectious complications after kidney transplantation: current epidemiology and associated risk factors. *Clin Transplant* 2006; **20**: 401-409 [PMID: 16842513 DOI: 10.1111/j.1399-0012.2006.00519.x]
67 **Pugliese F**, Ruberto F, Cappannoli A, Perrella SM, Bruno K, Martelli S, Marcellino V, D'Alio A, Diso D, Rossi M, Corradini SG, Morabito V, Rolla M, Ferretti G, Venuta F, Berloco PB, Coloni GF, Pietropaoli P. Incidence of fungal infections in a solid organ recipients dedicated intensive care unit. *Transplant Proc* 2007; **39**: 2005-2007 [PMID: 17692677 DOI: 10.1016/j.transproceed.2007.05.060]
68 **Ong CS**, Keogh AM, Kossard S, Macdonald PS, Spratt PM. Skin cancer in Australian heart transplant recipients. *J Am Acad Dermatol* 1999; **40**: 27-34 [PMID: 9922009 DOI: 10.1016/S0190-9622(99)70525-6]

P- Reviewer: Brenneisen P, Caulin C, Francisco G
S- Editor: Tian YL **L- Editor**: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.20

World J Meta-Anal 2015 February 26; 3(1): 20-25
ISSN 2308-3840 (online)

Fate of meta-analyses: The case of *Helicobacter pylori*

György Miklós Buzás

György Miklós Buzás, Department of Gastroenterology, Ferencváros Health Centre, 1095 Budapest, Hungary

Author contributions: Buzás GM designed the manuscript, acquired and analysed the data and drafted the paper independently.
Conflict-of-interest: I disclose any financial or personal relationship with other people/organization that could inappropriately influence their work (employment, consultancies, stock ownerships, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding.
Data sharing: I declare that the data the present manuscript is based on meta-analysis published in the literature, there are no personal data concerning the patients or any other person, and the manuscript was not shared with any unauthorized person. There were no persons/participants who would have given informed consent, so there are definitely no harms outweighing the potential benefits.

Correspondence to: György Miklós Buzás, MD, PhD, Department of Gastroenterology, Ferencváros Health Centre, Mester utca 45, 1095 Budapest, Hungary. drbgym@gmail.com
Telephone: +36-1-4554571
Fax: +36-1-4554504
Received: October 2, 2014
Peer-review started: October 6, 2014
First decision: October 28, 2014
Revised: November 20, 2014
Accepted: December 3, 2014
Article in press: December 10, 2014
Published online: February 26, 2015

Abstract

AIM: To overview the current diversity of meta-analysis and the implementation of their results in international guidelines.

METHODS: Relevant meta-analysis were identified from PubMed/Medline. The topics of meta-analyses were determined. Some topics (genetics, extragastric tumors) were analysed separately. Core journals publishing meta-analyses on *Helicobacter pylori* were ranked. The rate of citation of meta-analysis in major guidelines was calculated.

RESULTS: Between 1992 and 2014, some 356 meta-analyses were published on PubMed. These mainly appeared in core journals, but were also found in 128 other journals. Eradicating of the infection was the most addressed topic with 134 articles. Meta-analyses were rarely used in formulating statements and recommendations in the international guidelines. In other topics - genetics, extraintestinal manifestations - meta-analyses were rather overused.

CONCLUSION: The implementation of meta-analysis in current guidelines is rather rare, while other topics benefit from many studies. A more extensive use of meta-analyses in evidence-based medicine is recommended in the future, otherwise their continuous proliferation will lose reason and scientific significance.

Key words: Consensus guidelines; *Helicobacter pylori*; Meta-analysis; Randomised controlled trials; Systematic review

Core tip: The article provides a subjective overview of the meta-analysis published on the subject of *Helicobacter pylori*, profiling the topic, their distribution in literature, giving examples of over- and underuse, and revealing a discordance between the low implementation of meta-analysis in guidelines and their importance as top-level evidence.

Buzás GM. Fate of meta-analyses: The case of *Helicobacter pylori*. *World J Meta-Anal* 2015; 3(1): 20-25 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/20.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.20

INTRODUCTION

The discovery of *Helicobacter pylori* (*H. pylori*) had a tremendous impact on the clinical practice, public health and basic research, leading to an unsurpassed proliferation of written and electronic literature[1]. Besides 35472 articles published in peer-reviewed journals (http://www.pubmed.com, accessed on September 5 and 30, 2014), dozens of printed and some e-books have been published in the past 30 years. The plethora of literature created confusion, as readers were faced with many contradictory results and statements. The general purpose of this subjective overview is to present the development from a historical viewpoint and the current state of meta-analysis in the field of *H. pylori;* specifically, to analyse the use and implementation of meta-analytical results in current international guidelines for diagnosing and treating of the infection.

Historical background

The history of meta-analysis differs according to the source: according to the anonymous writer of Wikipedia's entry, the first meta-analysis was performed by Chinese philosopher Chu Hsi[2] (1130-1200) by simply summarising data from literature of his time. Scholars date the roots of meta-analysis back to the 17th century, when Blaise Pascal (1623-1662) approached games of chance mathematically[3]. The first medical meta-analysis was published in 1904: Karl Pearson[4] (1857-1936) summarised data on the effect of enteric fever bacteria inoculation in volunteer soldiers across the British Empire and studied the association of infection, mortality and inoculation. Considerable progress was subsequently made by the works of Ronald Fisher (1890-1962) and Frank Yates (1902-1994), although they were active in the agricultural field. William Gemmel Cochran (1909-1980) stressed the need for randomised controlled trials and studied the results of the then in vogue vagotomy for curing peptic ulcer. In the modern era, the first meta-analysis was performed by Gene V Glass[5], a psychologist at the University of Colorado, in 1976. He also coined the term "meta-analysis", which later gained several entries and definitions in dictionaries (Merriam-Webster's, Dorland's Medical Dictionary, A Dictionary of Epidemiology, *etc*.).

Meta-analysis is a rapidly evolving field of statistics and over the past 3 decades increasingly sophisticated methods have been developed: these are available in books[6,7], online courses are also accessible and included in statistical packages and software programmes. It became clear that robust meta-analytical data could only be obtained by using (1) a selection of high-quality trials; and (2) a complex statistical workup of the data, including an assessment of heterogeneity, effect sizes, random or fixed effects, subgroup analysis, meta-regression, publication bias, *etc*. Specific statistical methods were introduced from 2002, when Higgins *et al*[8] from Cambridge University elaborated methods to identify heterogeneity between studies. The QUORUM and PRISMA statements were proposed in 2006 and 2010 respectively, as a uniform reporting mode for meta-analysis: unfortunately, only a small number of authors report their results according to these statements[9]. Weak data leads to uncertain results and doubtful conclusions: mixing of good and bad studies is an early mistake and is increasingly avoided in recent studies; on the other hand, weak data is perhaps better than no data at all. For reasons unknown to the author, there are no mega-trials on *H. pylori* including thousands of patients as in the case of hypertension, diabetes or hyperlipidaemia treatment. Most of the studies on *H. pylori* included a rather small number of cases and under these circumstances, assessing heterogeneity and selecting adequate statistical methods are of pivotal importance. This was not always the case. In the meantime, other more sophisticated methods emerged, like network- and combinatorial meta-analysis: both are only starting to be used in *H. pylori* research.

It must be emphasised that meta-analyses are (1) retrospective; and (2) they include studies on populations with different ethnic and genetic backgrounds, mostly geographically remote from each other, and probably infected with different strains of *H. pylori*, resulting in a "mixed bug". Therefore, meta-analysis do not rule out the need for local, well-designed, prospective and adequately sized controlled trials[6,9].

Systematic reviews are structured studies of a focused subject-*H. pylori*, in our case-aiming to synthesize the evidence from the literature based on the most relevant publications. They may or may not use statistics to combine the results of the selected studies (both full-length articles or abstracts). The PRISMA statement standardised the requirements the complete reporting requirements for systematic reviews[9]. In practice, meta-analysis and systematic reviews are often performed and reported together.

The fate of meta-analysis in H. pylori research

The first meta-analysis on *H. pylori* was published 10 years after the discovery of the bacterium: Chiba *et al*[10] from the McMaster University, Canada calculated the pooled eradication rates of single, double and triple therapies against *H. pylori* from 27 studies. In 1996, Scandinavian authors assessed the efficiency of omeprazole-based and bismuth-based triple therapies in the same way[11]. Obviously, these studies are no longer valid today because of the simplified methodology, and many other regimens against the infection have been proposed in the meantime[12].

MATERIALS AND METHODS

Using the MESH terms "*Helicobacter pylori*" AND "meta-analysis" AND "systematic review", 504 articles were found in Medline/PubMed (accessed on September 5 and 30, 2014). After reviewing the abstracts, 148 were found to be irrelevant to our subject and 356 eligible meta-analyses/systematic reviews were identified. This is a fairly low compared to other fields (Table 1) (PubMed, accessed on September 30, 2014), but comparable with

Table 1 Number of meta-analyses published on selected topics (from PubMed, accessed on September 30, 2014)

Topic/field	No. of meta-analyses
Diabetes mellitus	3245
Hypertension	2964
Coronary heart disease	3159
Gastrointestinal cancer	2787
Statins	957
Hepatitis C	550
Liver cirrhosis	435
Proton pump inhibitors	356
Helicobacter pylori	356
Peptic ulcer	395
Gastroesophageal reflux	231

other gastrointestinal diseases. The articles were classified according to their topic and method of study (meta-analysis, systematic review or combined) and total percentages were calculated (Table 2).

The spectrum of journals publishing meta-analyses and systematic reviews on *H. pylori* was also studied and a group of core journals was selected, defined arbitrarily as those publishing > 10 meta-analyses and/or systematic reviews (Table 3).

The reference list of the main consensus meetings between 2007 and 2013 (Table 4) was searched for citations of meta-analysis and a similarity analysis was performed[13-19].

To assess the average citation rates, five meta-analyses published in core journals between 2006 and 2010 were randomly selected and their citation was searched on the Web of Science (accessed on September 4, 2014)[19-23]. The reference list of 5 randomly selected expert review articles from special issues on *H. pylori*, published with the 20th anniversary of the *World Journal of Gastroenterology* was also analysed[24-28].

RESULTS

Our search identified 356 studies. Most of the authors (75%) preferred to use meta-analysis, the rest of the studies were either systematic reviews (11.3%) or a combination of the two methods (13.4%). This preference for meta-analysis was maintained in almost every one of the 14 topics (Table 2).

The topic addressed most often was that of eradication therapy: 134 (37.6%) papers analysed the efficiency of antimicrobial regimens against *H. pylori,* followed by the extraintestinal manifestations of the infection (49 studies, 13.7%) and genetics (32 articles, 8.9%). The association of the infection with tumours other than gastric cancer also elicited high interest with 26 studies (7.3%) (laryngeal cancer: 1, oesophageal: 8, pancreatic: 5, colon: 7, liver and biliary tract: 2, lung: 1). Although peptic ulcer disease is the most important complication of *H. pylori* infection, it merited only 9 studies (2.5%).

Of the 356 studies, 153 (42.97%) were published in 7 core journals (Table 3). The rest of the articles (203, 57.29%) were found in 128 journals, mostly publishing 1-2

Table 3 Core journals publishing meta-analyses and systematic reviews on *Helicobacter pylori*

Title	No. of meta-analyses	Impact factor (2013)
Alim Pharmacol Ther	45	5.478
Helicobacter	27	2.993
World J Gastroenterol	25	2.433
Am J Gastroenterol	21	9.131
Plos ONE	14	3.534
BMJ	11	16.378
Eur J Gastroenterol Hepatol	10	2.152
Total	153	Not applicable

meta-analyses on *H. pylori*. Top-ranked journals such as *Gastroenterology*, *Gut* and *Lancet* published a small number of studies on this topic (editorial policy? high rate of rejection?). Impact factors and the number of meta-analysis published were seemingly not related.

The citation rate of meta-analysis in recent reviews on *H. pylori* is also low, achieving a mean of only 7.5%/article. By contrast, the citation of meta-analysis in journals of gastroenterology published between 2006 and 2010 is fairly high (43-172, with a mean of 97 ± 28 citations).

DISCUSSION

Inclusion of meta-analysis in the consensus statements

In biomedical research, meta-analyses are considered the highest level of evidence. The importance of these studies was recently summarised by Gisbert[12] of Madrid, who performed 36 meta-analyses and systematic reviews with his team between 2003 and 2013, concluding that "meta-analysis provides a means of combining raw statistical data from all eligible primary studies addressing an identical question of interest to arrive at conclusions that are more precise and reliable than those presented in a single study." By analysing all regimens against the bacterium historically, he stated that "meta-analysis has contributed in a relevant way to our understanding of the management of patients with *H. pylori* infection".

It could be expected that their results would be included in the recommendations of expert panels. Surprisingly, meta-analysis and systematic reviews represents only 10.6% of the citations in international guidelines (Table 4), and 34% of the cited meta-analyses are identical (*i.e.*, cited in ⩾ 3 consensus materials). One can conclude, that meta-analyses are underused in formulating consensus statements. Experts probably prefer to express their opinion based on randomised controlled trials and basic science.

In some areas, meta-analysis seems to be overused (extraintestinal manifestations of *H. pylori* infection, its associations with extragastric cancers, genetics) resulting in little practical use. Their release in the medical press could be explained by publication pressure too. According to most consensus statements, however, eradication of the infection is only recommended in cases of iron deficiency anaemia and idiopathic thrombcytopenic purpura. Although genetics was studied extensively,

Table 2 Profile and No. of published meta-analyses and systematic reviews on *Helicobacter pylori*

Topic	Total No. of publications	Meta-analyses	Systematic reviews	Meta-analysis + systematic review
Epidemiology	6	5	1	0
Diagnosis	23	13	4	6
Antibiotic resistance	7	5	2	0
Genetics	32	30	0	2
Eradication regimens	134	98	17	19
Extragastric manifestations	49	35	2	12
Probiotics	11	11	0	0
Peptic ulcer	9	7	1	1
Gastric cancer	16	12	0	4
Pathogenesis	22	21	1	1
Other cancers (oesophagus, colon, pancreas, liver, biliary, lung)	26	24	0	2
Children	8	7	0	1
Methodological issues	12	0	12	0
Traditional Chinese medicine	1	0	0	1
Total	356	268 (75.2%)	40 (11.3%)	48 (13.4%)

Table 4 Implementation of meta-analyses in international consensus guidelines

Year	Consensus meeting	No. of ref.	No. of meta-analyses/systematic reviews cited
2007	Maastricht Ⅲ consensus	99	10 (9.9%)
2007	Cervia Ⅱ Working Group guideline	72	5 (6.5%)
2007	American College of Gastroenterology guideline	175	23 (13.1%)
2009	Second Asia-Pacific Consensus Guidelines	118	12 (10.1%)
2012	Maastricht-Florence 4 guideline	325	36 (11.0%)
2013	3rd Brazilian Consensus	216	25 (11.3%)
2013	Revised Korean consensus	208	19 (9.3%)
	Total	1223	130 (10.6%)

Table 5 Citation of meta-analyses in recent expert reviews

Ref.	Year	Journal	No. ref.	No. and % of meta-analyses cited
[20]	2014	*World J Gastroenterol*	115	5 (4.3)
[21]	2014	*World J Gastroenterol*	137	8 (5.8)
[22]	2014	*World J Gastroenterol*	158	14 (8.8)
[23]	2014	*World J Gastroenterol*	69	1 (1.4)
[24]	2014	*World J Gastroenterol*	79	14 (17.7)
		Total	558	42 (7.5)

genetic counselling and tests are neither available nor recommended in diseases associated with *H. pylori* infection.

The association of extragastric cancer with the infection is largely documented, but there is no recommendation to screen and treat the infection in high risk patients, as it is in first-degree relatives of gastric carcinoma patients. In all these cases, however, association does not mean causation, further studies are necessary to see if the associations are casual or causal.

In a random selection of recent expert reviews, meta-analyses are again barely cited[20-23] (Table 5), excepting the Spanish team, which is the most active in this field[29].

In contrast with this, meta-analyses are adequately cited generally speaking. The data suggests, that meta-analyses are as frequently cited as other clinical studies in the literature, but not in consensus materials, where they really should be (Table 6)[24-28]. The reason for this discordance is not known.

In conclusion, meta-analysis represent the highest level of evidence in medical research and are themselves under continuous mathematical and statistical development. In the field of *H. pylori* research, 356 meta-analyses and systematic reviews or both were published between 1992 and 2014. Although these studies are widely cited in literature, their implementation in the national/international consensus quidelines is rather rare. Other topics, of less practical importance, benefit from many meta-analyses. In the future, a more extensive use of meta-analyses would be welcome, to maintain the scientific significance of the guidelines and statements: otherwise, they will proliferate simply as a result of publication pressure and will progressively loss their scientific significance.

ACKNOWLEDGMENTS

The English correction of the manuscript by Mr. Douglas Arnott (EDMF Language Services Ltd., Budapest), and

Table 6 Citation of randomly selected meta-analyses on *Helicobacter pylori* (Web of Science, accessed on September 4, 2014)

Ref.	Year	Journal	Total citations	Independent citations
[24]	2006	*Aliment Pharmacol Ther*	172	172
[25]	2009	*Am J Gasteroenterol*	127	127
[26]	2009	*Helicobacter*	84	84
[27]	2010	*Am J Gastroenterol*	62	62
[28]	2010	*Am J Gastroenterol*	43	43

the secretarial work and Web of Science search by Jolán Józan (Semmelweis University, Department of Physiology, Budapest, Hungary) is highly acknowledged.

COMMENTS

Background

Meta-analyses have come strongly to the fore in the past 3 decades and are considered the highest grade of evidence in medical research. Their further use and implementation in the guidelines and consensus statements is unknown.

Research frontiers

The article provides an analysis of the spectrum of meta-analysis published between 1992 and 2014 in the field of *Helicobacter pylori* (*H. pylori*) research, providing the distribution of topics, ranking of core journals publishing meta-analysis, giving examples of under- and overuse of meta-analysis in some areas. The author's main conclusion is that meta-analysis are underused in the formulation of statements from recent international guidelines for diagnosing and treating the infection.

Applications

The article suggests that meta-analysis must be more widely read, used and cited, especially when experts formulate their opinions/recommendations for treating the *H. pylori* infection. On the other side, their overuse in some topics (genetics, extraintestinal manifestations) did not resulted any benefit.

Peer review

The manuscript "Fate of meta-analysis: The case of *Helicobacter pylori*" is very interesting and original in its contents.

REFERENCES

1 **Marshall BJ**, Warren JR. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* 1984; **1**: 1311-1315 [PMID: 6145023 DOI: 10.1016/S0140-6736(84)91816-6]

2 Definition of Meta-analysis. [Accessed 2014 July 8]. Available from: URL: http: //www.wikipedia.org/wiki/Meta-analysis

3 **O'Rourke K**. An historical perspective on meta-analysis: dealing quantitatively with varying study results. *J R Soc Med* 2007; **100**: 579-582 [PMID: 18065712 DOI: 10.1258/jrsm.100.12.579]

4 Report on Certain Enteric Fever Inoculation Statistics. *Br Med J* 1904; **2**: 1243-1246 [PMID: 20761760]

5 **Glass GV**. Primary, secondary and meta-analysis of research. *Educ Researcher* 1976; **10**: 3-8

6 **Borenstein M**, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Chichester: John Wiley and Sons, 2009

7 **Leandro G**. Meta-analysis in Medical Research. Oxford: Blackwell Publishing, 2005

8 **Higgins JP**, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539-1558 [PMID: 12111919 DOI: 10.1002/sim.1186]

9 **Moher D**, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009; **62**: 1006-1012 [PMID: 19631508 DOI: 10.1016/j.jclinepi.2009.06.005]

10 **Chiba N**, Rao BV, Rademaker JW, Hunt RH. Meta-analysis of the efficacy of antibiotic therapy in eradicating Helicobacter pylori. *Am J Gastroenterol* 1992; **87**: 1716-1727 [PMID: 1449132]

11 **Unge P**, Berstad A. Pooled analysis of anti-Helicobacter pylori treatment regimens. *Scand J Gastroenterol Suppl* 1996; **220**: 27-40 [PMID: 8898433 DOI: 10.3109/00365529609097 47]

12 **Gisbert JP.** The contribution of meta-analyses to the treatment of Helicobacter pylori. In: Buzás GM. Helicobacter pylori: a worldwide perspective 2014. United Sharjah, United Arabian Emirates, Bentham Scientific Publications, 2014: 279-316

13 **Caselli M**, Zullo A, Maconi G, Parente F, Alvisi V, Casetti T, Sorrentino D, Gasbarrini G. "Cervia II Working Group Report 2006": guidelines on diagnosis and treatment of Helicobacter pylori infection in Italy. *Dig Liver Dis* 2007; **39**: 782-789 [PMID: 17606419 DOI: 10.1016/j.dld.2007.05.016]

14 **Chey WD**, Wong BC; Practice Parameters Committee of the American College of Gastroenterology. American College of Gastroenterology guideline on the management of Helicobacter pylori infection. *Am J Gastroenterol* 2007; **102**: 1808-1825 [PMID: 17608775 DOI: 10.1111/j.1572-0241.2007.01393.x]

15 **Malfertheiner P**, Megraud F, O'Morain C, Bazzoli F, El-Omar E, Graham D, Hunt R, Rokkas T, Vakil N, Kuipers EJ. Current concepts in the management of Helicobacter pylori infection: the Maastricht III Consensus Report. *Gut* 2007; **56**: 772-781 [PMID: 17170018 DOI: 10.1136/gut.2006.101634]

16 **Fock KM**, Katelaris P, Sugano K, Ang TL, Hunt R, Talley NJ, Lam SK, Xiao SD, Tan HJ, Wu CY, Jung HC, Hoang BH, Kachintorn U, Goh KL, Chiba T, Rani AA. Second Asia-Pacific Consensus Guidelines for Helicobacter pylori infection. *J Gastroenterol Hepatol* 2009; **24**: 1587-1600 [PMID: 19788600 DOI: 10.1111/j.1440-1746.09582.x]

17 **Malfertheiner P**, Megraud F, O'Morain CA, Atherton J, Axon AT, Bazzoli F, Gensini GF, Gisbert JP, Graham DY, Rokkas T, El-Omar EM, Kuipers EJ; European Helicobacter Study Group. Management of Helicobacter pylori infection--the Maastricht IV/ Florence Consensus Report. *Gut* 2012; **61**: 646-664 [PMID: 22491499 DOI: 10.1136/gutjnl-2012-302084]

18 **Coelho LG**, Maguinilk I, Zaterka S, Parente JM, do Carmo Friche Passos M, Moraes-Filho JP. 3rd Brazilian Consensus on Helicobacter pylori. *Arq Gastroenterol* 2013; **50**: pii: S0004-28032013005000113 [PMID: 23748591 DOI: 10.1590/S0004-28032013005000001]

19 **Kim SG**, Jung HK, Lee HL, Jang JY, Lee H, Kim CG, Shin WG, Shin ES, Lee YC; Korean College of Helicobacter and Upper Gastrointestinal Research. [Guidelines for the diagnosis and treatment of Helicobacter pylori infection in Korea, 2013 revised edition]. *Korean J Gastroenterol* 2013; **62**: 3-26 [PMID: 23954956]

20 **Yang JC**, Lu CW, Lin CJ. Treatment of Helicobacter pylori infection: current status and future concepts. *World J Gastroenterol* 2014; **20**: 5283-5293 [PMID: 24833858 DOI: 10.3748/wjg.v20.i18.5283]

21 **Hagymási K**, Tulassay Z. Helicobacter pylori infection: new pathogenetic and clinical aspects. *World J Gastroenterol* 2014; **20**: 6386-6399 [PMID: 24914360 DOI: 10.3748/wjg.v20.i21.6386]

22 **Shiotani A**, Haruma K, Graham DY. Metachronous gastric cancer after successful Helicobacter pylori eradication. *World J Gastroenterol* 2014; **20**: 11552-11559 [PMID: 25206262 DOI: 10.3748/wjg.v20.i33.11522]
23 **Wong F**, Rayner-Hartley E, Byrne MF. Extraintestinal manifestations of Helicobacter pylori: a concise review. *World J Gastroenterol* 2014; **20**: 11950-11961 [PMID: 25232230 DOI: 10.3748/wjg.v20.i34.11950]
24 **Gisbert JP**, Morena F. Systematic review and meta-analysis: levofloxacin-based rescue regimens after Helicobacter pylori treatment failure. *Aliment Pharmacol Ther* 2006; **23**: 35-44 [PMID: 16393278]
25 **Gatta L**, Vakil N, Leandro G, Di Mario F, Vaira D. Sequential therapy or triple therapy for Helicobacter pylori infection: systematic review and meta-analysis of randomized controlled trials in adults and children. *Am J Gastroenterol* 2009; **104**: 3069-3679; quiz 1080 [PMID: 19844205 DOI: 10.1038/ajg.2009.555]
26 **Essa AS**, Kramer JR, Graham DY, Treiber G. Meta-analysis: four-drug, three-antibiotic, non-bismuth-containing "concomitant therapy" versus triple therapy for Helicobacter pylori eradication. *Helicobacter* 2009; **14**: 109-118 [PMID: 19298338 DOI: 10.1111/j.1523-5378.2009.00671x]
27 **Luther J**, Higgins PD, Schoenfeld PS, Moayyedi P, Vakil N, Chey WD. Empiric quadruple vs. triple therapy for primary treatment of Helicobacter pylori infection: Systematic review and meta-analysis of efficacy and tolerability. *Am J Gastroenterol* 2010; **105**: 65-73 [PMID: 19755966 DOI: 10.1038/ajg.2009.508]
28 **Yaghoobi M**, Farrokhyar F, Yuan Y, Hunt RH. Is there an increased risk of GERD after Helicobacter pylori eradication?: a meta-analysis. *Am J Gastroenterol* 2010; **105**: 1007-1013; quiz 1006, 1014 [PMID: 20087334 DOI: 10.1038/ajg.2009.734]
29 **Molina-Infante J**, Gisbert JP. Optimizing clarithromycin-containing therapy for Helicobacter pylori in the era of antibiotic resistance. *World J Gastroenterol* 2014; **20**: 10338-10347 [PMID: 25132750 DOI: 10.3748/wjg.v20.i30.10338]

P- Reviewer: Guardia-Olmos J, Macedo CR, Nath G, Specchia ML
S- Editor: Song XX **L- Editor**: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.26

World J Meta-Anal 2015 February 26; 3(1): 26-35
ISSN 2308-3840 (online)

Effect of institutional volume on laparoscopic cholecystectomy outcomes: Systematic review and meta-analysis

Muireann Murray, Donagh A Healy, John Ferguson, Khalid Bashar, Seamus McHugh, Mary Clarke Moloney, Stewart R Walsh

Muireann Murray, Donagh A Healy, Khalid Bashar, Seamus McHugh, Department of Surgery, University Hospital Limerick, Limerick, Ireland
John Ferguson, Department of Nephrology, Graduate Entry Medical School, University of Limerick, Limerick, Ireland
Mary Clarke Moloney, Health Research Institute, University of Limerick, Limerick, Ireland
Stewart R Walsh, Department of Surgery, University College Hospital, Galway, Ireland
Author contributions: Walsh SR, Clarke Moloney M and McHugh S designed the research; Murray M identified eligible studies; Murray M and Healy DA extracted data; Healy DA, Ferguson J and Bashar K performed analysis; all authors contributed to drafting the manuscript and revising it critically; all authors approved the final draft.
Conflict-of-interest: The authors state that they have no conflicts of interest to declare.
Data sharing: Dataset and statistical code are available from the cooresponding author at the email address mentioned above.

Correspondence to: Dr. Muireann Murray, Department of Surgery, University Hospital Limerick, Castletroy, Limerick, Ireland. muireannm80@gmail.com
Telephone: +353-61-482761
Fax: +353-61-233778
Received: April 11, 2014
Peer-review started: April 12, 2014
First decision: May 1, 2014
Revised: December 18, 2014
Accepted: December 29, 2014
Article in press: December 31, 2014
Published online: February 26, 2015

Abstract

AIM: To determine whether institutional laparoscopy cholecystectomy (LC) volume affects rates of mortality, conversion to open surgery, bile leakage and bile duct injury (BDI).

METHODS: Eligible studies were prospective or retrospective cohort studies that provided data on outcomes from consecutive LC procedures in single institutions. Relevant outcomes were mortality, conversion to open surgery, bile leakage and BDI. We performed a Medline search and extracted data. A regression analysis using generalized estimating equations were used to determine the influence of annual institutional LC caseload on outcomes. A sensitivity analysis was performed including only those studies that were published after 1995.

RESULTS: Seventy-three cohorts (127404 LC procedures) were included. Average complication rates were 0.06% for mortality, 3.23% for conversion, 0.44% for bile leakage and 0.28% for bile duct injury. Annual institutional caseload did not influence rates of mortality (P = 0.142), bile leakage (P = 0.111) or bile duct injury (P = 0.198) although increasing caseload was associated with reduced incidence of conversion (P = 0.019). Results from the sensitivity analyses were similar.

CONCLUSION: Institutional volume is a determinant of LC complications. It is unclear whether volume is directly linked to complication rates or whether it is an index for protocolised care.

Key words: Abdominal; Cholecystectomy; Quality control; Systematic review; Meta-analysis

Core tip: We performed a meta-analysis to determine whether institutional laparoscopy cholecystectomy (LC) volume affects rates of mortality, conversion to open surgery, bile leakage and bile duct injury. Annual institutional caseload did not influence rates of mortality (P = 0.142), bile leakage (P = 0.111) or bile duct injury (P = 0.198) although increasing caseload was associated with reduced incidence of conversion (P = 0.019). Our results suggest that institutional LC volume may be a determinant of LC complications. It is unclear whether institutional LC volume is directly linked to complication rates or whether its influence is a surrogate for improved quality of care.

Murray M, Healy DA, Ferguson J, Bashar K, McHugh S, Clarke Moloney M, Walsh SR. Effect of institutional volume on laparoscopic choleystectomy outcomes: Systematic review and meta-analysis. *World J Meta-Anal* 2015; 3(1): 26-35 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/26.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.26

INTRODUCTION

Laparoscopic cholecystectomy (LC) is one of the most commonly performed operations-close to 400000 procedures are performed annually in non-federal community hospitals in the United States[1] and around 50000 procedures are performed annually in the United Kingdom[2]. LC is preferred over open cholecystectomy as it leads to a shorter hospital stay and a quicker recovery[3]. However, there are risks of serious complications with LC such as biliary leaks (0.4%-1%)[2,4], bile duct injury (BDI) (0.2%-0.3%)[3,5] and mortality (0.1%-0.4%)[3,5]. Conversion rates vary from about 15%-5%[5].

An expanding body of evidence suggests that outcomes in a variety of conditions are improved when patients are managed in high-volume centres or by high-volume healthcare providers[6]. High-volume centres dramatically improve the management of pancreatic cancer (≥ 20 cases per year), oesophageal cancer (≥ 30 cases per year), paediatric cardiac conditions (≥ 300 cases per year), unruptured abdominal aortic aneurysms (AAA) (≥ 36 cases per year) and acquired immune deficiency syndrome (≥ 100 cases per year)[6]. Similarly, high-volume surgeons or physicians dramatically improve the management of pancreatic cancer (10-42 cases per year), ruptured AAAs (≥ 10 cases per year), paediatric cardiac conditions (≥ 75 cases per year), colorectal cancer (≥ 22 cases per year), carotid endarterectomy (≥ 30 cases per year) and coronary artery bypass grafting (≥ 150 cases per year)[6]. In contrast, no proven volume-outcome relationships exist for conditions such as diabetes, cystic fibrosis, rheumatoid arthritis, appendicitis and hernias[7,8].

Recently, data have emerged confirming that high-volume surgeons improve outcomes following LC[2,4,5,9-12]. Giger *et al*[5] found improved results with surgeons who performed > 100 LCs per year, Nuzzo *et al*[10] found improved results with surgical teams who performed > 450 LCs in three years, Csikesz *et al*[11] found improved results with surgeons who performed > 15 LCs per year and McMahon *et al*[12] found improved results for surgeons who had performed more than 200 cases. Andrews *et al*[2] and Hobbs *et al*[4] did not specify thresholds although they identified significantly reduced complications with increasing surgeon volume. However, it is unclear whether a volume-outcome relationship exists for LC at institutional level. If such an institutional relationship can be proven and understood, the creation of high-volume LC centres may become a priority. Therefore we performed a systematic review and meta-analysis focusing on institutional volume/outcome relationships for LC. The aim was to determine whether institutional LC volume affects rates of mortality, conversion to open surgery, bile leakage and bile duct injury.

MATERIALS AND METHODS

This systematic review was performed in accordance with the PRISMA guidelines[13]. These guidelines are an evidence-based set of items that aim to enhance methodological and reporting clarity.

The Medline electronic database was searched from 1st January 1990 to 9th April 2014 using the free text "laparoscopic cholecystectomy".

Eligible studies were prospective or retrospective cohort studies that provided details on outcomes from consecutive LC procedures in single institutions. The relevant outcomes were the incidences of conversion to open surgery, bile leakage, BDI or mortality. The definitions and timeframes of these outcomes were those specified in retrieved manuscripts. There were no limitations on cohort sizes or on recruitment dates of studies. Studies reporting combined results from multiple centres were eligible provided that data were provided separately for individual centres. Studies were excluded if results did not allow the calculation of institutional complication rates. This lead to the exclusion of studies that reported on selected LCs rather than all consecutive LCs and studies that did not specify study start and finish dates. Case reports, narrative reviews and non-English language studies were also excluded.

One author (Murray M) identified eligible studies. Firstly, titles and abstracts were screened. Full-text manuscripts of potentially relevant studies were examined to finalise eligibility. Uncertainties regarding eligibility were discussed with a second author (Healy DA). For each included study, the following data were extracted independently by two authors (Murray M and Healy DA): author, publication date, study design, the institution's name, start and finish dates, duration, number of LCs, number of mortalities, number of conversions to open surgery, number of bile leaks and the number of cases of

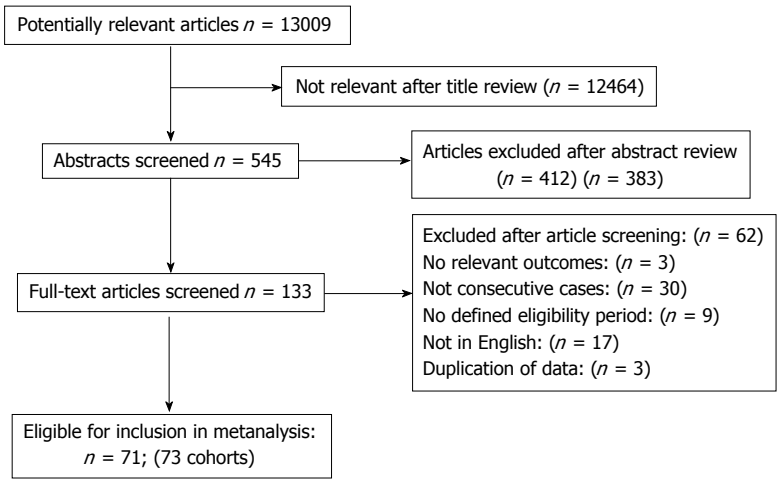


Figure 1 Summary of the results of the search.

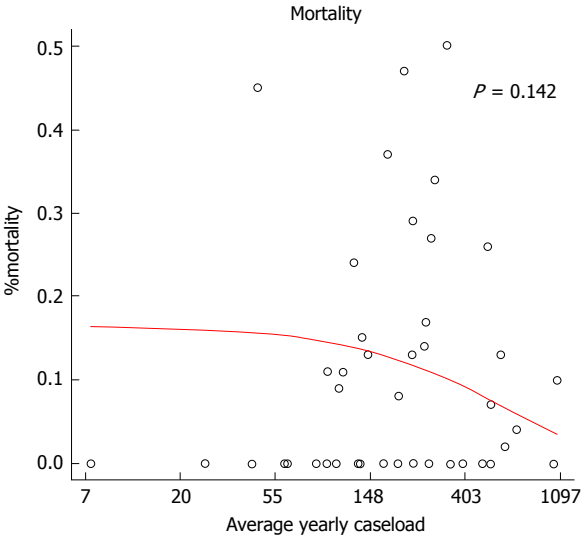


Figure 2 Scatterplot with regression line demonstrating the relationship between percentage mortality rate and average annual institutional volume of laparoscopic cholecystectomy (average caseload is plotted on a logarithmic scale).

BDI. Percentage complication rates were calculated for each outcome. Disagreements regarding data extraction were resolved by discussion with a third author (Walsh SR). Data were entered into a computerised spreadsheet for analysis.

All analyses were designed and performed by a biomedical statistician (JF). Scatterplots were used to summarise the relationships between numbers of LC procedures per year and percentage complication rates. Regression analyses were performed using generalized estimating equations. The generalized estimation equations were fit using a variance structure based on the binomial distribution. The response was the percentage of complications out of all procedures performed. A robust variance was used to account for extra variance around the regression line because of center specific effects. The function "gee" in the statistical language R was used. A sensitivity analysis was performed that was limited to studies that were published after 1995. This time point was chosen with the aim of eliminating the effects of learning curves and improvements in perioperative care. Significance was set at 5%.

Statistical analysis

The authors state that all statistical analyses were designed and performed by a biomedical statistician. A statement to this effect is included in the methods section.

RESULTS

Figure 1 summarises the results of the search. 13009 citations were identified and 12876 were excluded based on titles and abstracts. 133 full text manuscripts were retrieved and 71 articles (corresponding to 73 cohorts) were finally eligible for inclusion.

Table 1 provides a summary of the 73 eligible cohorts[14-84]. Most were retrospective and some were prospective cohorts. Study recruitment periods varied from 1990 to 2013. The total number of LC procedures was 127404.

Forty-three studies (71305 patients) provided data on mortality (43 cases of mortality; average mortality was 0.06%). Figure 2 displays the relationship between average annual number of LC procedures and institutional mortality rates as percentages. There was no significant relationship between mortality and annual number of procedures ($P = 0.142$). When only those cohorts published after 1995 were included (32 cohorts, 64273 patients) there was no significant relationship ($P = 0.168$).

Fifty-eight studies (87840 patients) provided data on conversion rates (2835 cases of conversion; average conversion rate was 3.23%). Figure 3 displays the relationship between average annual number of LC procedures and institutional percentage conversion rates. Increasing caseload was associated with lower conversion rates ($P = 0.019$). When only those studies that were published after 1995 were included (43 studies, 79311 patients) the result remained significant ($P = 0.019$).

Table 1 Characteristics and results of included studies

Ref.	Publication year	Study duration in months	Design	Total LC number	Average annual LC number	Mortality incidence	Percentage mortality rate (%)	Conversion to open surgery incidence	Percentage conversion rate (%)	Bile leak incidence	Percentage bile leak rate (%)	Bile duct injury incidence	Percentage bile duct injury rate (%)
Szego *et al*[14]	1991	6	Retrospective	31	62.00	0	0.00	2	6.45	0	0	0	0
Bailey *et al*[15]	1991	16	Prospective	375	281.25	1	0.27	20	5.33	1	0.27	1	0
Peters *et al*[16]	1991	6	Prospective	100	200.00	0	0.00	4	4.00	2	2	1	1
Rees *et al*[17]	1992	12	Retrospective	155	155.00	N/A	N/A	8	5.16	2	1.29	1	1
Huang *et al*[18]	1992	11	Retrospective	200	218.18	N/A	N/A	N/A	N/A	1	0.5	1	1
Davis *et al*[19]	1992	24	Retrospective	622	311.00	N/A	N/A	26	4.18	N/A	N/A	1	0
Fielding *et al*[20]	1992	12	Retrospective	220	220.00	N/A	N/A	8	3.64	2	0.91	N/A	N/A
Soper *et al*[21]	1992	21	Prospective	600	342.80	0	0.00	18	3.00	N/A	N/A	1	0
Périssat *et al*[22]	1992	32	Retrospective	700	262.50	1	0.14	41	5.86	N/A	N/A	3	0
Troidl *et al*[23]	1992	14.5	Prospective	400	331.00	2	0.50	20	5.00	3	0.75	N/A	N/A
Rubio *et al*[24]	1993	31	Retrospective	500	193.55	N/A	N/A	4	0.80	1	0.2	1	0
Huang *et al*[25]	1993	18	Retrospective	350	233.33	0	0.00	6	1.71	4	1.14	1	0
Cox *et al*[26]	1994	24	Prospective	410	205.00	N/A	N/A	N/A	N/A	N/A	N/A	1	0
Williams *et al*[27]	1994	27	Retrospective	600	266.67	1	0.17	24	4.00	N/A	N/A	3	1
Cappuccino *et al*[28]	1994	14	Retrospective	563	482.57	0	0.00	5	0.89	N/A	N/A	N/A	N/A
Bonatsos *et al*[29]	1995	41	Retrospective	1788	523.32	0	0.00	45	2.52	19	1.06	0	0
Newman *et al*[30]	1995	36	Retrospective	1525	508.33	4	0.26	34	2.23	0	0	0	0
Chen *et al*[31]	1996	42	Retrospective	2428	693.71	1	0.04	N/A	N/A	1	0.04	4	0
Bond *et al*[32]	1996	36	Retrospective	534	178.00	2	0.37	N/A	N/A	N/A	N/A	N/A	N/A
Sartori *et al*[33]	1996	14	Retrospective	322	276.00	0	0.00	N/A	N/A	N/A	N/A	N/A	N/A
Rather *et al*[34]	1997	24	Retrospective	340	170.00	0	0.00	26	7.65	6	1.76	2	1
Jan *et al*[35]	1997	60	Retrospective	1115	223.00	N/A	N/A	N/A	N/A	4	0.36	3	0
Ahmad *et al*[36]	1997	45	Retrospective	1300	346.67	0	0.00	40	3.08	6	0.46	0	0
Kok *et al*[37]	1998	58	Prospective	220	45.52	1	0.45	9	4.09	N/A	N/A	1	0
Targarona *et al*[38]	1998	61	Retrospective	1630	320.66	N/A	N/A	109	6.69	N/A	N/A	16	1
Kurauchi *et al*[39]	1998	32	Retrospective	1408	528.00	1	0.07	89	6.32	N/A	N/A	9	1
Jones-Monahan *et al*[40]	1998	60	Retrospective	2654	530.80	N/A	N/A	N/A	N/A	1	0.04	6	0
Matthews *et al*[41]	1999	53	Retrospective	1025	232.08	3	0.29	27	2.63	2	0.2	1	0
Calvete *et al*[42]	2000	72	Prospective	784	130.67	0	0.00	4	0.51	4	0.51	7	1
Patel *et al*[43]	2000	38	Prospective	135	42.63	0	0.00	7	5.19	2	1.48	N/A	N/A
Sikora *et al*[44]	2001	72	Retrospective	1200	200.00	N/A	N/A	N/A	N/A	N/A	N/A	16	1
Lichten *et al*[45]	2001	12	Retrospective	300	300.00	N/A	N/A	17	5.67	N/A	N/A	N/A	N/A
Miroshnik *et al*[46]	2002	110	Retrospective	1216	132.65	N/A	N/A	90	7.40	7	0.58	1	0
Hasaniah *et al*[47]	2002	84	Retrospective	2750	392.86	0	0.00	127	4.62	11	0.4	3	0
Fathy *et al*[48]	2003	93	Retrospective	2000	258.06	N/A	N/A	147	7.35	11	0.55	7	0
Duca *et al*[49]	2003	108	Retrospective	9542	1060.22	10	0.10	184	1.93	54	0.57	17	0
Mahatharadol *et al*[50]	2004	116	Retrospective	1522	157.45	N/A	N/A	N/A	N/A	N/A	N/A	9	1
Daradkeh *et al*[51]	2005	108	Retrospective	1208	134.22	0	0.00	32	2.65	N/A	N/A	0	0
Diamantis *et al*[52]	2005	132	Retrospective	2079	189.00	N/A	N/A	N/A	N/A	N/A	N/A	13	1
Söderlund *et al*[53]	2005	50	Prospective	1568	376.32	N/A	N/A	N/A	N/A	23	1.47	24	2
Baird[54]	2005	16	Prospective	782	586.50	1	0.13	18	2.30	N/A	N/A	0	0
Vagenas *et al*[55]	2006	156	Retrospective	1220	93.85	0	0.00	23	1.89	3	0.25	2	0
Tan *et al*[56]	2006	9	Prospective	202	269.33	N/A	N/A	14	6.93	3	1.49	1	0
Mufti *et al*[57]	2007	12	Retrospective	60	60.00	0	0.00	3	5.00	1	1.67	0	0

Brekalo *et al*[58]	2007	120	Retrospective	952	95.20	1	0.11	32	3.36	N/A	N/A	0	0
Marakis *et al*[59]	2007	144	Retrospective	1225	102.08	N/A	N/A	91	7.43	1	0.08	2	0
Herve *et al*[60]	2007	120	Retrospective	1255	125.50	3	0.24	25	1.99	N/A	N/A	12	1
Brekalo *et al*[58]	2007	120	Retrospective	1066	106.60	1	0.09	N/A	N/A	43	4.03	3	0
Shrestha *et al*[61]	2007	21	Prospective	140	80.00	N/A	N/A	13	9.29	N/A	N/A	N/A	N/A
Tantia *et al*[62]	2008	156	Retrospective	13305	1023.46	0	0.00	8	0.06	10	0.08	52	0
Priego *et al*[63]	2009	204	Retrospective	3933	231.35	5	0.13	331	8.42	17	0.43	13	0
Avgerinos *et al*[64]	2009	72	Prospective	1046	174.33	N/A	N/A	27	2.58	5	0.48	0	0
Clegg-Lamptey *et al*[65]	2010	24	Prospective	52	26.00	0	0.00	1	1.92	1	1.92	0	0
Al-Kubati *et al*[66]	2010	48	Retrospective	336	84.00	0	0.00	43	12.80	3	0.89	2	1
Ying *et al*[67]	2010	144	Retrospective	2400	200.00	2	0.08	11	0.46	7	0.29	3	0
Zha *et al*[68]	2010	156	Prospective	13000	1000.00	N/A	N/A	N/A	N/A	N/A	N/A	11	0
Wichmann *et al*[69]	2010	18	Prospective	140	93.33	N/A	N/A	11	7.86	3	2.14	N/A	N/A
Wichmann *et al*[69]	2010	18	Prospective	219	146.00	N/A	N/A	18	8.22	2	0.91	N/A	N/A
Kanakala *et al*[70]	2011	120	Retrospective	2117	211.70	10	0.47	133	6.28	31	1.46	7	0
Al-Mulhim *et al*[71]	2011	36	Prospective	968	322.67	N/A	N/A	5	0.52	3	0.31	N/A	N/A
Halilovic *et al*[72]	2011	12	Prospective	293	293.00	1	0.34	8	2.73	N/A	N/A	N/A	N/A
Hasbahceci *et al*[73]	2012	129	Retrospective	1557	144.84	2	0.13	39	2.50	10	0.64	4	0
Bekele *et al*[74]	2012	60	Retrospective	681	136.20	1	0.15	20	2.94	N/A	N/A	N/A	N/A
Le *et al*[75]	2012	24	Retrospective	3371	1685.50	N/A	N/A	86	2.55	N/A	N/A	N/A	N/A
Afuwape *et al*[76]	2012	20	Retrospective	13	7.80	0	0.00	1	7.69	N/A	N/A	1	8
Grbas *et al*[77]	2013	202	Retrospective	10317	612.89	2	0.02	220	2.13	52	0.5	25	0
Sultan *et al*[78]	2013	120	Retrospective	4434	443.40	N/A	N/A	234	5.28	N/A	N/A	N/A	N/A
Pulvirenti *et al*[79]	2013	120	Retrospective	882	88.20	N/A	N/A	51	5.78	N/A	N/A	N/A	N/A
Dip *et al*[80]	2014	5	Prospective	43	103.20	0	0.00	N/A	N/A	0	0	0	0
Comajuncosas *et al*[81]	2014	12	Prospective	276	276.00	N/A	N/A	26	9.42	N/A	N/A	N/A	N/A
Paajanen *et al*[82]	2014	204	Retrospective	1895	111.47	2	0.11	126	6.65	14	0.74	2	0
Alvarez *et al*[83]	2014	248	Retrospective	11423	552.73	N/A	N/A	N/A	N/A	5	0.04	20	0.18
Afaneh *et al*[84]	2014	44	Retrospective	1382	376.91	N/A	N/A	44	3.18	N/A	N/A	2	0.14

LC: Laparoscopic cholecystectomy; N/A: Not available.

Forty-four studies (86025 patients) provided data on bile leak rates (381 cases of bile leakage; average bile leak rate was 0.44%). Figure 4 displays the relationship between average annual number of LC procedures and the institutional percentage bile leak rate. There was no significant relationship between bile leak rates and annual number of procedures ($P = 0.111$). When only those studies that were published after 1995 were included (33 cohorts, 80381 patients) the result was similar ($P = 0.123$).

Fifty-six cohorts (113526 patients) provided data on bile duct injury rates (316 cases of bile duct injury; average bile duct injury rate was 0.28%). Figure 5 displays the relationship between average annual number of LC procedures and institutional percentage bile duct injury rate. There was no significant relationship between bile duct injury rates and annual number of procedures ($P = 0.198$). When only those studies that were published after 1995 were included (42 cohorts, 105570 patients) the result was similar ($P = 0.19$).

DISCUSSION

This systematic review examined the effect of institutional LC volume on LC outcomes-it identified 73 single centre cohorts involving 127404 patients. Using regression analyses based upon generalised estimating equations we found that there were no significant relationships between institutional LC volume and mortality ($P = 0.142$), bile leakage ($P = 0.111$) or bile duct injury ($P = 0.198$). However, increasing institutional LC volume was associated with reduced incidence of conversion ($P = 0.019$). These pooled results relate

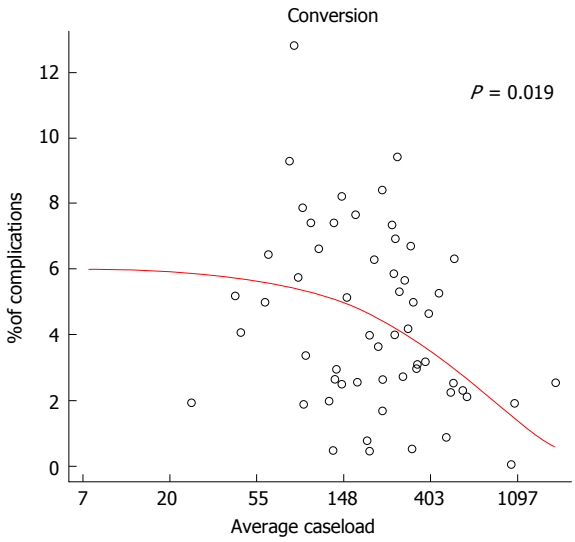


Figure 3 Scatterplot with regression line demonstrating the relationship between percentage conversion rate and average annual institutional volume of laparoscopic cholecystectomy (average caseload is plotted on a logarithmic scale).

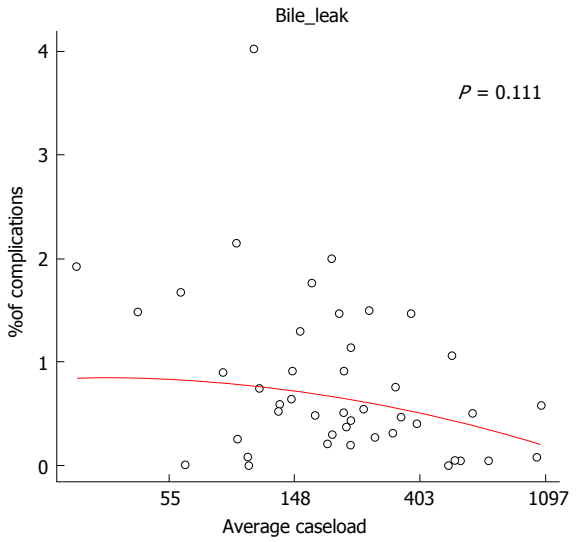


Figure 4 Scatterplot with regression line demonstrating the relationship between percentage bile leak rate and average annual institutional volume of laparoscopic cholecystectomy (average caseload is plotted on a logarithmic scale).

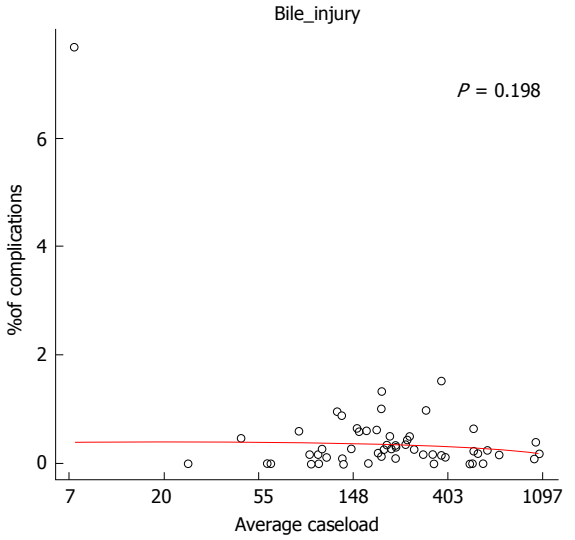


Figure 5 Scatterplot with regression line demonstrating the relationship between percentage bile duct injury rate and average annual institutional volume of laparoscopic cholecystectomy (average caseload is plotted on a logarithmic scale).

to cohorts that involved procedures performed between 1990 and 2013. Our sensitivity analysis was designed to limit the influence of the learning curve by excluding publications from before 1995-this analysis yielded similar results. Our findings are timely as mounting evidence confirms the importance of high-volume LC surgeons. Furthermore, evidence confirms the importance of high-volume centres and high-volume care providers in relation to other conditions[6]. Therefore the observation that LC complications may be influenced by institutional case load has implications for the future research and future service provision.

Our results are broadly consistent with previous studies that have examined the topic. The largest previous study was a retrospective population based study involving over one million patients from the United States National Inpatient Sample[85]. In a univariate analysis the authors of this study found that high-volume centres (≥ 225 LCs annually) had slightly improved major complication rates compared with lower-volume centres (6.4% *vs* 7.0%, $P < 0.0001$)[85]-significance was likely to have been related to the sample size and not to a clinically important effect. The effect on major complications was lost on multivariate testing. However an effect of hospital volume on conversion rates was present in a multivariable analysis-hospital volume of ≤ 120 cases per year was associated with an odds ratio (OR) for conversion of 1.32 (95%CI: 1.18-2.19) when compared with hospital volume of ≥ 225 per year. Another large population based study from Scotland involving 59918 procedures found higher mortality in lower volume (< 173 cases/year; OR = 1.45; 95%CI: 1.06-2.00; $P = 0.022$) and medium volume (173-244 cases/year; OR = 1.52; 95%CI: 1.11-2.08; $P = 0.01$) centres when high-volume centres (> 244 cases/year) were the reference group[86]. Although this again represents evidence for a hospital volume-outcome relationship for mortality, absolute effects were negligible for those patients at average risk-this suggests that the finding of significance may have simply been a reflection of the large sample size. In the late 1990s, another United States retrospective cohort study of 8602 procedures found no relationship between hospital volume and mortality[87] although a study from Norway on 4332 cases found a significant association between hospital volume and severe complications index[88]. Notably, the latter two studies only involved hospitals that nowadays would be deemed small volume.

From a wider perspective, patient safety is likely to have many underlying components and it is likely that hospital volume probably reflects clustering of these

factors[86]. In the future it is important that studies explore the possibility that "high volume" may be a surrogate for streamlined management and strict adherence to protocolised care. Equivalent outcomes may be achievable in smaller centres provided that a high quality of care is maintained. High volume LC centres should only be required if institutional volume is shown to have a clinically important effect that is independent of other aspects of quality of care. As mentioned previously, several studies suggest the existence of a surgeon volume-outcome relationship for LC[2,4,5,9-12]-this seems plausible given the high-volume but low-risk nature of gallbladder surgery. The relatively low overall complication rate of gallbladder surgery makes volume-related research difficult and therefore it is essential that high quality registries including case-mix data are maintained into the future. In the long term, this will be the only way to determine important patient, surgeon and hospital-related components of safety.

The chief strength of the current study relates to the inclusion of a large number of studies, including both small and large cohorts. Furthermore, we used an extensive search strategy and we focused on patient-important outcomes that are simply defined and easily diagnosed and are thus likely to be accurate even in retrospective studies. The external validity of the study is further enhanced by the finding of average complication rates that are quite similar to accepted published rates. The main limitation is the lack of data on case mix. Furthermore, as we included studies that spanned a twenty year period across all areas of the world, undoubtedly temporal and geographical variations in care would have existed. Notably, we declined to evaluate trends in outcomes over time as study inclusion periods were heterogenous (Table 1) and results were not provided by year but rather for entire study inclusion periods. Finally we were limited to univariate analyses, thereby restricting conclusions on other factors that influence safety. We also wish to highlight that we did not aim to estimate specific optimal volume thresholds but rather we aimed to measure the effect of institutional volume on outcomes using a regression analysis. Overall, we think that the results of our review are striking. We wish to encourage research on volume-outcome relationships in surgery, particularly through the use of large scale registries.

COMMENTS

Background

Laparoscopic cholecystectomy (LC) is one of the most commonly performed operations worldwide. It is preferred over open cholecystectomy as it offers a shorter length of hospital stay and a quicker recovery but it is associated with the chance of needing conversion to open surgery and the risks of bile leakage, bile duct injury and mortality.

Research frontiers

Studies have shown that institutional volume is an important determinant of outcome in a variety of conditions such as cancers, aortic aneurysms and cardiac surgery. Furthermore surgeon experience is an important factor in these conditions also. Although recent evidence suggests that surgeon volume is an important determinant of outcomes following LC, the authors have a poor understanding of the effect of institutional volume. Knowing the effect of institutional volume is important as it may influence how healthcare systems are organised.

Innovations and breakthroughs

Based on the authors review, they have identified that conversion rate is related to institutional volume. Increasing institutional LC volume leads to reduced incidence of conversion to opens surgery. The authors found no evidence to suggest the institutional volume influences mortality, bile leakage or bile duct injury rates.

Applications

Institutional volume is an important determinant of outcomes following LC. However, it is uncertain whether this is a direct effect or a surrogate for optimum standardised and protocolised care. Large scale prospective registries are needed to explore this topic further.

Terminology

Bile duct injury is a serious and potentially life-threatening complication of LC resulting from inadvertent damage to biliary system structures during the operation. Bile leakage refers is a serious complication that results to continued leakage of bile from the biliary system after the operation. Most bile leaks can be managed effectively but they contribute to morbidity and have economic implications.

Peer review

The current meta-analysis presents interesting.

REFERENCES

1 **National Centre for Health Statistics**. Health, United States, 2012: With Special Feature on Emergency Care. Hyattsville, MD: National Center for Health Statistics, 2013

2 **Andrews S**. Does concentration of surgical expertise improve outcomes for laparoscopic cholecystectomy? 9 year audit cycle. *Surgeon* 2013; **11**: 309-312 [PMID: 23916664 DOI: 10.1016/j.surge.2013.06.005]

3 **Keus F**, de Jong JA, Gooszen HG, van Laarhoven CJ. Laparoscopic versus open cholecystectomy for patients with symptomatic cholecystolithiasis. *Cochrane Database Syst Rev* 2006; (**4**): CD006231 [PMID: 17054285 DOI: 10.1002/14651858.CD006231]

4 **Hobbs MS**, Mai Q, Knuiman MW, Fletcher DR, Ridout SC. Surgeon experience and trends in intraoperative complications in laparoscopic cholecystectomy. *Br J Surg* 2006; **93**: 844-853 [PMID: 16671070 DOI: 10.1002/bjs.5333]

5 **Giger UF**, Michel JM, Opitz I, Th Inderbitzin D, Kocher T, Krähenbühl L. Risk factors for perioperative complications in patients undergoing laparoscopic cholecystectomy: analysis of 22,953 consecutive cases from the Swiss Association of Laparoscopic and Thoracoscopic Surgery database. *J Am Coll Surg* 2006; **203**: 723-728 [PMID: 17084335 DOI: 10.1016/j.jamcollsurg.2006.07.018]

6 **Halm EA**, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Ann Intern Med* 2002; **137**: 511-520 [PMID: 12230353 DOI: 10.7326/0003-4819-137-6-200209170-00012]

7 **Post PN**, Wittenberg J, Burgers JS. Do specialized centers and specialists produce better outcomes for patients with chronic diseases than primary care generalists? A systematic review. *Int J Qual Health Care* 2009; **21**: 387-396 [PMID: 19734175 DOI: 10.1093/intqhc/mzp039]

8 **Davoli M**, Amato L, Minozzi S, Bargagli AM, Vecchi S, Perucci CA. [Volume and health outcomes: an overview of systematic reviews]. *Epidemiol Prev* 2005; **29**: 3-63 [PMID: 16529350]

9 **Shi HY**, Lee KT, Chiu CC, Lee HH. The volume-outcome relationship in laparoscopic cholecystectomy: a population-based study using propensity score matching. *Surg Endosc* 2013; **27**: 3139-3145 [PMID: 23620382 DOI: 10.1007/s00464-013-2867-x]

10 **Nuzzo G**, Giuliante F, Giovannini I, Ardito F, D'Acapito F,

Vellone M, Murazio M, Capelli G. Bile duct injury during laparoscopic cholecystectomy: results of an Italian national survey on 56 591 cholecystectomies. *Arch Surg* 2005; **140**: 986-992 [PMID: 16230550 DOI: 10.1001/archsurg.140.10.986]
11 **Csikesz NG**, Singla A, Murphy MM, Tseng JF, Shah SA. Surgeon volume metrics in laparoscopic cholecystectomy. *Dig Dis Sci* 2010; **55**: 2398-2405 [PMID: 19911275 DOI: 10.1007/s10620-009-1035-6]
12 **McMahon AJ**, Fischbacher CM, Frame SH, MacLeod MC. Impact of laparoscopic cholecystectomy: a population-based study. *Lancet* 2000; **356**: 1632-1637 [PMID: 11089821 DOI: 10.1016/S0140-6736(00)03156-1]
13 **Moher D**, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151**: 264-269, W64 [PMID: 19622511 DOI: 10.7326/0003-4819-151-4-200908180-00135]
14 **Szego T**, Roll S, Nogueira Filho WS, Bensenor F. Video-laparoscopic cholecystectomy. Report of the first Brazilian series. *Arq Gastroenterol* 1991; **28**: 6-8 [PMID: 1843089]
15 **Bailey RW**, Zucker KA, Flowers JL, Scovill WA, Graham SM, Imbembo AL. Laparoscopic cholecystectomy. Experience with 375 consecutive patients. *Ann Surg* 1991; **214**: 531-540; discussion 540-541 [PMID: 1835346]
16 **Peters JH**, Ellison EC, Innes JT, Liss JL, Nichols KE, Lomano JM, Roby SR, Front ME, Carey LC. Safety and efficacy of laparoscopic cholecystectomy. A prospective analysis of 100 initial patients. *Ann Surg* 1991; **213**: 3-12 [PMID: 1824674]
17 **Rees BI**, Williams HR. Laparoscopic cholecystectomy: the first 155 patients. *Ann R Coll Surg Engl* 1992; **74**: 233-236 [PMID: 1416671]
18 **Huang CS**, Tai FC, Shi MY, Chen DF, Wang NY. Complications of laparoscopic cholecystectomy: an analysis of 200 cases. *J Formos Med Assoc* 1992; **91**: 785-792 [PMID: 1362118]
19 **Davis CJ**, Arregui ME, Nagan RF, Shaar C. Laparoscopic cholecystectomy: the St. Vincent experience. *Surg Laparosc Endosc* 1992; **2**: 64-68 [PMID: 1341504]
20 **Fielding GA**. Laparoscopic cholecystectomy. *Aust N Z J Surg* 1992; **62**: 181-187 [PMID: 1532305]
21 **Soper NJ**, Stockmann PT, Dunnegan DL, Ashley SW. Laparoscopic cholecystectomy. The new 'gold standard'? *Arch Surg* 1992; **127**: 917-921; discussion 921-923 [PMID: 1386505 DOI: 10.1001/archsurg.1992.01420080051008]
22 **Périssat J**, Collet D, Belliard R, Desplantez J, Magne E. Laparoscopic cholecystectomy: the state of the art. A report on 700 consecutive cases. *World J Surg* 1992; **16**: 1074-1082 [PMID: 1455876 DOI: 10.1007/BF0206704]
23 **Troidl H**, Spangenberger W, Langen R, al-Jaziri A, Eypasch E, Neugebauer E, Dietrich J. Laparoscopic cholecystectomy: technical performance, safety and patient's benefit. *Endoscopy* 1992; **24**: 252-261 [PMID: 1535313 DOI: 10.1055/s-2007-1010477]
24 **Rubio PA**. Laparoscopic cholecystectomy: experience in 500 consecutive cases. *Int Surg* 1993; **78**: 277-279 [PMID: 8175250]
25 **Huang SM**, Wu CW, Hong HT, Ming-Liu KL, Lui WY. Bile duct injury and bile leakage in laparoscopic cholecystectomy. *Br J Surg* 1993; **80**: 1590-1592 [PMID: 8298933]
26 **Cox MR**, Wilson TG, Jeans PL, Padbury RT, Toouli J. Minimizing the risk of bile duct injury at laparoscopic cholecystectomy. *World J Surg* 1994; **18**: 422-446; discussion 426-427 [PMID: 8091785 DOI: 10.1007/BF00316827]
27 **Williams GB**, Silverman RS. Laparoscopic cholecystectomy in a community hospital: experience with 600 laparoscopic cholecystectomies. *J Laparoendosc Surg* 1994; **4**: 101-107 [PMID: 8043916]
28 **Cappuccino H**, Cargill S, Nguyen T. Laparoscopic cholecystectomy: 563 cases at a community teaching hospital and a review of 12,201 cases in the literature. Monmouth Medical Center Laparoscopic Cholecystectomy Group. *Surg Laparosc Endosc* 1994; **4**: 213-221 [PMID: 8044366]
29 **Bonatsos G**, Leandros E, Dourakis N, Birbas C, Delibaltadakis G, Golematis B. Laparoscopic cholecystectomy. Intraoperative findings and postoperative complications. *Surg Endosc* 1995; **9**: 889-893 [PMID: 8525441 DOI: 10.1007/BF00768885]
30 **Newman CL**, Wilson RA, Newman L, Eubanks S, Duncan TD, Mason EM, Wilson JP, Lucas GW. 1525 laparoscopic cholecystectomies without biliary injury: a single institution's experience. *Am Surg* 1995; **61**: 226-228 [PMID: 7887534]
31 **Chen XR**, Lou D, Li SH, Mao JX, Zhou ZD, Yu SM, Duan ZW. Avoiding serious complications in laparoscopic cholecystectomy--lessons learned from an experience of 2428 cases. *Ann Acad Med Singapore* 1996; **25**: 635-639 [PMID: 8923993]
32 **Bond G**, De Costa A. Laparoscopic cholecystectomy: the experience of community hospital. *Aust N Z J Surg* 1996; **66**: 14-17 [PMID: 8629972 DOI: 10.1111/j.1445-2197.1996.tb00692.x]
33 **Sartori CA**, Franzato B. Laparoscopic cholecystectomy. Analysis of results of first 322 cases. *Minerva Gastroenterol Dietol* 1996; **42**: 201-205 [PMID: 17912211]
34 **Rather GM**, Ravi VK. Audit of laparoscopic cholecystectomies in a district general hospital. *Saudi J Gastroenterol* 1997; **3**: 15-21 [PMID: 19864808]
35 **Jan YY**, Chen HM, Wang CS, Chen MF. Biliary complications during and after laparoscopic cholecystectomy. *Hepatogastroenterology* 1997; **44**: 370-375 [PMID: 9164503]
36 **Ahmad SA**, Schuricht AL, Azurin DJ, Arroyo LR, Paskin DL, Bar AH, Kirkland ML. Complications of laparoscopic cholecystectomy: the experience of a university-affiliated teaching hospital. *J Laparoendosc Adv Surg Tech A* 1997; **7**: 29-35 [PMID: 9453862 DOI: 10.1089/lap.1997.7.29]
37 **Kok KY**, Mathew VV, Tan KK, Yapp SK. A prospective review of laparoscopic cholecystectomy in Brunei. *Surg Laparosc Endosc* 1998; **8**: 120-122 [PMID: 9566565]
38 **Targarona EM**, Marco C, Balagué C, Rodriguez J, Cugat E, Hoyuela C, Veloso E, Trias M. How, when, and why bile duct injury occurs. A comparison between open and laparoscopic cholecystectomy. *Surg Endosc* 1998; **12**: 322-326 [PMID: 9543521 DOI: 10.1007/s004649900662]
39 **Kurauchi N**, Kamii N, Kazui K, Saji Y, Uchino J. Laparoscopic cholecystectomy: a report on the community hospital experience in Hokkaido. *Surg Today* 1998; **28**: 714-718 [PMID: 9697264 DOI: 10.1007/BF02484617]
40 **Jones-Monahan K**, Gruenberg JC. Bile duct injuries during laparoscopic cholecystectomy: a community's experience. *Am Surg* 1998; **64**: 638-642 [PMID: 9655274]
41 **Matthews BD**, Williams GB. Laparoscopic cholecystectomy in an academic hospital: evaluation of changes in perioperative outcomes. *JSLS* 1999; **3**: 9-17 [PMID: 10323163]
42 **Calvete J**, Sabater L, Camps B, Verdú A, Gomez-Portilla A, Martín J, Torrico MA, Flor B, Cassinello N, Lledó S. Bile duct injury during laparoscopic cholecystectomy: myth or reality of the learning curve? *Surg Endosc* 2000; **14**: 608-611 [PMID: 10948294 DOI: 10.1007/s004640000103]
43 **Patel SC**, Bhatt JR. Laparoscopic cholecystectomy at the Aga Khan Hospital, Nairobi. *East Afr Med J* 2000; **77**: 194-198 [PMID: 12858902 DOI: 10.4314/eamj.v77i4.46621]
44 **Sikora SS**, Kumar A, Das NR, Sarkari A, Saxena R, Kapoor VK. Laparoscopic bile duct injuries: spectrum at a tertiary-care center. *J Laparoendosc Adv Surg Tech A* 2001; **11**: 63-68 [PMID: 11327128 DOI: 10.1089/109264201750162239]
45 **Lichten JB**, Reid JJ, Zahalsky MP, Friedman RL. Laparoscopic cholecystectomy in the new millennium. *Surg Endosc* 2001; **15**: 867-872 [PMID: 11443440 DOI: 10.1007/s004640080004]
46 **Miroshnik M**, Saafan A, Koh S, Farlow J, Neophyton J, Lizzio J, Yee F, Ethell T, Bean A, Fenton-Lee D. Biliary tract injury in laparoscopic cholecystectomy: results of a single unit. *ANZ J Surg* 2002; **72**: 867-870 [PMID: 12485222 DOI: 10.1046/j.1445-2197.2002.02587.x]
47 **Hasaniah WF**, Ghada I, Sabah AH, Sulaiman AH, Jamal MJ, Derar SA. Laparoscopic cholecystectomy in 2,750 cases in a teaching hospital in Kuwait. *Med Princ Pract* 2002; **11**:

176-179 [PMID: 12424410 DOI: 10.1159/000065811]
48 **Fathy O**, Zeid MA, Abdallah T, Fouad A, Eleinien AA, el-Hak NG, Eleibiedy G, el-Wahab MA, Sultan A, Anwar N, Ezzat F. Laparoscopic cholecystectomy: a report on 2000 cases. *Hepatogastroenterology* 2003; **50**: 967-971 [PMID: 12845960]
49 **Duca S**, Bălă O, Al-Hajjar N, Lancu C, Puia IC, Munteanu D, Graur F. Laparoscopic cholecystectomy: incidents and complications. A retrospective analysis of 9542 consecutive laparoscopic operations. *HPB* (Oxford) 2003; **5**: 152-158 [PMID: 18332976 DOI: 10.1080/13651820310015293]
50 **Mahatharadol V**. Bile duct injuries during laparoscopic cholecystectomy: an audit of 1522 cases. *Hepatogastroenterology* 2004; **51**: 12-14 [PMID: 15011821]
51 **Daradkeh S**. Laparoscopic cholecystectomy: analytical study of 1208 cases. *Hepatogastroenterology* 2005; **52**: 1011-1014 [PMID: 16001618]
52 **Diamantis T**, Tsigris C, Kiriakopoulos A, Papalambros E, Bramis J, Michail P, Felekouras E, Griniatsos J, Rosenberg T, Kalahanis N, Giannopoulos A, Bakoyiannis C, Bastounis E. Bile duct injuries associated with laparoscopic and open cholecystectomy: an 11-year experience in one institute. *Surg Today* 2005; **35**: 841-845 [PMID: 16175465 DOI: 10.1007/s00595-005-3038-z]
53 **Söderlund C**, Frozanpor F, Linder S. Bile duct injuries at laparoscopic cholecystectomy: a single-institution prospective study. Acute cholecystitis indicates an increased risk. *World J Surg* 2005; **29**: 987-993 [PMID: 15977078 DOI: 10.1007/s00268-005-7871-4]
54 **Baird DR**, Wilson JP, Mason EM, Duncan TD, Evans JS, Luke JP, Ruben DM, Lucas GW. An early review of 800 laparoscopic cholecystectomies at a university-affiliated community teaching hospital. *Am Surg* 1992; **58**: 206-210 [PMID: 1532704]
55 **Vagenas K**, Karamanakos SN, Spyropoulos C, Panagiotopoulos S, Karanikolas M, Stavropoulos M. Laparoscopic cholecystectomy: a report from a single center. *World J Gastroenterol* 2006; **12**: 3887-3890 [PMID: 16804976 DOI: 10.3748/wjg.v12.i24.3887]
56 **Tan JT**, Suyapto DR, Neo EL, Leong PS. Prospective audit of laparoscopic cholecystectomy experience at a secondary referral centre in South australia. *ANZ J Surg* 2006; **76**: 335-338 [PMID: 16768693 DOI: 10.1111/j.1445-2197.2006.03721.x]
57 **Mufti TS**, Ahmad S, Naveed D, Akbar M, Zafar A. Laparoscopic cholecystectomy: an early experience at Ayub Teaching Hospital Abbottabad. *J Ayub Med Coll Abbottabad* 2007; **19**: 42-44 [PMID: 18693596]
58 **Brekalo Z**, Innocenti P, Duzel G, Liddo G, Ballone E, Simunović VJ. Ten years of laparoscopic cholecystectomy: a comparison between a developed and a less developed country. *Wien Klin Wochenschr* 2007; **119**: 722-728 [PMID: 18157606 DOI: 10.1007/s00508-007-0906-0]
59 **Marakis GN**, Pavlidis TE, Ballas K, Aimoniotou E, Psarras K, Karvounaris D, Rafailidis S, Demertzidis H, Sakantamis AK. Major complications during laparoscopic cholecystectomy. *Int Surg* 2007; **92**: 142-146 [PMID: 17972469]
60 **Herve J**, Simoens Ch, Smets D, Ngongang Ch, Mendes da Costa P. Laparoscopic cholecystectomy; a retrospective 10-year study. *Hepatogastroenterology* 2007; **54**: 1326-1330 [PMID: 17708247]
61 **Shrestha S**, Pradhan G, Bhoomi K, Dhital A, Bhattachan CL. Review of laparoscopic cholecystectomy in Nepal Medical College Teaching Hospital. *Nepal Med Coll J* 2007; **9**: 32-35 [PMID: 17593675]
62 **Tantia O**, Jain M, Khanna S, Sen B. Iatrogenic biliary injury: 13,305 cholecystectomies experienced by a single surgical team over more than 13 years. *Surg Endosc* 2008; **22**: 1077-1086 [PMID: 18210186 DOI: 10.1007/s00464-007-9740-8]
63 **Priego P**, Ramiro C, Molina JM, Rodríguez Velasco G, Lobo E, Galindo J, Fresneda V. Results of laparoscopic cholecystectomy in a third-level university hospital after 17 years of experience. *Rev Esp Enferm Dig* 2009; **101**: 20-30 [PMID: 19335030]
64 **Avgerinos C**, Kelgiorgi D, Touloumis Z, Baltatzi L, Dervenis C. One thousand laparoscopic cholecystectomies in a single surgical unit using the "critical view of safety" technique. *J Gastrointest Surg* 2009; **13**: 498-503 [PMID: 19009323 DOI: 10.1007/s11605-008-0748-8]
65 **Clegg-Lamptey JN**, Amponsah G. Laparoscopic cholecystectomy at the Korle Bu Teaching Hospital, Accra, Ghana: an initial report. *West Afr J Med* 2010; **29**: 113-116 [PMID: 20544637]
66 **Al-Kubati WR**. Bile duct injuries following laparoscopic cholecystectomy: A clinical study. *Saudi J Gastroenterol* 2010; **16**: 100-104 [PMID: 20339179 DOI: 10.4103/1319-3767.61236]
67 **Ying F**, Shuodong W, Hong Y, Yang S, Jing K, Yu T, Amos SE. Lessons learnt after 12 years experience in laparoscopic cholecystectomy at a single center. *Hepatogastroenterology* 2010; **57**: 202-206 [PMID: 20583412]
68 **Zha Y**, Chen XR, Luo D, Jin Y. The prevention of major bile duct injures in laparoscopic cholecystectomy: the experience with 13,000 patients in a single center. *Surg Laparosc Endosc Percutan Tech* 2010; **20**: 378-383 [PMID: 21150413 DOI: 10.1097/SLE.0b013e3182008efb]
69 **Wichmann MW**, Lang R, Beukes E, Esufali ST, Jauch KW, Hüttl TK, Hüttl TP. Laparoscopic cholecystectomy--comparison of early postoperative results in an Australian rural centre and a German university hospital. *Langenbecks Arch Surg* 2010; **395**: 255-260 [PMID: 19937339 DOI: 10.1007/s00423-009-0569-6]
70 **Kanakala V**, Borowski DW, Pellen MG, Dronamraju SS, Woodcock SA, Seymour K, Attwood SE, Horgan LF. Risk factors in laparoscopic cholecystectomy: a multivariate analysis. *Int J Surg* 2011; **9**: 318-323 [PMID: 21333763 DOI: 10.1016/j.ijsu.2011.02.003]
71 **Al-Mulhim AS**, Amin TT. Outcome of laparoscopic cholecystectomy at a secondary level of care in Saudi Arabia. *Saudi J Gastroenterol* 2011; **17**: 47-52 [PMID: 21196653 DOI: 10.4103/1319-3767.74484]
72 **Halilovic H**, Hasukic S, Matovic E, Imamovic G. Rate of complications and conversions after laparoscopic and open cholecystectomy. *Med Arh* 2011; **65**: 336-338 [PMID: 22299293 DOI: 10.5455/medarh.2011.65.336-338]
73 **Hasbahceci M**, Uludag M, Erol C, Ozdemir A. Laparoscopic cholecystectomy in a single, non-teaching hospital: an analysis of 1557 patients. *J Laparoendosc Adv Surg Tech A* 2012; **22**: 527-532 [PMID: 22458833 DOI: 10.1089/lap.2012.0005]
74 **Bekele S**, Biluts H. Laparascopic cholecstectomy at Myungsung Christian Medical Center, Ethiopia: a five-years experience. *Ethiop Med J* 2012; **50**: 251-257 [PMID: 23409408]
75 **Le VH**, Smith DE, Johnson BL. Conversion of laparoscopic to open cholecystectomy in the current era of laparoscopic surgery. *Am Surg* 2012; **78**: 1392-1395 [PMID: 23265130]
76 **Afuwape OO**, Akute OO, Adebanjo AT. Preliminary experience with laparoscopic cholecystectomy in a nigerian teaching hospital. *West Afr J Med* 2012; **31**: 120-123 [PMID: 23208482]
77 **Grbas H**, Kunisek L, Zelić M, Petrosić N, Cepić I, Pirjavec A, Lovasić F, Uravić M. Outcome evaluation of 10,317 laparoscopic cholecystectomies: a 17-year experience at a single center. *Hepatogastroenterology* 2013; **60**: 1873-1876 [PMID: 24719920]
78 **Sultan AM**, El Nakeeb A, Elshehawy T, Elhemmaly M, Elhanafy E, Atef E. Risk factors for conversion during laparoscopic cholecystectomy: retrospective analysis of ten years' experience at a single tertiary referral centre. *Dig Surg* 2013; **30**: 51-55 [PMID: 23635600 DOI: 10.1159/000347164]
79 **Pulvirenti E**, Toro A, Gagner M, Mannino M, Di Carlo I. Increased rate of cholecystectomies performed with doubtful

or no indications after laparoscopy introduction: a single center experience. *BMC Surg* 2013; **13**: 17 [PMID: 23724992 DOI: 10.1186/1471-2482-13-17]

80 **Dip FD**, Asbun D, Rosales-Velderrain A, Lo Menzo E, Simpfendorfer CH, Szomstein S, Rosenthal RJ. Cost analysis and effectiveness comparing the routine use of intraoperative fluorescent cholangiography with fluoroscopic cholangiogram in patients undergoing laparoscopic cholecystectomy. *Surg Endosc* 2014; **28**: 1838-1843 [PMID: 24414461 DOI: 10.1007/s00464-013-3394-5]

81 **Comajuncosas J**, Hermoso J, Gris P, Jimeno J, Orbeal R, Vallverdú H, López Negre JL, Urgellés J, Estalella L, Parés D. Risk factors for umbilical trocar site incisional hernia in laparoscopic cholecystectomy: a prospective 3-year follow-up study. *Am J Surg* 2014; **207**: 1-6 [PMID: 24112669 DOI: 10.1016/j.amjsurg.2013.05.010]

82 **Paajanen H**, Suuronen S, Eskelinen M, Hytonen S, Juvonen P. Frequency of bile leak after laparoscopic cholecystectomy: audit of a surgical residency program. *Am Surg* 2014; **80**: 91-94 [PMID: 24401523]

83 **Alvarez FA**, de Santibañes M, Palavecino M, Sánchez Clariá R, Mazza O, Arbues G, de Santibañes E, Pekolj J. Impact of routine intraoperative cholangiography during laparoscopic cholecystectomy on bile duct injury. *Br J Surg* 2014; **101**: 677-684 [PMID: 24664658 DOI: 10.1002/bjs.9486]

84 **Afaneh C**, Abelson J, Rich BS, Dakin G, Zarnegar R, Barie PS, Fahey TJ, Pomp A. Obesity does not increase morbidity of laparoscopic cholecystectomy. *J Surg Res* 2014; **190**: 491-497 [PMID: 24636101 DOI: 10.1016/j.jss.2014.02.014]

85 **Murphy MM**, Ng SC, Simons JP, Csikesz NG, Shah SA, Tseng JF. Predictors of major complications after laparoscopic cholecystectomy: surgeon, hospital, or patient? *J Am Coll Surg* 2010; **211**: 73-80 [PMID: 20610252 DOI: 10.1016/j.jamcollsurg.2010.02.050]

86 **Harrison EM**, O'Neill S, Meurs TS, Wong PL, Duxbury M, Paterson-Brown S, Wigmore SJ, Garden OJ. Hospital volume and patient outcomes after cholecystectomy in Scotland: retrospective, national population based study. *BMJ* 2012; **344**: e3330 [PMID: 22623634 DOI: 10.1136/bmj.e3330]

87 **Khuri SF**, Daley J, Henderson W, Hur K, Hossain M, Soybel D, Kizer KW, Aust JB, Bell RH, Chong V, Demakis J, Fabri PJ, Gibbs JO, Grover F, Hammermeister K, McDonald G, Passaro E, Phillips L, Scamman F, Spencer J, Stremple JF. Relation of surgical volume to outcome in eight common operations: results from the VA National Surgical Quality Improvement Program. *Ann Surg* 1999; **230**: 414-429; discussion 429-432 [PMID: 10493488]

88 **Buanes T**, Mjåland O, Waage A, Langeggen H, Holmboe J. A population-based survey of biliary surgery in Norway. Relationship between patient volume and quality of surgical treatment. *Surg Endosc* 1998; **12**: 852-855 [PMID: 9602005 DOI: 10.1007/s004649900728]

P- Reviewer: Goetze TO, Peng B, Rabago L **S- Editor**: Ji FF
L- Editor: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.36

World J Meta-Anal 2015 February 26; 3(1): 36-42
ISSN 2308-3840 (online)

META-ANALYSIS

Antithrombotic treatment in chronic heart failure and sinus rhythm: Systematic review

Daniel Caldeira, Inês Cruz, Rita Calé, Cristina Martins, Helder Pereira, Joaquim J Ferreira, Fausto J Pinto, João Costa

Daniel Caldeira, Inês Cruz, Rita Calé, Cristina Martins, Helder Pereira, Cardiology Department, Hospital Garcia de Orta, 2805-267 Almada, Portugal
Daniel Caldeira, Joaquim J Ferreira, João Costa, Clinical Pharmacology Unit, Instituto de Medicina Molecular, 1649-028 Lisbon, Portugal
Daniel Caldeira, Joaquim J Ferreira, João Costa, Laboratory of Clinical Pharmacology, Faculty of Medicine, University of Lisbon, 1649-028 Lisbon, Portugal
Fausto J Pinto, Cardiology Department, CCUL, CAML, Faculty of Medicine, 1649-028 Lisbon, Portugal
João Costa, Evidence Based Medicine Centre, Faculty of Medicine, University of Lisbon, 1649-028 Lisbon, Portugal
João Costa, Portuguese Collaborating Centre of the Cochrane Iberoamerican Network, Faculty of Medicine, University of Lisbon, 1649-028 Lisbon, Portugal
Author contributions: Caldeira D and Costa J contributed to the concept and design of the study and acquisition, analysis and interpretation of the data, wrote the first draft of the manuscript, critically revised the manuscript, and gave final approval of the submitted manuscript; Ferreira JJ and Pinto FJ contributed to the concept and design of the study and interpretation of the data, critically revised the manuscript, and gave final approval of the submitted manuscript; Cruz I, Calé R, Martins C and Pereira H contributed to interpretation of data, critically revised the manuscript, and gave final approval of the submitted manuscript.

Correspondence to: Dr. Daniel Caldeira, Laboratory of Clinical Pharmacology, Faculty of Medicine, University of Lisbon, Av. Prof. Egas Moniz, 1649-028 Lisbon, Portugal. dgcaldeira@hotmail.com
Telephone: +351-21-7973453
Fax: +351-21-7819688
Received: May 20, 2014
Peer-review started: May 23, 2014
First decision: July 10, 2014
Revised: October 30, 2014
Accepted: November 17, 2014
Article in press: November 19, 2014
Published online: February 26, 2015

Abstract

AIM: To assess the efficacy and safety of antithrombotic drugs (antiplatelet or anticoagulant drugs) compared to no antithrombotic treatment or placebo in patients with heart failure (HF) and sinus rhythm.

METHODS: We searched Medline and Cochrane Library for randomized controlled trials evaluating antithrombotic treatment and no antithrombotic treatment in patients with HF and sinus rhythm. Risk ratio (RR) and 95%CIs were estimated performing meta-analysis with random effects method.

RESULTS: Two studies met the inclusion criteria: Heart failure Long-term Antithrombotic Study and Warfarin/Aspirin Study in Heart failure, with 336 patients and mean follow-up 1.8-2.25 years. Stroke risk was not reduced by acetylsalicylic acid (RR = 1.18, 95%CI: 0.17-8.15), oral anticoagulation (RR = 0.30, 95%CI: 0.03-2.65) or overall antithrombotic drugs (RR = 0.52, 95%CI: 0.10-2.74). Acetylsalicylic acid showed a significant increased risk of worsening HF (RR = 1.78, 95%CI: 1.08-2.92), while oral anticoagulation had no impact in this outcome (RR = 1.03, 95%CI: 0.61-1.75). Overall antithrombotic drugs showed a significant risk increase of major bleeding (RR = 6.99, 95%CI: 0.89-54.64).

CONCLUSION: Best available evidence does not support the routine use of antithrombotic drugs in patients with HF and sinus rhythm. These drugs, particularly oral anti-

coagulation has the hazard of increase significantly major bleeding risk.

Key words: Heart failure; Sinus rhythm; Platelet aggregation inhibitors; Anticoagulants

Core tip: In patients with atrial fibrillation, chronic heart failure (CHF) increases thromboembolic risk and oral anticoagulation is essential to decrease the risk of thromboembolic complications. Evidence suggests a positive association between CHF, impaired hemostasis and thromboembolic events. Whether antithrombotic drugs should be recommended for these patients (in sinus rhythm) is still debated. We looked for the best available evidence and we found 2 studies fulfilling the inclusion criteria. We performed a meta-analysis of antithrombotic drugs *vs* placebo and strengthened that antithrombotic drugs do not decrease the risk of stroke (fatal or non-fatal) and increase the risk of major bleeding.

Caldeira D, Cruz I, Calé R, Martins C, Pereira H, Ferreira JJ, Pinto FJ, Costa J. Antithrombotic treatment in chronic heart failure and sinus rhythm: Systematic review. *World J Meta-Anal* 2015; 3(1): 36-42 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/36.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.36

INTRODUCTION

Chronic heart failure (CHF) is an increasingly prevalent cardiovascular disease with significant associated morbidity and mortality[1]. CHF constitutes a significant economic burden[2,3], which is expected to increase over the next decades due to increasing prevalence of associated diseases and risk factors as well as population aging. Former observational studies suggest a positive association between CHF, impaired hemostasis and thromboembolic events[4,5]. In patients with atrial fibrillation (AF), CHF increases thromboembolic risk and oral anticoagulation is the cornerstone of AF treatment aiming to decrease the risk of thromboembolic complications[6]. The results from the WARCEF trial (Warfarin *vs* Aspirin in Reduced Cardiac Ejection Fraction) has highlighted the role of antithrombotic treatment in patients with CHF and sinus rhythm[7]. There were no differences between warfarin and acetylsalicylic acid in the primary outcome (time to the first event in a composite end point of ischemic stroke, intracerebral hemorrhage, or death from any cause). However, warfarin was associated with fewer stroke events (2.5% *vs* 4.7%) but also with a higher rate of major bleeding events (5.8% *vs* 2.7%). The clinical interpretation of these findings was that the choice between warfarin and aspirin should be made on the basis of the individual patient[8].

Previous systematic reviews with meta-analyses comparing oral anticoagulation (namely warfarin) and acetylsalicylic acid in patients with CHF and sinus rhythm reached conclusions overlapping those from the WARCEF study[9-13].

Although much effort have been done comparing and discussing the relative effectiveness of oral anticoagulation *vs* acetylsalicylic acid in patients with CHF and sinus rhythm, significantly less is known about the true efficacy of the overall antithrombotic treatment. Therefore, we aimed to perform a systematic review to better estimate the true clinical benefit of antithrombotic treatments (oral anticoagulation or antiplatelet drugs) against placebo, standard care or no treatment, in patients with CHF and sinus rhythm.

MATERIALS AND METHODS

Guidance

This work followed PRISMA guidelines for systematic reviews and meta-analyses promoted by the EQUATOR network[14].

Eligibility criteria

We have searched for all randomized controlled trials (RCTs) evaluating patients with CHF and sinus rhythm treated with oral antithrombotic therapy or control. We considered for antithrombotic treatments both oral anticoagulants (such as vitamin K antagonists, like warfarin, acenocoumarol or phenprocoumon) and antiplatelet drugs [such as acetylsalicylic acid (ASA), clopidogrel or ticlopidine]. We allowed controls under placebo, standard care or no antithrombotic treatment. Studies had to report clinical and/or echocardiographic features for the enrolled CHF patients, such as impaired left ventricle ejection fraction or shortening fraction.

Database and search method

Medline and Cochrane Library (CENTRAL) databases were searched from inception to November 2013 for eligible studies. The search strategy details are available at the Online Supplementary Material. We considered all studies irrespective of language. References of obtained studies were also comprehensively searched and cross-checked to identify possible missing studies.

Studies and data selection

Citations obtained from electronic search were independently screened by two authors, followed by full-text assessment of potentially eligible studies for inclusion in accordance with previously mentioned criteria.

Primary outcome was stroke (fatal or non-fatal). Secondary outcomes were all-cause mortality, myocardial infarction, worsening heart failure (HF), major bleeding and a composite of major adverse clinical events, defined as the combination of mortality, stroke, myocardial infarction and HF.

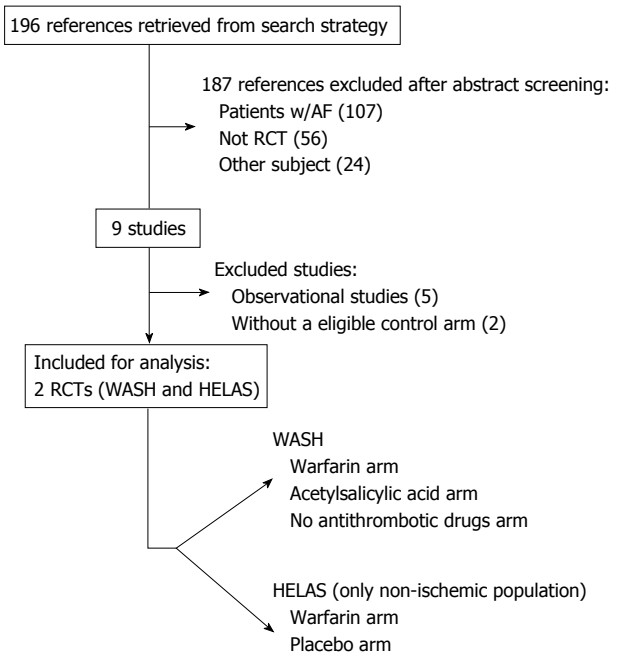


Figure 1 Flowchart of studies' selection. AF: Atrial fibrillation; RCT: Randomized controlled trial; WASH: Warfarin/Aspirin Study in Heart failure; HELAS: Heart failure Long-term Antithrombotic Study.

We extracted detailed data about demographics, comorbidities, interventions, follow-up and outcomes. Data extraction and data entry into software was double-checked. Disagreements were resolved by consensus.

Quality reporting assessment

Quality of reporting was analysed by using a qualitative classification according to risk of bias (high, unclear or low risk), adapted from Cochrane Collaboration's Tool[15]. Studies were not excluded based on quality of reporting.

Statistical analysis

Outcomes data were summarized as frequencies. Statistical analyses were performed using the RevMan version 5.2.6 (The Nordic Cochrane Centre, The Cochrane Collaboration, 2012) to derive forest plots with pooled estimates of risk ratios (RR) and their 95%CI. Statistical heterogeneity was assessed with χ^2 test and quantified with Higgins I^2 test[16]. Pooled results estimates were based on the random or fixed effects model according to the existence ($I^2 \geqslant 50\%$) or not ($I^2 < 50\%$) of significant heterogeneity[17]. Publication bias was assessed through visual inspection of funnel plots symmetry and Peters' regression tests[18,19]. Pooled results were evaluated for the overall antithrombotic treatment, as well separately for antiplatelet and anticoagulation groups.

RESULTS

Search

After title and abstract screening of citations obtained in Medline and Cochrane Library, 196 citations were retrieved. One-hundred and eighty seven studies did not meet inclusion criteria through initial assessment: 107 included AF patients; 56 studies were not randomized and 24 did not address the pretended topic (either different population and/or other interventions).

The remaining 9 studies were fully-evaluated, of which 7 were further excluded: 5 were observational studies, and 2 RCTs did not include a placebo, standard care or no antithrombotic treatment arm (WARCEF and WATCH trials)[5,20]. Therefore, 2 RCTs were eligible for the purpose of this systematic review[21,22]. The search of reference lists of review articles and included studies failed to identify any additional eligible study[23-27]. Figure 1 shows the flowchart of studies' selection.

Characteristics of obtained studies and quality of reporting

Studies Warfarin/Aspirin Study in Heart failure (WASH) and Heart failure Long-term Antithrombotic Study (HELAS) met the outlined inclusion criteria[21,22].

WASH study was an open-label RCT with blinded endpoint assessment, published in 2004. WASH enrolled 254 patients (80 warfarin; 80 ASA; 94 no anti-thrombotic treatment) with CHF and sinus rhythm and followed them for a mean period of 2.25 years. About 60% had CHF of ischemic etiology, 75% of the patients were male, mean age was 63 years old, and 30% were in New York Heart Association class Ⅲ/Ⅵ. About 34% of the patients had hypertension, and 20% had diabetes. In terms of echocardiography mean parameters, patients had a fractional shortening of 15% and a left-ventricular end-diastolic diameter of 66 mm. Regarding treatments, the daily dosage of acetylsalicylic acid was 300 mg and international normalized ratio (INR) target for warfarin-treated patients was 2.5 (range 2-3). Primary outcome was the composite of all-cause death, non-fatal myocardial infarction, or non-fatal stroke[21].

HELAS study was published in 2006 and included two comparisons: warfarin *vs* acetylsalicylic acid in patients with CHF of ischemic etiology (not evaluated in this review due to absence of a placebo/no treatment control arm); and warfarin *vs* placebo in 82 patients (38 *vs* 44) with dilated non-ischemic CHF in sinus rhythm. Study's mean follow was 1.8 years. Most of the patients were male and mean age was 55 years. Hypertension was present in 25% of the patients, and diabetes in 11%. No significant differences were noticed in the main baseline characteristics. Echocardiographic features of these patients were remarkable for a baseline ejection fraction of 28%, left ventricle end-systolic diameter of 58 mm and end-diastolic diameter of 70 mm. Target INR for warfarin treatment was 2-3. Primary outcome was the composite of all-cause mortality, non-fatal stroke, non-fatal myocardial infarction, peripheral or pulmonary embolism, hospitalisation, or HF worsening[22].

Quality of reporting assessment is available in Figure 2. The main methodological flaws were the open-label design of WASH and the unknown method of allocation concealment in HELAS.

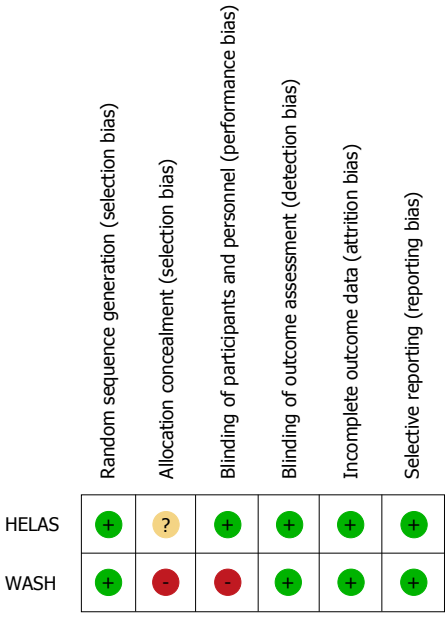


Figure 2 Studies quality of reporting. WASH: Warfarin/Aspirin Study in Heart failure; HELAS: Heart failure Long-term Antithrombotic Study.

Quantitative evaluation

Meta-analysis was performed for the following comparisons: antiplatelet drugs *vs* control, anticoagulant drugs *vs* control, and antithrombotic drugs (antiplatelet plus anticoagulant drugs) *vs* control.

While anticoagulation *vs* control data was derived from both WASH and HELAS studies[21,22], WASH study was the only that provided data for antiplatelet (acetylsalicylic acid) *vs* placebo[21]. For quantitative evaluation of overall antithrombotic treatment in this population, we considered both oral anticoagulation and antiplatelet from WASH study as a single arm and efficacy was directly obtained from WASH study[21].

Primary outcome

Antithrombotic drugs did not reduce stroke risk against placebo or no treatment, with RR = 1.18 (95%CI: 0.17-8.15) for antiplatelet drugs, RR = 0.30 (95%CI: 0.03-2.65) for anticoagulants, and RR = 0.52 (95%CI: 0.10-2.74) for overall antithrombotic drugs.

Secondary outcomes

Antithrombotic drugs showed an increased risk of CHF worsening (RR = 1.61, 95%CI: 1.04-2.48), mainly due to the single antiplatelet drug studied, acetylsalicylic acid, which had RR = 1.78 (95%CI: 1.08-2.92), while oral anticoagulants were not different from controls (RR = 1.03, 95%CI: 0.61-1.75).

Warfarin showed a significant increased risk of major bleeding (RR = 9.00, 95%CI: 1.14-70.90) and acetylsalicylic acid showed a non-significant trend (RR = 3.26, 95%CI: 0.13-79.04). The RR for overall major bleeding risk with antithrombotic drugs was 6.99 (95%CI: 0.89-54.64).

None of the antithrombotic drugs or overall antithrombotic treatment showed reduction of mortality and myocardial infarction risk in patients with systolic HF and sinus rhythm.

Antiplatelet drug/acetylsalicylic acid, but not warfarin, showed increased risk of the composite outcome of mortality, stroke, myocardial infarction, and worsening HF, most probably due to the increased risk of CHF worsening. Statistical heterogeneity was present in the evaluation of mortality when comparing antithrombotic drugs with control (I^2 = 58%). Figure 3 shows the pooled results. Publication bias was not evaluated due to the scarcity of studies[28].

DISCUSSION

Our main findings were the lack of proven efficacy of antithrombotic treatments, in patients with systolic HF and sinus rhythm, in the risk reduction of clinically important outcomes such as stroke, mortality and myocardial infarction; moreover, warfarin is associated to a significant 9-fold increased risk of major bleeding; and acetylsalicylic acid was associated with increased risk of CHF worsening.

The spotlight of this theme looks for Warfarin *vs* Acetylsalicylic acid comparison. By conducting this systematic review, the authors aimed to move back to the original problem and ask the question of whether antithrombotic treatments are, in the first place, effective in the treatment of CHF with sinus rhythm. If we accept that RCTs are the unique type of clinical studies that can prove causality with a reasonable margin of error, our results show that these interventions still have to prove their efficacy in this population, knowing that they owe an important bleeding risk. Furthermore, our attempt to perform a bayesian mixed treatment comparison meta-analysis, with data from clopidogrel arm from WATCH study[20], and warfarin *vs* acetylsalicylic acid presented in multiple systematic reviews and meta-analyses, failed due to high inconsistency in the statistical analysis of the network (data not shown). Although this inconsistence strongly compromises the results of such exercise, it is worth to report that placebo had a high probability of being the best treatment option. This reinforces the need of further trials to elucidate whether these interventions do/do not interfere with the prognosis, rather than have contradictory signs.

Accordingly, the 2012 consensus document of the HF Association of the European Society of Cardiology (ESC) and the ESC Working Group on Thrombosis corroborates our conclusions[29]. This consensus document stated that warfarin and acetylsalicylic acid should not be routinely used for thromboprophylaxis in patients with systolic HF and sinus rhythm, in the absence of concomitant comorbidities with clear indications for anticoagulation (*e.g.*, AF) or acetylsalicylic acid (*e.g.*, documented coronary artery disease).

Safety concerns regarding acetylsalicylic acid and HF (in patients with previously optimized background therapy with drugs such as angiotensin-converting enzyme inhibitors) were previously mentioned[30-32]. However if we

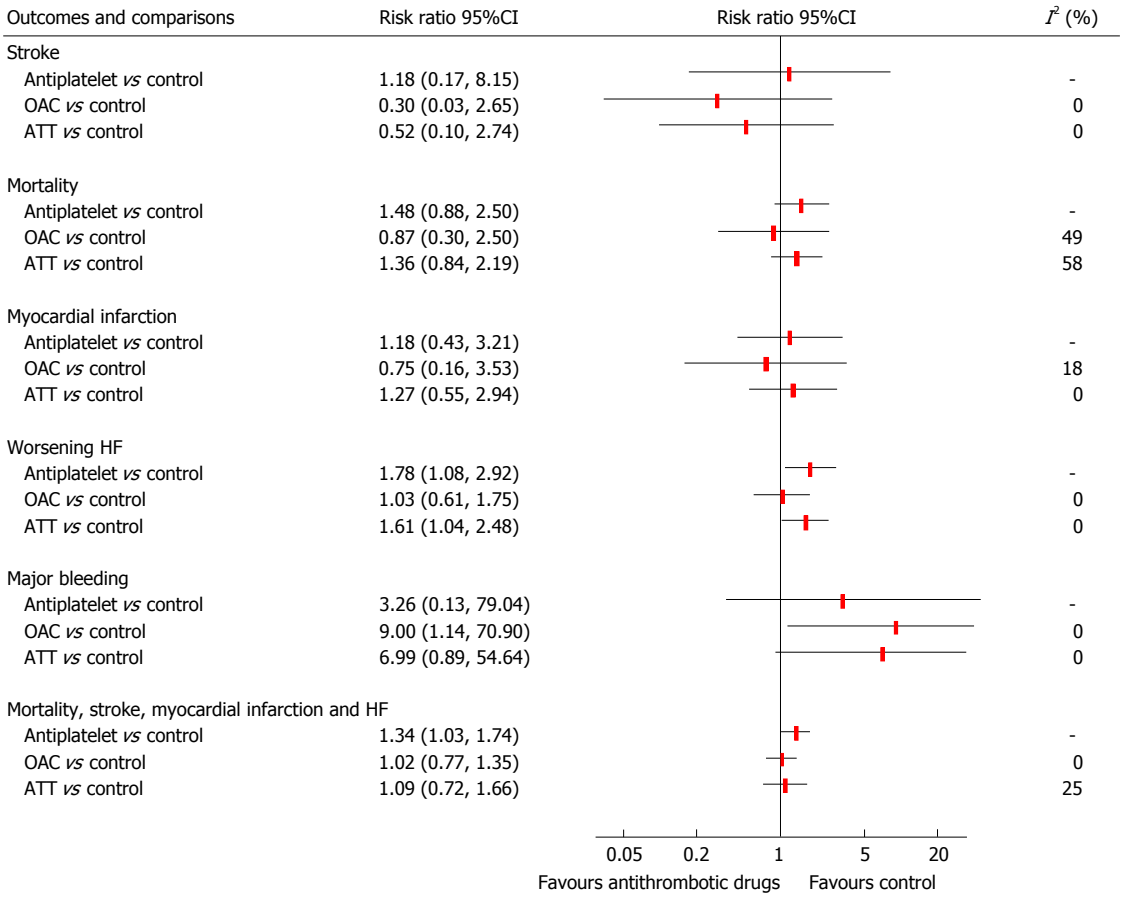


Figure 3 Forest plot evaluating antithrombotic drugs *vs* control. Data for "Antiplatelet *vs* control" comparison derived from WASH study. ATT: Antithrombotic treatment; OAC: Oral anticoagulation; HF: Heart failure; WASH: Warfarin/Aspirin Study in Heart failure.

consider warfarin as a "negative control", the pooled rates of HF worsening (after the WARCEF trial) were similar between acetylsalicylic acid and warfarin[7].

Along this century, antithrombotic treatment has gone forward in many therapeutic indications, but in patients with systolic HF and sinus rhythm the evidence to determine the prognostic importance of antithrombotic treatment (individually or globally) remained stationary and unsatisfactory for those who have to deal with CHF patients with sinus rhythm.

Limitations

This systematic review with meta-analysis has limitations attributed to included studies and analysis method.

As for included studies, WASH study had an open-label design; the control arm of this study was a no-antithrombotic treatment group (*i.e.*, not a placebo controlled trial), and included 7% of patients with AF that could not be excluded in the analyses. Furthermore the dosage of acetylsalicylic acid used in this trial was considerably higher than recommended[33].

Both studies had different proportions of HF etiologies. Although it can be important, particularly in ischemic HF cases where acetylsalicylic acid may play recognized prognostic role, here we aimed evaluate the thrombotic and embolic risk of patients with clinically important left ventricle impairment.

Major bleeding definitions were not common along the included trials. Worsening HF was defined by the investigator in WASH and no definition was provided in HELAS.

Periods of unrecognized paroxysmal AF could have biased of results. However it would bias favouring the antithrombotic drugs, which did not occur.

In conclusions, current evidence does not support the routine use of antithrombotic drugs (anticoagulant or antiplatelet drugs) for thromboprophylaxis in patients with systolic HF and sinus rhythm, as it carries a well known and documented bleeding risk without proven benefits compared to placebo or no antithrombotic treatment.

ACKNOWLEDGMENTS

Joaquim J Ferreira had speaker and consultant fees with GlaxoSmithKline, Novartis, TEVA, Lundbeck, Solvay, Abbott, Bial, Merck-Serono, Grunenthal, and Merck Sharp and Dohme. Fausto J Pinto had consultant and speaker fees with Astra Zeneca, Bayer and Boehringer Ingelheim. The referred companies were not directly or indirectly in this work.

COMMENTS

Background

In patients with atrial fibrillation (AF), chronic heart failure (CHF) increases

thromboembolic risk and oral anticoagulation is essential to decrease the risk of thromboembolic complications. Evidence suggests a positive association between CHF, impaired hemostasis and thromboembolic events. Whether antithrombotic drugs have an prognosis impact in patients with CHF in sinus rhythm (*i.e.*, without history of AF) is still very debated.

Research frontiers

Anticoagulation has been established as the gold standard treatment of stroke and embolism prevention in AF. The WARCEF trial did not show differences between warfarin and acetylsalicylic acid concerning major cardiovascular events in patients with CHF and sinus rhythm. Warfarin reduced the risk of ischemic stroke in this trial. However the efficacy of any of these drugs compared should be evaluated before drawing any conclusions and recommendations.

Innovations and breakthroughs

Based on the best available evidence (2 randomized controlled trials Warfarin/Aspirin Study in Heart failure and Heart failure Long-term Antithrombotic Study), this systematic review emphasizes the lack of efficacy of any antithrombotic drugs (individually or pooled together) in patients with CHF and sinus rhythm. In addition should be considered that these drugs increase significantly the risk of major bleeding.

Applications

Warfarin and acetylsalicylic acid should not be routinely used for thromboprophylaxis in patients with systolic HF and sinus rhythm, in the absence of concomitant comorbidities with clear indications for anticoagulation (*e.g.*, AF) or acetylsalicylic acid (*e.g.*, documented coronary artery disease).

Peer review

A systematic review and meta-analysis of two studies addressing antithrombotic drugs in patients with CHF and sinus rhythm. The manuscript is well written and adds new points to the discussion of anticoagulation.

REFERENCES

1 **Mosterd A**, Hoes AW. Clinical epidemiology of heart failure. *Heart* 2007; **93**: 1137-1146 [PMID: 17699180 DOI: 10.1136/hrt.2003.025270]

2 **Xuan J**, Duong PT, Russo PA, Lacey MJ, Wong B. The economic burden of congestive heart failure in a managed care population. *Am J Manag Care* 2000; **6**: 693-700 [PMID: 10977478]

3 **Braunschweig F**, Cowie MR, Auricchio A. What are the costs of heart failure? *Europace* 2011; **13** Suppl 2: ii13-ii17 [PMID: 21518742 DOI: 10.1093/europace/eur081]

4 **Hays AG**, Sacco RL, Rundek T, Sciacca RR, Jin Z, Liu R, Homma S, Di Tullio MR. Left ventricular systolic dysfunction and the risk of ischemic stroke in a multiethnic population. *Stroke* 2006; **37**: 1715-1719 [PMID: 16741172 DOI: 10.1161/01.STR.0000227121.34717.40]

5 **Loh E**, Sutton MS, Wun CC, Rouleau JL, Flaker GC, Gottlieb SS, Lamas GA, Moyé LA, Goldhaber SZ, Pfeffer MA. Ventricular dysfunction and the risk of stroke after myocardial infarction. *N Engl J Med* 1997; **336**: 251-257 [PMID: 8995087 DOI: 10.1056/NEJM199701233360403]

6 **Camm AJ**, Lip GY, De Caterina R, Savelieva I, Atar D, Hohnloser SH, Hindricks G, Kirchhof P; ESC Committee for Practice Guidelines-CPG; Document Reviewers. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation--developed with the special contribution of the European Heart Rhythm Association. *Europace* 2012; **14**: 1385-1413 [PMID: 22923145 DOI: 10.1093/europace/eus305]

7 **Homma S**, Thompson JL, Pullicino PM, Levin B, Freudenberger RS, Teerlink JR, Ammon SE, Graham S, Sacco RL, Mann DL, Mohr JP, Massie BM, Labovitz AJ, Anker SD, Lok DJ, Ponikowski P, Estol CJ, Lip GY, Di Tullio MR, Sanford AR, Mejia V, Gabriel AP, del Valle ML, Buchsbaum R. Warfarin and aspirin in patients with heart failure and sinus rhythm. *N Engl J Med* 2012; **366**: 1859-1869 [PMID: 22551105 DOI: 10.1056/NEJMoa1202299]

8 **Shah S**, Parra D, Rosenstein R. Warfarin versus aspirin in heart failure and sinus rhythm. *N Engl J Med* 2012; **367**: 771; author reply 772 [PMID: 22913690 DOI: 10.1056/NEJMc1207385]

9 **Hopper I**, Skiba M, Krum H. Updated meta-analysis on antithrombotic therapy in patients with heart failure and sinus rhythm. *Eur J Heart Fail* 2013; **15**: 69-78 [PMID: 23143796 DOI: 10.1093/eurjhf/hfs171]

10 **Kumar G**, Goyal MK. Warfarin versus aspirin for prevention of stroke in heart failure: a meta-analysis of randomized controlled clinical trials. *J Stroke Cerebrovasc Dis* 2013; **22**: 1279-1287 [PMID: 23182364 DOI: 10.1016/j.jstrokecerebrovasdis.2012.09.015]

11 **Lee M**, Saver JL, Hong KS, Wu HC, Ovbiagele B. Risk-benefit profile of warfarin versus aspirin in patients with heart failure and sinus rhythm: a meta-analysis. *Circ Heart Fail* 2013; **6**: 287-292 [PMID: 23264446 DOI: 10.1161/CIRCHEARTFAILURE.112.971697]

12 **Rengo G**, Pagano G, Squizzato A, Moja L, Femminella GD, de Lucia C, Komici K, Parisi V, Savarese G, Ferrara N, Perrone-Filardi P, Leosco D. Oral anticoagulation therapy in heart failure patients in sinus rhythm: a systematic review and meta-analysis. *PLoS One* 2013; **8**: e52952 [PMID: 23301006 DOI: 10.1371/journal.pone.0052952]

13 **Liew AY**, Eikelboom JW, Connolly SJ, O' Donnell M, Hart RG. Efficacy and safety of warfarin vs antiplatelet therapy in patients with systolic heart failure and sinus rhythm: a systematic review and meta-analysis of randomized controlled trials. *Int J Stroke* 2014; **9**: 199-206 [PMID: 23506160 DOI: 10.1111/ijs.12036]

14 **Moher D**, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535 [PMID: 19622551 DOI: 10.1136/bmj.b2535]

15 **Higgins JPT**, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from: URL: http:// www.cochrane-handbook.org

16 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327.7414.557]

17 **Lau J**, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. *Ann Intern Med* 1997; **127**: 820-826 [PMID: 9382404 DOI: 10.7326/0003-4819-127-9-199711010-00008]

18 **Sterne JA**, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol* 2001; **54**: 1046-1055 [PMID: 11576817 DOI: 10.1016/S0895-4356(01)00377-8]

19 **Peters JL**, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. *JAMA* 2006; **295**: 676-680 [PMID: 16467236 DOI: 10.1001/jama.295.6.676]

20 **Massie BM**, Collins JF, Ammon SE, Armstrong PW, Cleland JG, Ezekowitz M, Jafri SM, Krol WF, O'Connor CM, Schulman KA, Teo K, Warren SR; WATCH Trial Investigators. Randomized trial of warfarin, aspirin, and clopidogrel in patients with chronic heart failure: the Warfarin and Antiplatelet Therapy in Chronic Heart Failure (WATCH) trial. *Circulation* 2009; **119**: 1616-1624 [PMID: 19289640 DOI: 10.1161/CIRCULATIONAHA.108.801753]

21 **Cleland JG**, Findlay I, Jafri S, Sutton G, Falk R, Bulpitt C, Prentice C, Ford I, Trainer A, Poole-Wilson PA. The Warfarin/Aspirin Study in Heart failure (WASH): a randomized trial comparing antithrombotic strategies for patients with heart failure. *Am Heart J* 2004; **148**: 157-164 [PMID: 15215806 DOI: 10.1016/j.ahj.2004.03.010]

22 **Cokkinos DV**, Haralabopoulos GC, Kostis JB, Toutouzas PK; HELAS investigators. Efficacy of antithrombotic therapy in chronic heart failure: the HELAS study. *Eur J Heart Fail* 2006; **8**: 428-432 [PMID: 16737850 DOI: 10.1016/j.ejheart.2006.02.012]

23 **Lip GY**, Gibbs CR. Antiplatelet agents versus control or anticoagulation for heart failure in sinus rhythm. *Cochrane Database Syst Rev* 2001; (**4**): CD003333 [PMID: 11687189]

24 **Sirajuddin RA**, Miller AB, Geraci SA. Anticoagulation in patients with dilated cardiomyopathy and sinus rhythm: a critical literature review. *J Card Fail* 2002; **8**: 48-53 [PMID: 11862583 DOI: 10.1054/jcaf.2002.31907]

25 **Lip GY**, Gibbs CR. Anticoagulation for heart failure in sinus rhythm: a Cochrane systematic review. *QJM* 2002; **95**: 451-459 [PMID: 12096150 DOI: 10.1093/qjmed/95.7.451]

26 **Lip GY**, Gibbs CR. Antiplatelet agents versus control or anticoagulation for heart failure in sinus rhythm: a Cochrane systematic review. *QJM* 2002; **95**: 461-468 [PMID: 12096151]

27 **Lip GY**, Wrigley BJ, Pisters R. Anticoagulation versus placebo for heart failure in sinus rhythm. *Cochrane Database Syst Rev* 2012; **6**: CD003336 [PMID: 22696335 DOI: 10.1002/14651858.CD003336.pub2]

28 **Sterne JAC**, Egger M, Moher D (editors). Chapter 10: Addressing reporting biases. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Intervention. Version 5.1.0. [Updated 2011 March]. The Cochrane Collaboration, 2011. Available from: URL: http//www.cochrane-handbook.org

29 **Lip GY**, Ponikowski P, Andreotti F, Anker SD, Filippatos G, Homma S, Morais J, Pullicino P, Rasmussen LH, Marin F, Lane DA; ESC Task Force. Thrombo-embolism and antithrombotic therapy for heart failure in sinus rhythm. A joint consensus document from the ESC Heart Failure Association and the ESC Working Group on Thrombosis. *Eur J Heart Fail* 2012; **14**: 681-695 [PMID: 22611046 DOI: 10.1093/eurjhf/hfs073]

30 **Cleland JG**, Bulpitt CJ, Falk RH, Findlay IN, Oakley CM, Murray G, Poole-Wilson PA, Prentice CR, Sutton GC. Is aspirin safe for patients with heart failure? *Br Heart J* 1995; **74**: 215-219 [PMID: 7547012 DOI: 10.1136/hrt.74.3.215]

31 **Ahmed A**. Interaction between aspirin and angiotensin-converting enzyme inhibitors: should they be used together in older adults with heart failure? *J Am Geriatr Soc* 2002; **50**: 1293-1296 [PMID: 12133028 DOI: 10.1046/j.1532-5415.2002.50320.x]

32 **Teo KK**, Yusuf S, Pfeffer M, Torp-Pedersen C, Kober L, Hall A, Pogue J, Latini R, Collins R; ACE Inhibitors Collaborative Group. Effects of long-term treatment with angiotensin-converting-enzyme inhibitors in the presence or absence of aspirin: a systematic review. *Lancet* 2002; **360**: 1037-1043 [PMID: 12383982 DOI: 10.1016/S0140-6736(02)11138-X]

33 **Campbell CL**, Smyth S, Montalescot G, Steinhubl SR. Aspirin dose for the prevention of cardiovascular disease: a systematic review. *JAMA* 2007; **297**: 2018-2024 [PMID: 17488967 DOI: 10.1001/jama.297.18.2018]

P- Reviewer: Aronow WS, Rauch B, Skobel E
S- Editor: Song XX **L- Editor**: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.43

World J Meta-Anal 2015 February 26; 3(1): 43-53
ISSN 2308-3840 (online)

META-ANALYSIS

Effect of supervised exercise on aerobic capacity in cancer survivors: Adherence and workload predict variance in effect

Rhys Beaudry, Calvin Kruger, YuanYuan Liang, Matthew Parliament, Mark Haykowsky, Margaret L McNeely

Rhys Beaudry, Calvin Kruger, Mark Haykowsky, Margaret L McNeely, Department of Physical Therapy, University of Alberta, Edmonton, Alberta T6G 2G4, Canada
YuanYuan Liang, School of Medicine, Epidemiology and Biostatistics, University of Texas, San Antonio, TX 78229, United States
Matthew Parliament, Margaret L McNeely, Department of Oncology, University of Alberta and Cross Cancer Institute, Edmonton, Alberta T6G 1Z2, Canada
Author contributions: All authors contributed to the design of the research protocol and interpretation of the data; Beaudry R and Kruger C contributed equally to this work; Beaudry R, Kruger C and McNeely ML contributed to the acquisition of data as described in the methods, preliminary data analyses and wrote the paper; Liang Y performed the meta-regression analyses; Liang Y, Parliament M and Haykowsky M reviewed the paper critically for important intellectual content and approved the version to be published.
Conflict-of-interest: The authors declare no conflicts of interest financial or otherwise.
Data sharing: Not applicable as a review not basic or clinical research.

Correspondence to: Margaret L McNeely, Assistant Professor, Department of Physical Therapy, University of Alberta, 2-50 Corbett Hall, 116 St. and 85 Ave, Edmonton, Alberta T6G 2G4, Canada. ribeaudr@ualberta.ca
Telephone: +1-780-2481531
Fax: +1-780-4924429
Received: July 29, 2014
Peer-review started: July 29, 2014
First decision: September 16, 2014
Revised: December 23, 2014
Accepted: December 29, 2014
Article in press: December 31, 2014
Published online: February 26, 2015

Abstract

AIM: To examine the efficacy of supervised aerobic exercise training on aerobic capacity in survivors of cancer.

METHODS: We conducted a systematic search identifying randomized controlled trials of supervised aerobic exercise interventions among adult cancer survivors with aerobic capacity ($VO_{2max/peak}$) as the primary outcome. We calculated pooled effect sizes and performed multiple regression moderator analysis.

RESULTS: We identified 18 studies including 1149 survivors of cancer. Studies included mixed cancer groups (4 studies), breast cancer (10 studies), hematological cancers (2 studies), lung cancer (1 study) and liver cancer (1 study). Survivors of cancer who participated in supervised aerobic exercise training improved aerobic capacity (VO_{2peak}) more than controls (18 comparisons, 1093 participants; standardized mean effect: 0.74; 95%CI: 0.52, 0.96; $P < 0.001$). However, there was significant heterogeneity among the included trials (I^2: 63%; $P < 0.001$). Sixty-six percent of the between-study heterogeneity was explained by differences in exercise adherence and total exercise workload among studies (R^2: 65.8%; $P < 0.04$).

CONCLUSION: Supervised aerobic exercise training provides a moderate-to-large beneficial effect on aerobic capacity among survivors of cancer. Aerobic capacity was improved to a greater degree in exercise studies with better participant attendance and higher overall exercise workload.

Key words: Exercise; Neoplasms; Physical therapy modalities; Physical fitness; Meta-analysis

Core tip: The optimal exercise prescription for survivors of cancer is unknown and the effect of variations in exercise training parameters on cancer-specific outcomes are poorly understood. Therefore, questions remain over how to best tailor exercise prescriptions to optimize the health outcomes of survivors who are at different time points in their cancer care. We performed a meta-analysis of data from randomized controlled trials examining the effect of supervised aerobic exercise training on aerobic capacity in cancer survivors. We found that aerobic capacity was improved to a greater extent in exercise studies that prescribed a higher exercise workload and had better participant adherence.

Beaudry R, Kruger C, Liang Y, Parliament M, Haykowsky M, McNeely ML. Effect of supervised exercise on aerobic capacity in cancer survivors: Adherence and workload predict variance in effect. *World J Meta-Anal* 2015; 3(1): 43-53 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/43.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.43

INTRODUCTION

The burden of cancer continues to increase worldwide due to population growth and aging[1]. More effective cancer screening and novel treatment therapies have resulted in improved detection, earlier treatment and better disease free and overall survival, with the numbers of cancer survivors growing disproportionately to the number of new cancer cases and deaths[2]. Many cancer survivors experience symptoms and side effects related to their cancer or cancer treatment. As many of these effects go undetected and/or untreated, the survivor is placed at increased risk for other health issues such as declining functional status and cardiovascular disease[3,4]. As a result, there is an emerging need for the integration of services and interventions to address the long-term health of survivors[3].

Exercise training is gaining recognition as an important intervention to address acute, late and long-term effects of cancer, and is becoming more widely acceptable as confidence in safety is now established. Importantly, evidence is accumulating to support the benefit of exercise to improve the physical functioning and quality of life of survivors. Currently, the optimal exercise prescription is unknown and the effect of variations in exercise training parameters on cancer-specific outcomes are poorly understood[5]. Therefore, questions remain over how to best tailor exercise prescriptions to optimize the health outcomes of survivors at different times through the cancer continuum[5].

Cardiorespiratory fitness, measured objectively as the highest oxygen consumed during maximal aerobic exercise, provides a means to evaluate associations with disease outcomes. Aerobic capacity is inversely related to the risk of a cardiovascular event and all-cause mortality in healthy individuals and cancer patients[6-10]. Aerobic capacity is best increased by habitual aerobic exercise training that is of a moderate-to-vigorous intensity[11].

Aerobic capacity (VO_{2max}) is the maximum volume of oxygen that the body can consume during maximal exercise, using at least 60% of the musculature, and while breathing air at sea level[12]. This volume is expressed as an absolute rate in litres per minute (L/min) or as a relative rate in millilitres per kilogram of bodyweight per minute (mL/kg per minute). VO_{2peak} is the term used most commonly in clinical populations when a true maximal value is not attained[12]. For example, the test is described as VO_{2peak} rather than VO_{2max} when the test is carried out on a cycle ergometer (bike) rather than a treadmill, or when the highest value reached on the test is limited by the participant's symptoms.

A meta-analysis by Jones and colleagues included data from six randomized controlled trials (RCTs) and reported a significant benefit from supervised aerobic exercise training, compared with usual care, on VO_{2peak} (2.90 mL/kg per minute; 95%CI: 1.16, 4.64; $P = 0.01$)[13]. However, statistical and clinical heterogeneity was found among the exercise trials included in their review, leading them to recommend further research to build on and extend the current knowledge in the field. Since this publication, a number of newer studies have been published. Given this amount of new data, we contend that an updated review is warranted.

The primary purpose of this meta-analysis was to examine the efficacy of supervised aerobic exercise training programs on VO_{2peak} in survivors of cancer. Quality of life was analyzed as a secondary outcome measure. As well we aimed to explore heterogeneity in study findings through subgroup analyses and meta-regression where appropriate.

MATERIALS AND METHODS

The review conforms to the requirements of PRISMA reporting standards. The published protocol for the review can be found at: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013006215#.U1cOnle9aw).

Inclusion criteria

Studies were considered eligible for inclusion if they were RCTs comparing supervised aerobic exercise training with a placebo, controlled comparison or standard care. For the purposes of the review, exercise was defined as a form of leisure-time physical activity that was performed on a repeated basis over an extended period of time, with the intention of improving fitness, performance or health[14]. Studies with an additional treatment arm or combined intervention (*e.g.*, exercise with diet modification) were included only if the effects of exercise could be isolated. A priori, we excluded reports that were available only in abstract form.

Table 1 Example of medline search

(1) Exp neoplasms/
(2) (Cancer* or neoplasm* or (tumor* not tumor necrosis factor) or (tumour* not tumour necrosis factor) or malignan* or carcino* or leukaemia* or leukemi* or lymphoma* or myeloma* or adenocarcinoma*).mp.
(3) (1) or (2)
(4) Exercise therapy/or motion therapy, continuous passive/or muscle stretching exercises/or plyometric exercise/
(5) (Aerobic* or exercise or running or treadmill* or training).mp. [mp = title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
(6) (4) or (5)
(7) (3) and (6)
(8) (VO_2 or Aerobic capacity).mp. [mp = title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
(9) (7) and (8)
(10) Limit (9) to clinical trial, all

Trials were included if they involved adults (17 years and older) diagnosed with cancer who were actively receiving cancer treatment or off treatment. Included studies were required to measure maximal, peak, or estimated maximal oxygen consumption ($VO_{2max/peak}$) as a study outcome.

Systematic search

A search was performed of the databases including OVID MEDLINE (1948 to October 2013), PubMed (1975 to October 2013), SCOPUS (1950 to October 2013), Web of Science (1950 to October 2013), EMBASE (1988 to October 2013), Cochrane Central Registry of Controlled Trials (1991 to October 2013), and LILACS (1982-October 2013). The search strategy was developed and approved by a librarian with extensive database searching knowledge and experience. We searched terms related to cancer (*e.g.*, neoplasms, tumor), exercise (*e.g.*, exercise, exercise therapy/or motion therapy, aerobic training), publication type (*e.g.*, random allocation, clinical trial), and aerobic capacity (*e.g.*, VO_2). The search strategy was modified as necessary for each database. Non-English language publications were eligible for inclusion. To locate unpublished research, we reviewed clinical trial registries and websites housing theses and dissertations. Fourteen experts in the field of cancer and exercise were contacted in order to identify any research that was not published or was pending publication. Table 1 includes an example of the MEDLINE search strategy.

Coding and reliability

The titles and abstracts were screened for eligibility by two independent evaluators (C.K. and R.B.), and coded for exclusion or potential inclusion. Potentially eligible manuscripts were obtained and the same evaluators performed a second round of screening to evaluate full eligibility criteria. Any disagreements were resolved by consensus (C.K., R.B., and M.M.). The two evaluators (C.K. and R.B.) then independently abstracted data on study participants, the intervention and control (usual care) protocols, and study outcomes, and assessed for quality. Studies were evaluated using the quality assessment framework for RCTs developed by the Cochrane Collaboration[15] to assess risk of bias in the individual studies. Sensitivity analyses were conducted to examine the effect of including studies with high risk of bias.

For the purpose of evaluating exercise prescription variables, exercise intensity was standardized to a single $\%VO_{2max}$ value[16-18]. For studies that used $\%VO_{2max}$ as the intensity prescription the average of the range was used; time spent at different intensities was factored in to create the mean value. High intensity intervals were weighted at 50% of the contributing time. Resistance exercise was not included in intensity ratings. Total exercise workload, or intensity-minutes, was calculated by multiplying the exercise intensity by the prescribed exercise volume (program duration, minutes per session and sessions per week).

Study outcomes and effect size calculation

Study results were pooled using random effects models. For continuous outcomes, pooled statistics were calculated using mean differences (MD) when data were on a uniform scale and using standardized MD (SMD) when data were on different scales. All results were calculated with 95%CI. The SMD was interpreted as 0.2, 0.5 and 0.8 representing small, medium and large effects on outcomes respectively[19]. Statistical heterogeneity was assessed using a χ^2 test that considered a *P*-value of less than 0.10 to indicate significant heterogeneity. I^2 values, ranging from 0% (homogeneity) to 100% (heterogeneity) were also calculated to quantify variability in study effect and values of 25%, 50% and 75% were used to describe low, moderate and high heterogeneity respectively[20]. Subgroup analyses and multiple regression moderator analyses were performed to explore and explain heterogeneity among studies. A priori subgroup analyses included examining the pooled effect estimate by level of supervision of exercise (group or individual), the timing of the intervention (on or off treatment), and cancer type. Meta-regression was performed to explore exercise variables of frequency, time, intensity, duration and adherence on effect estimate.

Statistical analysis

A biomedical statistician (Y.L.) provided oversight on the statistical methods, and performed the meta-regression analyses. All data were entered into Review Manager 5.2 and analyzed with SPSS v15 software utilizing meta-regression scripts created by Lipsey and Wilson and Stata/SE (version 13.0)[21]. Figures were created using Comprehensive Meta-Analysis (version 3: http://www.meta-analysis.com/index.php).

RESULTS

Methodological characteristics

The search protocol yielded 1269 eligible studies; after

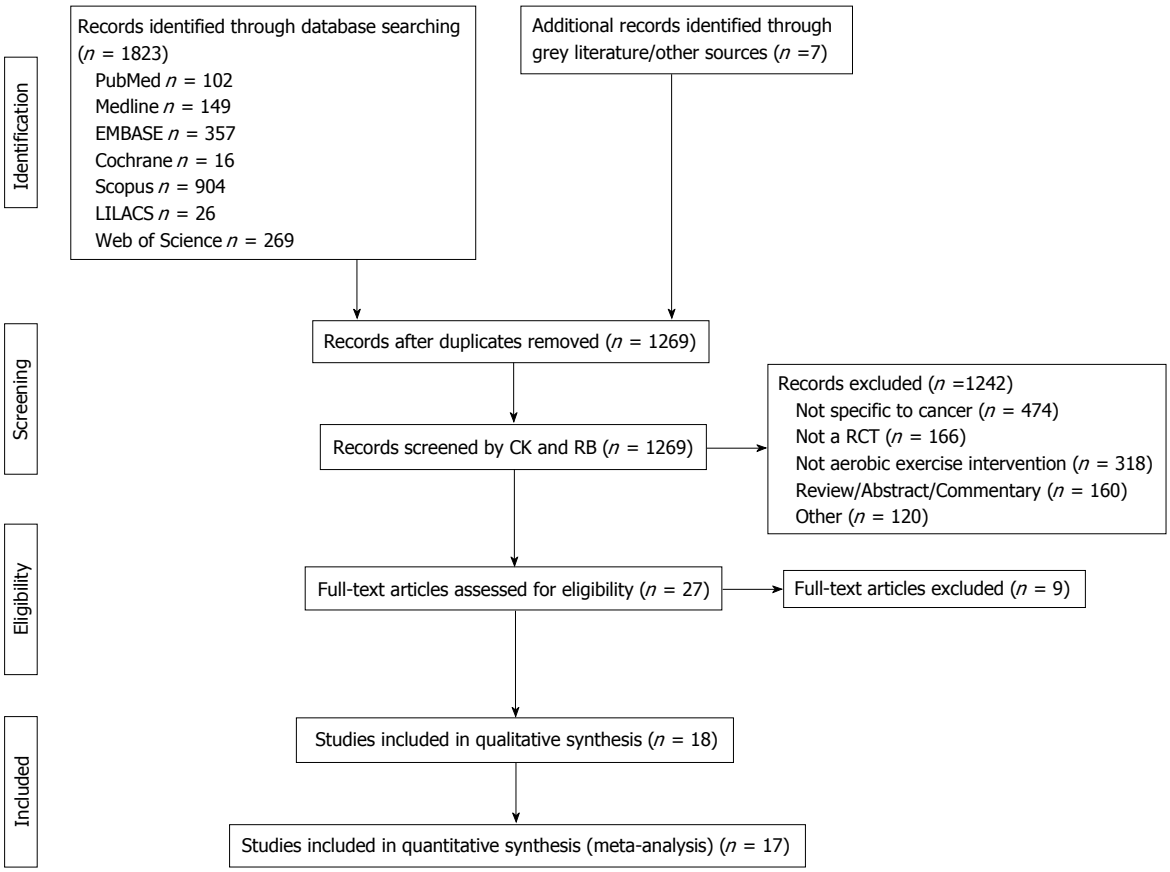


Figure 1 PRISMA flow diagram of study selection process.

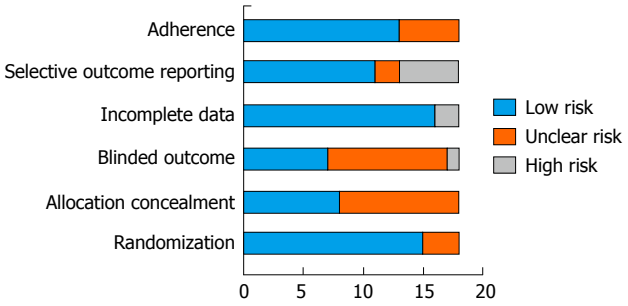


Figure 2 Risk of bias summary.

removal of duplicates and screening of abstracts, 23 studies remained. Reference tracking and contacting of experts accounted for 4 additional studies. Grey literature and trial register searches yielded no further articles. Full text review of the 27 studies excluded a further 9, leaving 18 studies for qualitative and quantitative synthesis[22-39]. One study was not used for the quantitative analyses due to missing data[32] and one study was divided into two comparison groups as it involved both on and off treatment subgroups[27] (unpublished data provided by author). The remaining 17 studies, generating 18 comparisons, were included in the meta-analyses (Figure 1). Kappa statistics for the inclusion of studies was 0.9 ($P < 0.001$). Following discussion there was 100% agreement in scores between evaluators.

Risk of bias

In general there was high or unclear risk of bias for selection (allocation concealment) and detection bias (lack of blinding of outcome assessors) and low risk of bias for attrition (handling of incomplete data) and reporting bias (outcome reporting) among the included studies (Figure 2). Sensitivity analyses were performed after excluding studies with a high or unclear risk of bias for allocation concealment ($n = 10$)[24,31-39] and for use of blinded outcome assessment ($n = 11$)[24,26,27,31-36,38,39]. The results showed minimal differences in the pooled effect estimates for aerobic capacity based on risk of bias. For allocation concealment, the pooled effect estimate increased by 0.6 (SMD: 0.80; 95%CI: 0.51, 1.25) whilst for blinding of outcome assessment the estimate decreased by 0.4 (SMD: 0.7; 95%CI: 0.35, 1.05). After excluding studies with a high or unclear risk of bias for any factor ($n = 13$), the pooled effect estimate decreased by 0.8 (SMD: 0.66, 95%CI: 0.22, 1.11).

Cancer survivor characteristics

The 18 included studies involved 1149 participants of which 576 were randomized to receive an aerobic exercise intervention and the remaining 573 received usual care or no exercise. Participants were on average 53 years of age and 76% were female. Survivors of breast cancer were most commonly studied in both breast cancer specific

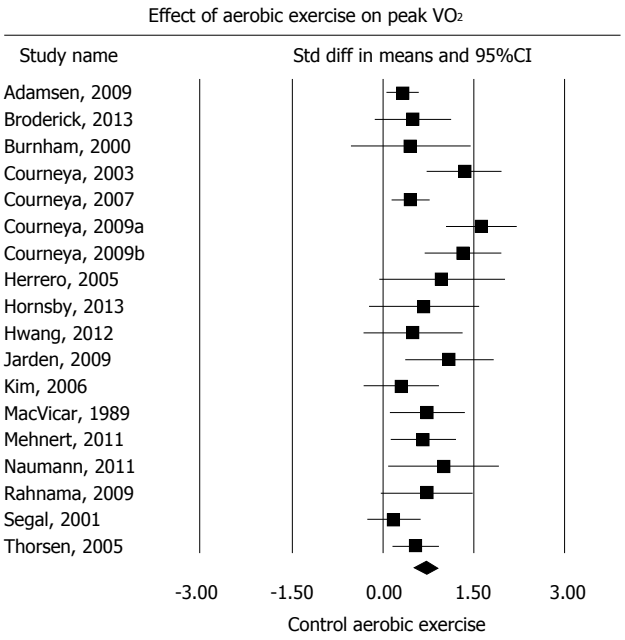


Figure 3 VO_2 effect size.

trials and mixed cancer type trials (14 studies)[22-26,28,29,33-39] accounting for 686 participants (60%) of the total participants in the review. Further details on the included studies are provided in Table 2.

Exercise intervention characteristics

Ten studies consisted exclusively of aerobic exercise training[23-27,29,30,33,34,38], six studies included a resistance exercise component with or without flexibility training[22,28,31,36,37,39], one included physiotherapy exercises and relaxation[35], and one included flexibility training plus a dietary intervention[32]. Exercise interventions consisted primarily of cycling[23-31,34,39] or walking/jogging[23,24,32,35,37-39]. Five studies[22,23,28,35,38] offered exercise programs in a class setting (group exercise format) and the remaining 13 studies[24-27,29-34,36,37,39] were individualized exercise programs, although further detail on the level of supervision was not often provided. Eight studies were carried out during active cancer treatment[22,26,29-31,33,34,38], nine in the post treatment phase[23-25,28,32,35-37,39] and one included participants both on and off treatment[27]. The duration of exercise programs ranged from 4-6 wk to 26 wk with individual exercise sessions ranging from 20-90 min including warm up and cool down. Seventeen studies prescribed aerobic exercise that was of moderate intensity with 4 of these studies[22,27,29,34] including high intensity intervals. One study combined both low and moderate intensity intervention groups into a single intervention group for their analysis due to the small sample size of the study[24]. Further information on the exercise prescription variables is provided in Table 3.

The effect of supervised aerobic exercise on aerobic capacity

All eighteen studies reported VO_{2peak}, with 13 studies (14 comparisons) indexing this outcome to body weight (mL/kg per minute)[23-30,35-39], 4 studies measuring absolute (L/min)[22,31,33,34], and 1 study measuring percent change in VO_{2peak} (mL/kg per minute)[32]. The study measuring percent change in VO_{2peak} was excluded from analysis due to insufficient data on measures of variability.

Pooling of all 18 comparisons showed a moderate-to-large effect estimate (SMD: 0.74; 95%CI: 0.52, 0.96; $P < 0.001$) in favour of supervised aerobic exercise training; however, moderate heterogeneity was found among the included studies ($I^2 = 63\%$; $P < 0.001$) (Figure 3). Pooling of the 13 studies (14 comparisons) reporting VO_{2peak} (mL/kg per minute) showed a statistically significant mean difference in VO_{2peak} of 3.13 mL/kg per minute (95%CI: 2.21, 4.05; $P < 0.001$) in favour of supervised aerobic exercise training; however, again moderate heterogeneity was found among the included studies ($I^2 = 58\%$; $P < 0.001$).

Subgroup analysis

Subgroup analyses were performed for level of supervision, treatment timing and cancer type (Table 4). A significantly smaller effect estimate ($P = 0.003$) was found for group/ class-led exercise studies[22,23,35,38] (SMD: 0.36; 95%CI: 0.17, 0.56) when compared to studies involving individualized exercise programs[24-31,33,34,36,37,39] (SMD: 0.87; 95%CI: 0.60, 1.15). Non-significant effects ($P = 0.11$) were observed between on and off treatment studies. Statistically significant differences in pooled effect estimates were observed between cancer types with a significantly larger beneficial effect found among studies including survivors with hematological cancers ($P < 0.001$)[27,31] when compared to other cancer tumor groups (breast cancer, lung cancer and mixed cancer).

Meta-regression

Meta regression was performed analyzing the effect estimate with exercise parameters of exercise workload and participant adherence as potential moderators. These two variables, workload and adherence, explained 65.8% ($P = 0.04$) of the between-study variance in effect estimate among the included studies (Figure 4).

Quality of life

Nine studies reported data for health-related quality of life as measured by the Functional Assessment of Cancer Therapy-General (FACT-G) scale[23,25,29,31], the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire: EORTC-QLQ-C30[22,28,30,39] and Medical Outcomes Survey: Short Form: SF36[38]. Pooling of all nine studies demonstrated a non-significant effect on quality of life (SMD: 0.3; 95%CI: -0.11, 0.71; $P = 0.16$), with high heterogeneity found among studies ($I^2 = 80\%$; $P < 0.001$). Further details are provided in Table 5.

DISCUSSION

This meta-analysis found that supervised aerobic exercise resulted in a moderate-to-large significant benefit on

Table 2 Description of Included Studies

Ref.	Sample size/ cancer type	Age (SD/range)	Gender (F/M)	Intervention group	Comparison group	Key outcomes	Adverse events
On treatment studies/subgroups							
Adamsen *et al*[22], 2009 Denmark	*n* = 117 Mixed Cancer Groups	47.2 (± 6.7) yr	F: 78 M: 39	Aerobic Training with High-intensity Intervals + Resistance Exercise + Relaxation + Massage	Usual care: allowed to freely increase physical activity	Estimated VO_{2max}	Seizure (*n* = 1)
Courneya *et al*[26], 2007 Canada	*n* = 133 Breast Cancer	49 yr (26-78)	F: 133	Aerobic Training	Usual care: continue usual activities	VO_{2peak} QoL: FACT-Anemia	Hypotension (*n* = 1) Dizziness (*n* = 1)
[1]Courneya *et al*[27], 2009[b] Canada	*n* = 54 NHL, HL	[2]53.2 yr (18-80)	[2]F: 50 M: 72	Aerobic Training with High-intensity Intervals	Usual Care: continue usual activities	VO_{2peak} QoL: FACT-B/Ac/An	Back (*n* = 1), hip (*n* = 1) and knee (*n* = 1) pain
Hornsby *et al*[29], 2013 United States	*n* = 20 Breast Cancer	51 (± 6) yr	F: 10	Aerobic Training with High-intensity Intervals	Control: Continue usual exercise levels	VO_{2peak} FACT-B Adverse Events	Leg pain (*n* = 1)
Hwang *et al*[30], 2012 Taiwan	*n* = 24 Lung	61 (± 6.3)	F: 12 M: 12	Aerobic Training	Usual Care: general patient education	VO_{2peak} QoL: EORTC	Not reported
Jarden *et al*[31], 2009 Denmark	*n* = 42 Mixed Cancer Groups	39.1 (12.2)	F: 16 M: 26	Aerobic Training + Resistance Exercise + Flexibility	Usual Care	Estimated VO_{2max} QoL: EORTC, FACT-An	None
Kim *et al*[33], 2006 United States	*n* = 41 Breast Cancer	51.3 (6.7) yr	F: 41	Aerobic Training	Waitlist Control	VO_{2peak}	Not reported
MacVicar *et al*[34], 1989 United States	*n* = 34 Breast Cancer	45.4 (10.2) yr	F: 34	Aerobic Training with High-intensity Intervals	Control: Continue normal activities	VO_{2max} L/min	Not reported
Segal *et al*[38], 2001 Canada	*n* = 66 Breast Cancer	51 (± 8.7) yr	F: 66	Aerobic Training	Control group encouraged to exercise	Estimated VO_{2max} QoL: SF36	Not reported
Off treatment studies/comparisons							
Broderick *et al*[23], 2013 Ireland	*n* = 43 Mixed Cancer Groups	52.3 (8.3) yr	F: 37 M: 6	Aerobic training	Usual Care	Estimated VO_{2max} QoL: FACT-G, SF36	Not reported
Burnham *et al*[24], 2000 United States	*n* = 18 Mixed Cancer Groups	54.2 (8.1) yr	F: 15 M: 3	Aerobic training	Control	VO_{2peak} QoL: LASA	Not reported
Courneya *et al*[25], 2003 Canada	*n* = 50 Breast Cancer	59 (± 6) yr	F: 54	Aerobic training	No exercise	VO_{2peak} QoL: FACT-Breast	Lymphedema (*n* = 3) Gynecological complication (*n* = 1)
[1]Courneya *et al*[27], 2009[a] Canada	*n* = 68 NHL, HL	[2]As per Courneya, 2009[b]	[2]As per Courneya, 2009[b]	[2]As per Courneya, 2009[b]	[2]As per Courneya, 2009[b]	[2]As per Courneya, 2009[b]	[2]As per Courneya, 2009[b]
Herrero *et al*[28], 2005 Spain	*n* = 16 Breast Cancer	51 (10) yr	F: 16	Aerobic plus Resistance Training	No Exercise	VO_{2peak} QoL: EORTC	Not reported
Kaibori *et al*[32], 2013 Japan	*n* = 51 Liver Cancer	68 (9.1) yr	F: 15 M: 36	Aerobic Training + Stretching + Diet Intervention	Diet Intervention	VO_{2peak}	Not reported
Mehnert *et al*[35], 2011 Germany	*n* = 58 Breast Cancer	53 (7.4) yr	F: 58	Aerobic Training + Physiotherapeutic Exercises + Relaxation	Waitlist Control	VO_{2max} QoL: BIQ	Not reported
Naumann *et al*[36], 2011 Australia	*n* = 21 Breast Cancer	49 (10) yr	F: 21	Aerobic Training + Resistance Exercise + Flexibility	Usual Care	Estimated VO_{2max} QoL: FACT-B	Not reported
Rahnama *et al*[37], 2010 Iran	*n* = 29 Breast Cancer	58.3 (6.3) yr	F: 29	Aerobic Training + Resistance Exercise	No exercise	Estimated VO_{2max}	Not reported
Thorsen *et al*[39], 2005 Norway	*n* = 111 Mixed Cancer Groups	39 (8.4) yr	F: 36 M: 75	Aerobic Training + Resistance Exercise	Usual Care	Estimated VO_{2max} QoL: EORTC	Not reported

[1]Courneya 2009 publication: Courneya 2009[b]-subgroup of participants on-treatment; Courneya 2009[a]-subgroup of participants off-treatment; [2]Data as per Courneya 2009[b]. QoL: Quality of life; FACT-G: Functional Assessment of Cancer Therapy-General scale; EORTC: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire: SF36: Medical Outcomes Survey Short Form; VO_{2max}: Maximal oxygen consumption; VO_{2peak}: Peak oxygen consumption.

Table 3 Exercise prescription variables

Ref.	Study duration (wk)	Days/week	Mins/session (mean)	Volume	Standardized intensity (mean)	Workload (intensity minutes)	Adherence (attendance)
Adamsen *et al*[22]	6	3	15	270	0.83	224	71%
Broderick *et al*[23]	8	2	30	480	0.57	274	78%
Burnham *et al*[24]	10	3	23	690	0.41	281	70%
Courneya *et al*[25]	15	3	25	1125	0.73	816	98%
Courneya *et al*[26]	12	3	30	1080	0.70	756	70%
Courneya *et al*[27] (1)	12	3	30	1080	0.79	858	84%
Courneya *et al*[27] (2)	12	3	30	1080	0.79	858	71%
Herrero *et al*[28]	8	3	25	600	0.59	351	91%
Hornsby *et al*[29]	12	3	23	828	0.79	657	82%
Hwang *et al*[30]	8	3	20	480	0.60	288	71%
Jarden *et al*[31]	5	5	22.5	563	0.72	405	80%
Kim *et al*[33]	8	3	30	720	0.65	468	78%
MacVicar *et al*[34]	10	3	NR	-	0.73	-	NR
Mehnert *et al*[35]	10	2	30	600	0.60	360	NR
Naumann *et al*[36]	8	3	53	1272	0.50	636	84%
Rahnama *et al*[37]	15	2	35	1050	0.28	289	NR
Segal *et al*[38]	26	3	NR	-	0.55	-	72%
Thorsen *et al*[39]	14	2	30	840	0.62	518	NR

NR: Not reported.

Table 4 Subgroup analyses

Subgroup category	Subgroup	No. studies	Mean Difference in mL/kg per minute (95%CI)	*P* value between subgroups	No. studies	Standardized mean difference (95%CI)	*P* value between subgroups
Level of exercise supervision	Group Exercise Class	3	1.77 (0.04, 3.51)	$P = 0.07$	4	0.36 (0.17, 0.56)	$P = 0.003$
	Individual Exercise	11	3.53 (2.64, 4.43)		14	0.87 (0.60, 1.15)	
Treatment status	On Treatment	5	2.59 (0.7, 4.48)	$P = 0.26$	9	0.56 (0.32, 0.81)	$P = 0.11$
	Off Treatment	9	3.74 (3.06, 4.42)		9	0.92 (0.56, 1.29)	
Cancer tumor group	Breast	8	2.41 (1.5, 3.31)		10	0.64 (0.34, 0.88)	
	Hematologic	3	5.08 (4.01, 6.16)		3	1.55 (1.09, 2.02)	
	Lung	1	2.10 (-1.36, 5.56)	$P = 0.002$	1	0.48 (-0.34, 1.30)	$P = 0.0002$
	Mixed Cancers	3	3.17 (1.34, 5.0)		4	0.41 (0.21, 0.61)	

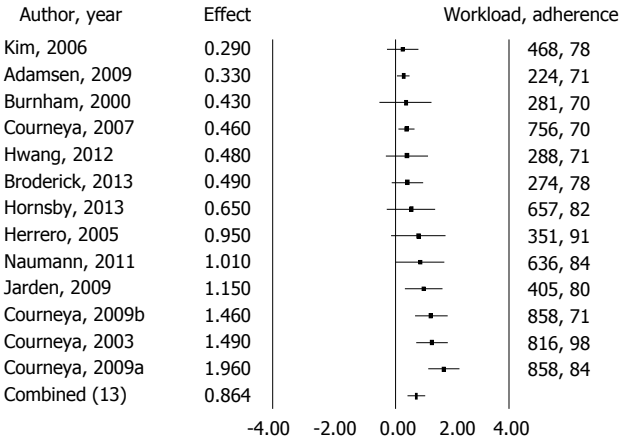


Figure 4 Meta-regression analysis: Workload, adherence.

VO_{2peak} in survivors of cancers. The pooled mean difference showed an improvement in VO_{2peak} of 3.13 mL/kg per minute, which is close to one metabolic equivalent (MET) improvement in fitness and similar to the 2.9 mL/kg per minute increase reported by Jones *et al*[13]. In the general population, each one MET increase in fitness has been found to translate to a 12% decrease in mortality in men[6] and a 17% decrease in women[40]. In the cancer population, a number of studies have reported an inverse correlation between VO_{2peak} and all-cause mortality, including cardiovascular, lung and breast cancer related deaths[41-43].

We did not find an overall significant effect of supervised aerobic exercise interventions on quality of life. Studies in our review used a variety of quality of life measures and when data were pooled significantly high heterogeneity was found. This finding suggests that the differences between study populations and/or differences inherent in the quality of life questionnaires may be factors. Supporting this premise, the pooled data from four studies using the FACT-General scale showed both statistical homogeneity and significant benefit on quality of life.

Our results showed that survivors of cancer participating in individually-based exercise experienced greater improvement in VO_{2peak} than those participating in group or class-led exercise. A reported advantage to group or class-led exercise is the social interaction and group

Table 5 Quality of life outcome

Quality of life measure	No. of studies	Mean difference (95%CI)	*P* value between groups	Standardized mean difference (95%CI)	*P* value between groups
All combined	9	Not applicable	-	0.3 (-0.12, 0.70)	$P = 0.16$
EORTC Global	4	1.45 (0.58, 2.32)	$P = 0.001$	0.13 (-0.06, 0.33)	$P = 0.17$
FACT-G	4	3.25 (-0.41, 6.92)	$P = 0.08$	0.47 (0.14, 0.79)	$P = 0.005$
MOS SF36	1	2.2 (1.34, 3.06)	$P < 0.001$	1.22 (0.69, 1.74)	$P < 0.001$

EORTC Global: European Organisation for Research and Treatment of Cancer Global Quality of Life Questionnaire; FACT-G: Functional Assessment of Cancer Therapy-General scale; MOS SF36: Medical Outcomes Survey Short Form.

support that may foster improvements in quality of life among survivors. Similar to our findings, a previous meta-analysis comparing group to individual exercise on quality of life in survivors of breast cancer reported that group exercise showed no benefit over individual exercise[44]. While the findings of our review appear to support individually based exercise programs for the outcome of aerobic capacity, we found that data were generally lacking on the ratio of the exercise participant to exercise specialist to allow for closer examination of impact of the level of supervision.

In contrast to the meta-analysis by Jones *et al*[13] we did not find a significant difference between groups based on the timing of the intervention relative to cancer treatment. Inspection of adherence across studies revealed a bimodal distribution with clusters in the 70-75 and 85-98 percent ranges. This bimodal distribution appeared to reflect on/off treatment status, as better adherence and larger effects were generally seen from exercise intervention studies carried out after completion of cancer treatment. Moreover, the direction of exercise effects compared to usual care may differ in relation to treatment status. For example, Jarden *et al*[31] demonstrated that exercise during active cancer treatment prevented a decline in VO_{2peak} when compared to usual care, whereas Kim *et al*[33] found that exercise following cancer treatment increased VO_{2peak} over usual care. More research is required to elucidate the influence of the timing of the exercise intervention through the continuum of cancer treatment and survivorship.

While our overall findings support the benefit of supervised aerobic exercise on VO_{2peak}, the relative benefit varied significantly across studies. As the number of research studies in the area has increased we were able to examine the influence of exercise prescription variables on aerobic capacity. Our analyses showed that VO_{2peak} improved to a larger extent in studies examining survivors of haematological cancers over other cancer groups. However, this finding was based on data from only 2 studies (3 comparisons) and thus, while compelling; further research is needed within this particular cancer subgroup. Of note, significant improvements were found within the subgroups of both breast cancer and mixed cancer groups; however, the effect was smaller.

Better participant adherence and overall exercise workload emerged as important predictors of intervention efficacy. Adherence, in this review, represented attendance to exercise sessions. Data on adherence to intensity and exercise volume were not reported in the majority of trials. Attendance to exercise sessions may reflect the impact of treatment-related side effects, patient motivation, or aspects of the study protocol such as opportunities for making up missed sessions. High adherence to the exercise prescription is critical for ensuring an adequate training stimulus to induce physiological change in cardiorespiratory function. Better reporting of adherence to prescription factors of intensity and duration would allow for more precise examination of the dose response to exercise[5].

Previous meta-analyses examining exercise interventions have reported benefit from more intense aerobic exercise interventions for both quality of life and depressive symptoms[45,46]. In the present meta-analysis, however, overall workload rather than intensity alone was found to predict response to exercise. We found that the majority of studies in the review prescribed moderate intensity exercise training, although some included high intensity interval work. Multiplying the exercise volume by the prescribed intensity provided a workload metric (*i.e.*, intensity-minutes) for discriminating between trials finding large effects from those with small effects. While some studies prescribing lower exercise volumes showed benefit, a target workload (intensity-minutes) of around 600 intensity-minutes (*e.g.*, 10 wk program of 90 min per week of supervised exercise at 70% VO_{2peak}) appears to represent the threshold workload required to obtain a clinically significant large improvement (effect size > 1.0) in VO_{2peak}. A recent meta-analysis by Carayol *et al*[47] examined the effect of exercise on fatigue and quality of life and found a workload in the range of 90-120 min of moderate intensity exercise was more beneficial in improving fatigue and quality of life than higher volumes of exercise. Our findings suggest that improvements in aerobic capacity can be attained at an exercise workload level that, in theory, should not negatively impact fatigue and quality of life.

Limitations

The major limitations of this meta-analysis were the assumptions revolving around exercise prescription factors. All intensity values represented average values obtained and were standardized to an estimated $\%VO_{2max}$ value. Conversions are imperfect as are average values created from studies using intervals and step protocols. Therefore we acknowledge that there is some associated error in our intensity estimates. As well, no data were provided

on actual adherence to intensity among participants in the individual studies to allow more precise estimation of intensity. Thus our crude estimates of targeted intensity functioned merely as a means to determine relative ranking for between study comparisons. Assumptions were also made that resistance exercise provided minimal contributions to VO_{2peak}. A further limitation of our meta-analysis was the small number of included studies, which permitted the analysis of only two moderator variables. Thus, further research is needed particularly in survivors of cancers other than breast cancer.

Studies included in this review were generally of good methodological quality with low risk of bias. However, further attention to study quality is needed, as many studies did not adequately report methods for allocation concealment and use of blinded assessment, limiting our ability to evaluate the impact of risk of bias across studies. Of note, the estimated effect size was lower when excluding studies at high risk of bias; thus, our findings may represent an overestimate of the effect of supervised exercise on aerobic capacity.

A final limitation is that the mechanism(s) responsible for the improvement in VO_{2peak} along the oxygen cascade were not studied in any of the studies included in our review; thus, the favourable finding in VO_{2peak} may be due to improved convective and/or diffusive oxygen transport coupled with improved oxygen utilization by the active muscles[48].

Supervised aerobic exercise training was found to have a moderate-to-large beneficial effect on VO_{2peak}. Aerobic capacity increased in a dose response fashion with overall workload, with larger effects found in studies prescribing a higher overall workload of aerobic exercise. Larger benefits were also seen in studies with better participant attendance and among survivors of haematological cancers. There is a need for further randomized controlled trials examining supervised aerobic exercise interventions in understudied but common cancers such as prostate, lung and colorectal cancer.

COMMENTS

Background

Evidence is accumulating to support the benefit of exercise to improve the physical functioning and quality of life of survivors. Currently, the optimal exercise prescription is unknown and the effect of variations in exercise training parameters on cancer-specific outcomes are poorly understood. Therefore, questions remain over how to best tailor exercise prescriptions to optimize the health outcomes of survivors at different times through the cancer continuum.

Research frontiers

A previous meta-analysis included data from six randomized controlled trials and reported a significant benefit from supervised aerobic exercise training, compared with usual care, on VO_{2peak} (2.90 mL/kg per minute; 95%CI: 1.16, 4.64; $P = 0.01$). However, statistical and clinical heterogeneity was found among the exercise trials included in their review, and therefore further research was indicated to build on and extend the current knowledge in the field.

Innovations and breakthroughs

Pooling of the 13 studies (14 comparisons) reporting VO_{2peak} (mL/kg per minute) showed a statistically significant mean difference in VO_{2peak} of 3.13 mL/kg per minute (95%CI: 2.21, 4.05; $P < 0.001$) in favour of supervised aerobic exercise training; however, again moderate heterogeneity was found among the included studies ($I^2 = 58\%$; $P < 0.001$). Meta-regression was performed analyzing the effect estimate with exercise parameters of exercise workload and participant adherence as potential moderators. These two variables, workload and adherence, explained 65.8% ($P = 0.04$) of the between-study variance in effect estimate among the included studies.

Applications

Supervised aerobic exercise training is an effective intervention to improve aerobic capacity in survivors of cancer. Aerobic capacity increased in a dose response fashion with overall workload, with larger effects found in studies prescribing a higher overall workload of aerobic exercise. Larger benefits were also seen in studies with better participant attendance and among survivors of haematological cancers.

Terminology

Aerobic capacity (VO_{2max}) is the maximum volume of oxygen that the body can consume during maximal exercise, using at least 60% of the musculature, and while breathing air at sea level. Aerobic capacity is best increased by habitual aerobic exercise training that is of a moderate-to-vigorous intensity.

Peer review

An excellent systematic review.

REFERENCES

1 **Jemal A**, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011; **61**: 69-90 [PMID: 21296855 DOI: 10.3322/caac.20107]

2 **Stubblefield MD**, Hubbard G, Cheville A, Koch U, Schmitz KH, Dalton SO. Current perspectives and emerging issues on cancer rehabilitation. *Cancer* 2013; **119** Suppl 11: 2170-2178 [PMID: 23695929 DOI: 10.1002/cncr.28059]

3 **Silver JK**, Baima J, Mayer RS. Impairment-driven cancer rehabilitation: an essential component of quality care and survivorship. *CA Cancer J Clin* 2013; **63**: 295-317 [PMID: 23856764 DOI: 10.3322/caac.21186]

4 **Schmitz KH**, Speck RM. Risks and benefits of physical activity among breast cancer survivors who have completed treatment. *Womens Health* (Lond Engl) 2010; **6**: 221-238 [PMID: 20187728 DOI: 10.2217/whe.10.1]

5 **Winters-Stone KM**, Neil SE, Campbell KL. Attention to principles of exercise training: a review of exercise studies for survivors of cancers other than breast. *Br J Sports Med* 2014; **48**: 987-995 [PMID: 23293010]

6 **Myers J**, Prakash M, Froelicher V, Do D, Partington S, Atwood JE. Exercise capacity and mortality among men referred for exercise testing. *N Engl J Med* 2002; **346**: 793-801 [PMID: 11893790 DOI: 10.1056/NEJMoa011858]

7 **Kaleth AS**, Kaleth J. Demonstrating functional outcomes for health and fitness. In: Myers J and Nieman D Acsm's resources for clinical exercise physiology. Baltimore, MD: Lippincott Williams & Wilkins, 2010

8 **Mora S**, Redberg RF, Cui Y, Whiteman MK, Flaws JA, Sharrett AR, Blumenthal RS. Ability of exercise testing to predict cardiovascular and all-cause death in asymptomatic women: a 20-year follow-up of the lipid research clinics prevalence study. *JAMA* 2003; **290**: 1600-1607 [PMID: 14506119 DOI: 10.1001/jama.290.12.1600]

9 **Jones LW**, Courneya KS, Mackey JR, Muss HB, Pituskin EN, Scott JM, Hornsby WE, Coan AD, Herndon JE, Douglas PS, Haykowsky M. Cardiopulmonary function and age-related decline across the breast cancer survivorship continuum. *J Clin Oncol* 2012; **30**: 2530-2537 [PMID: 22614980 DOI: 10.1200/JCO.2011.39.9014]

10 **Jones LW**. Physical activity and lung cancer survivorship. *Recent Results Cancer Res* 2011; **186**: 255-274 [PMID: 21113768 DOI: 10.1007/978-3-642-04231-7_11]

11 **Tabata I**, Nishimura K, Kouzaki M, Hirai Y, Ogita F, Miyachi M, Yamamoto K. Effects of moderate-intensity endurance and high-intensity intermittent training on anaerobic capacity and VO2max. *Med Sci Sports Exerc* 1996; **28**: 1327-1330 [PMID:

8897392 DOI: 10.1097/00005768-199610000-00018]
12 **Myers J**, Nieman DC, American College of Sports Medicine. Acsm's resources for clinical exercise physiology: Musculoskeletal, neuromuscular, neoplastic, immunologic, and hematologic conditions. 2nd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins Health, 2010
13 **Jones LW**, Liang Y, Pituskin EN, Battaglini CL, Scott JM, Hornsby WE, Haykowsky M. Effect of exercise training on peak oxygen consumption in patients with cancer: a meta-analysis. *Oncologist* 2011; **16**: 112-120 [PMID: 21212429 DOI: 10.1634/theoncologist.2010-0197]
14 **Shephard RJ**. Aerobic fitness and health/roy j. Shephard. Champaign, IL: Human Kinetics, c1994, 1994
15 **Higgins JPT**, Green S. Assessing risk of bias in included studies. Cochrane Handbook for Systematic Reviews of Interventions, 2008: 187
16 **Swain DP**, Abernathy KS, Smith CS, Lee SJ, Bunn SA. Target heart rates for the development of cardiorespiratory fitness. *Med Sci Sports Exerc* 1994; **26**: 112-116 [PMID: 8133731]
17 **Dalleck LC**, Kravitz L. Relationship Between %Heart Rate Reserve And %VO2 Reserve During Elliptical Crosstrainer Exercise. *J Sports Sci Med* 2006; **5**: 662-671 [PMID: 24357963]
18 **Arts FJ**, Kuipers H. The relation between power output, oxygen uptake and heart rate in male athletes. *Int J Sports Med* 1994; **15**: 228-231 [PMID: 7960315]
19 **Cohen J**. A power primer. *Psychol Bull* 1992; **112**: 155-159 [PMID: 19565683]
20 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327.7414.557]
21 **Lipsey MW**, Wilson DB. Practical meta-analysis. Thousand Oaks, Calif.: Sage Publications, 2001: 247
22 **Adamsen L**, Quist M, Andersen C, Møller T, Herrstedt J, Kronborg D, Baadsgaard MT, Vistisen K, Midtgaard J, Christiansen B, Stage M, Kronborg MT, Rørth M. Effect of a multimodal high intensity exercise intervention in cancer patients undergoing chemotherapy: randomised controlled trial. *BMJ* 2009; **339**: b3410 [PMID: 19826172 DOI: 10.1136/bmj.b3410]
23 **Broderick JM**, Guinan E, Kennedy MJ, Hollywood D, Courneya KS, Culos-Reed SN, Bennett K, O' Donnell DM, Hussey J. Feasibility and efficacy of a supervised exercise intervention in de-conditioned cancer survivors during the early survivorship phase: the PEACH trial. *J Cancer Surviv* 2013; **7**: 551-562 [PMID: 23749688 DOI: 10.2007/s11764-013-0294-6]
24 **Burnham TR**. The effects of exercise on physiological and psychological variables in cancer patients following clinical treatment. Oregon State University, 2000: 56
25 **Courneya KS**, Mackey JR, Bell GJ, Jones LW, Field CJ, Fairey AS. Randomized controlled trial of exercise training in postmenopausal breast cancer survivors: cardiopulmonary and quality of life outcomes. *J Clin Oncol* 2003; **21**: 1660-1668 [PMID: 12721239 DOI: 10.1200/JCO.2003.04.093]
26 **Courneya KS**, Segal RJ, Mackey JR, Gelmon K, Reid RD, Friedenreich CM, Ladha AB, Proulx C, Vallance JK, Lane K, Yasui Y, McKenzie DC. Effects of aerobic and resistance exercise in breast cancer patients receiving adjuvant chemotherapy: a multicenter randomized controlled trial. *J Clin Oncol* 2007; **25**: 4396-4404 [PMID: 17785708 DOI: 10.1200/JCO.2006.08.2014]
27 **Courneya KS**, Sellar CM, Stevinson C, McNeely ML, Peddle CJ, Friedenreich CM, Tankel K, Basi S, Chua N, Mazurek A, Reiman T. Randomized controlled trial of the effects of aerobic exercise on physical functioning and quality of life in lymphoma patients. *J Clin Oncol* 2009; **27**: 4605-4612 [PMID: 19687337 DOI: 10.1200/JCO.2008.20.0634]
28 **Herrero F**, San Juan AF, Fleck SJ, Balmer J, Pérez M, Cañete S, Earnest CP, Foster C, Lucía A. Combined aerobic and resistance training in breast cancer survivors: A randomized, controlled pilot trial. *Int J Sports Med* 2006; **27**: 573-580 [PMID: 16802254 DOI: 10.1055/S-2005-865848]
29 **Hornsby WE**, Douglas PS, West MJ, Kenjale AA, Lane AR, Schwitzer ER, Ray KA, Herndon JE, Coan A, Gutierrez A, Hornsby KP, Hamilton E, Wilke LG, Kimmick GG, Peppercorn JM, Jones LW. Safety and efficacy of aerobic training in operable breast cancer patients receiving neoadjuvant chemotherapy: a phase II randomized trial. *Acta Oncol* 2014; **53**: 65-74 [PMID: 23957716 DOI: 10.3109/0284186X.2013.781673]
30 **Hwang CL**, Yu CJ, Shih JY, Yang PC, Wu YT. Effects of exercise training on exercise capacity in patients with non-small cell lung cancer receiving targeted therapy. *Support Care Cancer* 2012; **20**: 3169-3177 [PMID: 22526147 DOI: 10.1007/S00520-012-1452-5]
31 **Jarden M**, Baadsgaard MT, Hovgaard DJ, Boesen E, Adamsen L. A randomized trial on the effect of a multimodal intervention on physical capacity, functional performance and quality of life in adult patients undergoing allogeneic SCT. *Bone Marrow Transplant* 2009; **43**: 725-737 [PMID: 19234513 DOI: 10.1038/bmt.2009.27]
32 **Kaibori M**, Ishizaki M, Matsui K, Nakatake R, Yoshiuchi S, Kimura Y, Kwon AH. Perioperative exercise for chronic liver injury patients with hepatocellular carcinoma undergoing hepatectomy. *Am J Surg* 2013; **206**: 202-209 [PMID: 23374372 DOI: 10.1016/j.amjsurg.2012.07.035]
33 **Kim CJ**, Kang DH, Smith BA, Landers KA. Cardiopulmonary responses and adherence to exercise in women newly diagnosed with breast cancer undergoing adjuvant therapy. *Cancer Nurs* 2006; **29**: 156-165 [PMID: 16565627 DOI: 10.1097/00002820-200603000-00013]
34 **MacVicar MG**, Winningham ML, Nickel JL. Effects of aerobic interval training on cancer patients' functional capacity. *Nurs Res* 1989; **38**: 348-351 [PMID: 2587289 DOI: 10.1097/00006199-198911000-00007]
35 **Mehnert A**, Veers S, Howaldt D, Braumann KM, Koch U, Schulz KH. Effects of a physical exercise rehabilitation group program on anxiety, depression, body image, and health-related quality of life among breast cancer patients. *Onkologie* 2011; **34**: 248-253 [PMID: 21577030 DOI: 10.1159/000327813]
36 **Naumann F**, Martin E, Philpott M, Smith C, Groff D, Battaglini C. Can counseling add value to an exercise intervention for improving quality of life in breast cancer survivors? A feasibility study. *J Support Oncol* 2012; **10**: 188-194 [PMID: 22169703 DOI: 10.1016/j.suponc.2011.09.004]
37 **Rahnama N**, Nouri R, Rahmaninia F, Damirchi A, Emami H. The effects of exercise training on maximum aerobic capacity, resting heart rate, blood pressure and anthropometric variables of postmenopausal women with breast cancer. *J Res Med Sci* 2010; **15**: 78-83 [PMID: 21526063]
38 **Segal R**, Evans W, Johnson D, Smith J, Colletta S, Gayton J, Woodard S, Wells G, Reid R. Structured exercise improves physical functioning in women with stages I and II breast cancer: results of a randomized controlled trial. *J Clin Oncol* 2001; **19**: 657-665 [PMID: 11157015]
39 **Thorsen L**, Skovlund E, Strømme SB, Hornslien K, Dahl AA, Fosså SD. Effectiveness of physical activity on cardiorespiratory fitness and health-related quality of life in young and middle-aged cancer patients shortly after chemotherapy. *J Clin Oncol* 2005; **23**: 2378-2388 [PMID: 15800330 DOI: 10.1200/JCO.2005.04.106]
40 **Gulati M**, Pandey DK, Arnsdorf MF, Lauderdale DS, Thisted RA, Wicklund RH, Al-Hani AJ, Black HR. Exercise capacity and the risk of death in women: the St James Women Take Heart Project. *Circulation* 2003; **108**: 1554-1559 [PMID: 12975254 DOI: 10.1161/01.CIR.0000091080.57509.E9]
41 **Peel JB**, Sui X, Adams SA, Hébert JR, Hardin JW, Blair SN. A prospective study of cardiorespiratory fitness and breast cancer mortality. *Med Sci Sports Exerc* 2009; **41**: 742-748 [PMID: 19276861 DOI: 10.1249/MSS.0b013e31818edac7]
42 **Burnett D**, Kluding P, Porter C, Fabian C, Klemp J. Car-

diorespiratory fitness in breast cancer survivors. *Springerplus* 2013; **2**: 68 [PMID: 23538987 DOI: 10.1186/2193-1801-2-6]

43 **Sui X**, Lee DC, Matthews CE, Adams SA, Hébert JR, Church TS, Lee CD, Blair SN. Influence of cardiorespiratory fitness on lung cancer mortality. *Med Sci Sports Exerc* 2010; **42**: 872-878 [PMID: 19996990 DOI: 10.1249/MSS.ob013e3181c47b65]

44 **Floyd A**, Moyer A. Group vs. individual exercise interventions for women with breast cancer: a meta-analysis. *Health Psychol Rev* 2009; **4**: 22-41 [PMID: 20607139 DOI: 10.1080/17437190903384291]

45 **Ferrer RA**, Huedo-Medina TB, Johnson BT, Ryan S, Pescatello LS. Exercise interventions for cancer survivors: a meta-analysis of quality of life outcomes. *Ann Behav Med* 2011; **41**: 32-47 [PMID: 20931309 DOI: 10.1007/s12160-010-9225-1]

46 **Brown JC**, Huedo-Medina TB, Pescatello LS, Ryan SM, Pescatello SM, Moker E, LaCroix JM, Ferrer RA, Johnson BT. The efficacy of exercise in reducing depressive symptoms among cancer survivors: a meta-analysis. *PLoS One* 2012; **7**: e30955 [PMID: 22303474 DOI: 10.1371/journal.pone.0030955]

47 **Carayol M**, Bernard P, Boiché J, Riou F, Mercier B, Cousson-Gélie F, Romain AJ, Delpierre C, Ninot G. Psychological effect of exercise in women with breast cancer receiving adjuvant therapy: what is the optimal dose needed? *Ann Oncol* 2013; **24**: 291-300 [PMID: 23041586 DOI: 10.1093/annonc/mds342]

48 **Jones LW**, Eves ND, Haykowsky M, Freedland SJ, Mackey JR. Exercise intolerance in cancer and the role of exercise therapy to reverse dysfunction. *Lancet Oncol* 2009; **10**: 598-605 [PMID: 19482248 DOI: 10.1016/S1470-2045(09)70031-2]

P- Reviewer: Lim ECW, Tokuhashi Y **S- Editor**: Ji FF
L- Editor: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.54

World J Meta-Anal 2015 February 26; 3(1): 54-60
ISSN 2308-3840 (online)

Hypertension in Zimbabwe: A meta-analysis to quantify its burden and policy implications

Mutsa Pamela Mutowo, John Chamunorwa Mangwiro, Paula Lorgelly, Alice Owen, Andre MN Renzaho

Mutsa Pamela Mutowo, Alice Owen, Andre MN Renzaho, School of Public Health and Preventive Medicine, Monash University, Melbourne 3004, Australia
John Chamunorwa Mangwiro, Zimbabwe Diabetes Association, PO Box 1797, Harare, Zimbabwe
Paula Lorgelly, Centre for Health Economics, Clayton Campus, Monash University, Melbourne 3800, Australia
Andre MN Renzaho, School of Social Science and Psychology, University of Western Sydney, Sydney 2751, New South Wales, Australia
Author contributions: Mutowo MP and Renzaho AMN performed the systematic review; Mutowo MP performed the meta-analysis, analyzed the data and wrote up results; Owen A, Mangwiro JC and Lorgelly P previewed retrieved papers and analyzed data; Renzaho AMN, Owen A, Lorgelly P and Mangwiro JC critically reviewed and edited the manuscript, and contributed towards sections in the paper.
Supported by An Australian Research Council Future Fellowship, No. FT110100345 (to Renzaho A); The sponsor had no role in study design, data collection, analysis, interpretation, or writing of the report.
Conflict-of-interest: The authors declare no conflict of interest.
Data sharing: No additional data is available.

Correspondence to: Andre MN Renzaho, PhD, MPH, MPHAA, Professor of Humanitarian and Development Studies, School of Social Science and Psychology, University of Western Sydney, James Ruse Drive, Sydney 2751, New South Wales, Australia. andre.renzaho@uws.edu.au
Telephone: +61-2-47360107
Received: October 24, 2014
Peer-review started: October 27, 2014
First decision: November 14, 2014
Revised: December 16, 2014
Accepted: December 29, 2014
Article in press: December 31, 2014
Published online: February 26, 2015

Abstract

AIM: To estimate the pooled prevalence of hypertension in Zimbabwe and describe its trend since independence in 1980 using secondary source data.

METHODS: MEDLINE, EMBASE and Scopus databases from April 1980 to December 2013 were searched for population and community based studies on the prevalence of hypertension among adults (≥ 18 years) in Zimbabwe. The key words used were "prevalence", "epidemiologic studies", "hypertension" or "high blood pressure", based on the cut-off (≥ 140 mmHg systolic blood pressure and/or ≥ 90 mmHg diastolic blood pressure). We conducted a meta-analysis on the published studies, using the random-effects model to estimate the pooled prevalence.

RESULTS: The search retrieved 87 publications, of which four studies met the selection criteria. The four studies had a total of 4829 study participants between 1997 and 2010 across 5 provinces in Zimbabwe. Two studies were in urban areas, while the other two had mixed study settings (urban and rural). The overall pooled prevalence of hypertension was 30% (95%CI: 19%, 42%, I^2= 98%, χ^2 = 164.15, P = 0.00).

CONCLUSION: Our results show a high prevalence of hypertension in Zimbabwe, with urban areas having higher prevalence than rural areas.

Key words: Hypertension; High blood pressure; Prevalence; Meta-analysis; Zimbabwe

Core tip: A systematic review and meta-analysis of studies on the prevalence of hypertension in Zimbabwe, from April 1980 to December 2013 reveals a high prevalence of 30%. Hypertension prevalence was higher

in studies in urban settings compared with studies in mixed settings (urban and rural), indicating the increase of cardiovascular risk factors associated with urbanization and economic progress. The development of national prevention policies and control strategies for hypertension are critical to reduce the increasing burden of hypertension in Zimbabwe.

Mutowo MP, Mangwiro JC, Lorgelly P, Owen A, Renzaho AMN. Hypertension in Zimbabwe: A meta-analysis to quantify its burden and policy implications. *World J Meta-Anal* 2015; 3(1): 54-60 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/54.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.54

INTRODUCTION

Hypertension-related conditions are the most common cause of death from non-communicable diseases (NCDs) in sub-Saharan Africa[1]. Hypertension is recognized as a global public health crisis due to it being asymptomatic and its high mortality rate[2]. The prevalence of hypertension is estimated at 22.9% in developing countries and 37.3% in developed countries[3]. Unfortunately, Zimbabwe faces the particular challenge of high morbidity and mortality from communicable diseases and increasing prevalence of NCDs[4]. NCDs accounted for an estimated 21% of total deaths in 2008 in Zimbabwe[5] and hypertension was ranked first amongst the NCD outpatient visits recorded in Zimbabwean public hospitals in 2006[4]. The limited data available suggests that there was a four-fold increase in the prevalence of hypertension from 1990 to 1997[6], and the age-standardized rate of hypertension in Zimbabwe (33.1%) was reported in one study to be higher than that seen in developed countries such as United States of America (20.3%), Canada (21.4%) and England (29.6%)[3].

Urbanization has resulted in the westernization of lifestyles in parts of Zimbabwe. In urban areas, diets high in refined, starchy carbohydrates are leading to high obesity rates and increased prevalence of hypertension, diabetes and cardiovascular diseases[7]. Hypertension awareness is low, resulting in inadequate treatment and management of hypertension in the Zimbabwean population, and hence there is an urgent need for a national policy for the prevention and control of hypertension in Zimbabwe[8]. This should include a major focus on prevention, as this may be more cost effective for a developing country with limited resources[9]. This will require development of evidence-based prevention strategies, which must be informed by a clear understanding of the hypertension burden across the country. However in Zimbabwe, as in many other resource-limited settings, the infrastructure available to enable detailed disease surveillance activities is lacking and no national studies on hypertension prevalence in Zimbabwe are available. The purpose of this study was to systematically review the epidemiological results of published studies and estimate the pooled prevalence of hypertension in Zimbabwe using meta-analysis.

MATERIALS AND METHODS

Search strategy

The systematic review and meta-analysis was conducted according to the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Group[10]. Published epidemiologic studies on the prevalence of hypertension we searched for between April 1980 and December 2013 in three electronic databases: MEDLINE, EMBASE and Scopus. The medical subject headings (MeSH) terms used in all databases were ("hypertension" OR "high blood pressure") AND ("prevalence" OR "epidemiological studies") AND ("Zimbabwe"). Prior to the national independence of Zimbabwe, on 18 April 1980, the nation had been known by several names including Rhodesia, Southern Rhodesia, and Zimbabwe-Rhodesia. We further searched the grey literature databases and individual Zimbabwean public health institute websites for relevant studies.

Criteria for inclusion and exclusion

Inclusion criteria for studies included studies on the prevalence of hypertension or high blood pressure, conducted among Zimbabwean residents (≥ 18 years old); population or community studies that were cross-sectional or cohort studies and cut off points for hypertension were systolic blood pressure (SBP) (≥ 140 mmHg) and/or diastolic blood pressure (DBP) (≥ 90 mmHg).

Studies had to abide by the hypertension diagnostic criteria of the Seventh Report of the Joint National committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC7)[11], and/or the 1999 World Health Organization (WHO)/International Society of Hypertension (WHO/ISH) classification of blood pressure levels[12], and/or the 2003 WHO/ISH Statement on Management of Hypertension[13], whose cut-off points are based on 140/90 mmHg. Studies conducted before 1999 had blood pressure cut-off points defined as ≥ 160/95 mmHg. Subgroup prevalence based on the cut-off point based on 140/90 mmHg was included from these studies. Articles were excluded if the participants were limited to gender (male or female only), pregnant participants, studies conducted on animals, editorial letters, abstracts, and reviews of original studies.

Study selection

Identified studies were screened by two independent reviewers (MM and AR) to confirm whether they satisfied the inclusion criteria. Lack of consensus about study selection was resolved through discussions with a third author (JC). Retrieved articles and their reference lists were searched for additional publications.

Data extraction

All data was independently extracted by the two reviewers

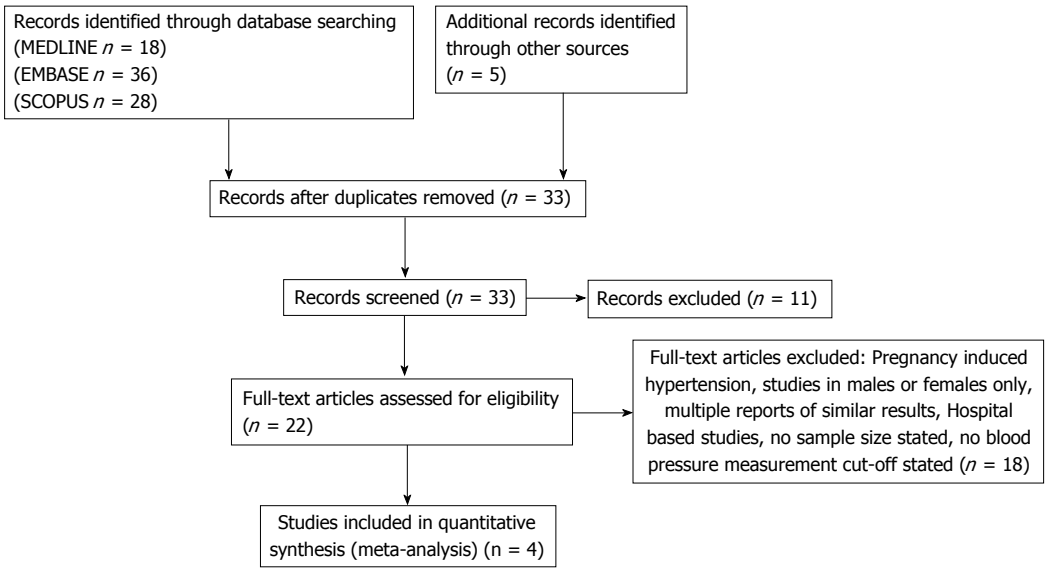


Figure 1 Flow diagram of the study selection process. As shown in Figure 1, our initial search yielded 87 citations: 18 from MEDLINE, 36 from EMBASE, 28 from Scopus, and 5 from grey literature. After screening titles and abstracts, 22 studies were considered potentially eligible and retrieved in full text. Of these, 18 studies were subsequently excluded because they did not satisfy the inclusion criteria. Thus, four fully eligible studies were identified.

(MM and AR), cross-checked and any disagreements were resolved by consensus. The following information was recorded from the included studies: author, year of publication, year of investigation, study period, study setting, sampling frame and method, sample size, age range of study population, reported prevalence, and diagnostic method and criteria used in the study.

Statistical analysis

The Cochran Q test or χ^2 and the I^2 statistic were used to evaluate and quantify statistical heterogeneity[14,15]. The values for χ^2 and I^2 (low is < 25%, moderate 25%-50%, high > 50%) are mentioned in the forest plot used to visualize the magnitude of heterogeneity among studies. As the differences between studies were very large (I^2 = 98%), we used a random-effects model to estimate the prevalence of hypertension and calculate the 95%CI[15]. All statistical analysis was done using MetaXL 1.4, Software[16]. The statistical methods of this study were reviewed by Dr. Baki Billah, Senior Biostatistician Consultant and Senior Lecturer in Biostatistics, from Monash University, Australia.

Dr. Baki Billah, Senior Biostatistician Consultant and Senior Lecturer in Biostatistics at Monash University reviewed and confirmed that the statistical approach reported in the manuscript was adequate and correct.

RESULTS

We initially identified 87 references from our search: 82 from electronic databases and 5 from other sources (Figure 1). After the application of inclusion and exclusion criteria, and removing duplications, as described in Methods, we selected four studies for the meta-analysis.

The four studies[8,17-19] were conducted across five provinces in Zimbabwe. The studies had a total of 4829 subjects and the enrollment years of the studies ranged from 1997 to 2010. Two studies[8,17] conducted in predominately urban areas, had a total sample size of 1077, while the other two studies[18,19], conducted in both urban and rural settings, had a total sample size of 3752. The four studies did not state age-specific data related to gender, and age was limited to above 25 years old in the four studies. Two studies[17,19] stated the use of JNC7 and WHO/ISH 2003 classifications, while the other two used cut-off points within the inclusion criteria.

Awareness of hypertension was found to be low and treatment and management of hypertension inadequate in one study sample[8]. One study reported a prevalence which was higher in females than in males and a family history of hypertension which was strongly associated with hypertension in participants in the study[19]. The commonly reported family members were mothers of participants and on stratified analysis, the association of hypertension and family history of hypertension was stronger in females than males. The study reported a high prevalence of abdominal obesity which is a powerful determinant of subsequent risk of hypertension[19]. Three studies reported the use of standardized measurement protocols, utilizing nurses or certified personnel for blood pressure measurement, with validity of readings done by a supervising physician[8,17,19]. Blood pressure was measured two times in a single visit in two studies[17,18], three times in a single visit in one study and the process for obtaining blood pressure readings was not reported in one study[8,19]. Two studies used standard mercury sphygmomanometer to measure blood pressure[8,17]; one study used digital blood pressure machines[19], while no specific instrument was reported for the remaining study[18].

Based on the reported hypertension prevalence in the included studies, Bulawayo (south Zimbabwe) had the highest prevalence of 38.4% (95%CI: 33%-44%)[19].

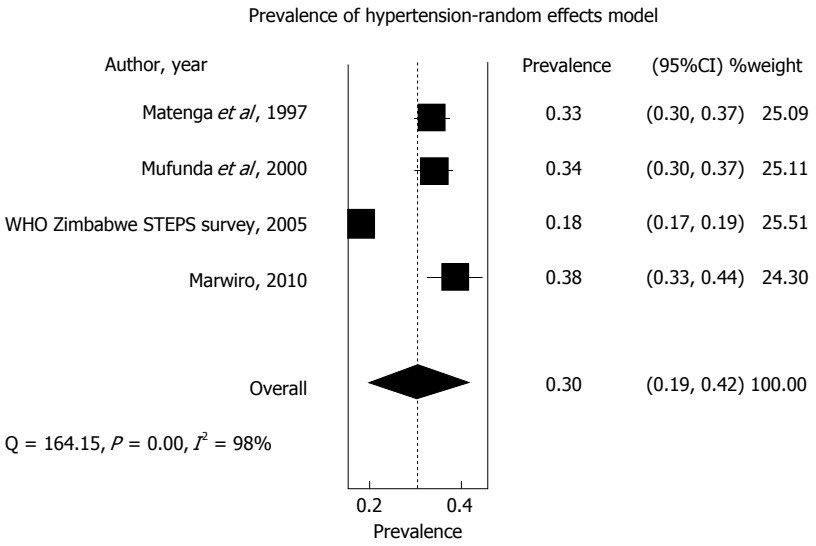


Figure 2 Forest plot of studies conducted from 1997 to 2010 on Hypertension prevalence. The forest plot depicted in Figure 2 (above) represents a meta-analysis of studies that measured the prevalence of hypertension in Zimbabwe from 1997 to 2010. Individual studies with their unadjusted prevalence are represented by a black square and a horizontal line, which corresponds to the point estimate and 95%CI of prevalence. The size of the black square reflects the weight of the study in the meta-analysis. The diamond at the bottom represents the pooled estimate of all studies with its 95% confidence interval. In this case, Figure 2 indicates the pooled estimated prevalence of hypertension is 30% (95%CI: 19-42). The test for overall prevalence also indicates statistical significance ($P < 0.0001$).

The lowest prevalence of 17.9% (95%CI: 17%-19%) was recorded across three provinces in mixed study setting (urban and rural)[17] (Table 1 summarizes the extracted data from included studies). Using the random-effects model for the meta-analysis, the overall hypertension prevalence is estimated to be 30% (95%CI: 19%-42%, I^2 = 98%, χ^2 = 164.15, P = 0.00) (Figure 2).

DISCUSSION

There is a shortage of national data on hypertension prevalence in Zimbabwe. This study summarized the prevalence of hypertension in Zimbabwe over a 14 year period (1997 to 2010). The estimated pooled prevalence for hypertension for the 14 year period was 30%, however as this was not age-standardized and is likely to be an under-estimate. The hypertension prevalence for Zimbabwe, estimated by the WHO was higher at 39% for both genders aged at least 25 years, 38.2% (95%CI: 29.9-46.9) in men and 39.9% (95%CI: 30.4-49.4) in women[20]. However, concerns remain over the different cut-off points used for hypertension measurement in prevalence studies, data sources and modelling methodology and assumptions used, so this creates difficulties in comparing prevalence rates across Africa[21-23].

Despite this, the observed trend towards increasing hypertension prevalence in our meta-analysis is congruent with the literature. Studies have indicated that the prevalence of hypertension has increased in developing countries over recent decades, with hypertension increasingly prevalent in lower socio-economic groups with limited access to essential treatment[24,25].

Hypertension was found to be higher in the urban Zimbabwe population[8,19]. Rapid urbanization and lifestyle changes have been implicated in the development of hypertension in African urban populations, notably adoption of Western-type diet, physical inactivity and increased psychosocial stress[6]. Hypertension was found to be prevalent in the lowest income groups, more common in women, linked with overweight and obesity and in heavy alcohol consumers in low income countries[25]. The Zimbabwe National Health Strategy reports that the prevalence of hypertension in Zimbabwe is increasing mainly because of physical inactivity, tobacco smoking, high salt diet and excessive alcohol consumption[4]. Therefore preventive measures need to take into account urban planning, whereby effective policies can promote physical activity through re-designing the landscape.

Hypertension is generally asymptomatic until chronic vascular disease develops, with the risk of disease doubling with each blood pressure reading increase of 20/10 mmHg, beginning at lower readings of 115/75 mmHg[21]. The lack of symptoms contributes not only to the lack of awareness of the condition in those who have it, but also reduces the levels of compliance and persistence with blood pressure lowering interventions, as an improvement in blood pressure control may not result in perceptible benefit to the individual[25]. The largest cause of years of life lost in low income countries is cardiovascular disease[25], and with a growing prevalence of hypertension, the burden of cardiovascular diseases in Zimbabwe is likely to increase, which has significant implications for healthcare, individual wellbeing and social stability.

The limited number of population-based studies on hypertension prevalence and risk factors may have contributed to its low priority as a public health problem in Zimbabwe, when compared to higher profile communicable diseases like human immunodeficiency virus infection and acquired immune deficiency syndrome (HIV/AIDS),

Table 1 Characteristics of studies included in the meta-analysis

Ref.	Study period	Setting	Sampling method	Sample size	Age range (yr)	Prevalence (cases)	Diagnostic criteria	Description of geographic area[1]
Matenga *et al*[8]	October to early December 1996	Community-household	Random	749	> 34	33.4% (250)	Hypertensive described as mean diastolic BP > 94 mmHg untreated or on antihypertensive medication, controlled BP described as mean DBP < 95 mmHg while on drug treatment	Marondera, Mashonaland East (mainly urban and unspecified rural area)
Hakim *et al*[17]	May to July 2005	Subnational-household	Multi-stage	3003	≥ 25	17.9% (538)	Systolic ≥ 140 and/or diastolic ≥ 90 mmHg	Urban and mainly rural communities in Midlands, Mashonaland Central, and Matebeleland South
Mufunda *et al*[18]	July-October 1995	Community-household	Cluster sampling	775	> 25	33.5% (260)	Systolic ≥ 140 mmHg and diastolic ≥ 90 mmHg and/or antihypertensive medication	Marondera, Mashonaland East (urban)
Marwiro[19]	June-July, 2010	Community-employee register	Systematic	302	> 25 to > 55	38.4% (116)	The 7th Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure. Hypertension stage 1: systolic 140-159 mmHg, diastolic 90-99 mmHg, Hypertension stage 2: systolic ≥ 160 mmHg, diastolic ≥ 100 mmHg	Bulawayo (urban)

[1]Geographic area refers to the geographic location where study took place in Zimbabwean urban or rural areas. DBP: Diastolic blood pressure.

malaria and tuberculosis. Unlike HIV/AIDS, hypertension is not considered a health priority in Zimbabwe, and no national hypertension program has been established to date. However the HIV/AIDS epidemic in Zimbabwe adds a new dimension to the hypertension burden. The use of highly active antiretroviral therapy to treat HIV is also associated with increased risk of high blood pressure[26-28].

National programs to diagnose and treat hypertension can lower cerebrovascular disease burden by at least one third[9]. A focus on primary prevention, through awareness and screening programs, training the health work force to deal with hypertension and its associated risk factors, and access to low-cost anti-hypertensive agents is likely to be more cost-effective for a developing country with limited resources[29]. Emphasis should be placed on modifiable behavioral factors, such as lifestyle behaviors of family environment, dietary changes, weight reduction and cessation of smoking, all potentially modifiable, and likely to yield greater impact than concentrating on genetic factors for hypertension[21,30]. Primary prevention of hypertension prevents and reduces the expensive management of hypertension and its ensuing complications[31].

Limitations

We followed the guidelines for reporting systematic reviews and meta-analysis[10], however certain drawbacks deserve attention.

Heterogeneity: The sample sizes in the studies used for the meta-analysis totalled a few hundred in three studies to a few thousands in one study. The number of included studies was very small, and various risk factors known to influence heterogeneity were not taken into account. The use of a few studies with large differences in sample size in a meta-analysis, results in pooled estimates with low precision and power, and higher χ^2 and I^2[32]. Due to insufficient data in the included studies, we were unable to perform sub-group analysis to assess the outcome of variations on the pooled prevalence.

Blood pressure measurement: The different methods of measuring blood pressure are documented in literature[22,23,33]. The World Health Organization recommends risk factor surveys measure blood pressure three times per single visit and use the average result[33], which was only done in two studies[8,18], as one measurement per single visit could result in overstated readings[34]. The number of blood pressure readings recorded has been found to determine whether a patient is classified as hypertensive[35].

Representativeness: A significant obstacle in developing effective national hypertension prevention programs is the lack of high quality health information systems to inform policy makers[36]. The burden of NCDs, such as hypertension, is not well documented in Zimbabwe, as its information system has communicable diseases as the main priority. Results from our meta-analysis indicate information on hypertension prevalence in Zimbabwe is limited with no studies providing age-standardized data, thus making direct comparison of results between studies difficult.

In conclusion, our study highlights that estimating the true prevalence of hypertension in Zimbabwe is a challenge due to methodological differences. Therefore, longitudinal national surveys using standardized methodologies are urgently needed in the future to further define the prevalence of hypertension and depict trends.

COMMENTS

Background

World Health Organization estimates the prevalence of hypertension in Sub-Saharan Africa to be 46%, making hypertension a major threat to public health. However the response of many governments and international aid agencies to hypertension has been described as similar to the "reaction to human immunodeficiency virus infection and acquired immune deficiency syndrome (HIV/AIDS) 20 years ago". Zimbabwe (like most sub-Saharan African countries), faces the dual challenge of communicable and non-communicable diseases, however funds donated to fight HIV/AIDS consistently exceed all other national healthcare expenditure. Many countries at a similar stage of development receive more than 50 percent of their total healthcare budgets from donors. Hypertension and its ensuing complications accounted for less than 3% of the global health assistance between 2001 and 2008, despite 80% of deaths from cardiovascular disease occurring in low and middle-income countries. Therefore an estimate of the magnitude of the burden caused by hypertension in Zimbabwe is required to enable the government and international organizations to work together to reduce risk factors for non-communicable diseases such as hypertension.

Research frontiers

There are very few studies on hypertension prevalence in Zimbabwe. This is the first meta-analysis, to the knowledge, to systematically review studies conducted in Zimbabwe, and provide a pooled estimate of hypertension prevalence in Zimbabwe, with the aim of promoting increased awareness of hypertension, and initiate a policy response in Zimbabwe.

Innovations and breakthroughs

By providing a pooled estimate for the prevalence of hypertension in Zimbabwe using studies conducted in Zimbabwe, can assist policy makers in preventive policies and strategies suited for the Zimbabwean urban and rural population.

Applications

The meta-analysis aimed to consolidate data on hypertension prevalence in Zimbabwean urban and rural areas to determine the burden of hypertension in the country.

Terminology

Meta-analysis combines results from independent studies and explores the heterogeneity, as some studies are affected by small sample size and the quality of data. Heterogeneity is the differences in methodology or study populations used in the different studies under examination. Sources of inconsistency include study design, various forms of bias, and how the outcome is measured. The random-effects model is applied when studies have different effects and different characteristics. Forest plots enable the reader to view all the studies at once. One axis of the Forest plot displays the effect estimates (prevalence of hypertension expressed as a percentage for each study in the meta-analysis) and corresponding confidence intervals. The overall pooled prevalence estimate (with 95%CI) is represented as a diamond and placed towards the bottom of the plot.

Peer review

This manuscript is a meta-analysis on the prevalence of hypertension in Zimbabwe. Its' results have provide evidences on policies and interventions hypertension. The results are interesting.

REFERENCES

1 **Peck RN**, Green E, Mtabaji J, Majinge C, Smart LR, Downs JA, Fitzgerald DW. Hypertension-related diseases as a common cause of hospital mortality in Tanzania: a 3-year prospective study. *J Hypertens* 2013; **31**: 1806-1811 [PMID: 23777761 DOI: 10.1097/HJH.0b013e328362bad7]

2 **World Health Organization**. A Global brief on Hypertension.Silent killer, global public health crisis. Geneva: World Health Organization, 2013

3 **Kearney PM**, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: analysis of worldwide data. *Lancet* 2005; **365**: 217-223 [PMID: 15652604 DOI: 10.1016/S0140-6736(05)17741-1]

4 **Ministry of Health and Child Welfare, Zimbabwe**. National Health Strategy for Zimbabwe 2009-2013: Equity and Quality in Health: A People's Right. MOHCW, Zimbabwe, 2009

5 **World Health Organization**. NCD Country Profiles. 2011. [Accessed 2014 March 6]. Available from: URL: http: //www.who.int/nmh/countries/zwe_en.pdf

6 **Mufunda J**, Chatora R, Ndambakuwa Y, Nyarango P, Chifamba J, Kosia A, Sparks HV. Prevalence of noncommunicable diseases in Zimbabwe: results from analysis of data from the National Central Registry and Urban Survey. *Ethn Dis* 2006; **16**: 718-722 [PMID: 16937610]

7 **Mathe S**, Matovu HL, Mossop RT. Nutritional status of an urban community in Zimbabwe. *Cent Afr J Med* 1985; **31**: 59-62 [PMID: 4016909]

8 **Matenga J**, Allain TJ, Wilson AO, Adamchak DJ, Senzanje B, Mushangi E, and Gomo Z. Hypertension management in Zimbabwe- awareness, treatment and blood pressure control. *S Afr J Med* 1997; **87**: 1371-1373

9 The costs and benefits of prevention. *J Public Health Policy* 1980; **1**: 285-292 [PMID: 6788800 DOI: 10.2307/3342171]

10 **Moher D**, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097 [PMID: 19621072 DOI: 10.1371/journal.pmed.1000097]

11 **Chobanian AV**, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, Jones DW, Materson BJ, Oparil S, Wright JT, Roccella EJ. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA* 2003; **289**: 2560-2572 [PMID: 12748199 DOI: 10.1001/jama.289.19.2560]

12 **Mabadeje AF**. WHO-ISH Guidelines for the management of hypertension: implementation in Africa--the Nigerian experience. *Clin Exp Hypertens* 1999; **21**: 671-681 [PMID: 10423091 DOI: 10.3109/10641969909060998]

13 **Whitworth JA**. 2003 World Health Organization (WHO)/International Society of Hypertension (ISH) statement on management of hypertension. *J Hypertens* 2003; **21**: 1983-1992 [PMID: 14597836 DOI: 10.1097/00004872-200311000-00002]

14 **Riley RD**, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011; **342**: d549 [PMID: 21310794 DOI: 10.1136/bmj.d549]

15 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327.7414.557]

16 **Barendregt JJ**, Doi SA. Epigear (2012) MetaXL. 2012. Available from: URL: http: //www.epigear.com/index_files/MetaXL User Guide.pdf

17 **Hakim JG**, Mujuru N, Rusakaniko S, Gomo ZAR. WHO STEPS Survey Zimbabwe. Available from: URL: http: //www.who.int/chp/steps/STEPS_Zimbabwe_Data.pdf 2005

18 **Mufunda J**, Scott LJ, Chifamba J, Matenga J, Sparks B, Cooper R, Sparks H. Correlates of blood pressure in an urban Zimbabwean population and comparison to other populations of African origin. *J Hum Hypertens* 2000; **14**: 65-73 [PMID: 10673734 DOI: 10.1038/sj.jhh.1000886]

19 **Marwiro A**. Prevalence and Risk Factors for Hypertension among Bulawayo City Council Employees. Institutional Repository at University of Zimbabwe, Faculty of Medicine e-Theses Collection 29, 2012

20 **World Health Organization**. Global status report on non-communicable diseases. Geneva: WHO, 2011

21 **World Health Organization**. STEPwise approach to Surveillance of Chronic Diseases and Risk Factors Instrument. [Accessed 2014 March 5]. Available from: URL: http: //www.who.int/chp/steps/Part1.pdf

22 **Croft PR**. Standardising blood pressure measurement in everyday practice: what's the gold standard? *J Hum Hypertens* 1999; **13**: 85-86 [PMID: 10100055 DOI: 10.1038/sj.jhh.1000733]

23 **McAlister FA**, Straus SE. Evidence based treatment of hypertension. Measurement of blood pressure: an evidence based review. *BMJ* 2001; **322**: 908-911 [PMID: 11302909 DOI: 10.1136/bmj.322.7291.908]

24 **Kearney PM**, Whelton M, Reynolds K, Whelton PK, He J. Worldwide prevalence of hypertension: a systematic review. *J Hypertens* 2004; **22**: 11-19 [PMID: 15106785 DOI: 10.1097/00004872-200401000-00003]

25 **Lloyd-Sherlock P**, Beard J, Minicuci N, Ebrahim S, Chatterji S. Hypertension among older adults in low- and middle-income countries: prevalence, awareness and control. *Int J Epidemiol* 2014; **43**: 116-128 [PMID: 24505082 DOI: 10.1093/ije/dyt215]

26 **Blair Research Institute, Zimbabwe**. The early socio-demographic impact of the HIV-1 epidemic in rural Zimbabwe. Harare, Zimbabwe, 1996

27 **Crane HM**, Van Rompaey SE, Kitahata MM. Antiretroviral medications associated with elevated blood pressure among patients receiving highly active antiretroviral therapy. *AIDS* 2006; **20**: 1019-1026 [PMID: 16603854 DOI: 10.1097/01.aids.0000222074.45372.00]

28 **Dillon DG**, Gurdasani D, Riha J, Ekoru K, Asiki G, Mayanja BN, Levitt NS, Crowther NJ, Nyirenda M, Njelekela M, Ramaiya K, Nyan O, Adewole OO, Anastos K, Azzoni L, Boom WH, Compostella C, Dave JA, Dawood H, Erikstrup C, Fourie CM, Friis H, Kruger A, Idoko JA, Longenecker CT, Mbondi S, Mukaya JE, Mutimura E, Ndhlovu CE, Praygod G, Pefura Yone EW, Pujades-Rodriguez M, Range N, Sani MU, Schutte AE, Sliwa K, Tien PC, Vorster EH, Walsh C, Zinyama R, Mashili F, Sobngwi E, Adebamowo C, Kamali A, Seeley J, Young EH, Smeeth L, Motala AA, Kaleebu P, Sandhu MS. Association of HIV and ART with cardiometabolic traits in sub-Saharan Africa: a systematic review and meta-analysis. *Int J Epidemiol* 2013; **42**: 1754-1771 [PMID: 24415610 DOI: 10.1093/ije/dyt198]

29 **Miranda JJ**, Kinra S, Casas JP, Davey Smith G, Ebrahim S. Non-communicable diseases in low- and middle-income countries: context, determinants and health policy. *Trop Med Int Health* 2008; **13**: 1225-1234 [PMID: 18937743 DOI: 10.1111/j.1365-3156.2008.02116.x]

30 **Addo J**, Smeeth L, Leon DA. Hypertension in sub-saharan Africa: a systematic review. *Hypertension* 2007; **50**: 1012-1018 [PMID: 17954720 DOI: 10.1161/HYPERTENSIONAHA.107.093336]

31 National High Blood Pressure Education Program Working Group report on primary prevention of hypertension. *Arch Intern Med* 1993; **153**: 186-208 [PMID: 8422207 DOI: 10.1001/archinte.1993.00410020042003]

32 **Thorlund K**, Imberger G, Johnston BC, Walsh M, Awad T, Thabane L, Gluud C, Devereaux PJ, Wetterslev J. Evolution of heterogeneity (I2) estimates and their 95% confidence intervals in large meta-analyses. *PLoS One* 2012; **7**: e39471 [PMID: 22848355 DOI: 10.1371/journal.pone.0039471]

33 **World Health Organization**. WHO STEPS Surveillance Manual. Geneva: WHO, 2008

34 **Bovet P**, Gervasoni JP, Ross AG, Mkamba M, Mtasiwa DM, Lengeler C, Burnier M, Paccaud F. Assessing the prevalence of hypertension in populations: are we doing it right? *J Hypertens* 2003; **21**: 509-517 [PMID: 12640244 DOI: 10.1097/00004872-200303000-00016]

35 **Birkett NJ**. The effect of alternative criteria for hypertension on estimates of prevalence and control. *J Hypertens* 1997; **15**: 237-244 [PMID: 9468450 DOI: 10.1097/00004872-199715030-00004]

36 **Beaglehole R**, Ebrahim S, Reddy S, Voûte J, Leeder S. Prevention of chronic diseases: a call to action. *Lancet* 2007; **370**: 2152-2157 [PMID: 18063026 DOI: 10.1016/S0140-6736(07)61700-0]

P- Reviewer: Tan XR, Zhao D **S- Editor**: Ji FF **L- Editor**: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.61

World J Meta-Anal 2015 February 26; 3(1): 61-71
ISSN 2308-3840 (online)

META-ANALYSIS

Operative *vs* nonoperative treatment of displaced intra-articular calcaneal fracture: A meta-analysis of randomized controlled trials

Nan Jiang, Hui-Juan Song, Guo-Ping Xie, Lei Wang, Chang-Xiang Liang, Cheng-He Qin, Bin Yu

Nan Jiang, Guo-Ping Xie, Lei Wang, Chang-Xiang Liang, Cheng-He Qin, Bin Yu, Department of Orthopaedics and Traumatology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong Province, China
Hui-Juan Song, Department of Nursing, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong Province, China
Chang-Xiang Liang, Department of Orthopaedics, Guangdong General Hospital Affiliated to Southern Medical University, Guangzhou 510515, Guangdong Province, China
Author contributions: Jiang N and Song HJ contributed equally to this study; Jiang N contributed to study design, data analysis and interpretation, manuscript drafting and revision, statistical analysis support; Song HJ contributed to literature search, data acquisition and analysis, manuscript drafting; Xie GP contributed to data acquisition, interpretation and statistical analysis; Wang L contributed to literature search and data acquisition; Liang CX contributed to literature search and methodology assessment of included studies; Qin CH contributed to data acquisition and interpretation, methodology assessment of included studies and statistical analysis; Yu B contributed to study design, data interpretation, manuscript drafting and revision, statistical analysis support and supervision.
Conflict-of-interest: The authors have declared that no competing interests exist.
Data sharing: Technical appendix, statistical code, and dataset available from the corresponding author at nanfanghot@126.com. No additional data are available.

Correspondence to: Dr. Bin Yu, Department of Orthopaedics and Traumatology, Nanfang Hospital, Southern Medical University, No.1838, Guangzhou Avenue North, Guangzhou 510515, Guangdong Province, China. nanfanghot@126.com
Telephone: +86-20-61641741
Fax: +86-20-61360066

Received: October 17, 2014
Peer-review started: October 18, 2014
First decision: December 3, 2014
Revised: December 8, 2014
Accepted: December 18, 2014
Article in press: December 19, 2014
Published online: February 26, 2015

Abstract

AIM: To investigate clinical efficacy of displaced intra-articular calcaneal fracture (DIACF) following operation and nonoperation.

METHODS: Literature search was performed of PubMed and Cochrane Library by two independent authors to identify randomized controlled trials (RCTs) comparing operative *vs* nonoperative treatment of DIACF from inception to December 31st, 2013. RCT quality was evaluated by the modified Jadad scale. Dichotomous variables were pooled using risk ratios by review manager 5.3 software. Fixed-effects or random-effects models were adopted with $P > 0.05$ or $P \leqslant 0.05$ for heterogeneity tests, respectively.

RESULTS: Eight RCTs comprising 767 cases met inclusion criteria. Results revealed that more surgically treated patients could resume pre-injury job ($P = 0.006$). No statistical differences were found between the two groups in residual pain ($P = 0.33$), shoe fitting problems ($P = 0.07$), limited walking distance ($P = 0.56$) or secondary late arthrodesis ($P = 0.38$). However, operative treatment was associated with a higher complication rate ($P = 0.003$). Subgroup analyses of specific complications revealed that except for a higher risk of superficial wound problems ($P < 0.0001$) in operative group, the two groups had similar complication rate in deep wound infection ($P = 0.34$),

compartment syndrome (P = 0.46), thromboembolism (P = 0.32), reflex sympathetic dystrophy (P = 0.51) or traumatic arthritis secondary to DIACF (P = 0.43).

CONCLUSION: Current evidence demonstrates that compared with operative treatment, conservative treatment of DIACF lead to similar clinical outcomes regarding residual pain, shoe fitting, walking distance and secondary subtalar arthrodesis but a significantly lower complication rate.

Key words: Displaced intra-articular calcaneal fracture; Surgery; Conservative treatment; Meta-analysis

Core tip: This updated meta-analysis regarding the optimal treatment of displaced intra-articular calcaneal fracture suggests that operative and nonoperative treatment have similar clinical outcomes in residual pain, shoe fitting, walking distance and secondary subtalar arthrodesis. However, operative treatment has a higher complication risk than nonoperative treatment.

Jiang N, Song HJ, Xie GP, Wang L, Liang CX, Qin CH, Yu B. Operative *vs* nonoperative treatment of displaced intra-articular calcaneal fracture: A meta-analysis of randomized controlled trials. *World J Meta-Anal* 2015; 3(1): 61-71 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/61.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.61

INTRODUCTION

Calcaneal fracture is the most frequent tarsal fracture in the human body[1,2] and approximately 75% are intra-articular[3,4]. Since they are mostly caused by high energy trauma[4,5], the great violence delivered to the foot usually causes displaced intra-articular calcaneal fracture (DIACF).

DIACF can be treated by operation as well as by nonoperation. However, conclusions by randomized controlled trials (RCTs) comparing clinical efficacy of the two methods were conflicting. In the year 1993, Parmar *et al*[6] showed no significant differences regarding outcomes between operation and nonoperation. However, subsequently in 1996, Thordarson *et al*[7] revealed a markedly superior functional score following operative treatment. In 2002, Buckley *et al*[8] found the two methods had equivalent functional outcomes without stratification of the groups but that operation was superior to non-operation only after exclusion of the data from patients who were receiving Workers' Compensation. In 2007, Ibrahim *et al*[9] reported similar clinical efficacy between the two strategies after 15-year follow-up.

Likewise, conflicting conclusions also existed in published meta-analyses. In a systematic review of three RCTs in 2000, Bridgman *et al*[10] found slightly better benefits following operative treatment, in consistent with a meta-analysis[11] published in the same year. However, both of them recommended further investigation because they believed the evidence was not strong enough to support operative treatment. In 2005, Bajammal *et al*[3] indicated there was no sufficient evidence to support with certainty that operation was better than nonoperation. This was also concluded by an updated systematic review[12] in 2009. In our meta-analysis[13] of RCTs and controlled clinical trials (CCTs) in 2012, we found that the data favored operative treatment of DIACF. However, in Jan, 2013, a systematic review[14] of four RCTs and quasi-RCTs up to 2011 concluded that operation and nonoperation could achieve similar clinical efficacy while it admitted insufficiency of the evidence.

In fact, although it is still problematic whether operative or nonoperative treatment is better for DIACF, the problem is clinically significant and warrants further study. It is also one of our chief concerns after we published our preliminary finding on this topic.

To our knowledge, there have been two more RCTs[15,16] comparing operative *vs* nonoperative treatment of DIACF since the year 2011. In addition, we believe the inclusion of four CCTs[17-20] in our previous meta-analysis[13] might have caused a bias which could have made our conclusions less reliable. Therefore, we decided to make a new meta-analysis of only and all retrieved RCTs until the most recently comparing clinical efficacy of operative and nonoperative treatment of DIACF.

MATERIALS AND METHODS

Ethics

No ethics approval was acquired.

Protocol

This study was reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement[21,22].

Search strategy

All RCTs comparing operative *vs* nonoperative treatment of DIACF were searched in electronic databases of PubMed and Cochrane library by two authors independently. A structured search was performed using the following search string: (displaced intra-articular) AND [(calcaneal fractures) OR (fractures of the calcaneus)] AND (operative OR operation OR surgical OR surgery OR conservative OR conservation). There was no restriction to publication language. The search time was set from inception to 31st December, 2013. We also consulted the references of published systematic reviews[10-12,14].

Eligibility criteria

Only RCTs and quasi-RCTs that reporting operation *vs* nonoperation for DIACF were taken for inclusion. CCTs, cohort studies and case reports were excluded. In addition, studies that did not report the primary outcomes were also

Table 1 Detailed assessment items of modified Jadad scale

Item assessed	Response	Score
Was the study described as randomized?	Yes	1
	No	0
Was the method of randomization appropriate?	Yes	1
	No	-1
	Not described	0
Was the study described as blinded?[1]	Yes	1
	No	0
Was the method of blinding appropriate?	Yes	1
	No	-1
	Not described	0
Was there a description of withdrawals and dropouts?	Yes	1
	No	0
Was there a clear description of the inclusion/exclusion criteria?	Yes	1
	No	0
Was the method used to assess adverse effects described?	Yes	1
	No	0
Was the method of statistical analysis described?	Yes	1
	No	0

[1]Double-blind RCTs 1 score; single-blind RCTs 0.5 score. RCTs: Randomized controlled trials.

excluded.

Study identification

Two independent authors viewed all titles of searched articles. Further review of article abstract was performed in those whose titles were relevant to the topic. If information from the abstract was inadequate, a full article was referred to. Disagreement on eligibility of included studies was resolved by the third author.

Risk-of-bias evaluation and scores of methodology

Risk-of-bias was assessed using the Cochrane Collaboration guidelines with seven items: generation of random sequence, allocation concealment, participants and personnel blinding, outcome assessment blinding, incomplete outcome data, selective reporting, and other bias graded by high, low or unclear bias risk[23].

In current study, the modified Jadad scale[24] was used to calculate the methodological scores of eligible studies. The scale includes eight items (Table 1) with scores range from 0 (lowest quality) to 8 (highest quality). The cut-off value between high quality and low quality was score 4. Scores higher than 4 mean high-quality trials while scores lower than 4 indicate low-quality trials. The methodological evaluation was performed by two independent reviewers and discrepancy was solved by discussion.

Data extraction

Two authors participated in data extraction independently. Discrepancies in outcome extraction were resolved by checking relevant studies until consensus was achieved.

Outcome measures

Primary outcomes covered assessment of resuming pre-injury job, residual pain, shoe fitting problems, limited walking distance and secondary late arthrodesis. Secondary outcomes were complication rate and subgroup analyses for specific complications.

Statistical analysis

Statistical heterogeneity was assessed using I^2 statistics, which can be calculated from the formula $I^2 = 100\% \times (Q - df)/Q$, (Q represents Cochrane's heterogeneity statistic, df represents the degrees of freedom)[25]. An I^2 value of 0% means no heterogeneity, with cut-off values of 25%, 50%, 75% or more as low, moderate and high risk of heterogeneity, respectively. For outcomes of heterogeneity test when $P > 0.05$, a fixed-effects model was used in the meta-analysis. Otherwise, a random-effects model was adopted for $P \leqslant 0.05$. Dichotomous variables are revealed as relative risk (RR) with 95%CIs. The data syntheses and publication bias were conducted using Review Manager 5.3 software (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014). The level of statistical significance was set at P value $\leqslant 0.05$.

RESULTS

Study selection and characteristics

A total of 195 potentially relevant articles were identified (Figure 1). After reference to titles, abstracts and even full texts, eight published RCTs[6-9,15,16,26,27] comprising 767 patients were included for analysis. General information of eligible studies were listed in Table 2.

During the process of eligibility assessment, we found that the two RCTs by Buckley *et al*[8] and Howard *et al*[28] shared the same groups of patients with the same base line characteristics but reported different measures and clinical outcomes. Therefore, the two studies were regarded as one trial for analysis though the data of outcomes were extracted separately. In addition, one study[29] was excluded because of the unavailability of

Table 2 General information of eligible randomized controlled trials

Ref.	location	Cases (O/N)	Sex ratio (M/F)	Mean age (O/N) (yr)	Follow-up time (O/N) (yr)	Main outcome measures
Parmar *et al*[6]	England	25/31	48/8	48.3/48.8	2.1/1.8	Pain level, site, pattern; walking problems; shoe wear; resuming pre-injury job; deformity; ankle and subtalar movement; foot function; complications
O'Farrell *et al*[27]	Ireland	12/12	20/4	33/38	1.3/1.2	Shoe wear; pain-free walking distance; resuming pre-injury job; restoration of BÖhler angle and Gissane angle; motion range of ankle, subtalar and calcaneocuboid
Chrintz *et al*[26]	Denmark	33/35	NR	NR	1.5/1.5	Radiography outcomes
Thordarson *et al*[7]	United States	15/11	21/5	35/36	1.4/1.2	Functional assessment scale; motion rage of subtalar and ankle; gait analysis; restoration of BÖhler angle; pain; daily activity; shoe wear; walking; exercise; work; complications
Buckley *et al*[8]	Canada	206/218	381/43	41/39	3.0/3.0	Complications; SF-36 scale; VAS; shoe wear; numbness
Ibrahim *et al*[9]	United Kingdom	15/11	21/5	61/58	15.2/14.8	AOFAS score; FFI score; calcaneal fracture score; restoration of BÖhler angle and calcaneal height; arthritic grading of the subtalar joint
Nouraei *et al*[16]	Iran	31/30	NR	46/52	3.0/3.0	Motion range of ankle and subtalar; X-ray findings; width of heel; pain in walking; shoe wear; swelling of foot and ankle; reflex systematic dystrophy
Agren *et al*[15]	Sweden	42/40	59/23	49/48	10 (8-12)[1]	VAS; SF-36 scale; AOFAS score; OM scale; complications

[1]Mean follow-up time was 10 yr with range of 8-12 yr. NR: Not reported; O/N: Operative group/non-operative group; M/F: Male/female; SF-36: Short-form-36 health survey; VAS: Visual analogue scale; AOFAS: American Orthopaedic Foot and Ankle Society; FFI: Foot function index; OM: Olerud-Molander.

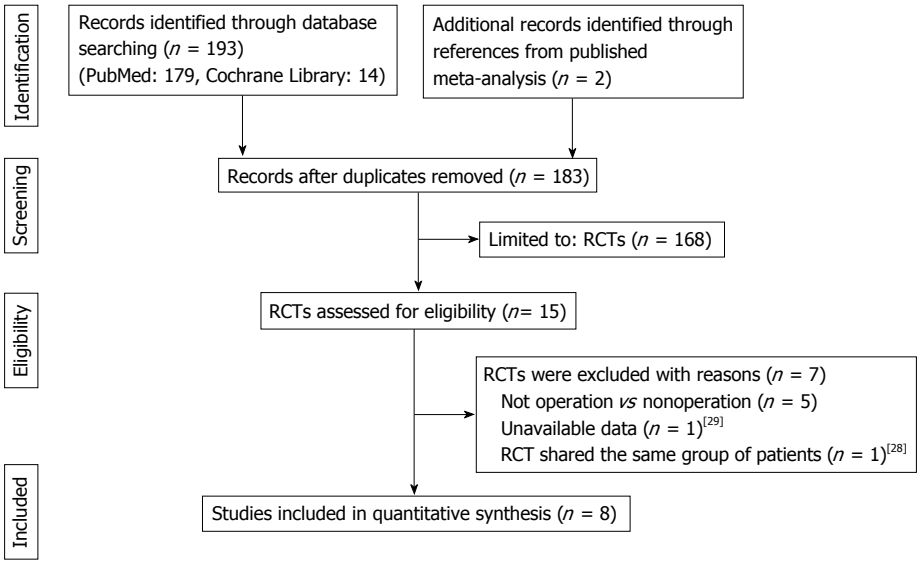


Figure 1 Flow chart of eligibility selection. RCT: Randomized controlled trial.

effective data.

Risk-of-bias evaluation and scores of methodology

Results of the bias risk was shown in Figure 2, indicating most of the eligible RCTs had low to moderate risk of bias. As revealed in Table 3, six[7-9,15,16,26] out of eight studies scored 4 or more than 4 by current rating scale, implying that most of the eligible RCTs were high quality studies. However, several problems were still existed in these studies. Firstly, none of the eligible studies provided detailed description regarding the blinding method. Moreover, most of the RCTs[6,7,9,15,16,26,27] failed to use method to assess adverse effects. In addition, some trials[6,9,16,26,27] still had problems in randomization and blinding. These disadvantages might cause biases.

Outcome measure reporting

Primary outcomes: As shown in Figure 3, 40 of 52 patients after operation compared with 28 of 54 patients after conservative treatment successfully resumed pre-injury work after treatment. No statistically significant difference was found between the two groups [RR = 1.53,

Table 3 Methodological assessment of eligible randomized controlled trials using modified Jadad scale

Item assessed	Parmar 1993	O'Farrell 1993	Chrintz 1993	Thordarson 1996	Buckley 2002	Ibrahim 2007	Nouraei 2011	Agren 2013
Was the study described as randomized?	√	×	√	√	√	√	√	√
Was the method of randomization appropriate?	?	?	?	√	√	?	?	√
Was the study described as blinded?	×	×	×	√	√	×	×	√
Was the method of blinding appropriate?	?	?	?	?	?	?	?	?
Was there a description of withdrawals and dropouts?	×	√	√	√	√	√	√	√
Was there a clear description of the inclusion/exclusion criteria?	×	×	√	√	√	√	√	√
Was the method used to assess adverse effects described?	×	×	×	×	√	×	×	×
Was the method of statistical analysis described?	√	√	√	√	√	√	√	√
Total score	2	2	4	5.5	6.5	4	4	5.5

√: Yes; ×: No; ?: Not described.

95%CI: (1.13, 2.07), $P = 0.006$].

Three RCTs[6,7,16] compared the number of patients who had residual pain during the follow-up period. But no statistical difference was identified [RR = 0.73, 95%CI: (0.40, 1.36), $P = 0.33$] (Figure 4).

With regard to shoe fitting problems after treatment, outcome based on six RCTs[6-8,15,16,27] indicated similar efficacy [RR = 0.61, 95%CI: (0.37, 1.04), $P = 0.07$] (Figure 5).

Two RCTs[6,7] reported the number of patients who had limited walking distance during follow-up time. As shown in Figure 6, no significant difference was found between operation and nonoperation groups [RR = 0.88, 95%CI: (0.57, 1.36), $P = 0.56$].

During the follow-up period, 12 of 248 surgically treated patients compared with 41 of 258 nonsurgically treated patients had secondary late arthrodesis. However, no significant group difference was identified [RR = 0.46, 95%CI: (0.08, 2.64), $P = 0.38$] (Figure 7).

Secondary outcomes: A total of 77 of 288 surgically treated patients compared with 51 of 300 nonsurgically treated patients had complications (26.74% *vs* 17.0%). The significant difference indicated a higher complication risk in operative group [RR = 1.60, 95%CI: (1.17, 2.18), $P = 0.003$] (Figure 8).

Subgroup analyses were performed to explore further differences between the two approaches regarding the specific complications. As revealed in Figure 9, except for a higher risk of superficial wound problems [RR = 30.64, 95%CI: (6.38, 147.29), $P < 0.0001$) after operative treatment, no significant differences were found in deep wound infection [RR = 3.01, 95%CI: (0.32, 28.60), $P = 0.34$], compartment syndrome [RR = 1.71, 95%CI: (0.42, 7.06), $P = 0.46$], thromboembolism [RR = 3.17, 95%CI: (0.33, 30.28), $P = 0.32$], reflex sympathetic dystrophy [RR = 0.68, 95%CI: (0.22, 2.11), $P = 0.51$] or traumatic arthritis secondary to DIACF [RR = 0.88, 95%CI: (0.64, 1.21), $P = 0.43$].

Sensitivity analysis: Sensitivity analysis was performed by excluding studies with Jadad score lower than 4. As shown in Table 3, we excluded two studies[6,27] of score < 4 and performed another meta-analysis. *P* values for outcome measures of residual pain, shoe fitting problems, limited walking distance and complications remained unchanged (Table 4). However, after excluding low quality studies[6,27], outcome regarding the number of patients who resumed pre-injury job showed insignificant difference between operative and nonoperative treatment (Table 4).

Publication bias: Publication bias was performed for incidence of shoe fitting problems and subgroup analyses of complications. Results indicated a potential publication bias of the above two outcome measures (Figures 10 and 11).

DISCUSSION

This updated meta-analysis with all retrieved RCTs suggests that compared with operative treatment, conservative

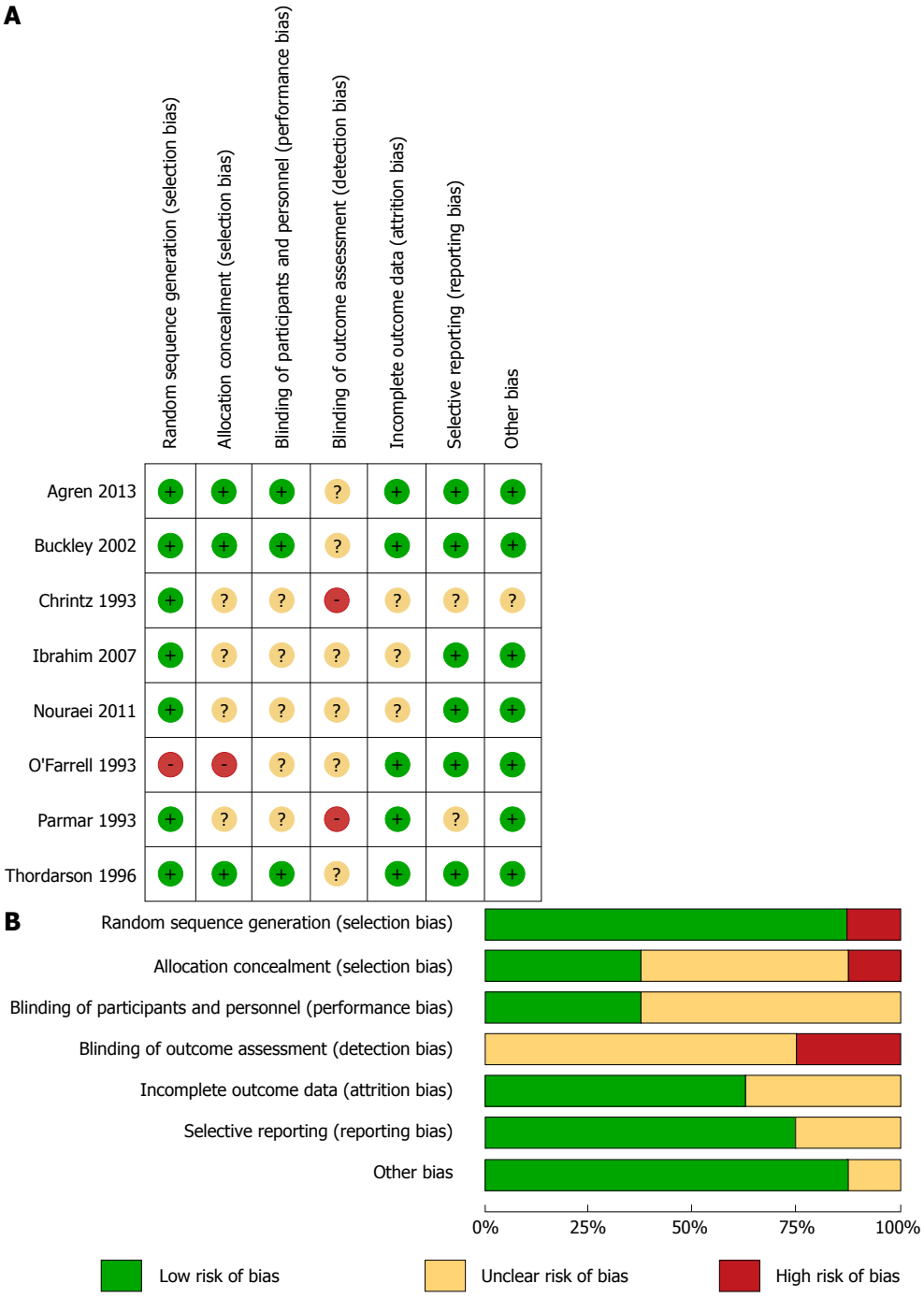


Figure 2 Risk of bias summary and graph. A: Risk of bias summary; B: Risk of bias graph.

Table 4 Results of sensitivity analysis

Outcomes	All eligible RCTs included					Only high score RCTs included				
	n	Patients	I^2	RR (95%CI)	*P* values	*n*	Patients	I^2	RR (95%CI)	*P* values
Resume pre-injury job	3	106	55%	1.53 (1.13, 2.07)	0.006	1	26	NA	2.20 (0.97, 5.00)	0.06
Residual pain	3	143	80%	0.73 (0.40, 1.36)	0.33	2	87	93%	0.63 (0.19, 2.11)	0.45
Shoe fitting problems	6	667	63%	0.61 (0.37, 1.04)	0.07	4	587	73%	0.57 (0.27, 1.21)	0.15
Limited walking distance	2	82	71%	0.88 (0.57, 1.36)	0.56	1	26	NA	0.42 (0.16, 1.08)	0.07
Complications	4	588	0%	1.60 (1.17, 2.18)	0.003	3	532	1%	1.59 (1.14, 2.22)	0.006

NA: Not applicable; RCTs: Randomized controlled trials.

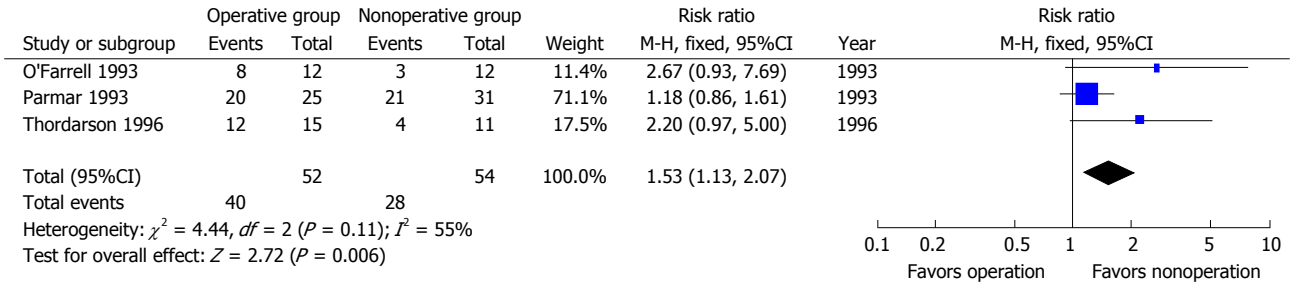


Figure 3 The number of patients who resumed pre-injury job after treatment.

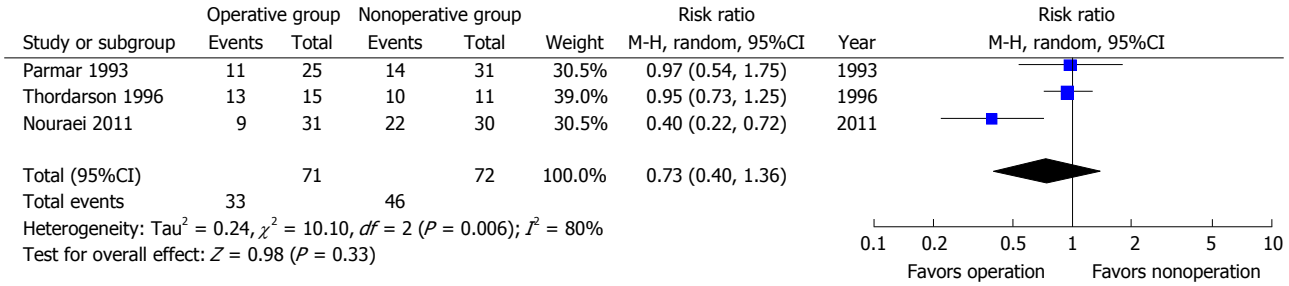


Figure 4 The number of patients who had residual pain after treatment.

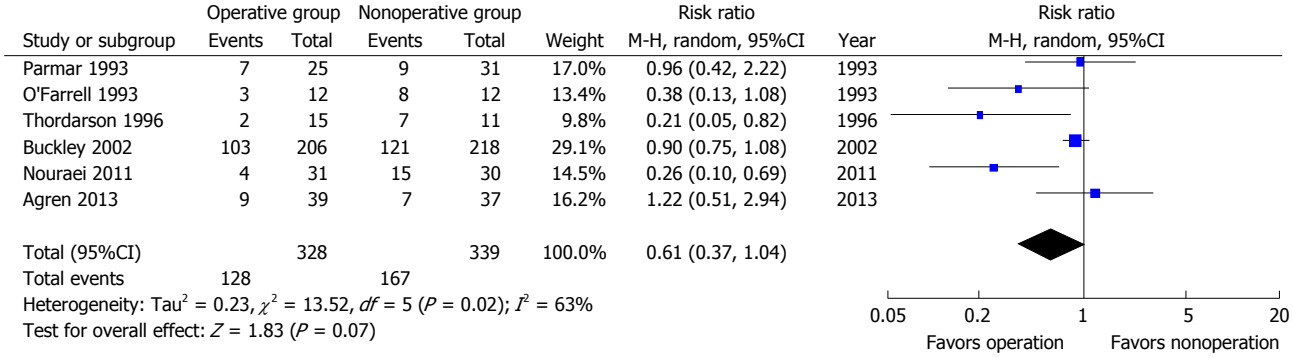


Figure 5 The number of patients who had shoe-fitting problems after treatment.

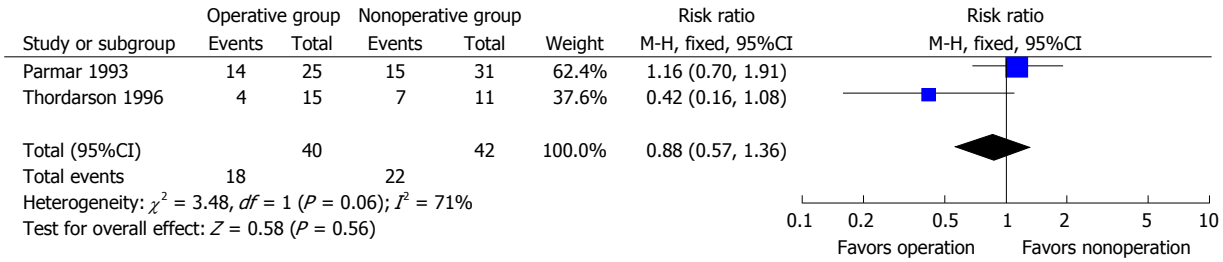


Figure 6 The number of patients who had limited walking distance after treatment.

treatment of DIACF can bring similar clinical outcomes regarding residual pain, shoe fitting, walking distance and secondary late arthrodesis. The only advantage following operative treatment was that more patients could resume pre-injury job after surgery. However, this superiority disappeared after sensitivity analysis by excluding low quality studies. In addition, operative treatment of DIACF elevated the risk of complications. Outcomes of the present study were different from historical meta-analyses, which was mainly because the inclusion of updated RCTs as well as only inclusion of RCTs for analysis.

The present study based on three RCTs[6,7,27] showed that more surgically treated patients could resume pre-injury job. However, Bruce *et al*[14] indicated that no significant differences were identified between operation and nonoperation, neither in returning to the same work nor to any work. Although result of the sensitivity analysis also revealed no statistical difference, cautious

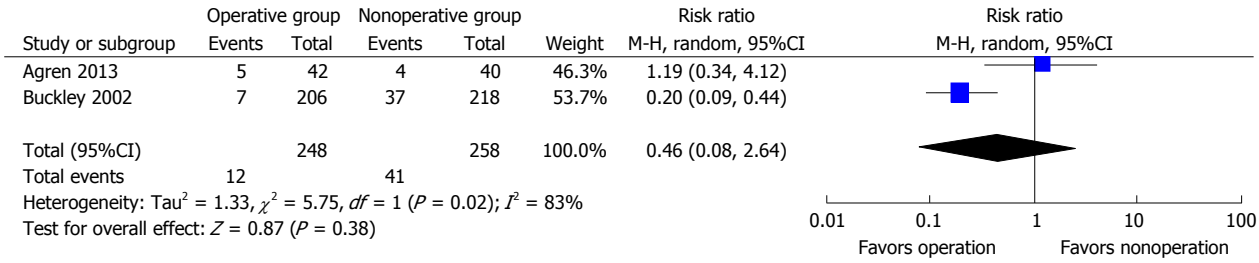


Figure 7 The number of patients who had secondary late arthrodesis.

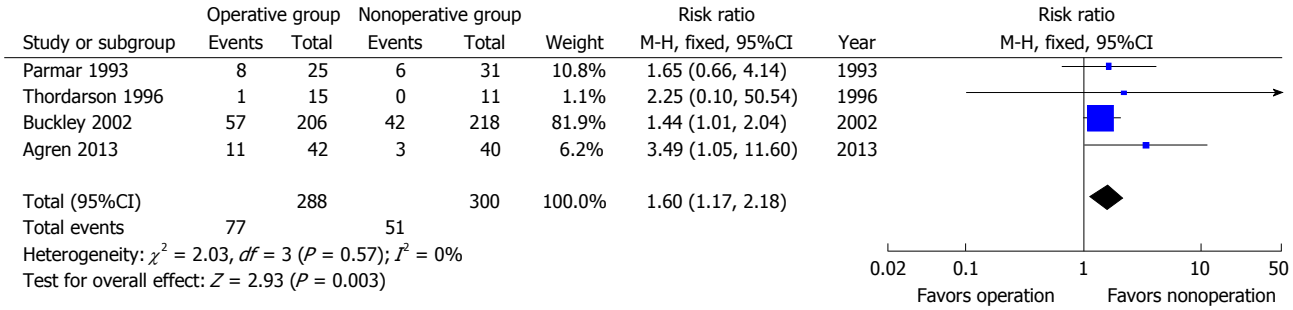


Figure 8 Incidence of complications after operative and nonoperative treatments.

attitude should be taken due to the following two reasons. On one hand, only one RCT was left for the analysis, making the outcome less reliable. On the other hand, different workload may also affect the result. Just as Buckley *et al*[8] analyzed, light or moderate workload might lead to better recovery from DIACF, but patients with heavy workload were unlikely to recover well regardless of treatment strategies.

No significant difference was identified regarding the number of patients with residual pain between the two methods. But the heterogeneity among the eligible studies was high (I^2 = 80%, P = 0.006), which was probably associated with several factors, such as pain tolerance, fracture type and analgesic strategy. Quite different from our previous study[13] of fewer shoe-wear problems in the operative group, the present study revealed no statistical difference between the two groups, which was probably because of the inclusion of two additional RCTs[8,15]. Howard *et al*[28] found no significant differences between the two approaches in the number of patients who required shoe-wear modifications at 2 wk, 6 wk, 3 mo, 1 year, 2 years or more than 2 years, respectively. Agren *et al*[15] reported the incidence of shoe-wear problems following operation and nonoperation was 23% and 19%. However, the authors of the two studies[15,28] did not give possible reasons for their findings.

The pooled result regarding the number of patients who had limited walking distance based on two RCTs[6,7] showed insignificant difference between operation and nonoperation. Parmar *et al*[6] only listed the percentage of patients without limited walking distance and did not provide the precise definition of the distance. Thordarson *et al*[7] defined the distance as six blocks. Therefore, the lack of consistent definition of limited walking distance might account for the high heterogeneity of included studies (I^2 = 71%, P = 0.06). With respect to the number of patients who had secondary late arthrodesis, outcome based on two studies[8,15] also revealed no statistical difference. One RCT[28] reported the incidence of arthrodesis in nonoperative group was significantly higher than operative group [16% *vs* 3%, RR = 0.20, 95%CI: (0.09, 0.44), P < 0.0001]. This was probably because the calcaneal geometry was comparatively better preserved after operation[30]. However, Agren *et al*[15] reported the arthrodesis rates for operative and nonoperative managements were 12% and 10%, respectively [RR = 1.19, 95%CI: (0.34, 4.12), P = 0.78]. The authors also did not give explanations for a relatively higher incidence of arthrodesis following operative treatment. We considered it might due to the slightly larger percentage of more severe types of fracture in the operative group.

The present meta-analysis supported that surgically treated patients had a significantly higher risk of complication than those in nonsurgical group. To investigate the detailed differences of complications between the two groups, subgroup analysis was further performed on specific complications. Outcomes of the subgroup analysis implied that superficial wound problems might be the main cause of a higher complication rate after operation. Although no significant differences were identified in the number of patients who had compartment syndrome, thromboembolism or reflex sympathetic dystrophy, they need to be reported so that patients treated for DIACF are fully informed of potential complications regardless of the treatment strategy chosen. It was interesting that the incidence of traumatic arthritis secondary to DIACF was similar between the two groups (operative group of 41.67% *vs* nonoperative group of 44.78%). However,

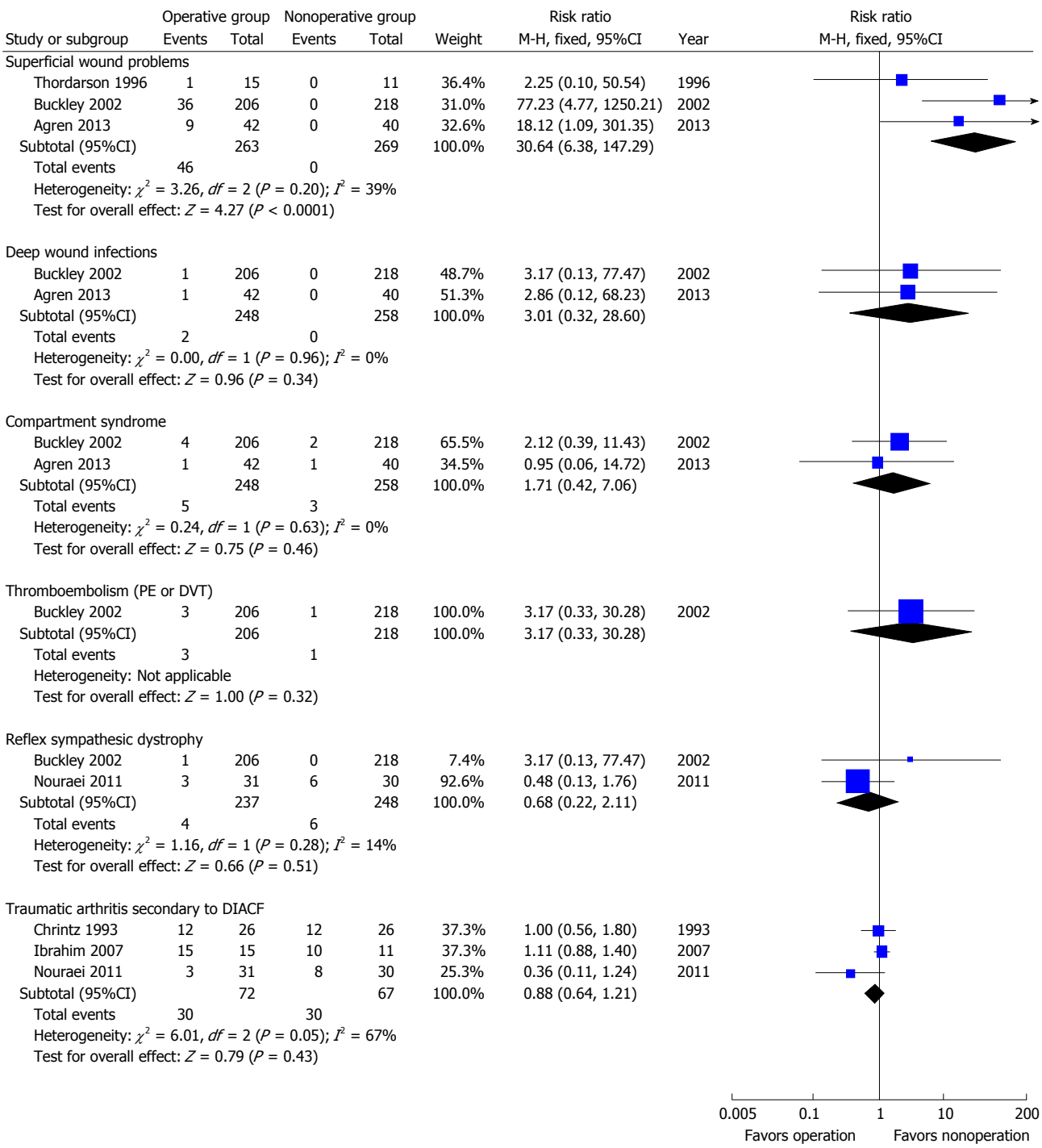


Figure 9 Subgroup analysis on specific complications. DIACF: Displaced intra-articular calcaneal fracture; PE: Pulmonary embolism; DVT: Deep venous thrombosis.

the outcome was based on three RCTs[9,16,26] with 139 participants, and it also might be affected by different follow-up time. Therefore, whether difference indeed exists requires more studies with adequate follow-up time.

Several scales or scores were adopted to evaluate clinical efficacy of the two methods in eligible RCTs. Ibrahim *et al*[9] showed that no significant differences were identified in total AOFAS score, total FFI score or calcaneal fracture score at 15 years' follow-up time. After analyzed outcomes of SF-36 and VAS scores, Howard *et al*[28] concluded that the functional outcomes were partly associated with treatment strategy and partly related to the complications. Agren *et al*[15] used several stratified scales to show clinical efficacy at one year follow-up and at eight to twelve years' follow-up, including visual analog scale (VAS) pain and function scoring by patients as well as by surgeon, VAS pain at rest and during weight-bearing, SF-36 physical and SF-36 mental scores, AOFAS and OM scores. However, outcomes from all these scales and scores were similar between the two methods. We did not pool these results for meta-analysis due to the following reasons: (1) not correct report form for

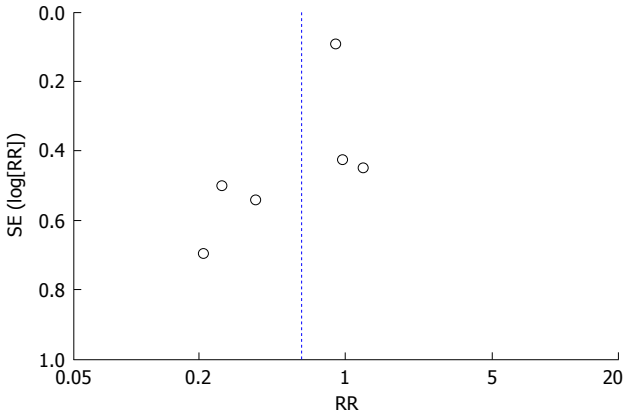


Figure 10 Funnel plot based on studies with data on incidence of shoe-fitting problems.

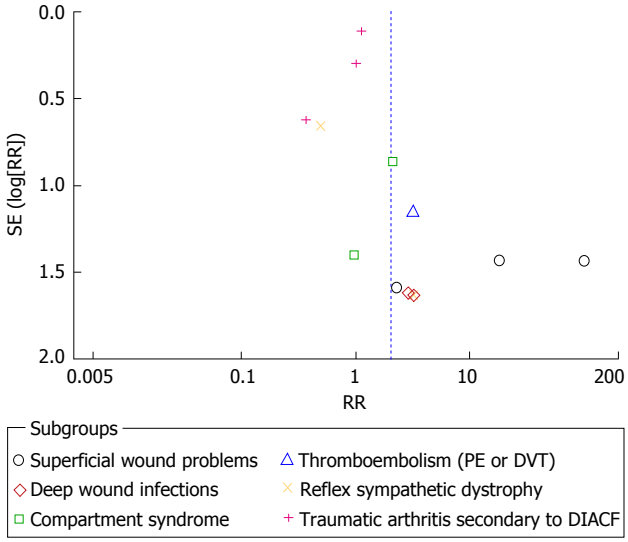


Figure 11 Funnel plot based on studies with data on subgroup analyses of complications. DIACF: Displaced intra-articular calcaneal fracture; SE: Standard error; PE: Pulmonary embolism; DVT: Deep venous thrombosis.

data synthesis in meta-analysis (correct form should be mean ± standard deviation); (2) a single report; and (3) outcomes were not reported at the same follow-up time.

The main limitation of the current study might be the still limited number of eligible RCTs with limited number of participants. Although a total of eight RCTs with 767 participants was included in our study, more than half of the participants were from one study[8], which may cause a bias. In addition, the current study was purely based on a methodological standpoint, which lacks practical information regarding treatment strategies on different fracture types, especially severe and challenging injuries. Therefore, conclusions of this analysis should be interpreted with caution and more high quality RCTs are needed in the future.

In summary, the current study indicates that compared with operative treatment, conservative treatment of DIACF lead to similar clinical outcomes regarding residual pain, shoe fitting, walking distance and secondary subtalar arthrodesis but a significantly lower complication rate.

ACKNOWLEDGMENTS

The authors thank Professor Allen P Liang for revising and editing this manuscript.

COMMENTS

Background

Displaced intra-articular calcaneal fracture (DIACF) can be managed by both operative and nonoperative strategies. However, up till now, controversy still exists regarding the optimal treatment of this fracture, which is mainly due to the conflicting outcomes derived from previous studies.

Research frontiers

It is generally believed that intra-articular fractures should be treated operatively as operative management can provide better fracture reduction, promote early functional rehabilitation and reduce the rate of traumatic arthritis. However, several studies showed that conservative treatment can achieve similar functional recovery as surgery but had a lower complication risk. Therefore, whether surgery is a must for DIACF treatment requires more investigations.

Innovations and breakthroughs

Compared with previous systematic reviews or meta-analyses, the present study included more studies with high quality in methodology and thus made the outcomes more reliable. In addition, the current study once again confirmed similar clinical efficacy following operation and nonoperation.

Applications

The present study provides evidence to support conservative treatment of DIACF. However, cautious attitude should be taken towards the conclusion because of the still limited number of randomized controlled studies (RCTs) and future more high quality surveys are warranted.

Terminology

Clinical RCT is a type of scientific experiment, where the people being studied are randomly allocated one or other of the different treatment methods under study. RCT is a golden standard for a clinical trial. However, the quality of an RCT is important, which will affect the reliability of the outcomes. Meta-analysis is a statistical method of combining different treatment outcomes derived from different studies to generate more conclusive and reliable conclusions.

Peer review

This is a well written meta-analysis which confirms with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

REFERENCES

1 **Daftary A**, Haims AH, Baumgaertner MR. Fractures of the calcaneus: a review with emphasis on CT. *Radiographics* 2005; **25**: 1215-1226 [PMID: 16160107 DOI: 10.1148/rg.255045713]
2 **Zwipp H**, Rammelt S, Barthel S. [Calcaneal fractures--the most frequent tarsal fractures]. *Ther Umsch* 2004; **61**: 435-450 [PMID: 15354753]
3 **Bajammal S**, Tornetta P, Sanders D, Bhandari M. Displaced intra-articular calcaneal fractures. *J Orthop Trauma* 2005; **19**: 360-364 [PMID: 15891550]
4 **Juliano P**, Nguyen HV. Fractures of the calcaneus. *Orthop Clin North Am* 2001; **32**: 35-51, viii [PMID: 11465132 DOI: 10.1016/S0030-5898(05)70192-9]
5 **Bakker B**, Halm JA, Van Lieshout EM, Schepers T. The fate of Böhler's angle in conservatively-treated displaced intra-articular calcaneal fractures. *Int Orthop* 2012; **36**: 2495-2499 [PMID: 23138968 DOI: 10.1007/s00264-012-1706-3]
6 **Parmar HV**, Triffitt PD, Gregg PJ. Intra-articular fractures of the calcaneum treated operatively or conservatively. A prospective study. *J Bone Joint Surg Br* 1993; **75**: 932-937 [PMID: 8245085]
7 **Thordarson DB**, Krieger LE. Operative vs. nonoperative

treatment of intra-articular fractures of the calcaneus: a prospective randomized trial. *Foot Ankle Int* 1996; **17**: 2-9 [PMID: 8821279 DOI: 10.1177/107110079601700102]

8 **Buckley R**, Tough S, McCormack R, Pate G, Leighton R, Petrie D, Galpin R. Operative compared with nonoperative treatment of displaced intra-articular calcaneal fractures: a prospective, randomized, controlled multicenter trial. *J Bone Joint Surg Am* 2002; **84-A**: 1733-1744 [PMID: 12377902]

9 **Ibrahim T**, Rowsell M, Rennie W, Brown AR, Taylor GJ, Gregg PJ. Displaced intra-articular calcaneal fractures: 15-year follow-up of a randomised controlled trial of conservative versus operative treatment. *Injury* 2007; **38**: 848-855 [PMID: 17445815 DOI: 10.1016/j.injury.2007.01.003]

10 **Bridgman SA**, Dunn KM, McBride DJ, Richards PJ. Interventions for treating calcaneal fractures. *Cochrane Database Syst Rev* 2000; (**2**): CD001161 [PMID: 10796422 DOI: 10.1002/14651858.CD001161]

11 **Randle JA**, Kreder HJ, Stephen D, Williams J, Jaglal S, Hu R. Should calcaneal fractures be treated surgically? A meta-analysis. *Clin Orthop Relat Res* 2000; (**377**): 217-227 [PMID: 10943205]

12 **Gougoulias N**, Khanna A, McBride DJ, Maffulli N. Management of calcaneal fractures: systematic review of randomized trials. *Br Med Bull* 2009; **92**: 153-167 [PMID: 19734165 DOI: 10.1093/bmb/ldp030]

13 **Jiang N**, Lin QR, Diao XC, Wu L, Yu B. Surgical versus nonsurgical treatment of displaced intra-articular calcaneal fracture: a meta-analysis of current evidence base. *Int Orthop* 2012; **36**: 1615-1622 [PMID: 22576080 DOI: 10.1007/s00264-012-]

14 **Bruce J**, Sutherland A. Surgical versus conservative interventions for displaced intra-articular calcaneal fractures. *Cochrane Database Syst Rev* 2013; **1**: CD008628 [PMID: 23440830 DOI: 10.1002/14651858.CD008628.pub2]

15 **Agren PH**, Wretenberg P, Sayed-Noor AS. Operative versus nonoperative treatment of displaced intra-articular calcaneal fractures: a prospective, randomized, controlled multicenter trial. *J Bone Joint Surg Am* 2013; **95**: 1351-1357 [PMID: 23925738 DOI: 10.2106/JBJS.L.00759]

16 **Nouraei MH**, Moosa FM. Operative compared to non-operative treatment of displaced intra-articular calcaneal fractures. *J Res Med Sci* 2011; **16**: 1014-1019 [PMID: 22279476]

17 **Järvholm U**, Körner L, Thorén O, Wiklund LM. Fractures of the calcaneus. A comparison of open and closed treatment. *Acta Orthop Scand* 1984; **55**: 652-656 [PMID: 6524335]

18 **Leung KS**, Yuen KM, Chan WS. Operative treatment of displaced intra-articular fractures of the calcaneum. Medium-term results. *J Bone Joint Surg Br* 1993; **75**: 196-201 [PMID: 8444936]

19 **Rodriguez-Merchan EC**, Galindo E. Intra-articular displaced fractures of the calcaneus. Operative vs non-operative treatment. *Int Orthop* 1999; **23**: 63-65 [PMID: 10192023 DOI: 10.1007/s002640050307]

20 **Xia S**, Shi D, Wang Z, Wang X, Lu Y, Wang H, Wu Z, Zhu H. Operative versus nonoperative managements for displaced intraarticular fractures of calcaneus. *Chin J Orthop Trauma* 2010; **12**: 1089-1091

21 **Liberati A**, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; **339**: b2700 [PMID: 19622552 DOI: 10.1136/bmj.b2700]

22 **Moher D**, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151**: 264-269, W64 [PMID: 19622511]

23 **Higgins JP**, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **343**: d5928 [PMID: 22008217 DOI: 10.1136/bmj.d5928]

24 **Oremus M**, Wolfson C, Perrault A, Demers L, Momoli F, Moride Y. Interrater reliability of the modified Jadad quality scale for systematic reviews of Alzheimer's disease drug trials. *Dement Geriatr Cogn Disord* 2001; **12**: 232-236 [PMID: 11244218 DOI: 10.1159/000051263]

25 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327]

26 **Chrintz H**, Sonne-Holm S. Radiograhic results after conservative versus operative treatment of dislocated intraarticular fractures of calcaneus. *Acta Orthopaedica Scandinavica-Supplementum* 1993; **251**: 63-64

27 **O'Farrell DA**, O'Byrne JM, McCabe JP, Stephens MM. Fractures of the os calcis: improved results with internal fixation. *Injury* 1993; **24**: 263-265 [PMID: 8325686 DOI: 10.1016/0020-1383(93)90183-7]

28 **Howard JL**, Buckley R, McCormack R, Pate G, Leighton R, Petrie D, Galpin R. Complications following management of displaced intra-articular calcaneal fractures: a prospective randomized trial comparing open reduction internal fixation with nonoperative management. *J Orthop Trauma* 2003; **17**: 241-249 [PMID: 12679683 DOI: 10.1097/00005131-200304000-00001]

29 **Meggitt B**. A multi-centre prospective randomised controlled trial to compare operative with non-operative treatment of displaced intra-articular fractures of the calcaneus. The national research register, 2001. Available from: URL: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/820/CN-00845820/frame.html

30 **Radnay CS**, Clare MP, Sanders RW. Subtalar fusion after displaced intra-articular calcaneal fractures: does initial operative treatment matter? *J Bone Joint Surg Am* 2009; **91**: 541-546 [PMID: 19255213 DOI: 10.2106/JBJS.G.01445]

P- Reviewer: Sodergren MH, Tsikouras P **S- Editor**: Ji FF
L- Editor: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i1.72

World J Meta-Anal 2015 February 26; 3(1): 72-81
ISSN 2308-3840 (online)

META-ANALYSIS

Association between SHBG (TAAAA)n and AR (CAG)n polymorphisms and PCOS risk: A meta-analysis

Jie-Wen Jin, Shi-Lin Chen, Zhan-Tao Deng

Jie-Wen Jin, Shi-Lin Chen, Zhan-Tao Deng, Center for Translational Medicine and Jiangsu Key Laboratory of Molecular Medicine, Medical School of Nanjing University, Nanjing 210093, Jiangsu Province, China
Author contributions: Jin JW designed the study, analyzed the data and wrote the manuscript; Chen SL contributed to discussion and wrote the manuscript; Deng ZT designed the study, contributed to discussion and edited the manuscript.
Supported by Medical School of Nanjing University.

Correspondence to: Zhan-Tao Deng, PhD, Center for Translational Medicine and Jiangsu Key Laboratory of Molecular Medicine, Medical School of Nanjing University, Gulou District, Nanjing 210093, Jiangsu Province, China. 15298386724@163.com
Telephone: +86-25-83594755
Fax: +86-25-83594755
Received: May 21, 2014
Peer-review started: May 23, 2014
First decision: July 10, 2014
Revised: July 28, 2014
Accepted: October 31, 2014
Article in press: November 3, 2014
Published online: February 26, 2015

Abstract

AIM: To systematically assess the association between sex hormone-binding globulin (SHBG) (TAAAA)n and androgen receptor (AR) (CAG)n polymorphisms and polycystic ovarian syndrome (PCOS) risk.

METHODS: We searched MEDLINE (PubMed), EMBASE and Web of Science database from inception to May 2014. To avoid missing any additional studies, we looked through all the references of relevant articles. Case-control studies concerning the (CAG)n variants in the *AR* gene or the (TAAAA)n polymorphism in the *SHBG* gene in PCOS patients were included. Five studies regarding the (TAAAA)n polymorphism in the *SHBG* gene and 14 studies regarding the (CAG)n polymorphism in the AR gene met our criteria. Odd ratio (OR) and weighted mean difference (WMD) were selected as the effect size measurements to evaluate the influence of the (TAAAA)n polymorphism and (CAG)n variants on PCOS risk. Begg's test was used for the evaluation of publication bias.

RESULTS: With respect to the relationship between the (TAAAA)n polymorphism and PCOS risk, the statistical results showed that there was no significant difference between PCOS patients and controls in the alleles of TAAAA (S: OR = 0.91, 95%CI: 0.78-1.05; L: OR = 1.10, 95%CI: 0.95-1.27). Subgroup analyses of the combination of alleles indicated similar results (short-short: OR = 0.87, 95%CI: 0.66-1.14; short-long: OR = 1.12, 95%CI: 0.86-1.46; long-long: OR = 1.03, 95%CI: 0.72-1.47). As for the relationship between the (CAG)n polymorphism and PCOS risk, we found no association between CAG repeat variants and PCOS risk (WDM = 0.03, 95%CI: -0.13-0.08). Subgroup analyses by race and diagnosis criteria indicated the same results (Asian: WMD = -0.03, 95%CI: -0.14-0.07; Caucasian: WMD = -0.02, 95%CI: -0.24-0.21; the criteria of Rotterdam: WMD = 0.01, 95%CI: -0.01-0.03).

CONCLUSION: There is no association between (TAAAA)n polymorphism in SHBG gene, (CAG)n repeat variants in *AR* gene and PCOS.

Key words: Sex hormone-binding globulin; TAAAA; Androgen receptor; CAG; Polycystic ovarian syndrome

Core tip: Our study investigated the association between sex hormone-binding globulin (SHBG) (TAAAA)n and androgen receptor (AR) (CAG)n polymorphisms and polycystic ovarian syndrome (PCOS) risk. Five studies regarding the (TAAAA)n polymorphism in the *SHBG* gene and 14 studies regarding the (CAG)n polymorphism in the *AR* gene were included based on the strict inclusion criteria. The overall meta-analysis, as well as the subgroup analysis, showed that there was no association between PCOS risk and the SHBG (TAAAA)n polymorphism or AR (CAG)n repeat variants.

Jin JW, Chen SL, Deng ZT. Association between SHBG (TAAAA)n and AR (CAG)n polymorphisms and PCOS risk: A meta-analysis. *World J Meta-Anal* 2015; 3(1): 72-81 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i1/72.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i1.72

INTRODUCTION

The morbidity of polycystic ovarian syndrome (PCOS) is estimated to be 7%[1] in women of reproductive age and it is characterized by hyperandrogenism, chronic anovulation and polycystic ovaries[2]. There are three main diagnostic guides for PCOS, including the 1990 National Institutes of Health criteria, the 2003 Rotterdam criteria and the 2006 Androgen Excess and PCOS Society criteria[3-5]. Hyperandrogenism plays an important part in these PCOS diagnostic features and is increasingly considered a main pathogenic factor of PCOS[4]. Recently, the genetic aspects of PCOS have been clarified to prove the presence of genetic abnormalities in PCOS patients. Although the specific genetic alterations that contribute to the development of PCOS remain unclear, several candidate genes have been proposed, including the sex hormone-binding globulin (*SHBG*) gene and the androgen receptor (*AR*) gene[6].

The *SHBG* gene (17p13-p12) encodes a 373 amino acid polypeptide that regulates the bioavailability of sex steroids by binding androgens, particularly testosterone and estrogens[7,8]. The free SHBG levels frequently diminish in patients with hyperandrogenism, especially in those who have PCOS, which may result in an increase in free androgen levels and magnify the biological impact of androgens. SHBG can be influenced by many factors, including gender, age, metabolic, genetic and nutritional factors, with genetic factors being more important[9,10]. A (TAAAA)n repeat variant in the 5' non-coding region of SHBG promoter has been described and its influence on *in vitro* transcriptional activity has been reported[11]. Compared with normal women, those with PCOS tend to have a significantly greater frequency of longer (TAAAA)n alleles (more than eight repeats)[10]. However, the genetic association studies between (TAAAA)n repeat polymorphism of SHBG and PCOS risk show controversial results, which make it difficult to judge PCOS by the number of SHBG (TAAAA)n repeats.

The *AR* gene (Xq11-q12)[12] consists of eight exons and seven introns. The CAG trinucleotide repeats in exon 1 ranged in length from 8 to 35 in healthy individuals and have been reported to influence the transcriptional activity of AR[13]. Chamberlain *et al*[13] reported that there was a negative correlation between the number of CAG repeats and the AR activity, which means that a higher number of CAG repeats is associated with a lower AR biological activity. There are several studies focusing on the relationship between CAG repeat number and PCOS risk, but inconsistent results make it hard to assess the importance of CAG repeat number in PCOS.

Currently, there is no consensus regarding the relationship between SHBG (TAAAA)n polymorphism, AR CAG length and PCOS, although this relation may influence the time to diagnosis and drug intervention. For this reason, we conducted this meta-analysis to address such inconsistency.

MATERIALS AND METHODS

Data sources and searches

We underwent a systematic search of MEDLINE (PubMed), EMBASE, and Web of Science database with the assistance of computer from inception to May 2014, attempting to find all publications about the relationship between (TAAAA)n SHBG and (CAG)n AR polymorphisms and PCOS. Key words for the search of MEDLINE were as follows: ("sex hormone-binding globulin" or "SHBG") and "TAAAA" and ("polycystic ovarian syndrome" or "PCOS") for the *SHBG* gene and ("androgen receptor" or "AR") and "CAG" and ("polycystic ovarian syndrome" or "PCOS") for the *AR* gene. We used similar strategies to search EMBASE. The abstracts of additional meetings were mainly from Web of Science. To avoid missing any additional studies, we looked through all the references of relevant articles.

Study selection

We skimmed titles and abstracts of identified papers to exclude studies that clearly not meeting the inclusion criteria and retrieved the full texts of selected studies for further review and evaluation.

The inclusion criteria for studies were as follows: (1) studies concerning the association between the (CAG)n polymorphism in the *AR* gene or (TAAAA)n variants in the *SHBG* gene and PCOS risk; (2) independent case-control study; (3) specific diagnosis criteria for PCOS; (4) hospital-based healthy women were selected as controls; and (5) data were enough for our further analysis. In order to avoid overlapping data, only the latest study or the study having the most sufficient data was enrolled in our analysis if several studies were conducted by the same author.

Data extraction

Two authors (Jin JW and Chen SL) extracted data from

Table 1 Characteristics of studies on (TAAAA)n sex hormone-binding globulin polymorphism and polycystic ovarian syndrome

Ref.	Year	Country	Race	PCOS women	Controls	PCOS alleles (2 n)		Controls alleles (2 n)		PCOS women genotype			Controls genotype			PCOS diagnostic criteria
						S (< 8 repeats)	L (≥ 8 repeats)	S (< 8 repeats)	L (≥ 8 repeats)	SS	SL	LL	SS	SL	LL	
Xita *et al*[10]	2003	Greece	Caucasian	185	324	230	140	446	202	NR	NR	NR	NR	NR	NR	A
Zhao *et al*[24]	2005	China	Asian	157	156	180	134	175	137	48	84	25	48	79	29	B
Ferk *et al*[26]	2007	Slovenia	Caucasian	123	110	155	91	151	69	54	48	21	52	47	11	C
Liu *et al*[27]	2008	China	Asian	187	176	216	158	210	142	59	96	32	66	78	32	C
Díaz *et al*[25]	2010	Spain	Caucasian	70	107	102	38	139	75	NR	NR	NR	NR	NR	NR	D

A: The 1990 National Institutes of Health–National Institute of Child Health and Human Development conference on PCOS; B: (1) Amenorrhoea or oligomenorrhoea; (2) LH/FSH ⩾ 2.5 or T ⩾ 1.56; (3) More than 10 follicles measuring 2-8 mm in diameter at least in one ovary; and (4) Exclusion of congenital adrenal hyperplasia, Cushing's syndrome, androgen-secreting tumor, hyperprolactinaemia and thyroid dysfunction; C: The criteria of Rotterdam Revised 2003 (two of three) diagnosis; D: (1) Hirsutism; (2) Amenorrhoea or oligomenorrhoea; (3) Increased serum T and/or androstenedione and 17OH-P hyperresponse to GnRH agonist; and (4) Hyperinsulinaemia during an oral glucose tolerance test. PCOS: Polycystic ovarian syndrome; S: Short alleles; L: Long alleles; SS: Short-short genotype; SL: Short-long genotype; LL: Long-long genotype; NR: Non-reported.

each selected article independently and the specific items were as follows: (1) first author; (2) year of publication; (3) regions of the population investigated; (4) diagnosis criteria for PCOS; (5) size of controls and PCOS patients; (6) the number of cases and controls for SHBG (TAAAA)n alleles and genotype; and (7) the mean and standard deviation of cases and controls for AR (CAG)n repeats. We extracted quantitative data directly from articles or using original information provided in the tables and figures[14].

Statistical analysis

We used STATA Statistical Software for all the analyses (version 12.0, STATA Corporation, United States). The evaluation indicators were odd ratio (OR) with 95%CI for the *SHBG* gene and weighted mean difference (WMD) with 95%CI for the *AR* gene.

Meta-analysis

We used two models to calculate the pooled OR and WMD estimates with 95%CI: a fixed-effects model known as Mantel Haenszel method[15] or a random-effects model known as Der Simonian-Laird method[16]. We used the χ^2 test to evaluate the heterogeneity of the studies[17] and the quantity I^2 was also calculated[18,19]. I^2 is the percentage of between-study variation in total variation. The value of 25% is regarded as low heterogeneity while the value of 75% represents high heterogeneity. While I^2 was over 50%, the random-effect model was used instead of the fixed-effect model.

Publication bias was evaluated to find whether the results of the studies were homogeneous[20], and the Egger regression asymmetry test[21] and the Begg-Mazumdar adjusted rank correlation test[22] were used. When the *P*-value of the Egger's test or Begg's test was < 0.05, we considered significant bias among the studies.

RESULTS

Search results

For the (TAAAA)n polymorphism in the *SHBG* gene, 22 records were found in electronic databases, including 8 records in MEDLINE, 11 records in Web of Science database and 3 records in EMBASE. According to the selection criteria, we ultimately identified 5 studies for our final statistical analysis (Figure 1). Table 1 summarizes the characteristics of all the included studies. For the (CAG)n repeats in the *AR* gene, a total of 65 studies were found, including 26 records in MEDLINE, 37 records in Web of Science database and 2 references from reference lists. According to the selection criteria, we identified 14 studies for our meta-analysis (Figure 2) and present their characteristics in Table 2.

Meta-analysis of the SHBG (TAAAA)n polymorphism and PCOS risk

We involved 5 studies, a total of 722 cases and 873 controls, to compare short (S) alleles and long (L) alleles in PCOS patients with those in controls. Because heterogeneity

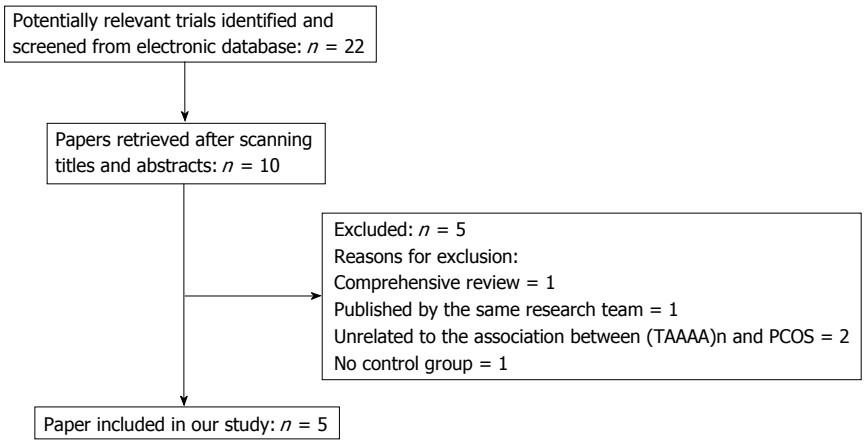


Figure 1 Strategy for searching studies concerning the association between the (TAAAA)n sex hormone-binding globulin polymorphism and polycystic ovarian syndrome risk. PCOS: Polycystic ovarian syndrome.

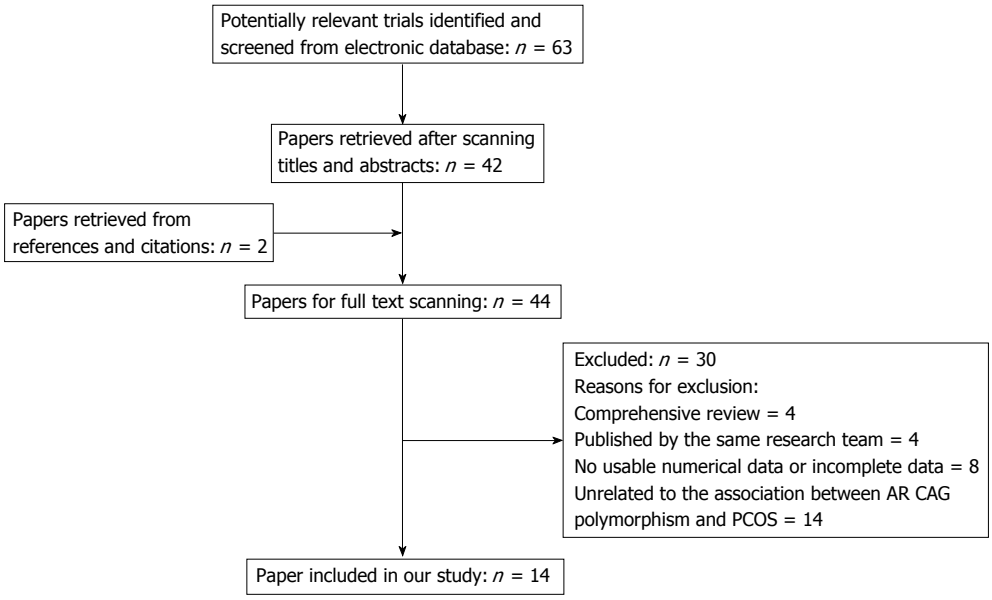


Figure 2 Strategy for searching studies concerning the association between the (CAG)n androgen receptor polymorphisms and polycystic ovarian syndrome risk. AR: Androgen receptor; PCOS: Polycystic ovarian syndrome.

was moderate (I^2 = 47.8% < 50%), we calculate the pooled OR estimates with 95%CI using the fixed-effects model (S: OR = 0.91, 95%CI: 0.78-1.05; L: OR = 1.10, 95%CI: 0.95-1.27) (Figure 3 and Table 3). We considered there was no significant association between PCOS and (TAAAA)n SHBG allele. No obvious publication bias was found in all the selective studies.

On the other hand, we identified 3 studies to compare short-short (SS) genotype, short-long (SL) genotype and long-long (LL) genotype in PCOS patients with those in controls. Because heterogeneity was low (SS: I^2 = 0; SL: I^2 = 0; LL: I^2 = 29.6%), we calculated OR using the fixed-effects model (SS: OR = 0.87, 95%CI: 0.66-1.14; SL: OR = 1.12, 95%CI: 0.86-1.46; LL: OR = 1.03, 95%CI: 0.72-1.47) (Figure 4 and Table 3). Similar to the results of alleles, there was no association between PCOS and (TAAAA)n SHBG genotype. No obvious publication bias was found in all the selective studies.

Meta-analysis of the AR (CAG)n polymorphism and PCOS risk

We involved 14 studies (1882 cases and 1988 controls in total) to compare biallelic mean of CAG length in PCOS patients with controls. Because heterogeneity was moderate (I^2 = 51.0% > 50%), we calculate the pooled WMD estimates with 95%CI using the random-effects model (WMD = 0.03, 95%CI: -0.13-0.08) (Figure 5 and Table 4). Begg's test (P = 0.621) and Egger's test (P = 0.866) showed no obvious publication bias. Further, subgroup analyses were done by race and diagnosis criteria. Because the heterogeneity was high in subgroup by race (Asian: I^2 = 72.9%; Caucasian: I^2 = 41.9%) and low by diagnosis criteria (The criteria of Rotterdam: I^2 = 0), random and fixed models were used, respectively.

Table 2 Characteristics of studies on (CAG)n androgen receptor polymorphism and polycystic ovarian syndrome

Ref.	Year	Country	Race	PCOS			Controls			PCOS diagnostic criteria
				Size	Mean	Std dev	Size	Mean	Std dev	
Mifsud *et al*[23]	2000	Singapore	Asian	91	22.97	0.24	112	23.09	0.23	A
Hickey *et al*[37]	2002	Australia	Caucasian	122	23	2.025	83	22.34	2.094	B
Jääskeläinen *et al*[43]	2005	Finland	Caucasian	106	21.5	2.2	112	21.5	2.1	C
Kim *et al*[39]	2008	South Korea	Asian	114	23.3	1.8	205	23.1	2	D
Ferk *et al*[38]	2008	Slovene	Caucasian	102	22.4	3.5	110	21.9	3.5	E
Liu *et al*[27]	2008	China	Asian	148	22.88	1.76	104	22.85	1.6	D
Shah *et al*[36]	2008 (1)	America	Caucasian	270	21.8	3.1	165	22.3	3.11	B
Shah *et al*[36]	2008 (2)	America	Black	37	20.1	3.44	84	20.2	3.08	B
Van Nieuwerburgh *et al*[40]	2008	Belgium	Caucasian	97	21.93	2.122	31	21.823	3.112	NC
Dasgupta *et al*[41]	2010	India	Asian	250	18.74	0.13	299	18.73	0.12	D
Laisk *et al*[34]	2010	Estonia	Caucasian	32	21.5	1.6	79	21.6	1.8	D
Robeva *et al*[44]	2010	Bulgaria	Caucasian	52	21.6	2.62	41	21.3	3.71	D
Skrgatic *et al*[42]	2011	Croatia	Caucasian	214	22.1	3.4	209	21.9	3.2	D
Schüring *et al*[35]	2011	Germany	Caucasian	72	21.43	1.87	179	21.99	2.07	D
Rajender *et al*[47]	2013	India	Asian	169	17.39	2.29	175	17.43	2.43	D

A: (1) Proven fertility; (2) No history of subfertility treatment; and (3) Normal menstrual cycles (25-32 d); B: National Institutes of Health criteria; C: (1) Non-hirsute; (2) Proven fertility; (3) Regular menstrual cycles; and (4) Normal ovaries; D: The criteria of Rotterdam Revised 2003 (two of three); E: (1) Oligo-/amenorrhea; (2) Polycystic ovaries; and (3) Hyper-androgenism. PCOS: Polycystic ovarian syndrome; NC: Unclear.

Table 3 Meta-analysis of (TAAAA)n sex hormone-binding globulin polymorphism and polycystic ovarian syndrome

	No. of studies	OR (95%CI)	Heterogeneity			Publication bias	
			χ^2	I^2 (%)	*P*	Begg's *P*	Egger's *P*
S (< 8 repeats)	5	Fixed, 0.91 (0.78-1.05)	7.59	47.8	0.108	0.221	0.221
L (≥ 8 repeats)	5	Fixed, 1.10 (0.95-1.27)	7.30	45.2	0.121	0.221	0.225
SS	3	Fixed, 0.87 (0.66-1.14)	0.60	0.0	0.749	1.000	0.564
SL	3	Fixed, 1.12 (0.86-1.46)	1.64	0.0	0.441	0.296	0.072
LL	3	Fixed, 1.03 (0.72-1.47)	2.84	29.6	0.242	1.000	0.256

OR: Odd ratio; S: Short alleles; L: Long alleles; SS: Short-short genotype; SL: Short-long genotype; LL: Long-long genotype.

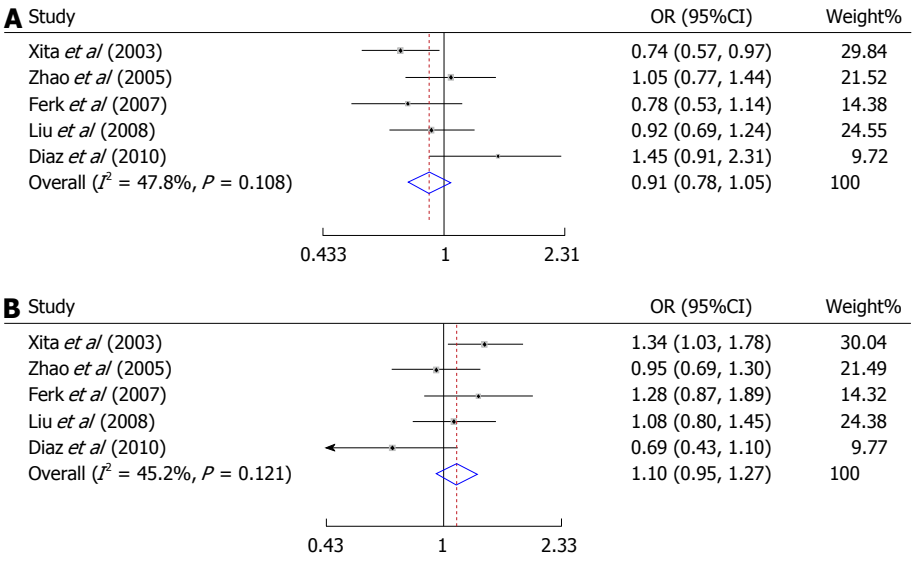


Figure 3 Association between polycystic ovarian syndrome risk and (TAAAA)n sex hormone-binding globulin alleles. A: Comparison of short alleles in the PCOS group with those in the control group; B: Comparison of long alleles in the PCOS group with those in the control group. PCOS: Polycystic ovarian syndrome.

We found that the (CAG)n repeats in race group (Asian: WMD = -0.03, 95%CI: -0.14-0.07; Caucasian: WMD = -0.02, 95%CI: -0.24-0.21) (Figure 6 and Table 4) and in diagnosis criteria group (the criteria of Rotterdam: WMD

Table 4 Meta-analysis of (CAG)n androgen receptor polymorphism and polycystic ovarian syndrome

		Number of studies	WMD (95%CI)	Heterogeneity			Publication bias	
				χ^2	I^2 (%)	*P*	Begg's *P*	Egger's *P*
All		15	Random, -0.03 (-0.13, 0.08)	28.55	51.0	0.012	0.621	0.866
Race	Asian	5	Random, -0.03 (-0.14, 0.07)	14.74	72.9	0.005	0.806	0.875
	Caucasian	9	Fixed, -0.02 (-0.24, 0.21)	13.77	41.9	0.088	0.175	0.596
The criteria of Rotterdam		8	Fixed, 0.01 (-0.01, 0.03)	5.91	0.0	0.550	1.000	0.784

WMD: Weighted mean difference.

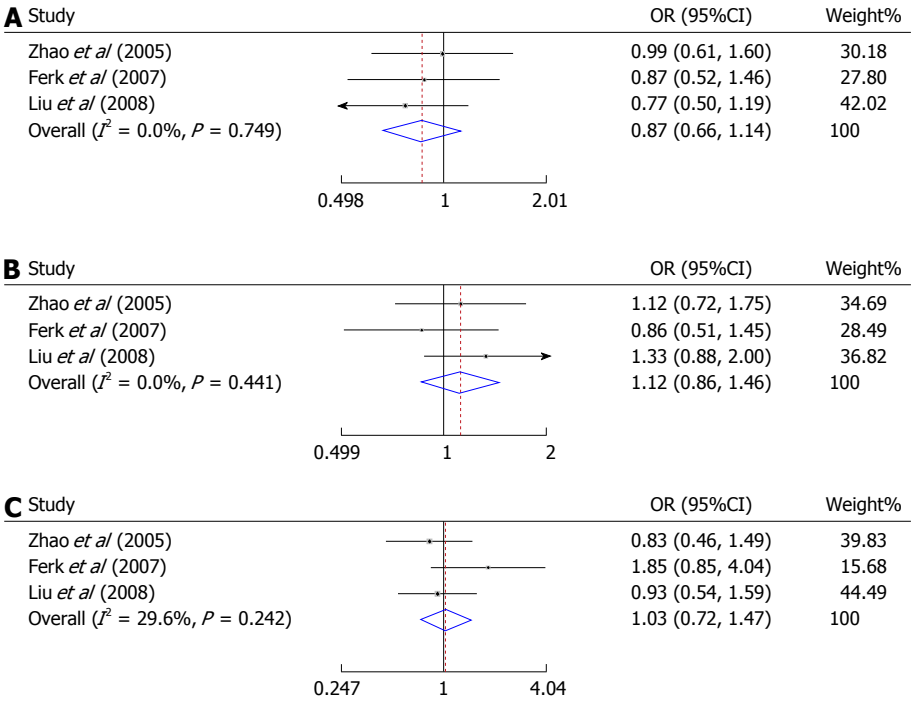


Figure 4 Association between polycystic ovarian syndrome risk and (TAAAA)n sex hormone-binding globulin genotypes. A-C: Comparison of short-short (A), short-long (B), long-long (C) genotype in the PCOS group with those in the control group. PCOS: Polycystic ovarian syndrome.

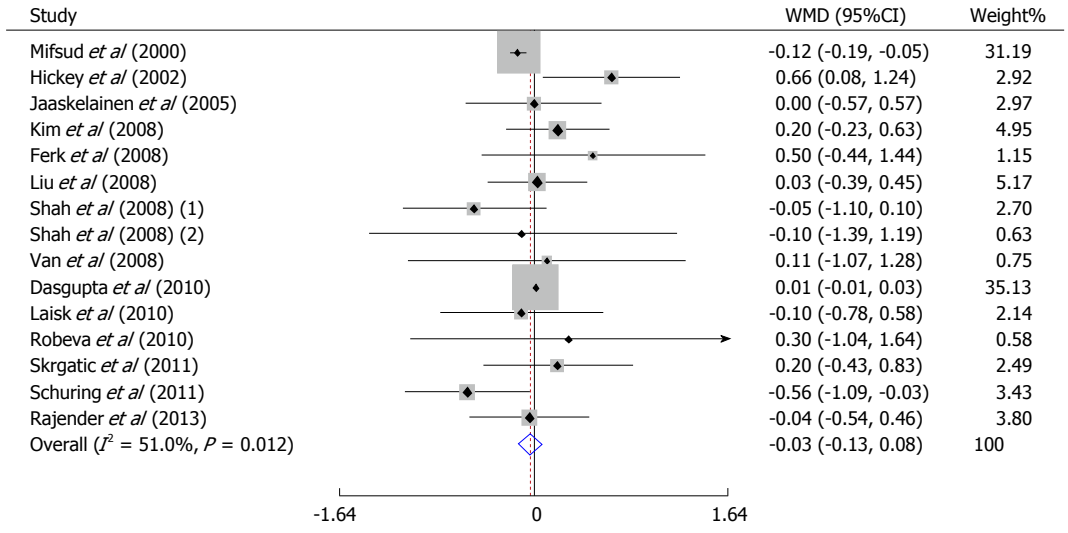


Figure 5 Association between polycystic ovarian syndrome risk and (CAG)n repeats in androgen receptor in all selected studies. WMD: Weighted mean difference.

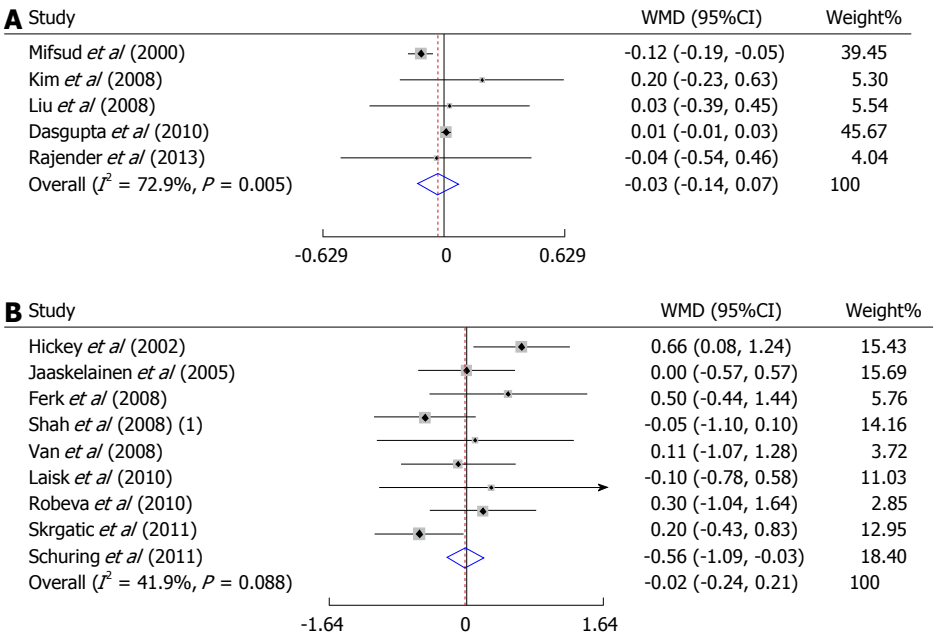


Figure 6 Association between polycystic ovarian syndrome risk and (CAG)n repeats in polycystic ovarian syndrome by race. A: Asian; B: Caucasian. WMD: Weighted mean difference.

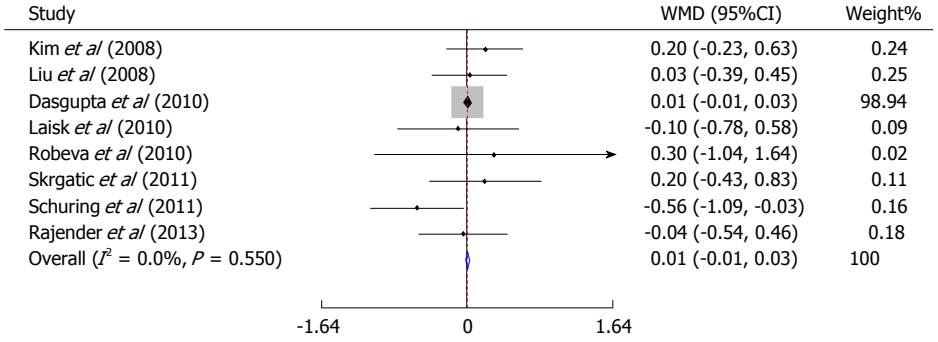


Figure 7 Association between polycystic ovarian syndrome diagnosed according to the criteria of Rotterdam Revised 2003 and (CAG)n repeats in androgen receptor. WMD: Weighted mean difference.

= 0.01, 95%CI: -0.01-0.03) (Figure 7 and Table 4) had no association with PCOS, which was similar to the previous results.

DISCUSSION

PCOS is a multifactorial disorder of unclear etipathogenesis. Hyperandrogenism is gradually being the hallmark of PCOS, and it is an item included in all the three worldwide diagnostic guides. SHBG regulates the bioavailability of sexual hormones to target tissues and is the primary plasma transport protein for those hormones, while AR is the protein to bind androgen and activate the downstream pathway in target cells. Since Xita *et al*[10] first reported PCOS risk in association with genetic variants in SHBG and Mifsud e*t al*[23] reported the relationship between PCOS risk and AR polymorphic CAG repeat, a series of following studies were performed. If there was a definite conclusion of PCOS risk with SHBG or AR polymorphism, PCOS could been caught earlier and prognosis would be better.

For the *SHBG* gene, Xita *et al*[10] discovered that PCOS women had a greater frequency of longer (TAAAA)n (more than 8 repeats) than normal women and proposed that (TAAAA)n repeat variants may be implicated in the development of PCOS. Whereafter, some case-control experiments proved this[24], but others did not[25-27]. In our meta-analysis, we selected OR as the effect size measurement to estimate the influence of (TAAAA)n repeat variants on PCOS risk. The summary ORs for TAAAA alleles (including S and L) indicated that there was no association between the TAAAA polymorphism and PCOS risk. Furthermore, the ORs of the combination of TAAAA alleles (SS, SL and LL) showed no differences between PCOS patients and controls either. These results indicate that the (TAAAA)n polymorphism has no influence on the development of PCOS. As Martínez-García *et al*[28] proposed SHBG as a candidate gene for PCOS, our result may be partly explained by the influence of other single nucleotide polymorphisms in the *SHBG*

gene, such as *rs1799941*[29], *rs2075230*[30], *rs6257* (T/C)[31], *rs727428*[32] and *rs6259*[28].

Our analysis on the *SHBG* gene has several strengths: (1) selection of different combinations of alleles; (2) comprehensive search for original case-control studies without limitation of language; and (3) adoption of bias measurements to avoid publication bias in study selection and data abstraction. On the other hand, limitations exist in this meta-analysis: (1) the numbers of studies and subjects included in this meta-analysis were small; (2) the diagnosis criteria for PCOS in selected studies were different and could not guarantee that involved PCOS cases had similar characteristics; and (3) only publications were enrolled in our meta-analysis, resulting in potential publication bias which was inevitable.

For the *AR* gene, there has been a study showing that higher AR activity correlated with shorter CAG and speculating that the CAG repeat variants were a sign in the development of PCOS[23]. A series of following case-control studies were performed to confirm this result. After scanning titles, abstracts and full texts, fourteen studies were included in our analysis. Among those, four showed that there were more short CAG alleles in PCOS patients compared with controls[33-36], while eight reported the opposite results[27,37-42] and the remaining two found no significant difference between the two groups[43,44]. We selected WMD as the effect size measurement to estimate the association between (CAG)n repeat variants and PCOS risk. The summary WMD for mean of CAG alleles showed that there was no statistical relationship between CAG repeat variants and PCOS risk. Furthermore, the WMD of CAG biallelic mean in subgroups (by race: Asian and Caucasian; by diagnosis criteria: the Rotterdam criteria) displayed no difference between cases and controls. These statistical results indicate that the CAG repeat variants have no influence on the development of PCOS, which was similar to the conclusions of three other meta-analysis[45-47].

Our meta-analysis on the *AR* gene has several strengths: (1) subgroup analysis was performed by race and diagnosis criteria; (2) comprehensive search for original case-control studies was done without limitation of language; and (3) most studies were proved to be homogeneous. Further, there are some limitations in our study: (1) only biallelic mean of CAG repeat variants was analyzed, without analysis on CAG alleles and genotype; (2) the number of studies focusing on Black was small, making it hard to analyze the association in the Black population; and (3) most results were not adjusted because of the inconsistent characteristics of participants among different studies.

In summary, there is no association between SHBG (TAAAA)n and AR (CAG)n polymorphisms and PCOS risk. In order to further understand the relationship between gene polymorphisms and PCOS risk, more studies should be launched to enlarge the sample size and variety of gene polymorphisms with unified diagnostic criteria, which will make the meta-analysis more convincing and useful.

COMMENTS

Background

Polycystic ovarian syndrome (PCOS) is one of the most common endocrine disorders in women of reproductive age and its risk is increasing in the association with genetic variants. The (TAAAA)n polymorphism in the sex hormone-binding globulin (*SHBG*) gene and the (CAG)n polymorphism in the androgen receptor (*AR*) gene are two hotspots, but there are no definite results regarding the association between those two genetic variants and PCOS risk.

Research frontiers

Over the past two decades, many studies have been performed to understand the associations between SHBG (TAAAA)n and AR (CAG)n repeat variants and PCOS risk. Moreover, several systematic reviews were recently performed to investigate these associations. However, the inclusion criteria varied greatly among those reviews and thus could not achieve a comprehensive conclusion.

Innovations and breakthroughs

Based on this meta-analysis, neither the TAAAA polymorphism in the *SHBG* gene nor the CAG polymorphism in the *AR* gene has no influence on the risk of PCOS. Similar results were indicated in subgroup analyses of the combination of alleles by race and diagnosis criteria, which were not presented clearly in previous reviews.

Applications

SHBG (TAAAA)n and AR (CAG)n repeat variants have no association with PCOS risk, which prompts a further investigation of other single nucleotide polymorphisms in those genes, including *rs1799941*, *rs2075230*, *rs6257* (T/C), *rs727428* and *rs6259*.

Terminology

Polymorphism is the regular and simultaneous occurrence in a single interbreeding population of two or more discontinuous genotypes. The concept includes differences in genotypes ranging in size from a single nucleotide site to large nucleotide sequences visible at the chromosomal level.

Peer review

This is a good meta-analysis in which the authors investigated the association among SHBG (TAAAA)n and AR (CAG)n polymorphisms and PCOS risk. The result definitely proved the results of previous reviews and informed that other polymorphisms in the *SHBG* and *AR* genes may contribute to PCOS risk. The meta-analysis is innovative and the manuscript is well written.

REFERENCES

1 **Azziz R**, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The prevalence and features of the polycystic ovary syndrome in an unselected population. *J Clin Endocrinol Metab* 2004; **89**: 2745-2749 [PMID: 15181052 DOI: 10.1210/jc.2003-032046]

2 **Azziz R**, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE, Witchel SF. The Androgen Excess and PCOS Society criteria for the polycystic ovary syndrome: the complete task force report. *Fertil Steril* 2009; **91**: 456-488 [PMID: 18950759 DOI: 10.1016/j.fertnstert.2008.06.035]

3 **Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group**. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril* 2004; **81**: 19-25 [PMID: 14711538]

4 **Azziz R**, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE, Witchel SF. Positions statement: criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an Androgen Excess Society guideline. *J Clin Endocrinol Metab* 2006; **91**: 4237-4245 [PMID: 16940456 DOI: 10.1210/jc.2006-0178]

5 **Trivax B**, Azziz R. Diagnosis of polycystic ovary syndrome. *Clin Obstet Gynecol* 2007; **50**: 168-177 [PMID: 17304034 DOI: 10.1097/GRF.0b013e31802f351b]

6 **Fan W**, Li S, Chen Q, Huang Z. Association between the (TAAAA)n SHBG polymorphism and PCOS: a systematic

review and meta-analysis. *Gynecol Endocrinol* 2013; **29**: 645-650 [PMID: 23772775 DOI: 10.3109/09513590.2013.797394]
7 **Bérubé D**, Séralini GE, Gagné R, Hammond GL. Localization of the human sex hormone-binding globulin gene (SHBG) to the short arm of chromosome 17 (17p12----p13). *Cytogenet Cell Genet* 1990; **54**: 65-67 [PMID: 2249477]
8 **Hammond GL**. Molecular properties of corticosteroid binding globulin and the sex-steroid binding proteins. *Endocr Rev* 1990; **11**: 65-79 [PMID: 2180688 DOI: 10.1210/edrv-11-1-65]
9 **Cousin P**, Calemard-Michel L, Lejeune H, Raverot G, Yessaad N, Emptoz-Bonneton A, Morel Y, Pugeat M. Influence of SHBG gene pentanucleotide TAAAA repeat and D327N polymorphism on serum sex hormone-binding globulin concentration in hirsute women. *J Clin Endocrinol Metab* 2004; **89**: 917-924 [PMID: 14764814 DOI: 10.1210/jc.2002-021553]
10 **Xita N**, Tsatsoulis A, Chatzikyriakidou A, Georgiou I. Association of the (TAAAA)n repeat polymorphism in the sex hormone-binding globulin (SHBG) gene with polycystic ovary syndrome and relation to SHBG serum levels. *J Clin Endocrinol Metab* 2003; **88**: 5976-5980 [PMID: 14671199 DOI: 10.1210/jc.2003-030197]
11 **Hogeveen KN**, Talikka M, Hammond GL. Human sex hormone-binding globulin promoter activity is influenced by a (TAAAA)n repeat element within an Alu sequence. *J Biol Chem* 2001; **276**: 36383-36390 [PMID: 11473114 DOI: 10.1074/jbc.M104681200]
12 **Brown CJ**, Goss SJ, Lubahn DB, Joseph DR, Wilson EM, French FS, Willard HF. Androgen receptor locus on the human X chromosome: regional localization to Xq11-12 and description of a DNA polymorphism. *Am J Hum Genet* 1989; **44**: 264-269 [PMID: 2563196]
13 **Chamberlain NL**, Driver ED, Miesfeld RL. The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucleic Acids Res* 1994; **22**: 3181-3186 [PMID: 8065934]
14 **Parmar MK**, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med* 1998; **17**: 2815-2834 [PMID: 9921604]
15 **Mantel N**, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; **22**: 719-748 [PMID: 13655060]
16 **DerSimonian R**, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177-188 [PMID: 3802833]
17 **Cochran WG**. The combination of estimates from different experiments. *Biometrics* 1954; **10**: 29 [DOI: 10.2307/3001666]
18 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327.7414.557]
19 **Higgins JP**, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539-1558 [PMID: 12111919 DOI: 10.1002/sim.1186]
20 **Bonovas S**, Filioussi K, Tsantes A. Diabetes mellitus and risk of prostate cancer: a meta-analysis. *Diabetologia* 2004; **47**: 1071-1078 [PMID: 15164171 DOI: 10.1007/s00125-004-1415-6]
21 **Egger M**, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-634 [PMID: 9310563]
22 **Begg CB**, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994; **50**: 1088-1101 [PMID: 7786990]
23 **Mifsud A**, Ramirez S, Yong EL. Androgen receptor gene CAG trinucleotide repeats in anovulatory infertility and polycystic ovaries. *J Clin Endocrinol Metab* 2000; **85**: 3484-3488 [PMID: 10999852 DOI: 10.1210/jcem.85.9.6832]
24 **Zhao JL**, Chen ZJ, Zhao YR, Zhao LX, Wang LC, Li Y, Tang R, Shi YH. [Study on the (TAAAA)n repeat polymorphism in sex hormone-binding globulin gene and the SHBG serum levels in putative association with the glucose metabolic status of Chinese patients suffering from polycystic ovarian syndrome in Shandong province]. *Zhonghua Yixue Yichuanxue Zazhi* 2005; **22**: 644-647 [PMID: 16331562]
25 **Díaz M**, López-Bermejo A, Petry CJ, de Zegher F, Ibáñez L. Efficacy of metformin therapy in adolescent girls with androgen excess: relation to sex hormone-binding globulin and androgen receptor polymorphisms. *Fertil Steril* 2010; **94**: 2800-2803.e1 [PMID: 20691435 DOI: 10.1016/j.fertnstert.2010.06.083]
26 **Ferk P**, Teran N, Gersak K. The (TAAAA)n microsatellite polymorphism in the SHBG gene influences serum SHBG levels in women with polycystic ovary syndrome. *Hum Reprod* 2007; **22**: 1031-1036 [PMID: 17189294 DOI: 10.1093/humrep/del457]
27 **Liu Q**, Gu W, Cui B, Hong J, Zhang Y, Chi Z, Su Y, Ning G. The association of TAAAAn repeat polymorphism in sex hormone-binding protein gene with polycystic ovary syndrome in Chinese population. *Endocrine* 2008; **34**: 62-67 [PMID: 18937076 DOI: 10.1007/s12020-008-9104-8]
28 **Martínez-García MÁ**, Gambineri A, Alpañés M, Sanchón R, Pasquali R, Escobar-Morreale HF. Common variants in the sex hormone-binding globulin gene (SHBG) and polycystic ovary syndrome (PCOS) in Mediterranean women. *Hum Reprod* 2012; **27**: 3569-3576 [PMID: 23001781 DOI: 10.1093/humrep/des335]
29 **Svartberg J**, Schirmer H, Wilsgaard T, Mathiesen EB, Njølstad I, Løchen ML, Jorde R. Single-nucleotide polymorphism, rs1799941 in the Sex Hormone-Binding Globulin (SHBG) gene, related to both serum testosterone and SHBG levels and the risk of myocardial infarction, type 2 diabetes, cancer and mortality in men: the Tromsø Study. *Andrology* 2014; **2**: 212-218 [PMID: 24327369 DOI: 10.1111/j.2047-2927.2013.00174.x]
30 **Chen Z**, Tao S, Gao Y, Zhang J, Hu Y, Mo L, Kim ST, Yang X, Tan A, Zhang H, Qin X, Li L, Wu Y, Zhang S, Zheng SL, Xu J, Mo Z, Sun J. Genome-wide association study of sex hormones, gonadotropins and sex hormone-binding protein in Chinese men. *J Med Genet* 2013; **50**: 794-801 [PMID: 24049095 DOI: 10.1136/jmedgenet-2013-101705]
31 **Sunbul M**, Eren F, Nacar C, Agirbasli M. Sex hormone binding globulin gene polymorphisms and metabolic syndrome in postmenopausal Turkish women. *Cardiol J* 2013; **20**: 287-293 [PMID: 23788303 DOI: 10.5603/cj.2013.0074]
32 **Pau C**, Saxena R, Welt CK. Evaluating reported candidate gene associations with polycystic ovary syndrome. *Fertil Steril* 2013; **99**: 1774-1778 [PMID: 23375202 DOI: 10.1016/j.fertnstert.2012.12.033]
33 **Ibáñez L**, Ong KK, Mongan N, Jääskeläinen J, Marcos MV, Hughes IA, De Zegher F, Dunger DB. Androgen receptor gene CAG repeat polymorphism in the development of ovarian hyperandrogenism. *J Clin Endocrinol Metab* 2003; **88**: 3333-3338 [PMID: 12843184 DOI: 10.1210/jc.2002-021791]
34 **Laisk T**, Haller-Kikkatalo K, Laanpere M, Jakovlev U, Peters M, Karro H, Salumets A. Androgen receptor epigenetic variations influence early follicular phase gonadotropin levels. *Acta Obstet Gynecol Scand* 2010; **89**: 1557-1563 [PMID: 21050150 DOI: 10.3109/00016349.2010.526182]
35 **Schüring AN**, Welp A, Gromoll J, Zitzmann M, Sonntag B, Nieschlag E, Greb RR, Kiesel L. Role of the CAG repeat polymorphism of the androgen receptor gene in polycystic ovary syndrome (PCOS). *Exp Clin Endocrinol Diabetes* 2012; **120**: 73-79 [PMID: 22068615 DOI: 10.1055/s-0031-1291343]
36 **Shah NA**, Antoine HJ, Pall M, Taylor KD, Azziz R, Goodarzi MO. Association of androgen receptor CAG repeat polymorphism and polycystic ovary syndrome. *J Clin Endocrinol Metab* 2008; **93**: 1939-1945 [PMID: 18303071 DOI: 10.1210/jc.2008-0038]
37 **Hickey T**, Chandy A, Norman RJ. The androgen receptor CAG repeat polymorphism and X-chromosome inactivation in Australian Caucasian women with infertility related to polycystic ovary syndrome. *J Clin Endocrinol Metab* 2002; **87**:

161-165 [PMID: 11788641 DOI: 10.1210/jcem.87.1.8137]
38 **Ferk P**, Perme MP, Teran N, Gersak K. Androgen receptor gene (CAG)n polymorphism in patients with polycystic ovary syndrome. *Fertil Steril* 2008; **90**: 860-863 [PMID: 18555222 DOI: 10.1016/j.fertnstert.2007.07.1291]
39 **Kim JJ**, Choung SH, Choi YM, Yoon SH, Kim SH, Moon SY. Androgen receptor gene CAG repeat polymorphism in women with polycystic ovary syndrome. *Fertil Steril* 2008; **90**: 2318-2323 [PMID: 18191848 DOI: 10.1016/j.fertnstert.2007.10.030]
40 **Van Nieuwerburgh F**, Stoop D, Cabri P, Dhont M, Deforce D, De Sutter P. Shorter CAG repeats in the androgen receptor gene may enhance hyperandrogenicity in polycystic ovary syndrome. *Gynecol Endocrinol* 2008; **24**: 669-673 [PMID: 19172534 DOI: 10.1080/09513590802342841]
41 **Dasgupta S**, Sirisha PV, Neelaveni K, Anuradha K, Reddy AG, Thangaraj K, Reddy BM. Androgen receptor CAG repeat polymorphism and epigenetic influence among the south Indian women with Polycystic Ovary Syndrome. *PLoS One* 2010; **5**: e12401 [PMID: 20865044 DOI: 10.1371/journal.pone.0012401]
42 **Skrgatic L**, Baldani DP, Cerne JZ, Ferk P, Gersak K. CAG repeat polymorphism in androgen receptor gene is not directly associated with polycystic ovary syndrome but influences serum testosterone levels. *J Steroid Biochem Mol Biol* 2012; **128**: 107-112 [PMID: 22107839 DOI: 10.1016/j.jsbmb.2011.11.006]
43 **Jääskeläinen J**, Korhonen S, Voutilainen R, Hippeläinen M, Heinonen S. Androgen receptor gene CAG length polymorphism in women with polycystic ovary syndrome. *Fertil Steril* 2005; **83**: 1724-1728 [PMID: 15950642 DOI: 10.1016/j.fertnstert.2004.11.080]
44 **Robeva R**, Dobreva D, Kirilov G, Mechandjiev T, Tomova A. CAG repeat polymorphism in women with PCOS and healthy controls. *Endocrine Abstracts* 2010; **22**: 459
45 **Wang R**, Goodarzi MO, Xiong T, Wang D, Azziz R, Zhang H. Negative association between androgen receptor gene CAG repeat polymorphism and polycystic ovary syndrome? A systematic review and meta-analysis. *Mol Hum Reprod* 2012; **18**: 498-509 [PMID: 22695532 DOI: 10.1093/molehr/gas024]
46 **Zhang T**, Liang W, Fang M, Yu J, Ni Y, Li Z. Association of the CAG repeat polymorphisms in androgen receptor gene with polycystic ovary syndrome: a systemic review and meta-analysis. *Gene* 2013; **524**: 161-167 [PMID: 23628801 DOI: 10.1016/j.gene.2013.04.040]
47 **Rajender S**, Carlus SJ, Bansal SK, Negi MP, Sadasivam N, Sadasivam MN, Thangaraj K. Androgen receptor CAG repeats length polymorphism and the risk of polycystic ovarian syndrome (PCOS). *PLoS One* 2013; **8**: e75709 [PMID: 24116069 DOI: 10.1371/journal.pone.0075709]

P- Reviewer: Mais V, Yokoyama Y **S- Editor**: Gong XM
L- Editor: Wang TQ **E- Editor**: Liu SQ

Published by **Baishideng Publishing Group Inc**
8226 Regency Drive, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

ISSN 2308-3840 (online)

World Journal of *Meta-Analysis*

World J Meta-Anal 2015 April 26; 3(2): 82-132

Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

A peer-reviewed, online, open-access journal of meta-analysis

Editorial Board

2013-2018

The *World Journal of Meta-Analysis* Editorial Board consists of 400 members, representing a team of worldwide experts in clinical meta-analysis research. They are from 41 countries, including Argentina (2), Australia (3), Austria (1), Belgium (5), Brazil (11), Canada (16), Chile (3), China (121), Croatia (1), Egypt (1), Finland (4), France (2), Germany (9), Greece (11), Hungary (2), India (12), Iran (2), Ireland (1), Israel (2), Italy (41), Japan (5), Lithuania (1), Malaysia (1), Netherlands (8), New Zealand (1), Norway (1), Peru (1), Poland (4), Portugal (6), Romania (1), Saudi Arabia (4), Singapore (3), South Africa (1), South Korea (7), Spain (8), Sri Lanka (2), Sudan (1), Switzerland (2), Thailand (4), Turkey (3), United Kingdom (22), and United States (64).

Alessandro Cucchetti, *Bologna*
Gianfranco Damiani, *Rome*
Fabrizio D'Ascenzo, *Turin*
Massimo Del Fabbro, *Milano*
Valeria Fadda, *Florence*
Alessandro Fancellu, *Sassari*
Giuseppe Ferrante, *Milan*
Virginia Festa, *Rome*
Francesco Fiorica, *Ferrara*
Guglielmo Giraldi, *Rome*
Jenny Guidi, *Bologna*
Lorenzo Loffredo, *Rome*
Andrea Messori, *Firenze*
Eliano Pio Navarese, *Bydgoszcz*
Stefano Omboni, *Solbiate Arno*
Alvisa Palese, *Udine*
Stefano Palomba, *Reggio Emilia Stefano*
Carlo Perricone, *Rome*
Mario Petretta, *Naples*
Alessandro Pezzini, *Brescia*
Gianluca Pontone, *Milan*
Paolo Emilio Puddu, *Rome*
Andrea Rognoni, *Novara*
Giuseppe Scalabrino, *Milan*
Città della Salute e della Scienza, *Turin*
Fabrizio Sgolastra, *L'Aquila*
Maria Lucia Specchia, *Rome*
Stefano Trastulli, *Terni*
Fabio Tine, *Palermo*
Nereo Vettoretto, *Chiari*
Alberto Vianello, *Perugia*
Luigi Zorcolo, *Cagliari*

Japan

Nguyen Tien Huy, *Nagasaki*
Hiroharu Kamioka, *Tokyo*
Koji Kawakami, *Kyoto*
Keitaro Matsuo, *Nagoya*
Kazushi Okamoto, *Nagoya*

Lithuania

Edmundas Kadusevicius, *Kaunas*

Malaysia

SP Pani, *Ipoh*

Netherlands

Michel van den Bekerom, *Amsterdam*
Dimitra Dodou, *Delft*
D Haverkamp, *Amsterdam*
Vassilios Koussoulas, *Drachten*
BJ Polder, *Nijmegen*
Theo Stijnen, *Leiden*
RNM Weijers, *Amsterdam*
Joost de Winter, *Delft*

New Zealand

Shaofeng Li, *Auckland*

Norway

Eivind Berge, *Oslo*

Peru

Rafael Bolaños Díaz, *Lima*

Poland

Maciej Banach, *Lodz*
Krzysztof Jonderko, *Sosnowiec*
Jolanta Lissowska, *Warsaw*
Maciej Plaszewski, *Warsaw*

Portugal

Daniel Caldeira, *Lisboa*
J Costa, *Lisbon*
Ana Miguel, *Coimbra*
Manuel Morgado, *Covilh*
Bárbara Peleteiro, *Porto*
Rui Torres, *Paredes*

Romania

Fratila Ovidiu, *Oradea*

Saudi Arabia

Hazem M Al-Mandeel, *Riyadh*
Ezzeldin M Ibrahim, *Jeddah*
Mutahir A Tunio, *Riyadh*
Hayfaa A Wahabi, *Riyadh*

Singapore

Nikos LD Chatzisarantis, *Singapore*
Edwin Choon Wyn Lim, *Singapore*
Roger Ho, *Singapore*

South Africa

Alaine Umubyeyi Nyaruhirira, *Pretoria*

South Korea

Jung-Hee Kim, *Cheonan*
Hyangsook Lee, *Seoul*
Myeong Soo Lee, *Daejeon*
Chi-Un Pae, *Bucheon*
Yong Hyun Park, *Yangsan*
Jae Hong Seo, *Seoul*
Yong Sang Song, *Seoul*

Spain

Pablo Avanzas, *Oviedo*
Joan Cid, *Barcelona*
Joaquin de Haro, *Madrid*
Joan Guardia-Olmos, *Barcelona*
Nabil Halaihlel, *Zaragoza*
JA Monge-Argilés, *Alicante*
Raul Moreno, *Madrid*
Inés Velasco, *Aracena*

Sri Lanka

Ranil Jayawardena, *Colombo*
Priyanga Ranasinghe, *Colombo*

Switzerland

Jay P Singh, *Zurich*
Giorgio Treglia, *Bellinzona*

Sudan

Samir MH Shaheen, *Khartoum*

Thailand

Chuenjid Kongkaew, *Phitsanulok*
Manop Pithukpakorn, *Bangkok*
Piyamitr Sritara, *Bangkok*
Surasak Saokaew, *Phayao*

Turkey

Nese Demirturk, *Afyonkarahisar*
Nilüfer Ozabaci, *Eskisehir*
Ilke Sipahi, *Istanbul*

United Kingdom

Omar M Aboumarzouk, *Wales*
Abeer Al-Namankany, *London*
Ernest A Azzopardi, *Cardif*
Umberto Benedetto, *Cambridge*
Joanne Brooke, *London*
Noriko Cable, *London*
David Chan, *Cardiff*
YC Cheong, *Southampton*
Andrew Currie, *Harrow*
G Nabi, *Dundee*

Lesley A Anderson, *Belfast*
Valentina Gallo, *London*
Gianpiero Gravante, *Leicester*
Peter N Lee, *Surrey*
Igho Onakpoya, *Oxford*
Ashish Pradhan, *Huntingdon*
Evridiki Patelarou, *London*
Jian-Qing Shi, *Newcastle*
Surendra P Singh, *Wolverhampton*
Natalie Taylor, *Leeds*
Yousef Shahin, *Hull*
Zheng Ye, *Cambridge*

United States

L Joseph Su, *Rockville*
Olusola Adesope, *Pullman*
Mike Allen, *Milwaukee*
Bhupinder Anand, *Bellaire*
Stephen Aronoff, *Philadelphia*
KoKo Aung, *San Antonio*
William L Baker, *Farmington*
Moritz C Wyler von Ballmoos, *Milwaukee*
Matthew L Bechtold, *Columbia*
Atul Bhardwaj, *Hershey*
Somjot S Brar, *Los Angeles*
Hui Cai, *Nashville*
Subhash Chandra, *Towson*
Wen-Pin Chang, *Omaha*
Yong Chen, *North Wales*
Myunghan Choi, *Phoenix*
John Coverdale, *Houston*
Prakash C Deedwania, *Fresno*
Eugene Demidenko, *Hanover*
Hong-Wen Deng, *New Orleans*
Eric M Deshaies, *Syracuse*
Tao Fan, *Whitehouse Station*
Shinga Feresu, *Bloomington*
Janvier Gasana, *Miami*
Kaveh Hajifathalian, *Boston*
Mohammad Obaidul Hoque, *Baltimore*
Larissa R Brunner Huber, *Charlotte*
Imran H Iftikhar, *Columbia*
Vijayvel Jayaprakash, *Buffalo*
Xuezhi Jiang, *Weat Reading*
Shuo Jiao, *Seattle*
Evelyn Johnson, *Boise*
Le Kang, *Silver Spring*
SR Kapadia, *Cleveland*
Lior Katz, *Houston*
Daniel M Laskin, *Richmond*
Yu Liang, *Foster City*
Paul Ellis Marik, *Norfolk*
Lynne V McFarland, *Seattle*
Marcovalerio Melis, *New York*
Brian J Miller, *Augusta*
Pavlos Msaouel, *New York*
Joshua E Muscat, *Hershey*
Chee Yuan Ng, *Loma Linda*
Nghi C Nguyen, *Saint Louis*
Brandi S Niemeier, *Whitewater*
Thomas D Parsons, *Denton*
Nidal Abi Rafeh, *New Orleans*
Praveen Roy, *Marshfield*
Ali Salavati, *Philadelphia*
Ankur Sethi, *Chicago*
Tatyana A Shamliyan, *Minneapoli*
Qian Shi, *Rochester*
Zhongjie Shi, *Philadelphia*
Param Puneet Singh, *Chicago*
KV Slavin, *Chicago*
Ali El Solh, *Buffalo*
Jielin Sun, *Winston-Salem*
Richard G Trohman, *Chicago*
Laurah Turner, *Cincinnati*
Sheila Wilhelm, *Detroit*
Alex K Wong, *Los Angeles*
Xiaohui Xu, *Gainesville*
Lu Yin, *Nashville*

WJMA World Journal of Meta-Analysis

Contents

Bimonthly Volume 3 Number 2 April 26, 2015

EDITORIAL

SYSTEMATIC REVIEWS

META-ANALYSIS

NAME OF JOURNAL
World Journal of Meta-Analysis

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Bimonthly

EDITOR-IN-CHIEF
Giuseppe Biondi-Zoccai, MD, Assistant Professor, Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
World Journal of Meta-Analysis
Room 903, Building D, Ocean International Center, No. 62 Dongsihuan Zhonglu, Chaoyang District, Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
8226 Regency Drive,
Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
April 26, 2015

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i2.82

World J Meta-Anal 2015 April 26; 3(2): 82-88
ISSN 2308-3840 (online)

EDITORIAL

Arthroscopic capsular release and manipulation under anaesthesia for frozen shoulders: A hot topic

Tim Kraal, Lijkele Beimers

Tim Kraal, Lijkele Beimers, Department of Orthopaedic Surgery, Slotervaartziekenhuis, 1066 EC Amsterdam, The Netherlands

Author contributions: Kraal T wrote this article with support of Beimers L who helped with the design of the article structure and reviewed the manuscript several times.

Conflict-of-interest: All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Correspondence to: Lijkele Beimers, MD, PhD, Department of Orthopaedic Surgery, Slotervaartziekenhuis, Louwesweg 6, 1066 EC Amsterdam, Noord-Holland,
The Netherlands. lijkele.beimers@slz.nl
Telephone: +31-20-5124554
Fax: +31-20-512339

Received: October 14, 2014
Peer-review started: October 14, 2014
First decision: December 17, 2014
Revised: December 31, 2014
Accepted: January 15, 2015
Article in press: January 19, 2015
Published online: April 26, 2015

Abstract

A frozen shoulder is a common cause of shoulder pain and stiffness. The etiology and pathology of frozen shoulders is not fully understood yet. Frozen shoulder is characterized by a decrease in intra-articular volume and capsular compliance. This can lead to significant limitations in daily life. The majority of the patients can be treated conservatively, with functional recovery to be expected in two to three years. However, if conservative treatment fails, manipulation under anaesthesia and arthroscopic capsular release can both be considered as appropriate treatments. Manipulation is a traditionally well-established technique but in recent years it seems that arthroscopic capsular release has gained popularity. Manipulation is a relative time efficient and technically low-demanding procedure in which the glenohumeral joint is forced into different directions under general anaesthesia to release the capsular contracture, thereby increasing the range of motion of the joint. In arthroscopic capsular release the glenohumeral capsule can be released in a more controlled manner under direct vision. There are no prospective comparative trials available to display superiority of one procedure over the other. In addition, the optimal timing of both these interventions still has to be determined. An overview of the literature concerning this topic and a description of both procedures with its own advantages and disadvantages is provided.

Key words: Frozen shoulder; Adhesive capsulitis; Manipulation; Arthroscopy; Capsular release; Shoulder; Shoulder stiffness

Core tip: A frozen shoulder is a common cause of shoulder pain and stiffness, which is characterized by a decrease in intra-articular volume and capsular compliance. If conservative treatment fails, manipulation under anaesthesia and arthroscopic capsular release can both be considered as appropriate treatments. An overview of the literature concerning this topic and a description of both procedures with its own advantages

and disadvantages is provided.

Kraal T, Beimers L. Arthroscopic capsular release and manipulation under anaesthesia for frozen shoulders: A hot topic. *World J Meta-Anal* 2015; 3(2): 82-88 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i2/82.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i2.82

EPIDEMIOLOGY

A frozen shoulder is a commonly encountered condition in the orthopaedic surgeons' practice. Pain and restricted range of motion of the shoulder may lead to disability and a decrease in quality of life. In 1872 Duplay[1] described a painful stiffening of the shoulder, which he named humero-scapular periarthritis. Codman[2] was the first to coin the term frozen shoulder in 1934, for a condition which was characterized by painful restriction of shoulder motion. Neviaser *et al*[3] suggested the term adhesive capsulitis after a cadaveric study and intra-operative findings of a thickened capsule, adherent to the humeral head. Both terms, frozen shoulder and adhesive capsulitis, are now used interchangeable in the literature for the same condition.

Frozen shoulder affects approximately 2%-4% of the general population[4,5]. The peak incidence is mainly between the age of 40 and 65 years, slightly more frequent in women than in men[6]. The most important associated systemic condition is diabetes mellitus, followed by thyroid disorders. The prevalence of frozen shoulder increases to 10%-20% in diabetic patients. These patients seem to have a prolonged course of the disease, less response to conservative treatment and bilateral involvement is seen more frequently[7,8].

The natural history of a frozen shoulder is described in a relative limited amount of studies. In the majority of patients, it seems to be a self-limited condition with functional recovery after 2-3 years[9]. However, some patients experience continued pain and limited range of motion. After recovery, recurrence of a frozen shoulder is extremely rare[10].

DIAGNOSIS

Although frozen shoulder is a well-known clinical entity, there are still many controversies existing about the definition, the different phases and certainly about the optimal treatment regimen. Zuckerman *et al*[11] proposed this descriptive consensus definition, which was agreed by 82% of members of the American Shoulder and Elbow Surgeons: A condition characterized by functional restriction of both active and passive shoulder motion for which radiographs of the glenohumeral joint are essentially unremarkable except for the presence of osteopenia or calcific tendinitis. Commonly clinical findings consist of: painful stiff shoulder for at least 4 wk; severe shoulder pain that interferes with activities of daily living or work; night pain; painful restriction of both passive and active shoulder range of motion and normal radiographic appearance[12,13]. With physical examination, the selective loss of passive external rotation is typical[14].

Frozen shoulder is usually categorized in primary (or idiopathic) and secondary frozen shoulder. In a primary (idiopathic) frozen shoulder, an underlying aetiology cannot be found. In secondary frozen shoulder, local or intrinsic factors (such as proximal humeral fracture, rotator cuff disorders, biceps tendonitis), remote or extrinsic factors (*e.g.,* ipsilateral breast surgery, cervical radiculopathy, cerebrovascular accident, postoperative immobilization) or systemic pathology (including diabetes mellitus, thyroid disorders, hypo-adrenalism) may be related to the disease[11,12].

In 1975 Reeves[15] believed the condition to involve three separate phases. Phase one, the painful phase followed by phase two, the frozen phase in which pain persists and stiffness is aggravated. Phase three is named the thawing phase, where joint motion and pain gradually improve[15]. A wide variety in the duration of each phase is described, but most authors agree with spontaneous functional recovery after 2-3 years[14,16].

Frozen shoulder is a clinical entity which can generally be diagnosed after a thorough history and physical examination. Plain radiographs are typical without abnormalities. Osteoarthritis of the glenohumeral joint can easily be ruled out. Calcifications in the rotator cuff is a common incidental finding. Ultrasonography is not required for the diagnosis but is appropriate to screen for rotator cuff or biceps tendon abnormalities when suspected. Magnetic resonance imaging (MRI) arthrography can show thickening of the coracohumeral ligament and joint capsule in the rotator interval. Also, synovial thickening in the axillary pouch correlates with the stage of adhesive capsulitis[17]. However, magnetic resonance imaging should not be routinely ordered in the evaluation of the frozen shoulder.

PATHOPHYSIOLOGY

In a secondary frozen shoulder, a local or remote factor that leads to immobilisation of the limb or a systemic condition is an underlying cause to be held accountable for the development of a frozen shoulder. However, most cases of frozen shoulder are primary or idiopathic in which the pathophysiology is not yet fully understood. White *et al*[18] suggest an increase in sedentary jobs with a low level of activity as a possible explanation for the increasing occurrence of a frozen shoulder. A decrease in intra-articular volume and capsular compliance was already described in 1969[19]. An inflammatory contracture of the anterosuperior capsule, the glenohumeral ligaments and the coracohumeral ligament is demonstrated in cadaveric studies and MRI studies[14,20]. This

corresponds with the characteristic clinical finding of loss of external rotation in adduction with physical examination. Significant synovial hypertrophy and neovascular proliferation, especially in the rotator interval is often observed during arthroscopy. A histologic study of Bunker *et al*[21] demonstrates that the predominant cells involved are fibroblasts and myofibroblasts in the joint capsule that produce the extracellular matrix. The produced type Ⅲ collagen matrix is packed more densely, causing the shoulder capsule to contract. This excess of extracellular matrix is characteristic for fibroproliferative disorders. Other histologic changes consist of chronic inflammation and perivascular infiltration and fibrosis[22]. On a cellular level, the extracellular matrix turnover (production, degradation and remodelling) is involved by matrix metalloproteinases and their inhibitors. An imbalance can lead to fibroproliferation, which is demonstrated in frozen shoulder patients[23]. The microscopic changes in the anterior capsule and the coracohumeral ligament are very similar to the changes seen in Dupuytrens' disease in the hand. Dupuytrens' disease is frequently observed in patients with a frozen shoulder[21]. Smith *et al*[24] report an incidence of Dupuytrens' disease of 52% in a cohort of patients with a primary frozen shoulder. Although frozen shoulder has a different natural history than Dupuytrens' disease (self-limiting versus progressive), a common biochemical pathway of both fibroproliferative disorders that leads to contracture is suggested[24].

More recently, the role of inflammatory cytokines and growth factors in the pathogenesis of a frozen shoulder is investigated, because they regulate the growth and function of fibroblasts. The study of Lho *et al*[25] confirmed the overexpression of inflammatory cytokines (such as interleukin 1-α, tumor necrosis factor-α and cyclooxygenase-2) in the joint capsule of patients with a frozen shoulder compared to a control group. Also, an overexpression of these inflammatory mediators was found in tissue samples of the subacromial bursa in frozen shoulder patients[25], possibly contributing to the cascade of inflammation eventually leading to fibrosis.

A future better understanding of the pathophysiology of a frozen shoulder on a cellular level can possibly lead to targeted therapy with anti-inflammatory medication[26].

MANAGEMENT

There are many different strategies in the treatment of a frozen shoulder: including but not limited to supervised neglect[9], physiotherapy[27,28], corticosteroid infiltration[29,30], manipulation under anaesthesia (MUA)[31], arthroscopic capsular release (ACR)[32], arthrographic capsular distension[33] and stretching devices[34]. The optimal treatment regimen has not yet been established. Systematic reviews point to a lack of good quality evidence to give evidence based supported recommendations[35,36]. Non-steroidal anti-inflammatory drugs, intra-articular corticosteroid injections and physiotherapy are among the most widely used treatment modalities in the treatment of a frozen shoulder, in both primary and secondary healthcare settings[35,37]. Because the natural history of a frozen shoulder develops in different phases, it is suggested that the timing of different treatment modalities might be important in this regard. However, there is only a limited amount of good quality studies that have investigated this matter. The positive effect of intra-articular corticosteroid injections appears to be most obvious at an early painful phase of the condition[38,39]. Shin *et al*[40] found a similar positive effect of a subacromial corticosteroid injection compared to an intra-articular injection. The role of physiotherapy is still controversial[41]. Most authors are convinced that the physiotherapy protocol must be adjusted to the phase of the condition with a more important role for physiotherapy in later, less painful phases of the condition. Hanchard *et al*[42] suggest different physiotherapy modalities for a pain-predominant or stiffness-predominant frozen shoulder. Kelley *et al*[43] distinguishes three levels of tissue irritability (high, moderate or low irritability) in frozen shoulder patients to adjust the physiotherapy protocol. Furthermore, other than a primary (idiopathic) frozen shoulder, secondary frozen shoulders after trauma or surgery are often more resistant to conservative treatment[44,45].

Taking above into account, conservative treatment seems to be sufficient for most cases, and almost full recovery takes place in two or three years[14]. Most authors state that failure of at least 6 to 12 mo of appropriate non-operative treatment is an indication for more invasive interventions[46]. However, it is questioned if the course of the disease can be shortened when more invasive interventions are undertaken earlier on in the disease[47]. On the other hand, early surgical intervention for symptomatic frozen shoulder may lead to overtreatment in patients with a mild, self-limiting natural course. It might be interesting to know if it is possible to identify which patients will develop a prolonged course, thus could benefit from early invasive treatment. Prospective studies of non-operative treatment showed that approximately 10% of the patients with an idiopathic frozen shoulder develop a refractory frozen shoulder in which further intervention such as MUA or ACR should be considered[6,9]. MUA is a traditionally well-established technique. However, according to the number of publications on this subject in recent years, ACR is gaining more attention. Both procedures have their own specific advantages and disadvantages.

MVA

The same Duplay[1] who described painful stiffening of the shoulder as humero-scapular periarthritis in 1872 suggested MUA as an appropriate treatment

for frozen shoulder[1]. Before the improvement in arthroscopic shoulder surgery, MUA was the standard treatment of a frozen shoulder if conservative treatment had failed.

Different techniques have been described, but a fixed order of manipulations is recommended. The use of a small lever arm and scapular stabilization is recommended to prevent fractures and brachial plexus traction injuries[48]. First the arm is brought in to full flexion, then cross body adduction followed by external rotation with the elbow adducted against the trunk. Then the arm is abducted and moved into internal and finally external rotation. A characteristic crepitus can be heard and felt by the surgeon as the contracted capsule is ruptured. The addition of an intra-articular injection with corticosteroids and local anaesthetic agent is often used at the end of the procedure.

Consistently satisfactory results in both short- and long-term follow-up are reported with MUA. A significant improvement in range of motion and an overall satisfaction rate of 94% at short term is reported by Dodenhoff *et al*[48]. A major cause of satisfaction was to regain the ability to perform normal daily tasks within days of the manipulation. Long term results confirm that the results do not deteriorate after 15 years[49]. Equal range of motion to the contralateral shoulder and no pain was reported in 90% of the patients after 23 years of follow up in a small cohort[50].

ACR

ACR has gained popularity over the years[51]. The first ACR was described by Conti[52] in 1979. The exact procedure and the magnitude of the capsular release differs between various authors. Earlier techniques describe an anterior and inferior release[46,53]. More recent articles favour a complete circumferential (360 degrees) release[32,54,55].

Both beach chair and the lateral decubitus position with the arm suspended in traction are possible to perform an ACR. However, in the beach chair position it is easier to assess the range of motion of the shoulder during surgery. A pressure pump system and a vasoconstrictive agent (*e.g.,* adrenaline or epinephrine) in the irrigation solution are recommended to improve visibility. The capsular release is performed with a radiofrequency probe. The structures in the rotator interval and the anterior capsule must be released first. Ogilvie-Harris *et al*[46] and Pearsall *et al*[56] recommend to release the intra-articular portion of the subscapularis tendon, however, several studies show excellent results without sacrificing the subscapularis[46,51,55,56]. The superior capsule can be released parallel to the joint surface until the muscular fibres of the supraspinatus are visible. It is also possible to release the posterior inferior aspect of the capsule. However, the benefit of this posterior release could not be confirmed in a recent level 1 randomized controlled trial[57]. A gentle manipulation can be performed to assess the obtained range of motion. Some authors infiltrate the shoulder joint with corticosteroids at this point[54]. Good to excellent results with regard to function and pain at both short and long term after ACR are reported. A large prospective study of smith at all reported good pain relief in 80% of the patients within six weeks[55]. Le Lievre *et al*[54] demonstrated that the obtained improvements of pain and patient reported shoulder function maintained after a mean follow up of seven years. In addition, the shoulder range of motion was comparable with that of the contralateral shoulder at time of follow up.

Postoperative treatment and pain management

Similar rehabilitation protocols after MUA and ACR are described. An important aspect after both MUA and ACR is to start physiotherapy immediately, from day one after the surgical intervention. Postoperative pain management must be adequate to tolerate early physiotherapy treatment. This can be achieved in several ways. Pre-operative regional interscalene block[53], a local intra-articular analgesic injection with or without corticosteroid, an indwelling pain pump in the subacromial space, oral analgetics and icepacks have all been described. Immobilisation in a sling must be discouraged at all times to prevent the shoulder joint from getting stiff again[54]. With adequate pain management, both procedures are assumed to be very well tolerated with minimal postoperative pain[48,51]. Most authors agree on intensive supervised physiotherapy twice or three times a week, possibly supplemented by a home exercise program[53,55].

Pros and cons for manipulation under anaesthesia or arthroscopic capsular release

Comparable satisfactory results are reported by many authors for MUA as well as for ACR. To our knowledge there are no randomized controlled trials comparing manipulation with capsular release for frozen shoulder. A comparison between both procedures was attempted in a recent systematic review primarily based on level Ⅳ evidence. With caution, this study slightly favoured ACR over MUA in recalcitrant idiopathic or diabetic frozen shoulders[12]. The need for prospective higher level evidence is emphasized. The overall complication rate for both procedures is rather low with 0.5% complications reported. The advantages and risks of MUA and ACR are listed in Table 1.

One of the most important arguments used by opponents of MUA, is that it is a fairly uncontrollable procedure. You can not see what is released, or torn within or around the shoulder joint. The potential risks of MUA are wide-ranging. Reported iatrogenic injuries are: proximal humerus or humeral shaft fractures[58], brachial plexus traction injury[59], glenohumeral ligament tears, rotator cuff tears, labral lesions, osteochondral fractures of the anterior glenoid rim[60]. Significant osteopenia can be considered as a relative

Table 1 Advantages and risks of manipulation under anaesthesia and arthroscopic capsular release in the treatment of a frozen shoulder

Advantages	Disadvantages/risks
Manipulation under anaesthesia	
Time efficient	Fracture of humeral shaft or neck
Cost efficient	Rotator cuff tearing
Technical easy procedure	Brachial plexus nerve injury
	Labral lesions
	(Osteo)chondral fracture (glenoid rim)
Arthroscopic capsular release	
Visually controlled capsular release	Less time and cost efficient compared to MUA
Identification and treatment of associated intra-articular pathology	Can be technically more demanding
No excessive bleeding in the joint	Cartilage damage when introducing the arthroscope
	Axillary nerve injury
	Chondrolysis due to heat generation
	Extravasation of fluid in surrounding tissues
	Infection

MUA: Manipulation under anaesthesia.

contra indication to MUA. Although a lot of articles address the risk of a humeral fracture and the use of a short lever arm is emphasized, the complication itself is seldom reported[58,61]. Loew *et al*[60] performed an arthroscopy directly after MUA in 30 persons to investigate the intra-articular damage. Hemarthrosis was found in all patients. The anterior capsule was ruptured in 22 out of 30 shoulders, mostly adjacent to and parallel to the labrum, where it is intended to tear. Unequivocal lesions were found in 12 out of 30 shoulders, this involved the anterior and superior labrum, partial tears of the subscapularis tendon, the supraspinatus tendon, the long head of the biceps and one small osteochondral fracture[60]. An evident advantage of MUA in comparison to ACR is that it is more time efficient and that it is associated with substantial lower costs.

Proponents of the ACR procedure believe that a complete release of the capsule can be achieved in a more controlled way. Associated intra-articular pathology can be identified and treated simultaneous. The risks are fairly low, with a documented complication rate of 0.5%[12,45]. However, serious complications as axillary nerve injury, chondrolysis and skin burns due to heat generation or infection are documented[3,45,62]. Nowadays, temperature controlled diathermal probes are commercially available, possibly preventing overheating of the fluids in the joint during surgery. Different from MUA, ACR can be a more technical demanding procedure. Some authors even state that ACR should only be done when MUA has failed[14].

Another option is to combine ACR with manipulation. The manipulation can be a gentle one only to release the capsule where it is difficult to reach or risky to release arthroscopically (for example in the area of the axillary nerve). Early significant improvement in shoulder range of motion with relief of pain and maintenance of these results at long term are reported[41,54,55].

CONCLUSION

A frozen shoulder is a common cause of shoulder pain and stiffness. The majority of the patients can be treated conservatively, with functional recovery to be expected after two to three years. However, if conservative treatment fails, MUA and ACR can both be considered as appropriate treatments. MUA is an easy, time- and cost-efficient technique, but is accompanied by the risk of iatrogenic damage. ACR seems to be a safer way to release the joint capsule. Associated intra-articular pathology can be identified and bleeding can be controlled. However, ACR is technically more demanding, and is also accompanied by the risk of damage to the cartilage or the axillary nerve. Both procedures are performed in large numbers and are considered safe and beneficial for the patient. Superiority of one technique over the other can't be supported by randomized trials comparing both techniques. In addition, the optimal timing of any surgical intervention for frozen shoulder has to be determined yet. Therefore, the decision for either one procedure to treat a frozen shoulder is made by the orthopaedic surgeon and the individual patient together.

REFERENCES

1 **Duplay S.** De la peri-arthrite scapula-humerale et des raideurs de l' epaule qui ensont la consequence. *Arch Gen Med* 1872; **20**: 513

2 **Codman RA**. The Shoulder. Boston, MA: Thomas Todd co, 1934: 216-224

3 **Neviaser RJ**, Neviaser TJ. The frozen shoulder. Diagnosis and management. *Clin Orthop Relat Res* 1987; **223**: 59-64 [PMID: 3652593]

4 **van der Windt DA**, Koes BW, de Jong BA, Bouter LM. Shoulder disorders in general practice: incidence, patient characteristics, and management. *Ann Rheum Dis* 1995; **54**: 959-964 [PMID: 8546527 DOI: 10.1136/ard.54.12.959]

5 **Shah N**, Lewis M. Shoulder adhesive capsulitis: systematic review of randomised trials using multiple corticosteroid injections. *Br J Gen Pract* 2007; **57**: 662-667 [PMID: 17688763]

6 **Griggs SM**, Ahn A, Green A. Idiopathic adhesive capsulitis. A prospective functional outcome study of nonoperative treatment. *J Bone Joint Surg Am* 2000; **82-A**: 1398-1407 [PMID: 11057467]

7 **Dehghan A**, Pishgooei N, Salami MA, Zarch SM, Nafisi-Moghadam R, Rahimpour S, Soleimani H, Owlia MB. Comparison between NSAID and intra-articular corticosteroid injection in frozen shoulder of diabetic patients; a randomized clinical trial. *Exp Clin Endocrinol Diabetes* 2013; **121**: 75-79 [PMID: 23426700 DOI: 10.1055/s-0032-1333278]

8 **Wang K**, Ho V, Hunter-Smith DJ, Beh PS, Smith KM, Weber AB. Risk factors in idiopathic adhesive capsulitis: a case control study. *J Shoulder Elbow Surg* 2013; **22**: e24-e29 [PMID: 23352186 DOI: 10.1016/j.jse.2012.10.049]

9 **Diercks RL**, Stevens M. Gentle thawing of the frozen shoulder: a prospective study of supervised neglect versus intensive physical therapy in seventy-seven patients with frozen shoulder syndrome followed up for two years. *J Shoulder Elbow Surg* 2004; **13**: 499-502 [PMID: 15383804 DOI: 10.1016/j.jse.2004.03.002]

10 **Cameron RI**, McMillan J, Kelly IG. Recurrence of a "primary frozen shoulder": a case report. *J Shoulder Elbow Surg* 2000; **9**: 65-67 [PMID: 10717864 DOI: 10.1016/S1058-2746(00)90011-9]

11 **Zuckerman JD**, Rokito A. Frozen shoulder: a consensus definition. *J Shoulder Elbow Surg* 2011; **20**: 322-325 [PMID: 21051244 DOI: 10.1016/j.jse.2010.07.008]

12 **Grant JA**, Schroeder N, Miller BS, Carpenter JE. Comparison of manipulation and arthroscopic capsular release for adhesive capsulitis: a systematic review. *J Shoulder Elbow Surg* 2013; **22**: 1135-1145 [PMID: 23510748 DOI: 10.1016/j.jse.2013.01.010]

13 **Brue S**, Valentin A, Forssblad M, Werner S, Mikkelsen C, Cerulli G. Idiopathic adhesive capsulitis of the shoulder: a review. *Knee Surg Sports Traumatol Arthrosc* 2007; **15**: 1048-1054 [PMID: 17333122 DOI: 10.1007/s00167-007-0291-2]

14 **Robinson CM**, Seah KT, Chee YH, Hindle P, Murray IR. Frozen shoulder. *J Bone Joint Surg Br* 2012; **94**: 1-9 [PMID: 22219239 DOI: 10.1302/0301-620X.94B1.27093]

15 **Reeves B**. The natural history of the frozen shoulder syndrome. *Scand J Rheumatol* 1975; **4**: 193-196 [PMID: 1198072 DOI: 10.3109/03009747509165255]

16 **Schultheis A**, Reichwein F, Nebelung W. [Frozen shoulder. Diagnosis and therapy]. *Orthopade* 2008; **37**: 1065-1066, 1068-1072 [PMID: 18825364 DOI: 10.1007/s00132-008-1305-6]

17 **Harris G**, Bou-Haidar P, Harris C. Adhesive capsulitis: review of imaging and treatment. *J Med Imaging Radiat Oncol* 2013; **57**: 633-643 [PMID: 24283550 DOI: 10.1111/1754-9485.12111]

18 **White D**, Choi H, Peloquin C, Zhu Y, Zhang Y. Secular trend of adhesive capsulitis. *Arthritis Care Res* (Hoboken) 2011; **63**: 1571-1575 [PMID: 22034118 DOI: 10.1002/acr.20590]

19 **Lundberg BJ**. The frozen shoulder. Clinical and radiographical observations. The effect of manipulation under general anesthesia. Structure and glycosaminoglycan content of the joint capsule. Local bone metabolism. *Acta Orthop Scand Suppl* 1969; **119**: 1-59 [PMID: 4952729 DOI: 10.3109/ort.1969.40.suppl-119.01]

20 **Ozaki J**. Pathomechanics and operative management of chronic frozen shoulder. *Ann Chir Gynaecol* 1996; **85**: 156-158 [PMID: 8817053]

21 **Bunker TD**, Anthony PP. The pathology of frozen shoulder. A Dupuytren-like disease. *J Bone Joint Surg Br* 1995; **77**: 677-683 [PMID: 7559688]

22 **Rodeo SA**, Hannafin JA, Tom J, Warren RF, Wickiewicz TL. Immunolocalization of cytokines and their receptors in adhesive capsulitis of the shoulder. *J Orthop Res* 1997; **15**: 427-436 [PMID: 9246090 DOI: 10.1002/jor.1100150316]

23 **Lubis AM**, Lubis VK. Matrix metalloproteinase, tissue inhibitor of metalloproteinase and transforming growth factor-beta 1 in frozen shoulder, and their changes as response to intensive stretching and supervised neglect exercise. *J Orthop Sci* 2013; **18**: 519-527 [PMID: 23604641]

24 **Smith SP**, Devaraj VS, Bunker TD. The association between frozen shoulder and Dupuytren's disease. *J Shoulder Elbow Surg* 2001; **10**: 149-151 [PMID: 11307078 DOI: 10.1067/mse.2001.112883]

25 **Lho YM**, Ha E, Cho CH, Song KS, Min BW, Bae KC, Lee KJ, Hwang I, Park HB. Inflammatory cytokines are overexpressed in the subacromial bursa of frozen shoulder. *J Shoulder Elbow Surg* 2013; **22**: 666-672 [PMID: 22999851 DOI: 10.1016/j.jse.2012.06.014]

26 **Schydlowsky P**, Szkudlarek M, Madsen OR. Treatment of frozen shoulder with subcutaneous TNF-alpha blockade compared with local glucocorticoid injection: a randomised pilot study. *Clin Rheumatol* 2012; **31**: 1247-1251 [PMID: 22562389 DOI: 10.1007/s10067-012-1993-5]

27 **Doner G**, Guven Z, Atalay A, Celiker R. Evalution of Mulligan's technique for adhesive capsulitis of the shoulder. *J Rehabil Med* 2013; **45**: 87-91 [PMID: 23037929 DOI: 10.2340/16501977-1064]

28 **Russell S**, Jariwala A, Conlon R, Selfe J, Richards J, Walton M. A blinded, randomized, controlled trial assessing conservative management strategies for frozen shoulder. *J Shoulder Elbow Surg* 2014; **23**: 500-507 [PMID: 24630545 DOI: 10.1016/j.jse.2013.12.026]

29 **Carette S**, Moffet H, Tardif J, Bessette L, Morin F, Frémont P, Bykerk V, Thorne C, Bell M, Bensen W, Blanchette C. Intraarticular corticosteroids, supervised physiotherapy, or a combination of the two in the treatment of adhesive capsulitis of the shoulder: a placebo-controlled trial. *Arthritis Rheum* 2003; **48**: 829-838 [PMID: 12632439 DOI: 10.1002/art.10954]

30 **Ryans I**, Montgomery A, Galway R, Kernohan WG, McKane R. A randomized controlled trial of intra-articular triamcinolone and/or physiotherapy in shoulder capsulitis. *Rheumatology* (Oxford) 2005; **44**: 529-535 [PMID: 15657070 DOI: 10.1093/rheumatology/keh535]

31 **Kivimäki J**, Pohjolainen T, Malmivaara A, Kannisto M, Guillaume J, Seitsalo S, Nissinen M. Manipulation under anesthesia with home exercises versus home exercises alone in the treatment of frozen shoulder: a randomized, controlled trial with 125 patients. *J Shoulder Elbow Surg* 2007; **16**: 722-726 [PMID: 17931902 DOI: 10.1016/j.jse.2007.02.125]

32 **Beimers L**, Murell GAC. Arthroscopic capsular release for idiopathic adhesive capsulitis. *J Bone Joint Surg Am* 2013; **94**: 1208-1216 [DOI: 10.2106/JBJS.ST.L.00024]

33 **Buchbinder R**, Green S, Youd JM, Johnston RV, Cumpston M. Arthrographic distension for adhesive capsulitis (frozen shoulder). *Cochrane Database Syst Rev* 2008; **(1)**: CD007005 [PMID: 18254123 DOI: 10.1002/14651858]

34 **Ibrahim M**, Donatelli R, Hellman M, Echternach J. Efficacy of a static progressive stretch device as an adjunct to physical therapy in treating adhesive capsulitis of the shoulder: a prospective, randomised study. *Physiotherapy* 2014; **100**: 228-234 [PMID: 24211154 DOI: 10.1016/j.physio.2013.08.006]

35 **Maund E**, Craig D, Suekarran S, Neilson A, Wright K, Brealey S, Dennis L, Goodchild L, Hanchard N, Rangan A, Richardson G, Robertson J, McDaid C. Management of frozen shoulder: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2012; **16**: 1-264 [PMID: 22405512 DOI: 10.3310/hta16110]

36 **Rookmoneea M**, Dennis L, Brealey S, Rangan A, White B, McDaid C, Harden M. The effectiveness of interventions in the management of patients with primary frozen shoulder. *J Bone Joint Surg Br* 2010; **92**: 1267-1272 [PMID: 20798446 DOI: 10.1302/0301-620X.92B9.24282]

37 **van der Windt DA**, Koes BW, Devillé W, Boeke AJ, de Jong BA, Bouter LM. Effectiveness of corticosteroid injections versus physiotherapy for treatment of painful stiff shoulder in primary care: randomised trial. *BMJ* 1998; **317**: 1292-1296 [PMID: 9804720 DOI: 10.1136/bmj.317.7168.1292]

38 **Neviaser AS**, Hannafin JA. Adhesive capsulitis: a review of current treatment. *Am J Sports Med* 2010; **38**: 2346-2356 [PMID: 20110457 DOI: 10.1177/0363546509348048]

39 **Song A**, Higgins LD, Newman J, Jain NB. Glenohumeral corticosteroid injections in adhesive capsulitis: a systematic search and review. *PM R* 2014; **6**: 1143-1156 [PMID: 24998406 DOI: 10.1016/j.pmrj.2014.06.015]

40 **Shin SJ**, Lee SY. Efficacies of corticosteroid injection at different sites of the shoulder for the treatment of adhesive capsulitis. *J Shoulder Elbow Surg* 2013; **22**: 521-527 [PMID: 22999847 DOI: 10.1016/j.jse.2012.06.015]

41 **Blanchard V**, Barr S, Cerisola FL. The effectiveness of corticosteroid injections compared with physiotherapeutic interventions for adhesive capsulitis: a systematic review. *Physiotherapy* 2010; **96**: 95-107 [PMID: 20420956 DOI: 10.1016/j.physio.2009.09.003]

42 **Hanchard NC**, Goodchild L, Thompson J, O'Brien T, Davison D, Richardson C. Evidence-based clinical guidelines for the diagnosis, assessment and physiotherapy management of contracted (frozen) shoulder: quick reference summary. *Physiotherapy* 2012; **98**: 117-120 [PMID: 22507361 DOI: 10.1016/j.physio.2010.08.012]

43 **Kelley MJ**, Shaffer MA, Kuhn JE, Michener LA, Seitz AL, Uhl TL, Godges JJ, McClure PW. Shoulder pain and mobility deficits: adhesive capsulitis. *J Orthop Sports Phys Ther* 2013; **43**: A1-31 [PMID: 23636125 DOI: 10.2519/jospt.2013.0302]

44 **Vezeridis PS**, Goel DP, Shah AA, Sung SY, Warner JJ. Postarthroscopic arthrofibrosis of the shoulder. *Sports Med Arthrosc* 2010; **18**: 198-206 [PMID: 20711052]

45 **Jerosch J**, Nasef NM, Peters O, Mansour AM. Mid-term results following arthroscopic capsular release in patients with primary and secondary adhesive shoulder capsulitis. *Knee Surg Sports Traumatol Arthrosc* 2013; **21**: 1195-1202 [PMID: 22763569 DOI: 10.1007/s00167-012-2124-1]

46 **Ogilvie-Harris DJ**, Biggs DJ, Fitsialos DP, MacKay M. The resistant frozen shoulder. Manipulation versus arthroscopic release. *Clin Orthop Relat Res* 1995; **(319)**: 238-248 [PMID: 7554636]

47 **Thomas WJ**, Jenkins EF, Owen JM, Sangster MJ, Kirubanandan R, Beynon C, Woods DA. Treatment of frozen shoulder by manipulation under anaesthetic and injection: does the timing of treatment affect the outcome? *J Bone Joint Surg Br* 2011; **93**: 1377-1381 [PMID: 21969438 DOI: 10.1302/0301-620X.93B10.27224]

48 **Dodenhoff RM**, Levy O, Wilson A, Copeland SA. Manipulation under anesthesia for primary frozen shoulder: effect on early recovery and return to activity. *J Shoulder Elbow Surg* 2000; **9**: 23-26 [PMID: 10717858 DOI: 10.1016/S1058-2746(00)90005-3]

49 **Farrell CM**, Sperling JW, Cofield RH. Manipulation for frozen shoulder: long-term results. *J Shoulder Elbow Surg* 2005; **14**: 480-484 [PMID: 16194738 DOI: 10.1016/j.jse.2005.02.012]

50 **Vastamäki H**, Vastamäki M. Motion and pain relief remain 23 years after manipulation under anesthesia for frozen shoulder. *Clin Orthop Relat Res* 2013; **471**: 1245-1250 [PMID: 22907476 DOI: 10.1007/s11999-012-2542-x]

51 **Tasto JP**, Elias DW. Adhesive capsulitis. *Sports Med Arthrosc* 2007; **15**: 216-221 [PMID: 18004221 DOI: 10.1097/JSA.0b013e3181595c22]

52 **Conti V**. Arthroscopy in rehabilitation. *Orthop Clin North Am* 1979; **10**: 709-711 [PMID: 460843]

53 **Warner JJ**, Allen A, Marks PH, Wong P. Arthroscopic release for chronic, refractory adhesive capsulitis of the shoulder. *J Bone Joint Surg Am* 1996; **78**: 1808-1816 [PMID: 8986657]

54 **Le Lievre HM**, Murrell GA. Long-term outcomes after arthroscopic capsular release for idiopathic adhesive capsulitis. *J Bone Joint Surg Am* 2012; **94**: 1208-1216 [PMID: 22760389 DOI: 10.2106/JBJS.J.00952]

55 **Smith CD**, Hamer P, Bunker TD. Arthroscopic capsular release for idiopathic frozen shoulder with intra-articular injection and a controlled manipulation. *Ann R Coll Surg Engl* 2014; **96**: 55-60 [PMID: 24417832 DOI: 10.1308/003588414X13824511650452]

56 **Pearsall AW**, Osbahr DC, Speer KP. An arthroscopic technique for treating patients with frozen shoulder. *Arthroscopy* 1999; **15**: 2-11 [PMID: 10024027 DOI: 10.1053/ar.1999.v15.0150002]

57 **Kim YS**, Lee HJ, Park IJ. Clinical outcomes do not support arthroscopic posterior capsular release in addition to anterior release for shoulder stiffness: a randomized controlled study. *Am J Sports Med* 2014; **42**: 1143-1149 [PMID: 24585363 DOI: 10.1177/0363546514523720]

58 **Amir-Us-Saqlain H**, Zubairi A, Taufiq I. Functional outcome of frozen shoulder after manipulation under anaesthesia. *J Pak Med Assoc* 2007; **57**: 181-185 [PMID: 17489525]

59 **Anil Kumar PG**, Jacob MB, Newton J, Stewart MPM. Transient brachial plexus palsy following manipulation and local anaesthetic infiltration of a 'primary frozen shoulder'. *CME Orthopaedics* 2007; **4**: 26-27

60 **Loew M**, Heichel TO, Lehner B. Intraarticular lesions in primary frozen shoulder after manipulation under general anesthesia. *J Shoulder Elbow Surg* 2005; **14**: 16-21 [PMID: 15723009 DOI: 10.1016/j.jse.2004.04.004]

61 **Jacobs LG**, Smith MG, Khan SA, Smith K, Joshi M. Manipulation or intra-articular steroids in the management of adhesive capsulitis of the shoulder? A prospective randomized trial. *J Shoulder Elbow Surg* 2009; **18**: 348-353 [PMID: 19393928]

62 **Jerosch J**, Aldawoudy AM. Chondrolysis of the glenohumeral joint following arthroscopic capsular release for adhesive capsulitis: a case report. *Knee Surg Sports Traumatol Arthrosc* 2007; **15**: 292-294 [PMID: 16799827 DOI: 10.1007/s00167-006-0112-z]

P- Reviewer: Fenichel I, Rosenberg N **S- Editor**: Ji FF **L- Editor**: A **E- Editor**: Lu YJ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i2.89

World J Meta-Anal 2015 April 26; 3(2): 89-92
ISSN 2308-3840 (online)

EDITORIAL

Unsaturation index and type 2 diabetes: Unknown, unloved

Rob NM Weijers

Rob NM Weijers, Teaching Hospital, Onze Lieve Vrouwe Gasthuis, 1090 HA Amsterdam, The Netherlands

Author contributions: Weijers RNM solely contributed to this paper.

Conflict-of-interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

Correspondence to: Rob NM Weijers, PhD, Member of Teaching Hospital, Onze Lieve Vrouwe Gasthuis, Oosterparkstraat 9, PO Box 95500, 1090 HA Amsterdam, The Netherlands. robw01@xs4all.nl
Telephone: +31-30-6362637

Received: December 14, 2014
Peer-review started: December 17, 2014
First decision: January 20, 2015
Revised: January 28, 2015
Accepted: April 1, 2015
Article in press: April 7, 2015
Published online: April 26, 2015

Abstract

A useful parameter for interpreting analyses of membrane fatty-acid composition is the unsaturation index (UI), a measure of unsaturation that is calculated as the mean number of cis double bonds per fatty-acid residue multiplied by 100. The UI is a fundamental parameter that contains information about many membrane biophysical properties and behavior. UI is a crucial index for type 2 diabetes (T2D) and other disorders, yet it is not properly considered in the scientific community. The goal of the present editorial is to familiarize the scientific T2D community with the UI. The idea of early systemic cell-membrane disease necessitates new thinking and suggests that UI should feature prominently on the research agenda.

Key words: Type 2 diabetes; Unsaturation index; Phospholipid; Cell membrane; Fatty acid

Core tip: A useful parameter for interpreting analyses of membrane fatty-acid composition is the unsaturation index (UI), a measure of unsaturation that is calculated as the mean number of *cis* double bonds per fatty-acid residue multiplied by 100. The UI is a fundamental parameter that contains information about many membrane biophysical properties and behaviour. UI is a crucial index for type 2 diabetes (T2D) and other disorders, yet it is not properly considered in the scientific community. The goal of the present editorial is to familiarize the scientific T2D community with the UI.

Weijers RNM. Unsaturation index and type 2 diabetes: Unknown, unloved. *World J Meta-Anal* 2015; 3(2): 89-92 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i2/89.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i2.89

INTRODUCTION

A useful parameter for interpreting analyses of membrane fatty-acid composition is the unsaturation index (UI), a measure of unsaturation that is calculated as the mean number of *cis* double bonds per fatty-acid residue multiplied by 100[1]. This parameter characterizes a phospholipid bilayer and describes the fluidity, or flexibility, of a biological membrane. As the UI increases, so does the distance between plasma-membrane fatty-acyl chains, decreasing their mutual attraction energy and thus increasing membrane flexibility, which promotes an increase in the number of all functional Class I glucose transporters per

Table 1 Unsaturation index of erythrocyte membrane fatty-acid composition of controls and type 2 diabetes patients without diabetic retinopathy

Fatty acids	Controls (n = 18)		T2D patients without diabetic retinopathy (n = 14)	
	% of total fatty acids[1]	× number of double bonds	% of total fatty acids[1]	× number of double bonds
C14:0	0.43	-	0.48	-
C15:0	0.17	-	0.22	-
C16:0	22.41	-	23.75	-
C17:0	0.38	-	0.44	-
C18:0	17.22	-	17.72	-
C20:0	0.15	-	0.18	-
C22:0	0.33	-	0.42	-
C24:0	0.74	-	0.78	-
C16:1	0.73	0.73	1.03	1.03
C18:1 (trans)	0.22	-	0.2	-
C18:1	17.16	17.16	19.15	19.15
C20:1 n-9	0.21	0.21	0.26	0.26
C20:3 n-9	0.19	0.57	0.2	0.6
C22:1 n-9	0.06	0.06	0.09	0.09
C24:1 n-9	0.69	0.69	1.15	1.15
C18:2 n-6	12.87	25.74	10.58	21.16
C18:3 n-6	0.09	0.27	0.12	0.36
C20:2 n-6	0.2	0.4	0.23	0.46
C20:3 n-6	1.3	3.9	1.5	4.5
C20:4 n-6	13.04	52.16	11.33	45.32
C22:4 n-6	2.0	8	2.01	8.04
C22:5 n-6	0.35	1.75	0.3	1.5
C18:3 n-3	0.22	0.66	0.21	0.63
C20:5 n-3	0.97	4.85	1.03	5.15
C22:5 n-3	2.23	11.15	1.56	7.8
C22:6 n-3	4.51	27.06	2.85	17.1
Unsaturation index[2]		155.36		134.3

[1]Data published by Koehrer *et al*[5]; [2]The unsaturation index was calculated as the mean number of double bonds per fatty acid residue multiplied by 100.

membrane surface area[2]. At the most basic level, the basal metabolic rate of a cell is directly linked to its cell membrane acyl composition, and thus to its UI[3]. To date, this relationship has not received due attention in the treatment for type 2 diabetes (T2D).

The UI is a fundamental parameter that contains information about many membrane biophysical properties and behavior. Arachidonic acid and docosahexaenoic acid are key fatty acids; a minimal increase in the percentage of arachidonic acid in phospholipid tails improves membrane flexibility due to its four double bonds. A similar effect is seen for docosahexaenoic acid, with its six unsaturated bonds. UI is a crucial index for T2D and other disorders, yet it is not properly considered in the scientific community[4]. The goal of the present editorial is to familiarize the scientific T2D community with the UI.

In the September issue of *PLoS ONE*, Koehrer *et al*[5] reported the erythrocyte phospholipid and polyunsaturated fatty-acid composition in diabetic retinopathy. Several points in this article require additional clarification. Given that the study consisted nearly exclusively T2D patients, the reported observations are likely to be restricted to this type of diabetes. In contrast to one previous publication, Koehrer *et al*[5] presented measurements of total phospholipids from red blood cells, with fatty-acid composition specified for a total of 26 fatty acids. Based on the presented data, we calculated the UIs of membrane phospholipids from control subjects, T2D patients with and without retinopathy, and patients with gestational diabetes mellitus[6]. For example, Table 1 describes the calculation of the UIs for controls and T2D patients without retinopathy included in the study of Koehrer *et al*[5].

The UIs based on the erythrocyte membrane fatty-acid compositions reported in these studies[5,6] yielded novel information (Table 2). First, although phosphatidylcholine and phosphatidylethanolamine comprise about 60% of the total phospholipid in the bilayer membrane of human erythrocytes, the red cell phosphatidylcholine and phosphatidylethanolamine UI of subjects with normal glucose tolerance in the gestational diabetes mellitus study[6] are in line with the total phospholipid UI of the reference population in the diabetic retinopathy study[5] (162.8 and 155.4, respectively; Δ = 4.5%). Second, the decrease in the UI of phosphatidylcholine and phosphatidylethanolamine for gestational diabetes mellitus patients relative to controls was substantially higher than the total phospholipid UI decrease for T2D individuals without diabetic retinopathy compared with controls (16.3% and 13.5%, respectively; Δ = 17.2%), due to two underlying phenomena, *i.e.,* a temporary gestational and a chronic prediabetic increase in plasma FFA[7]. Third, the total phospholipid UI was substantially lower in T2D individuals than in healthy controls (134.3 and 155.4, respectively; Δ = 13.5%). Finally, the mean total phospholipid UI was substantially lower in T2D individuals with mild, moderate, and severe diabetic retinopathy than in T2D individuals without diabetic retinopathy (123.4 and 134.3, respectively; Δ = 8.1%). These experimental outcomes indicate that membrane flexibility plays an important role in microvascular complications of T2D. Further, these data support our working hypothesis: a gradual elevation of the plasma levels of saturated and monounsaturated free fatty acids causes a decrease in the number of polyunsaturated fatty-acyl chains in membrane phospholipids[2], a classical principle of membrane biogenesis[3,8]. In this context, it is noteworthy that our working hypothesis predicts that the transition from a healthy condition to a state with T2D will be matched by a decrease in UI, as will the transition from T2D without diabetic retinopathy to T2D with retinopathy[2].

In a study of the relationship between insulin sensitivity and the fatty-acid composition of skeletal-muscle phospholipids, Borkman *et al*[9] concluded that reduced levels of unsaturated fatty acids in the membrane may be due to a net reduction in the action of insulin, as a consequence of either insulin resistance or insulin deficiency or, alternatively, as a consequence

Table 2 Calculated unsaturation indices based on erythrocyte fatty-acid compositions reported by several studies

Disease	Participants (*n*)		Erythrocyte membrane	Unsaturation index[1]		Decrease (%)	Ref.
	Controls	Disease		Controls	Diabetic subjects		
T2D/T1D	18	13/1	Total phospholipid	155.4	134.3	13.5	[5]
T2D/T1D + mild DR	18	11/1	Total phospholipid	155.4	125.9	19.0	[5]
T2D/T1D + moderate DR	18	11/1	Total phospholipid	155.4	119.5	23.1	[5]
T2D/T1D + severe DR	18	19/3	Total phospholipid	155.4	124.7	19.7	[5]
T2D/T1D + proliferative DR	18	17/7	Total phospholipid	155.4	136.9	11.9	[5]
Gestational diabetes	61	53	PC + PE	162.8	137.1	16.3	[6]

[1]The unsaturation index was calculated as the mean number of double bonds per fatty acid residue multiplied by 100[4]. DR: Diabetic retinopathy; PC: Phosphatidylcholine; PE: Phosphatidylethanolamine.

of hyperinsulinaemia. This interpretation seems unlikely for the following reasons: first, gestational diabetes mellitus is a marker of a prediabetic phase characterized by time-dependent increase in insulin levels[10,11], where the T2D phase is marked by a time-dependent decrease in insulin levels[11]. Second, we demonstrated that patients in both phases were associated with lower UIs than were healthy controls, which suggests that insulin levels do not have an important causative role in lowering the UI. We hypothesize that a gradual increase in plasma free fatty-acid concentration during the prediabetic phase and after overt T2D[12,13] decreases the UI[7].

A well-known characteristic of the euglycaemic hyperinsulinaemic clamp is its wide inter-subject variability in insulin sensitivity. In a study of metabolic effects of lacidipine: a placebo-controlled study using the euglycaemic hyperinsulinaemic clamp, Morris *et al*[14] reported that even amongst non-diabetic subjects who were homogeneous for age, sex and body weight there was a wide inter-subject variability in insulin sensitivity, *i.e.*, 5.6-16.2 mg/kg per minute where the intra-subject variability in insulin sensitivity on the two placebo study days was 9%. Since physical activity and caloric intake are individual entities, which significantly affect a persons' free fatty acid concentration, we suggest that the wide inter-subject variability may be attributable to the inter-subject variability in free fatty acid concentration, and thus in the individual UI[13].

Despite extensive guidelines for managing T2D, in the United States during the years 2005-2008, 28.5% of adults with diabetes aged 40 years or older had diabetic retinopathy and 4.4% had advanced diabetic retinopathy[15]. These incidences are probably due to a longstanding period of decreased UI, increasing the stiffness of both the erythrocyte and plasma membranes and, as a consequence, decreasing microcirculatory flow, ultimately leading to chronic tissue hypoxia, insufficient tissue nutrition, and diabetes-specific microvascular pathology[2]. Thus, the idea of early systemic cell-membrane disease necessitates new thinking and suggests that UI should feature prominently on the research agenda.

REFERENCES

1 **Baur LA**, O'Connor J, Pan DA, Storlien LH. Relationships between maternal risk of insulin resistance and the child's muscle membrane fatty acid composition. *Diabetes* 1999; **48**: 112-116 [PMID: 9892230 DOI: 10.2337/diabetes.48.1.112]

2 **Weijers RN**. Lipid composition of cell membranes and its relevance in type 2 diabetes mellitus. *Curr Diabetes Rev* 2012; **8**: 390-400 [PMID: 22698081 DOI: 10.2174/157339912802083531]

3 **Hulbert AJ**. Life, death and membrane bilayers. *J Exp Biol* 2003; **206**: 2303-2311 [PMID: 12796449 DOI: 10.1242/jeb.00399]

4 **Chatgilialogou A**. U.I. arm in arm with Diabetes. Remembrane Blog. Available from: URL: http: //www.remembrane.com/en/n//u-i-arm-in-arm-with-diabetes/

5 **Koehrer P**, Saab S, Berdeaux O, Isaïco R, Grégoire S, Cabaret S, Bron AM, Creuzot-Garcher CP, Bretillon L, Acar N. Erythrocyte phospholipid and polyunsaturated fatty acid composition in diabetic retinopathy. *PLoS One* 2014; **9**: e106912 [PMID: 25188352 DOI: 10.1371/journal.pone.0106912]

6 **Min Y**, Ghebremeskel K, Lowy C, Thomas B, Crawford MA. Adverse effect of obesity on red cell membrane arachidonic and docosahexaenoic acids in gestational diabetes. *Diabetologia* 2004; **47**: 75-81 [PMID: 14634727 DOI: 10.1007/s00125-003-1275-5]

7 **Weijers RNM**. Membrane flexibility and cellular energy management in type 2 diabetes, gestational diabetes, and obesity. *EMJ Diabet* 2014; **2**: 65-72 [DOI: 10.13140/2.1.1027.6803]

8 **Gramling C**. Geochemistry. Low oxygen stifled animals' emergence, study says. *Science* 2014; **346**: 537 [PMID: 25359946 DOI: 10.1126/science.346.6209.537]

9 **Borkman M**, Storlien LH, Pan DA, Jenkins AB, Chisholm DJ, Campbell LV. The relation between insulin sensitivity and the fatty-acid composition of skeletal-muscle phospholipids. *N Engl J Med* 1993; **328**: 238-244 [PMID: 8418404 DOI: 10.1056/NEJM199301283280404]

10 **Weijers RN**, Bekedam DJ, Smulders YM. Determinants of mild gestational hyperglycemia and gestational diabetes mellitus in a large dutch multiethnic cohort. *Diabetes Care* 2002; **25**: 72-77 [PMID: 11772904 DOI: 10.2337/diacare.25.1.72]

11 **Yki-Järvinen H**. Pathogenesis of non-insulin-dependent diabetes mellitus. *Lancet* 1994; **343**: 91-95 [PMID: 7903784 DOI: 10.1016/S0140-6736(94)90821-4]

12 **Laws A**, Hoen HM, Selby JV, Saad MF, Haffner SM, Howard BV. Differences in insulin suppression of free fatty acid levels by gender and glucose tolerance status. Relation to plasma triglyceride and apolipoprotein B concentrations. Insulin Resistance Atherosclerosis Study (IRAS) Investigators. *Arterioscler Thromb Vasc Biol* 1997; **17**: 64-71 [PMID: 9012639 DOI: 10.1161/01.ATV.17.1.64]

13 **Reaven GM**, Hollenbeck C, Jeng CY, Wu MS, Chen YD. Measurement of plasma glucose, free fatty acid, lactate, and insulin for 24 h in patients with NIDDM. *Diabetes* 1988; **37**: 1020-1024 [PMID: 3292322 DOI: 10.2337/diab.37.8.1020]

14 **Morris AD**, Donnelly R, Connell JM, Reid JL. Metabolic effects of lacidipine: a placebo-controlled study using the euglycaemic hyper-

insulinaemic clamp. *Br J Clin Pharmacol* 1993; **35**: 40-45 [PMID: 8448066]

15 **Centers for disease control and prevention**. National diabetes statistics report; estimates of diabetes and its burden in the United States. Available from: URL: http: //www.cdc.gov/diabetes/pubs/estimates14.htm

P- Reviewer: Biondi-Zoccai G, Brigo F, Loffredo L **S- Editor**: Ma YJ
L- Editor: A **E- Editor**: Lu YJ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i2.93

World J Meta-Anal 2015 April 26; 3(2): 93-96
ISSN 2308-3840 (online)

EDITORIAL

Low back pain-related meta-analysis: Caution is needed when interpreting published research results

Christophe Demoulin, Olivier Bruyère, Pierre-René Somville, Marc Vanderthommen

Christophe Demoulin, Olivier Bruyère, Marc Vanderthommen, Department of Sport and Rehabilitation Sciences, University of Liege, 4000 Liege, Belgium
Christophe Demoulin, Pierre-René Somville, Marc Vanderthommen, Department of Physical Medicine and Rehabilitation, University Hospital Center of Liege (CHU), 4000 Liege, Belgium
Olivier Bruyère, Department of Public Health, Epidemiology and Health Economics, University of Liege, 4000 Liege, Belgium
Author contributions: All the authors contributed substantially to the writing and revising the manuscript.
Conflict-of-interest: The authors declare that there are no conflicts of interest.

Correspondence to: Christophe Demoulin, PT, PhD, Department of Sport and Rehabilitation Sciences, University of Liege, ISEPK, Bat B21, Allée des Sports, 4000 Liege, Belgium. christophe.demoulin@ulg.ac.be
Telephone: +32-43-663895
Fax: +32-43-662901
Received: January 20, 2015
Peer-review started: January 24, 2015
First decision: February 7, 2015
Revised: February 21, 2015
Accepted: April 1, 2015
Article in press: April 7, 2015
Published online: April 26, 2015

Abstract

The systematic reviews (SRs) including a meta-analysis are considered as the top level of evidence. Although the existence of more than a hundred of low back pain (LBP)-related SRs seems very appealing and might therefore suggest significant evidence on the topic, a deep analysis indicates that several of these SRs included only very few studies. Other SRs raise concerns because they included some randomized controlled trials which had a low methodological quality, or some studies which differed significantly regarding the studied populations and/or the experimental procedure. The sometimes controversial results of different SRs conducted on the same topic also highlight the significant influence of the inclusion/exclusion criteria used in the SRs to select the articles. To conclude, although meta-analysis is at the top of the evidence pyramid and have several strengths, the conclusions drawn from SRs should always be interpreted with caution because they can also have weaknesses. This is true, whether it be for LBP-related SRs including a meta-analysis, or any other. Therefore a critical analysis of any SR is always needed before integrating the results of the SR in its own clinical practice. Furthermore, clinical reasoning remains crucial, especially to consider the potential differences between one's patient and the patients included in the meta-analysis.

Key words: Meta-analysis; Systematic review; Spine; Back pain; Limitations; Recommendations; Evidence-based practice

Core tip: Although meta-analysis are at the top of the evidence pyramid and have several strengths, the conclusions drawn from systematic reviews combined to a meta-analysis should always be interpreted with caution because they can also have weaknesses. This is true, whether it be for low back pain-related systematic reviews including a meta-analysis, or any other. Therefore, a critical analysis of a systematic review is always needed before integrating the results in its own clinical practice. Furthermore, clinical reasoning remains crucial, especially to consider the potential differences between one's patient and the patients included in the

meta-analysis.

Demoulin C, Bruyère O, Somville PR, Vanderthommen M. Low back pain-related meta-analysis: Caution is needed when interpreting published research results. *World J Meta-Anal* 2015; 3(2): 93-96 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i2/93.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i2.93

INTRODUCTION

Nowadays, the relevance and necessity to treat patients according to the evidence-based medicine is widely recognized[1]. This approach aims to integrate clinical expertise, patient values and the best research evidence[1,2]. Regarding the latter component, it has become impossible for clinicians, and even scientists, to read all the papers on a specific topic considering the constant increasing number of scientific studies conducted every year[3]. Therefore, narrative and systematic reviews (SRs) have become very popular "easy access" methods for clinicians and researchers to help them to overview the scientific literature[1]. In contrast to the narrative reviews, which are rather qualitative summaries based on the knowledge of an experienced author, the SRs are performed according to a complex but very transparent process of search[4]. Furthermore, its combination with a meta-analysis provides a statistical summary of the evidence (*e.g.,* treatment effect, diagnostic method, prognosis, *etc.*) regarding a given topic, thereby facilitating the integration of the best evidence into practice. Meta-analysis also allow increasing the power, improving precision, and analyzing the consistency of effects[4,5]. Because the systematic reviews with a meta-analysis are considered as the top level of evidence, their results can greatly influence the guidelines and decisions related to a specific topic.

The number of meta-analysis published in the scientific literature is increasing in an exponential way[6]; the low back pain (LBP) field is no exception. This is well illustrated by a quick search of Pubmed with the generic terms "low back pain" and "meta-analysis" at the beginning of January 2015 which resulted in 377 hits, among which half of them (*n* = 166) were published over the last five years. An analysis of the hits reveals that two thirds of these 166 hits were effectively SRs combined to a meta-analysis related to LBP. The main topics concerned surgery, exercises, injections, pharmacological treatments, risk factors, spinal manual/manipulative therapies and imaging. Most of them were meta-analyses of randomized controlled trials (RCTs). There were also an overview of systematic reviews[7] and a SR of systematic reviews[8] reflecting the expansion of the SRs.

Although the existence of more than a hundred of LBP-related SRs seems very appealing and might therefore suggest significant evidence on the topic, a deep analysis indicates that several of these SRs combined to a meta-analysis included only very few studies (because of the low number of studies on the topic and/or of the low methodological quality of several studies) (*e.g.,*[9,10]). In some SRs, a meta-analysis was planned but was not conducted due to a lack of articles on the topic (*e.g.,*[11]) or a high clinical heterogeneity between studies regarding the intervention (*e.g.,*[12,13]) or the functional outcome measures (*e.g.,*[14]). The choice to conduct or not a meta-analysis appears subjective and differed between authors (*e.g.,* Hansen *et al*[11] decided to perform it only if the literature search resulted in at least 5 RCTs meeting the predefined inclusion criteria and if they were homogeneous whereas other performed a meta-analysis based only on three studies (*e.g.,*[10]).

Other SRs raise concerns because they included some RCTs which had a low methodological quality (*e.g.,* 5 out of the 7 RCTs included in the review of Ebadi *et al*[15] on the effectiveness of therapeutic ultrasound for chronic LBP had a score < 6/12 when using the 12 criteria recommended by the Cochrane Back Review Group) or some studies which differed significantly regarding the experimental procedure (*e.g.,* treatment provided). The pooling of spinal mobilizations and spinal manipulations[16], which are two different techniques regarding the indications, effects, *etc.*, or of a specific technique (manipulation) and a manual therapy concept (combination of techniques)[17], are good examples of mixing heterogeneous studies[18]. Another example of not ideal pooling is taken from meta-analysis regarding the effectiveness of "exercises" for LBP. Indeed, some conducted a calculation of a summary estimate[19] although there exists so many different types of exercise and although some parameters might influence the treatment outcomes[20] (*e.g.,* the number of sessions[21]). Another review on the topic tried to distinguish the different types of exercises[22], but the pooling was not always relevant[23]. Thus, reviewing the effectiveness of some LBP-related treatments (*e.g.,* physical therapy) is much more complex than for other treatments (*e.g.,* drug therapy) which are less heterogeneous between studies.

The characteristics of the participants are other crucial parameters to consider when conducting and interpreting a SR related to LBP. Indeed, in contrast to specific diseases (*e.g.,* influenza), "non-specific LBP" has no identifiable cause and is rather a symptom. Furthermore, according to the bio-psycho-social model and the literature on the topic, numerous (individual, psychosocial, work-related) factors influence the outcome/prognosis of musculoskeletal pain and should be taken into account when treating a patient, especially in case of chronic pain[24,25]. Therefore, lots of subgroups of patients with LBP have been described in literature[26]. A meta-analysis about the effectiveness of classification-based interventions reveals that such interventions

seem more effective for reducing pain and disability than "standard" treatments. This highlights the potential bias when pooling studies with different populations. To solve this problem, a subgroup analysis can be conducted to study the influence of several parameters. However, one has to keep in mind that an effect of chance can occur when analyzing too many subgroups[27]; moreover, a low number of studies included in the SRs prevents to achieve such an analysis.

The inclusion/exclusion criteria used in the SRs to select the articles are also important to consider when interpreting the results of a meta-analysis because they can have a significant influence[6]. Indeed, a SR of the SRs conducted on the effectiveness of the Pilates exercises in patients with chronic LBP revealed that the 5 SRs which had been published on the topic by that time had different conclusions although they had similar research objectives[8]; furthermore, only two out of the ten primary studies were included in the 5 papers[8]. Besides, some SRs considered only quantitative results from questionnaires while ignoring qualitative studies which might therefore introduced also a bias[28].

Another point to be noticed about the LBP-related meta-analyses is that most of them have been conducted on aggregate-level data whereas only a few have been performed on individual data (*e.g.,*[29]), although the latter analysis appears very relevant[30].

To conclude, although meta-analyses are at the top of the evidence pyramid and have several strengths, the conclusions drawn from SRs should always be interpreted with caution because they can also have weaknesses. This is true, whether it be for LBP-related SRs including a meta-analysis, or any other. Interpreting results of a meta-analysis is not easy, as evidenced by the letters to the editor[18,31] related to the SR of Licciardone *et al*[17]. Using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis checklist can help readers for critical appraisal purposes. Besides, analyzing the methodology of the SR, examining the studies included (*e.g.,* population, intervention, outcome) and thinking about the possible sources of heterogeneity of the results (in case it occurs) are really necessary before integrating the results of the SR in its own clinical practice[2,6,27]. Furthermore, clinical reasoning remains crucial[32], especially to consider the potential differences between one's patient and the patients included in the meta-analysis[27]. Only reading the abstract of a SR combined to a meta-analysis is clearly not good enough to do so.

REFERENCES

1 **Mellis C**. Evidence-based medicine: what has happened in the past 50 years? *J Paediatr Child Health* 2015; **51**: 65-68 [PMID: 25536873 DOI: 10.1111/jpc.12800]

2 **Manchikanti L**, Boswell MV, Giordano J. Evidence-based interventional pain management: principles, problems, potential and applications. *Pain Physician* 2007; **10**: 329-356 [PMID: 17387356]

3 **Fraser AG**, Dunstan FD. On the impossibility of being expert. *BMJ* 2010; **341**: c6815 [PMID: 21156739 DOI: 10.1136/bmj.c6815]

4 **Akobeng AK**. Understanding systematic reviews and meta-analysis. *Arch Dis Child* 2005; **90**: 845-848 [PMID: 16040886 DOI: 10.1136/adc.2004.058230]

5 **Ioannidis JP**, Patsopoulos NA, Rothstein HR. Reasons or excuses for avoiding meta-analysis in forest plots. *BMJ* 2008; **336**: 1413-1415 [PMID: 18566080 DOI: 10.1136/bmj.a117]

6 **Greco T**, Zangrillo A, Biondi-Zoccai G, Landoni G. Meta-analysis: pitfalls and hints. *Heart Lung Vessel* 2013; **5**: 219-225 [PMID: 24364016]

7 **Jacobs WC**, Rubinstein SM, Willems PC, Moojen WA, Pellisé F, Oner CF, Peul WC, van Tulder MW. The evidence on surgical interventions for low back disorders, an overview of systematic reviews. *Eur Spine J* 2013; **22**: 1936-1949 [PMID: 23681497 DOI: 10.1007/s00586-013-2823-4]

8 **Wells C**, Kolt GS, Marshall P, Hill B, Bialocerkowski A. Effectiveness of Pilates exercise in treating people with chronic low back pain: a systematic review of systematic reviews. *BMC Med Res Methodol* 2013; **13**: 7 [PMID: 23331384 DOI: 10.1186/1471-2288-13-7]

9 **Michaleff ZA**, Kamper SJ, Maher CG, Evans R, Broderick C, Henschke N. Low back pain in children and adolescents: a systematic review and meta-analysis evaluating the effectiveness of conservative interventions. *Eur Spine J* 2014; **23**: 2046-2058 [PMID: 25070788 DOI: 10.1007/s00586-014-3461-1]

10 **McGregor AH**, Probyn K, Cro S, Doré CJ, Burton AK, Balagué F, Pincus T, Fairbank J. Rehabilitation following surgery for lumbar spinal stenosis. A Cochrane review. *Spine* (Phila Pa 1976) 2014; **39**: 1044-1054 [PMID: 24732858 DOI: 10.1097/BRS.0000000000000355]

11 **Hansen H**, Manchikanti L, Simopoulos TT, Christo PJ, Gupta S, Smith HS, Hameed H, Cohen SP. A systematic evaluation of the therapeutic effectiveness of sacroiliac joint interventions. *Pain Physician* 2012; **15**: E247-E278 [PMID: 22622913]

12 **Slater SL**, Ford JJ, Richards MC, Taylor NF, Surkitt LD, Hahne AJ. The effectiveness of sub-group specific manual therapy for low back pain: a systematic review. *Man Ther* 2012; **17**: 201-212 [PMID: 22386046 DOI: 10.1016/j.math.2012.01.006]

13 **Surkitt LD**, Ford JJ, Hahne AJ, Pizzari T, McMeeken JM. Efficacy of directional preference management for low back pain: a systematic review. *Phys Ther* 2012; **92**: 652-665 [PMID: 22247407 DOI: 10.2522/ptj.20100251]

14 **Richards E**, van Kessel G, Virgara R, Harris P. Does antenatal physical therapy for pregnant women with low back pain or pelvic pain improve functional outcomes? A systematic review. *Acta Obstet Gynecol Scand* 2012; **91**: 1038-1045 [PMID: 22583125 DOI: 10.1111/j.1600-0412.2012.01462.x]

15 **Ebadi S**, Henschke N, Nakhostin Ansari N, Fallah E, van Tulder MW. Therapeutic ultrasound for chronic low-back pain. *Cochrane Database Syst Rev* 2014; **3**: CD009169 [PMID: 24627326 DOI: 10.1002/14651858]

16 **Rubinstein SM**, Terwee CB, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for acute low back pain: an update of the cochrane review. *Spine* (Phila Pa 1976) 2013; **38**: E158-E177 [PMID: 23169072 DOI: 10.1097/BRS.0b013e31827dd89d]

17 **Licciardone JC**, Brimhall AK, King LN. Osteopathic manipulative treatment for low back pain: a systematic review and meta-analysis of randomized controlled trials. *BMC Musculoskelet Disord* 2005; **6**: 43 [PMID: 16080794 DOI: 10.1186/1471-2474-6-43]

18 **Franke H**. Why reservations remain: a critical reflection about the systematic review and meta-analysis "Osteopathic manipulative treatment for low back pain" by Licciardone et al. *J Bodyw Mov Ther* 2012; **16**: 411-415 [PMID: 23036874 DOI: 10.1016/j.jbmt.2012.05.002]

19 **Oesch P**, Kool J, Hagen KB, Bachmann S. Effectiveness of exercise on work disability in patients with non-acute non-specific low back pain: Systematic review and meta-analysis of randomised controlled trials. *J Rehabil Med* 2010; **42**: 193-205 [PMID: 20411212 DOI: 10.2340/16501977-0524]

20 **Hayden JA**, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Ann Intern Med* 2005; **142**: 776-785 [PMID: 15867410 DOI: 10.7326/0003-4819-142-9-200505030-00014]

21 **Ferreira ML**, Smeets RJ, Kamper SJ, Ferreira PH, Machado LA.

Can we explain heterogeneity among randomized clinical trials of exercise for chronic back pain? A meta-regression analysis of randomized controlled trials. *Phys Ther* 2010; **90**: 1383-1403 [PMID: 20671101 DOI: 10.2522/ptj.20090332]

22 **van Tulder M**, Malmivaara A, Esmail R, Koes B. Exercise therapy for low back pain: a systematic review within the framework of the cochrane collaboration back review group. *Spine* (Phila Pa 1976) 2000; **25**: 2784-2796 [PMID: 11064524]

23 **May S**. Re: Exercise therapy for low back pain. A systematic review within the framework of the Cochrane Collaboration Back review Group. *Spine* (Phila Pa 1976) 2001; **26**: 1829 [PMID: 11493862]

24 **Gatchel RJ**, Peng YB, Peters ML, Fuchs PN, Turk DC. The biopsychosocial approach to chronic pain: scientific advances and future directions. *Psychol Bull* 2007; **133**: 581-624 [PMID: 17592957 DOI: 10.1037/0033-2909.133.4.581]

25 **Fersum KV**, Dankaerts W, O'Sullivan PB, Maes J, Skouen JS, Bjordal JM, Kvåle A. Integration of subclassification strategies in randomised controlled clinical trials evaluating manual therapy treatment and exercise therapy for non-specific chronic low back pain: a systematic review. *Br J Sports Med* 2010; **44**: 1054-1062 [PMID: 19996331 DOI: 10.1136/bjsm.2009.063289]

26 **Kamper SJ**, Maher CG, Hancock MJ, Koes BW, Croft PR, Hay E. Treatment-based subgroups of low back pain: a guide to appraisal of research studies and a summary of current evidence. *Best Pract Res Clin Rheumatol* 2010; **24**: 181-191 [PMID: 20227640 DOI: 10.1016/j.berh.2009.11.003]

27 **Yuan Y**, Hunt RH. Systematic reviews: the good, the bad, and the ugly. *Am J Gastroenterol* 2009; **104**: 1086-1092 [PMID: 19417748 DOI: 10.1038/ajg.2009.118]

28 **Brox JI**. Current evidence on catastrophizing and fear avoidance beliefs in low back pain patients. *Spine J* 2014; **14**: 2679-2681 [PMID: 25441973 DOI: 10.1016/j.spinee.2014.08.454]

29 **Griffith LE**, Shannon HS, Wells RP, Walter SD, Cole DC, Côté P, Frank J, Hogg-Johnson S, Langlois LE. Individual participant data meta-analysis of mechanical workplace risk factors and low back pain. *Am J Public Health* 2012; **102**: 309-318 [PMID: 22390445 DOI: 10.2105/AJPH.2011.300343]

30 **Garg AX**, Hackam D, Tonelli M. Systematic review and meta-analysis: when one study is just not enough. *Clin J Am Soc Nephrol* 2008; **3**: 253-260 [PMID: 18178786 DOI: 10.2215/CJN.01430307]

31 **Licciardone JC**. Systematic review and meta-analysis conclusions relating to osteopathic manipulative treatment for low back pain remain valid and well accepted. *J Bodyw Mov Ther* 2013; **17**: 2-4 [PMID: 23294676 DOI: 10.1016/j.jbmt.2012.10.003]

32 **Sniderman AD,** LaChapelle KJ, Rachon NA, Furberg CD. The necessity for clinical reasoning in the era of evidence-based medicine. *Mayo Clin Proc* 2013; **88**: 1108-1114 [PMID: 24079680 DOI: 10.1016/j.mayocp.2013.07.012]

P- Reviewer: Nikolopoulos D, Recnik G **S- Editor**: Gong XM
L- Editor: A **E- Editor**: Lu YJ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i2.97

World J Meta-Anal 2015 April 26; 3(2): 97-117
ISSN 2308-3840 (online)

SYSTEMATIC REVIEWS

Meta-analysis of single strain probiotics for the eradication of *Helicobacter pylori* and prevention of adverse events

Lynne V McFarland, Peter Malfertheiner, Ying Huang, Lin Wang

Lynne V McFarland, Department of Medicinal Chemistry, University of Washington, Seattle, WA 98195, United States
Lynne V McFarland, Health Services Research and Development, VA Puget Sound Health Care System, Department of Veterans Affairs, Seattle, WA 98101, United States
Peter Malfertheiner, Universitatsklinikum Magdeburg AOR, 39120 Magdeburg, Germany
Ying Huang, Lin Wang, Children's Hospital of Fudan University, Shanghai 201102, China
Author contributions: McFarland LV, Malfertheiner P, Huang Y and Wang L reviewed, scored papers and were involved in the manuscript preparation; McFarland LV designed the study, did literature search and analyzed the data.
Conflict-of-interest: Lynne V McFarland has received fees as a speaker (Biocodex, France and Lallemand, France) and is on the scientific advisory board of BioK+, Canada. Peter Malfertheiner has received speaker fees from Biocodex and is on the scientific advisor board of Danone. Ying Huang and Lin Wang have no conflicts of interest. No authors are employed by or have stock or equity in any company manufacturing probiotics.
Data sharing: No additional data are available.

Correspondence to: Lynne V McFarland, PhD, Health Services Research and Development, VA Puget Sound Health Care System, Department of Veterans Affairs, Metropolitan Park West, 1100 Olive Way #1400, Seattle, WA 98101, United States. lvmcfarl@u.washington.edu
Telephone: +1-206-2771780
Fax: +1-206-7642935
Received: October 22, 2014
Peer-review started: October 24, 2014
First decision: December 26, 2014
Revised: January 3, 2015
Accepted: January 15, 2015
Article in press: January 19, 2015
Published online: April 26, 2015

Abstract

AIM: To assess the efficacy and safety of single strain probiotics for the: (1) eradication of *Helicobacter pylori* (*H. pylori*); (2) prevention of adverse events; and (3) prevention of antibiotic-associated diarrhea associated with eradication therapy.

METHODS: We searched PubMed (1960-2014), EMBASE (1974-2014), Cochrane Database of Systematic Reviews (1990-2014), and ISI Web of Science (2000-2014). Additionally, we conducted a grey literature search including contact with National Institutes of Health Clinical Trials Registry, abstracts from annual infectious disease and gastroenterology meetings, experts in the field and correspondence with authors. Randomized controlled trials of *H. pylori* positive adults or children treated with eradication therapy and assessing the adjunctive therapy with a single strain of probiotics were included. The primary outcomes were the rates of eradication of *H. pylori* and frequency of patients with adverse events or antibiotic-associated diarrhea. Outcomes were pooled using fixed or random-effects models to calculate the relative risk and corresponding 95%CI and weighted on study size. To explore possible explanations for heterogeneity, a priori subgroup analyses were conducted on daily probiotic dose, study population, and quality of the study. The overall quality of the evidence for each probiotic strain was assessed using the GRADE criteria.

RESULTS: A total of 25 randomized controlled trials (28 treatment arms, with a total of 3769 participants) assessed one of six single probiotic strains as adjunctive treatments to standard eradication therapy. Only one probiotic strain significantly improved *H. pylori* eradication rates: *Saccharomyces boulardii* (*S. boulardii*) CNCM I-745 [pooled relative risks (pRR) = 1.11, 95%CI: 1.07-1.16]. Only one probiotic strain (*S. boulardii* CNCM I-745) significantly prevented any adverse events (pRR = 0.42, 95%CI: 0.28-0.62). Both *S. boulardii* CNCM I-745 and *Lactobacillus rhamnosus* GG significantly

reduced antibiotic-associated diarrhea (pRR = 0.47, 95%CI: 0.37-0.60 and pRR = 0.29, 95%CI: 0.17-0.48, respectively) associated with *H. pylori* eradication therapy. Meta-regression of sub-groups did not detect significant differences by dose, adult *vs* pediatric, symptom status, or study quality, but did find significant differences by the strain of probiotic. Potential mild publication bias was found for antibiotic-associated diarrhea, but not for eradication or adverse event outcomes. Analysis of the study quality illuminated areas for improvement in future studies (use of placebos, study size calculations, attrition reasons and discussion of limitations and generalizability).

CONCLUSION: The pooled evidence suggests that the adjunctive use of a few probiotic strains may improve *H. pylori* eradication rates and prevent the development of adverse events and antibiotic-associated diarrhea in those treated with standard eradication therapies. The type of probiotic strain was the most important factor in predicting efficacy.

Key words: Probiotics; Safety; *Saccharomyces boulardii*; *Helicobacter pylori*; Meta-analysis; Adverse reactions; Diarrhea; *Lactobacillus rhamnosus*; Randomized clinical trials

Core tip: A meta-analysis was conducted (1960-2014) for randomized clinical trials testing single strained probiotics as an adjunct to standard *Helicobacter pylori* (*H. pylori*) eradication therapy. Of the single strains with multiple trials, only one significantly improved *H. pylori* eradication rates {*Saccharomyces boulardii* (*S. boulardii*) I-745 [pooled relative risks (pRR) = 1.11, 95%CI: 1.07-1.16)]}, while two strains significantly reduced the rate of antibiotic-associated diarrhea [*S. boulardii* I-745 (pRR = 0.47, 95%CI: 0.37-0.60) and *Lactobacillus rhamnosus* GG (pRR = 0.29, 95%CI: 0.17-0.48)]. None of the other four probiotic strains improved *H. pylori* therapy (*C. butyricum*, *L. acidophilus*, *L. reuteri*, *L. casei*).

McFarland LV, Malfertheiner P, Huang Y, Wang L. Meta-analysis of single strain probiotics for the eradication of *Helicobacter pylori* and prevention of adverse events. *World J Meta-Anal* 2015; 3(2): 97-117 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i2/97.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i2.97

INTRODUCTION

Helicobacter pylori (H. pylori) was first associated with chronic gastritis, duodenal and peptic ulcers by Marshall and Warren[1] in 1984. Surveillance studies since that time have found *H. pylori* colonization is a global concern with a prevalence ranging from 70%-90% in developing countries and 25%-50% in developed countries[2]. *H. pylori* is typically acquired during childhood from other humans and transmitted by the oral-oral or oral-fecal route or by ingestion of contaminated water. *H. pylori* infection in childhood may lead to chronic gastritis, but only 20% will develop clinical symptoms[2]. Prolonged carriage may result in an onset of symptoms in adults, which include dyspepsia, peptic or duodenal ulcers, gastric adenocarcinoma, B-cell lymphoma and rarely extragastric complications[3]. Current guidelines from the Maastricht Ⅳ consensus for the eradication of *H. pylori* include triple therapy [typically two antibiotics and a proton-pump inhibitor (PPI) for 7-14 d], with eradication rates ranging from 71% to 81%, sequential therapy (with slightly improved *H. pylori* eradication rates from 85% to 84%) and, more recently, bismuth-based quadruple therapy (with 90% efficacy)[4-7]. However, the common development of adverse events [such as antibiotic-associated diarrhea (AAD), nausea, *etc.*] from the eradication therapies cause many patients to prematurely discontinue their treatments, leading to plummeting eradication rates and the development of antibiotic resistance[8-10]. The development of antibiotic resistant strains of *H. pylori* varies by country and type of antibiotic exposure (ranging 11%-29% for clarithromycin, 17%-86% for metronidazole, levofloxacin 14%-24%)[11,12]. In addition, relapses of symptoms occur > 40% in patients within 32 wk after triple therapy eradication therapy[13]. Recently several alternative treatments, including probiotics, have been tested to improve eradication rates, prevent the development of antibiotic resistant strains and to prevent the development of adverse events[14].

Probiotics (defined as living microbes, given in adequate doses, with proven health effects) have been shown to be effective in many diseases and may be useful as an adjunct to eradication therapy. Probiotics are known to be effective for the prevention of side-effects of antibiotic use, typically antibiotic-associated diarrhea[15]. Several studies have also shown some probiotic strains [*Saccharomyces boulardii* (*S. boulardii*), *Lactobacillus acidophilus* (*L. acidophilus*) or mixtures of strains, *etc.*] have specific mechanisms of action against *H. pylori,* including inhibiting *H. pylori* attachment to mucosal cells[16-18], regulation of the immune response to *H. pylori*[19], or direct physiologic effects[20]. Probiotics may also restore the normal microbiota disrupted by antibiotic exposure (causing diarrhea or colitis) and thus prevent *H. pylori*-associated adverse events[21,22].

Choosing the appropriate probiotic can be challenging as the choice must be matched to both probiotic strain and the disease being treated (or prevented), based on the strength of evidence-based clinical trials. Different mechanisms of action are strain-specific, therefore it is necessary to analyze the efficacy by similar probiotic strains whenever possible[23-25]. Most meta-analyses of probiotics for *H. pylori* infections have not done this. Probiotics are also available as single strain products or in mixtures of two or more probiotic strains. This paper will focus only on single strains tested in at least two

randomized, controlled trials.

The aims of this meta-analysis are to analyze the effectiveness of adjunctive single strain probiotics for the: (1) eradication of *H. pylori*; (2) reduction of adverse events; and (3) reduction of antibiotic-associated diarrhea commonly linked with eradication therapy.

MATERIALS AND METHODS

Study objectives

Primary aims: (1) To systematically assess whether single strain probiotics (given as an adjunct with *H. pylori* eradication therapy) could improve the eradication rate of *H. pylori*; and to systematically assess whether probiotics could reduce the frequency of: (2) any types of adverse events; or (3) antibiotic-associated diarrhea associated with *H. pylori* eradication therapy.

Secondary aims: To systematically assess if differences in effect were associated with specific subgroups, defined by: daily dose effect of probiotics, type of study population (adult versus pediatric, asymptomatic versus symptomatic), study quality and strain of probiotic used.

Search strategy

As shown in Table 1, this meta-analysis followed PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analysis) statement guidelines[26] and guidelines using clearly delineated parameters, a *priori* inclusion and exclusion criteria and standardized data extraction tools[27,28]. We undertook systematic searches of PubMed (1960-2014), EMBASE (1974-2014), Cochrane Database of Systematic Reviews (1990-2014), ISI Web of Science (2000-2014) and three on-line clinical trial registries: Cochrane Central Register of Controlled trials (http://www.cochrane.org), MetaRegister of Controlled Trials (http:www.controlled-trials.com/mrct) and National Institutes of Health (http://www.clinicaltrials.gov). We used bibliographies of all relevant studies to do a recursive search. Additionally, we conducted an extensive grey literature search including abstracts from annual infectious disease and gastroenterology meetings, probiotic product websites, experts in the field and communication with published authors on *H. pylori* infections. Search terms included: *H. pylori*, randomized controlled trial and probiotics and specific probiotic strains. Search strategies were broad-based initially, then narrowed to the disease and population of interest. Abstracts of all citations and retrieved studies were reviewed and rated for inclusion. Full articles were retrieved if probiotics were given to treat *H. pylori* infections or carriage or to prevent adverse events associated with *H. pylori* eradication therapies.

Inclusion and exclusion criteria

Inclusion criteria included randomized (well described or partially) controlled trials (RCT), blinded or open trials, in pediatric or adult populations (inpatient or outpatients), published in peer-reviewed journals or on clinical trial websites, or as meeting abstracts. All participants were required to have received *H. pylori* eradication therapy (double, triple, quadruple or sequential therapy) that included at least one antibiotic and one PPI. Non-English language trials were translated and included whenever possible. Exclusion criteria included pre-clinical studies, safety, kinetic or formulation phase 2 studies, case reports or case series, duplicate reports, trials of unspecified types of probiotics, non-randomized trials, incomplete or no outcomes reported, or if translation could not be obtained. Trials which did not assess either *H. pylori* eradication rates or the incidence of adverse events were excluded. Probiotic strains with only one randomized controlled trial (lacking at least one other confirmatory trial) were also excluded. Randomized controlled trials testing probiotic products with a mixture of different probiotic strains were reviewed, but will be presented elsewhere.

Data extraction

Each article was reviewed and scored independently by at least two reviewers. One reviewer (LVM) screened all abstracts, extracted and scored all articles using pre-constructed and piloted, data extraction forms (see Figure 1). Each of three other reviewers (PM, YH, LW) independently extracted data and assessed risk of bias from one-third of the articles (each sent different articles). Any disagreements were resolved by a third reviewer. For articles published in abstract form only or for any missing significant data in full articles, further information was sought by contacting authors or by the company manufacturing the probiotic product. Using a standardized data extraction form, we systematically collected the following data: authors, year of publication and journal, population data (age range, setting, types of eradication therapy given), study aims and outcomes, study methods (study design, eligibility criteria, sample size calculations, interim analysis, statistical methods used, recruitment methods, subgroup analysis done), randomization (method of randomization allocation, randomization method), degree of blinding (open, single or double), intervention data (probiotic strains used, daily dose, duration of treatment, duration of follow-up, type of control used, treatment concealment), results (balanced randomization achieved, attrition rate and reasons, comparison of treatment groups by demographics, *etc.*, CONSORT flow-chart provided), outcome data [by group, intent-to-treat (ITT) or as-per-protocol (APP) analysis], safety data (adverse events reported by group), discussion points (limitations, generalizability and comparison of study results to published papers), clinical trial registration, location of protocol, and

Table 1 Preferred reporting items for systematic reviews and meta-analyses checklist 2009[26]

Item	Topic	Reported on page
Title		
1	Title includes systematic review or meta-analysis or both	97
Abstract		
2	Structured abstract/summary background, objectives, data sources, eligibility criteria, participant, interventions, appraisal and synthesis methods, results, limitation, conclusions and implication of key findings, systematic review registration number	97
Introduction		
3	Rationale for review, what is already known	98
4	Objectives: Specific questions addressed: (PICOS)-participants, interventions, comparisons, outcomes, study design	99
Methods		
5	If review protocol (location and accessed URL, registration number)	NA
6	Eligibility criteria (study characteristics (PICOS, follow-up, *etc*.) and report characteristics (years searched, language, publication status), provide rationale	99
7	Information sources (databases with dates of coverage, contact with study authors to identify additional studies, date last searched)	99
8	Search strategy: Full search strategy for at least one database, including any limits used, such that it could be repeated	99
9	Study selection: (process for screening, eligibility)	99
10	Data collection process: Method of data extraction (piloted forms, independently, in duplicate) and any processes of obtaining and confirming data from investigators)	99
11	Data items: List and define all variables sought (*e.g*., PICOS, funding sources, *etc*.) and any assumptions	99-102
12	Risk of bias in individual studies: (Describe methods used for assessing risk of bias (at study or outcome level), how this info is to be used in any data synthesis)	102
13	Summary Measures: State principal summary outcome measures (RR or Difference in means) for pooled estimates of risk	103
14	Synthesis of results: Describe method of handling data and pooling data (measures of consistency with I2 for each meta-analysis)	103
15	Risk for bias across studies: (publication bias)	103
16	Additional analysis: Any subgroup or sensitivity analysis, meta-regression and if pre-specified	103
Results		
17	Study selection: N of RCT screened, # assessed for eligibility, reasons for exclusions, with flow diagram	103-104, Figure 2
18	Study characteristics (for each study: study size, PICOS, follow-up with citations)	Table 4
19	Risk of bias within studies: Data on risk for bias and if there, any outcome level assessment (see #12, study quality)	107
20	Results of individual studies: Simple summary data for txt arm, effect estimates and confidence intervals for each study, with forest plot	Figures 3-5 Tables 4, 5
21	Synthesis of results: Data on each meta-analysis, pooled data, 95%CI and measures of consistency	Figures 3-5
22	Risk of bias across studies: results of any assessment of risk across studies (see #15)	Figures 6-8
23	Additional analysis data: if done (see #16, sub-groups)	107
Discussion		
24	Summary of evidence: Summarize main findings, strength of evidence for each main outcome. Relevance to key groups (providers, users, policy makers)	110
25	Limitations: Limitations at study level and outcome level (risk of bias), at review-level (incomplete retrieval of identified research, reporting bias)	113
26	Conclusions: General interpretation of results compared to other evidence, implications for future research.	113
Funding		
27	Funding: describe funding sources	Not found

The PRISMA Statement. Available from: URL: http://prisma-statement.org/statement.htm. Accessed 7/25/2014.

source of funding.

Interventions

Included trials had participants who were randomized to either an adjunctive probiotic group or a control group. The type of control group may have included either a placebo (blinded study) or no treatment (open study) in addition to the eradication therapy currently used as standard practice. The type of probiotic intervention included probiotics in any form (*e.g.*, capsule, sachet, tablets, drink, *etc.*) given in conjunction with the *H. pylori* eradication therapy. Trials investigating non-specific probiotics or yogurts [*e.g.*, articles not providing the probiotic strain(s) used] were excluded. Trials combining probiotics with prebiotics were included if the prebiotic dose was less than 2.5 g/d, as this was judged to be of limited impact to alter the intestinal microflora[29,30]. The most recent probiotic strain designations are presented in this study for those strains whose names have changed over time (older articles may have reported a different strain designation). The taxonomy of the probiotic strain type was confirmed by correspondence with authors or the manufacturing companies.

Reference:

First reviewer: _____ Second reviewer: _____

Study design (methodological)

____ 1. □ Randomized or Controlled Trial in title?

Introduction/aims

_____ 2. □ Background and rationale described: Yes/No

_____ 3. □ Aims given: 1° outcome(s): _____
2° outcome(s): _____

Study Population:

_____ 4. □ Setting (Inpatient or outpatient, number of sites, etc. or any of below):
Disease (condition) ____ PUD _____ Gastritis ____ Dyspepsia _____Mixed
or _____ asymptomatic carrier
Adult or pediatric, or mixed
Age range:
Country:

_____ 5. □ If recruitment/study stopped early (reason given?, na if not stopped early)

Methods

_____ 6. □ Prospective study design

_____ 7. □ Eligibility/exclusion criteria described

_____ 8. □ Sample size calculations given

_____ 9. □ Interim Analysis (yes/no or na, if not done)

_____ 10. □ Statistical methods described (yes/no)

_____ 11. □ Recruitment methods or population described, referred from hosp/clinic? (yes/no)

_____ 12. □ Subgroup analysis methods described a priori or na (if no-sub-group done)

Intervention:

_____ 13. □ Intervention well described (strain, dose, duration) (+1 if most done below)
Probiotic strain(s):
Daily dose (cfu/d):
Duration intervention period (txt time):
Duration follow-up (post-intervention):
Formulation (capsule, yogurt, milk/drink, sachet, tablet, other, not described)
Type of control (placebo, no placebo/eradication therapy only, other):

Hp eradication therapy given (double/triple/quadruple/sequential/none): Duration:

Randomization (selection bias)

_____ 14. □ Method to generate random numbers described (blocked, computer)

_____ 15. □ Balanced randomization allocation achieved (yes/no)
Probiotic group: *n* = _____ Control group: *n* =_____

Blinding (detection bias)

_____ 16. □ Blinded (single or double = +1 point) versus an open study (0 points)

_____ 17. □ Control concealment done (yes/no) [same appearance, taste, *etc*.]
Allocation concealment method described

Results: Attrition (attrition bias)

_____ 18. □ Attrition rates given by group (yes/no)

_____ 19. □ Reasons for attrition described by group (yes/no)

Outcomes (reporting bias)

_____ 20. □ Data or text comparing baseline of two groups (demographics, *etc*.)

_____ 21. □ Consort Flow-chart figure done (required post-2006)

Our primary outcome: Hp eradication

_____ 22. □ Primary-Intention to treat analysis? (+1) *vs* As-per-protocol (excludes drop-outs) (0) ?
How was primary outcome assessed?
(___13 C-urea breath test, ___histology, ___ serology, __culture, ___other)

_____23. □ Primary outcome data provided (see table below) (+1 if provided, 0 if not done)

outcome	Probiotic-arm #1	Probiotic-arm #2	Probiotic-arm #3	Probiotic-arm #4	Control	power
HP eradication (Hp negative)						
still Hp+						
totals						
P value:						

Our secondary outcome: Prevention of any adverse events

_____ 24. □ Was either AE or AAD Intention to treat analysis (+1) vs As-per-protocol (excludes drop-outs) (0) ?
How were Adverse events assessed? ____Diary ____Survey Other:_____

_____ 25. □ Outcome data provided (see table below) (+1 if provided, 0 if not done)

	Probiotic-arm #1	Probiotic-arm #2	Probiotic-arm #3	Probiotic-arm #4	Control	power
Any AE:						
No AEs noted						
Totals						
P value:						

Our Secondary outcome: Prevention of antibiotic-associated diarrhea (AAD)

_____ 26. □ AAD data given per group (+1 if provided, 0 if not done)

	Probiotic-arm #1	Probiotic-arm #2	Probiotic-arm #3	Probiotic-arm #4	Control	power
AAD						
No AAD						
Totals						
P value:						

or Description of adverse events:

Types of Adverse Events	Probiotic	Control	power

Sub-group analysis (if done)

_____ 27. □ Sub-group analysis results presented? (n/a if not done)

What were they?_____

Other bias: Discussion:

_____ 28. □ Limitations discussed

Types of limitations found: _____

_____ 29. □ Generalisability discussed (yes/no)

_____ 30. □ Compare these results to other studies (yes/no)

_____ 31. □ Trial registration number/trial registry given (for United States or European studies. post-2006)

_____ 32. □ Location where protocol can be found described (post-2006)

_____ 33. □ Source of funding given (in acknowledgements, elsewhere, or if none)

_____

Quality score (of 33 items):

Reviewer #1 _____: _____# items present (#p), _____ #items absent (#a) _____#n/a (not applicable)

Total score (#p/#p + #a) = _____

Reviewer #2 _____: _____# items present (#p), _____ #items absent (#a) _____#n/a (not applicable)

Total score (#p/#p + #a) = _____

% agreement: _____%

Figure 1 Standardized data extraction form. Scoring: For each of 33 items: +1 if numbered item is present, 0 if absent, or na (not applicable).

Outcomes and definitions

Three outcomes were assessed by this meta-analysis review: (1) eradication rates of *H. pylori*; (2) frequency of adverse events; and (3) frequency of AAD. The outcome for *H. pylori* eradication was defined by having a positive assay (pre-intervention) and a negative *H. pylori* assay done after the intervention was completed. *H. pylori* infection was diagnosed using at least one of the following assays: ^{14}C urea breath test, histology, serology, rapid urease test, stool test or culture[7]. The outcome for adverse events (AE) included any symptoms associated with eradication therapy (nausea, bloating, vomiting, diarrhea, metallic taste) were grouped as "any AE". The outcome for AAD was defined as reported diarrhea or colitis, which developed during the intervention or during the follow-up periods.

Assessment of methodological quality

Quality components for each trial were assessed for selection, detection, performance, reporting and loss to follow-up bias. Each of the included studies was evaluated using 33 items collected with the standardized data extraction form. Each item was graded as: present, absent, or not applicable (for example studies done in countries not requiring clinical trial registration, CONSORT flow-chart not present if trial was published before this became a standard, *etc.*)[28]. The overall quality score for the trial was calculated as the percent of items present divided by the total items present and absent (not applicable items were excluded from the calculation). Each of the 33 quality items were analyzed within one of six categories of potential of bias: study design bias (trial title, setting, early stoppage, background, study aims, prospective design, eligibility criteria, sample size calculation, interim analysis, statistical methods, recruitment methods, subgroup methods, probiotic well described by strain, daily dose and duration), selection bias (randomization allocation method, balanced groups resulted), detection bias (double blinded, treatments concealment), attrition bias (rates provided and reasons by each group), reporting bias (baseline group comparison, CONSORT flow-chart, intent to treat analysis done for each outcome, incidence of each outcome provided, adverse event data provided and sub-group analysis

provided, if applicable) and miscellaneous sources of bias (limitations, generalizability and comparison with other studies in discussion, trial registration, location of protocol for access and source of funding, if appropriate). Trials were classified as high quality if > 75% of the quality items were present, moderate quality if 50%-75% were present and low quality of < 50% were present. Each trial was scored for the 33 items of quality independently by at least two reviewers and a kappa statistic was applied to test for the degree of concordance.

We also employed the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system for rating overall quality of evidence for each of the outcomes by probiotic strain or type[31,32]. Recommendation for use of each probiotic strain can be assessed by the overall strength of the evidence ["strong", many randomized controlled trials show significant protection, more benefit than risk, cost-effective or "weak", only case series or reports, limited number of small trials, *etc.*]. Quality of the evidence is graded as "high quality" (further research is unlikely to change our confidence in the estimate of the effect), or "moderate quality" (further research is likely to have an important impact on our confidence and may change the estimate of the effect), or "low quality" (further research is very likely to change our confidence in the estimate and may change the direction of the estimate of the effect).

Statistical analysis

The statistical methods of this study were reviewed by Lynne McFarland from University of Washington, who holds a PhD in Epidemiology. Statistical analysis was performed using Stata software version 12 (Stata Corporation, College Station, Texas) to calculate pooled relative risks (pRR), bias estimates and number-needed-to-treat statistics. Univariate analysis results were analyzed using χ^2 test or Fisher's exact test for small cell sizes (< 5) with a significance level of $P < 0.05$. Meta-analysis was conducted for primary outcomes (*e.g.*, eradication frequency of *H. pylori* or the rate of adverse events or AAD) using models to calculate the pooled relative risk and corresponding 95%CI using the DerSimonian Laird method. Heterogeneity across trials was evaluated using Cochran Q test based on pooled relative risks by the Mantel-Haenazel method[33]. If the studies were homogenous, a fixed effects model was used; if studies were heterogeneous, a random effect model was employed. A *P*-value < 0.05 is considered statistically significant and *P*-values between 0.05 and 0.1 had a significant trend. The models used in this analysis were weighted by sample size, as study quality did not improve the fit.

If significant heterogeneity was found, subgroup analyses were conducted to determine the potential sources of heterogeneity. To explore possible explanations for heterogeneity, *a priori* subgroup analyses were conducted on study population (adult *vs* pediatric and asymptomatic versus symptomatic), daily dose [$\geqslant 1 \times 10^9$ colony-forming units (cfu) per day or $< 1 \times 10^9$ cfu/d] and study quality. A meta-regression was done without the subgroup indicator and compared to a model with the subgroup indicator included. The difference in tau^2 estimates from the two models indicates the proportion of study heterogeneity explained by the subgroup covariate (between study variance).

Publication bias

To assess for publication bias, a funnel plot, as well as a weighted regression (Egger's test) and a rank correlation test (Begg's test for small study effects) were conducted[27,34]. Funnel plots show graphically that as sample sizes of trials increase, the precision is estimating the underlying treatment effect increases, which results in the effect estimates (relative risks) from small trials scattering more widely at the bottom of the graph and narrower scattering among larger studies. In the absence of publication bias, the funnel plot resembles a symmetrical inverted funnel. Reporting bias (smaller studies showing no protective effect) often are not published, and are indicated by an asymmetrical appearance with a gap in the bottom left of a funnel plot[35,36].

RESULTS

Initial screening of data search

The literature review yielded 301 abstracts relating to probiotics and *H. pylori* that were screened for inclusion. Of those, 225 were excluded after initial screening according to our exclusion criteria (Figure 2): reviews (n = 143), pre-clinical animal models or phase two studies for pharmacokinetics, formulation or safety (n = 67), no control group (n = 6), not randomized (n = 5) or other miscellaneous reasons (n = 4). The literature search for probiotics and *H. pylori* infections found the earliest randomized, controlled efficacy trial was published in 2000. Literature from 1994-1999 only included early investigative studies (mechanism of action, dose-ranging and safety studies) and no clinical trials were found published before 1994.

Secondary screening of full articles

Of the 76 full articles or meeting abstracts retrieved, an additional 35 were excluded: just one RCT found, *i.e.*, no confirmatory RCTs for probiotic strain found (n = 18), no *H. pylori* eradication therapy given with probiotic (n = 11), undefined probiotic product with no species and strain identification (n = 3), no *H. pylori* assays done (n = 1) and two RCTs assessed the burden of *H. pylori* reduced by probiotics but did not document eradication rates nor the frequency of adverse events. Of these trials assessing probiotics

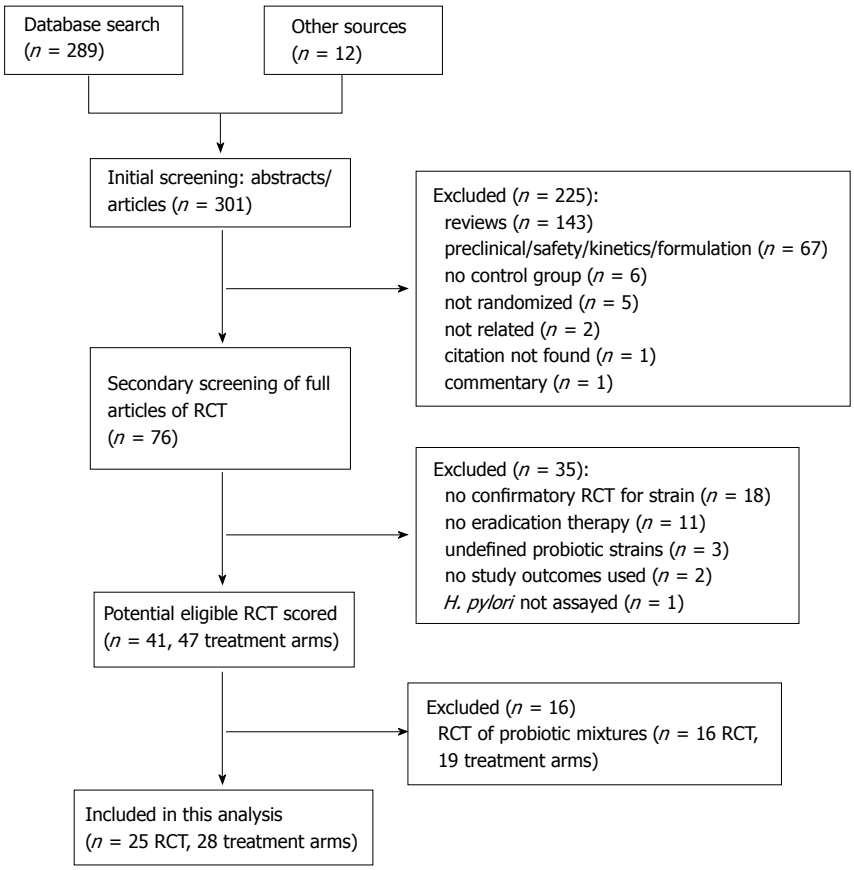


Figure 2 Flow chart of included and excluded trials for Helicobacter pylori eradication/adverse events. RCT: Randomized controlled trials.

and *H. pylori* eradication and/or side effects, 25 (61%) were testing a single strain of probiotic and were included in this analysis and 16 (39%) used multiple strains of probiotics and will be addressed elsewhere. Data extraction was performed independently by co-authors on the remaining 25 RCTs. Examples of RCTs included in prior published meta-analyses, but excluded in our analysis, are shown in Table 2. Reasons for excluding RCTs included: other types of outcomes were assessed[37-39], no concurrent *H. pylori* eradication therapy given[40-46], only one RCT for a specific strain was found[47-52].

Included trials

Of the 25 randomized controlled trials included[53-77], several had multiple treatment arms[53,57,65], resulting in 28 treatment arms, totaling 3769 participants. The sample sizes of the trials ranged from 12 to 991, with a mean number per trial of 68 ± 84 in probiotic arms and 66 ± 83 in control arms. Three articles were translated from their original languages into English: Chinese[62,63] or Spanish[69]. Only two articles were from published meeting abstracts[72,77] with no subsequent full article publications found, the remaining were peer-reviewed full articles.

Patient population

The characteristics of the enrolled study populations by trial arm are presented in Table 3. Of the 28 treatment arms, most enrolled adult participants (*n* = 24, 86%) and four (14%) enrolled children and all trials included both genders. Race or ethnicity was not reported in most clinical trials. The trials were carried out in a wide array of countries: Italy (40%), Turkey (12%), China (12%), Japan (8%), South Korea (8%) and one trial each (4%) for the following: Greece, Iran, Poland, Romania and Venezuela. All treatment arms enrolled *H. pylori* positive participants who were either symptomatic (*n* = 21, 75%), or asymptomatic carriers (*n* = 5, 18%), or had a mixed population (*n* = 1) but one RCT did not report symptom status at enrollment.

Study design

Randomization: All 28 RCT were randomized, but only 12 (43%) provided the method used to randomize patients (*e.g.*, computer random number generator, random block design).

Degree of blinding: Of the 28 treatment arms, only seven arms (25%) were double-blinded (used placebos that were of identical appearance as the probiotic formulation)[53,55,67-69,73], four arms (14%) were single blinded (either participants were unaware of the other treatment arm[61] or outcome assessor was blinded)[56,57]. Most, 17 (61%) of the treatment arms were open trials (no placebos and participants were aware that there was another treatment arm),

Table 2 Excluded randomized controlled trials

Probiotic strain	Reason for exclusion	Ref.
L. gasseri OLL2716	Study quality poor for treatment arm	Boonyaritichaikij *et al*[37]
L. rhamnosus GG	No *H. pylori* assay done	Gawrońska *et al*[38]
L. reuteri ATCC 55730	Outcome was *H. pylori* burden	Francavilla *et al*[39]
S. boulardii I-745 or *L. acidophilus* Lb	No eradication therapy given with probiotic	Gottleland *et al*[40]
Bifido. bifidum YIT4007	No eradication therapy given with probiotic	Miki *et al*[41]
L. casei Shirota	No eradication therapy given with probiotic	Cats *et al*[42]
L. gasseri OLL2716	No eradication therapy given with probiotic	Takagi *et al*[43]
L. johnsonii Lj1	No eradication therapy given with probiotic	Pantoflickova *et al*[44]
L. johnsonii Lj1	No eradication therapy given with probiotic	Gottleland *et al*[45]
L. reuteri ATCC 55730	No eradication therapy given with probiotic	Saggioro *et al*[46]
Bacillus clausii nr	< 2 RCT with eradication therapy	Nista *et al*[47]
Bifido. animalis DN173010	< 2 RCT with eradication therapy	Yaşar *et al*[48]
Bifido. infantis 2036	< 2 RCT with eradication therapy	Dajani *et al*[49]
L. johnsonii Lc-1	< 2 RCT with eradication therapy	Felley *et al*[50]
L. casei DN 114001	< 2 RCT with eradication therapy	Sýkora *et al*[51]
L. casei Shirota	< 2 RCT with eradication therapy	Sahagún-Flores *et al*[52]

nr: Strain not reported; RCT: Randomized controlled trials.

Table 3 Characteristics of enrolled populations in patients receiving eradication therapy by 28 treatment arms

Probiotic strain	Country	Population	Symptoms	Blinding	Eradication therapy	Duration eradication (d)	Ref.
S. boulardii I-745	Italy	Adults	Asymptomatic	Placebo	CTR	7	Cremonini *et al*[53]
S. boulardii I-745	Turkey	Adults	Symptomatic	None	ACO	14	Duman *et al*[54]
S. boulardii I-745	Turkey	Adults	Symptomatic	Placebo	ACL	14	Cindoruk *et al*[55]
S. boulardii I-745	Romania	Pediatric	Symptomatic	Single	AC + O/E	7-21	Hurduc *et al*[56]
S. boulardii I-745	South Korea	Adults	Symptomatic	Single	ACO	7	Song *et al*[57]
S. boulardii I-745 + MPA	South Korea	Adults	Symptomatic	Single	ACO	7	Song *et al*[57]
S. boulardii I-745	Turkey	Adults	Symptomatic	None	ACL	14	Ozdil *et al*[58]
S. boulardii I-745	China	Adults	Symptomatic	None	ACO	14	Chu *et al*[59]
S. boulardii I-745	Iran	Adults	Symptomatic	None	ACO	14	Zojaji *et al*[60]
S. boulardii I-745	Greece	Adults	Symptomatic	Single	ACO	14	Kyriakos *et al*[61]
S. boulardii I-745	China	Pediatric	Symptomatic	None	ACO	14	Zhao *et al*[62]
Clost. butyricum 588	China	Adults	Symptomatic	None	AFO	7	Guo *et al*[63]
Clost. butyricum 588	Japan	Adults	Symptomatic	None	ACL	7	Shimbo *et al*[64]
Clost. butyricum 588 (low dose)	Japan	Adults	Symptomatic	None	ACL	7	Imase *et al*[65]
Clost. butyricum 588 (high dose)	Japan	Adults	Symptomatic	None	ACL	7	Imase *et al*[65]
L. rhamnosus GG	Italy	Adults	Asymptomatic	None	CPT	7	Armuzzi *et al*[66]
L. rhamnosus GG	Italy	Adults	Asymptomatic	Placebo	CRT	7	Armuzzi *et al*[67]
L. rhamnosus GG	Italy	Adults	Asymptomatic	Placebo	CRT	7	Cremonini *et al*[53]
L. rhamnosus GG	Poland	Pediatric	Asymptomatic	Placebo	ACO	7	Szajewska *et al*[68]
L. rhamnosus GG	Venezuela	Adults	Symptomatic	Placebo	ACO	7	Padilla Ruiz *et al*[69]
L. acidophilus Lb	Italy	Adults	Symptomatic	None	ACR	7	Canducci *et al*[70]
L. acidophilus Lb	Italy	Adults	Symptomatic	None	AO	7-30	De Francesco *et al*[71]
L. acidophilus nr	South Korea	Adults	Mixed	None	ACO	7	Yeom *et al*[72]
L. reuteri 55730	Italy	Pediatric	Symptomatic	Placebo	AO, COT	15	Lionetti *et al*[73]
L. reuteri 55730	Italy	Adults	Symptomatic	None	ACT	7	Scaccianoce *et al*[74]
L. reuteri 55730	Italy	Adults	Symptomatic	None	AELe	7	Ojetti *et al*[75]
L. casei DG	Italy	Adults	Symptomatic	None	ART (E/P)	10	Tursi *et al*[76]
L. casei DG	Italy	Adults	nr	None	ACE	7	Giovannone *et al*[77]

This strain is now designated: Saccharomyces boulardii CNCM I-745. Clostridium butyricum 588 (MIYAIRI). Placebo indicates double-blinded design, single indicates either just patient or outcome assessor was blinded and none indicates an open study. A: Amoxicillin; C: Clarithromycin; E: Esomeprazole; F: Furazolidone; L: Lansoprazole; Le: Levofloxacin; MPA: Mucoprotective agent; nr: Not reported in paper/abstract; O: Omeprazole; P: Pantoprazole; R: Randazole; T: Tindazole.

as shown in Table 3.

***H. pylori* eradication therapy**: All trials were required to use an *H. pylori* eradication therapy, which included at least one antibiotic and one PPI for both the probiotic and control group (Table 3). Of the 28 treatment arms, only 1 (4%) used double therapy (amoxicillin and omeprazole)[71]. Most used triple therapy (n = 25, 89%), which most commonly included two antibiotics (amoxicillin and clarithyromycin) combined with a PPI

(omeprazole). Less commonly used were quadruple therapy (n = 1 arm, 4%) or sequential therapy (n = 1 arm, 4%). Overall, the duration of eradication therapy ranged from one week (61% of treatment arms), to 10 d (3%), to two weeks (29%) or varied from 1-4 wk (7%).

Attrition: Attrition ranged from 0%-27% in the 28 treatment arms, usually due to drop-outs due to adverse events or loss to follow-up. Fourteen treatment arms (50%) reported no attrition, 10 (36%) had attrition frequencies from 1%-10% and only three (11%) reported higher attrition (11%-27%), while one trial did not document attrition rates. Of the 28 treatment arms, 24 (86%) used ITT analysis and four (14%) used APP analysis. However, only three of the trials reported how the ITT analysis incorporated the missing data (treated all missing outcomes as failures)[61,63,70].

Intervention

Details of the intervention for the 25 RCT (28 treatment arms) are given in Tables 4 and 5.

Type of probiotic strain(s): In the 28 treatment arms, six different single strain probiotic types were assessed (Tables 3-5) by at least two RCTs that met our eligibility criteria. The most commonly tested strain is *S. boulardii* CNCM I-745, with 11 (39% of RCT arms). *Lactobacillus rhamnosus* (*L. rhamnosus*) GG was tested in five arms (18%), *Clostridium butyricum* 588 was tested in four arms (14%), *L. reuteri* ATCC 55730 and *L. acidophilus* Lb were each tested in two (7%) treatment arms and *L. casei* DG was tested in two treatment arms (7%), one strain of *L. acidophilus* could not be determined.

Newer strain designations for several probiotics and the retrospective review of older studies may have used different strain designations, but were, in fact, the same strain. The most recent strain designations are used in this study. The most current strain designation for *S. boulardii* is CNCM I-745, the registration number at the Pasteur Institute[78], but older studies also refer to this strain as *S. boulardii* lyo, or *S. boulardii*, with no strain designation. *Clostridium butyricum* 588 was also known as *C. butyricum* MIYAIRI. The strain of *L. acidophilus* in one study was referred to only by the brand name (Antibio, China) in the meeting abstract and correspondence with authors and manufacturers were unproductive, but this strain was included in the analysis to illustrate the importance of providing strain designations[72].

Probiotic dose: The daily dose of probiotics varied widely from 1×10^6 to 2×10^{10} colony-forming units (cfu) per day. The *a priori* subgroup analyses on dose compared high dose probiotic ($\geqslant 1 \times 10^9$ cfu/d) versus low dose ($< 1 \times 10^9$ cfu/d). Nineteen (68%) of the treatment arms used the higher daily dose of probiotics and nine (32%) used lower doses (Table 4). The daily dose was reported in all trials, but in some cases the dose was reported as mg/d not cfu/d and required conversion.

Formulation used: Most of the 28 treatment arms used a capsule formulation (12 arms, 43%), while six (21%) used sachets, six (21%) used tablets, two (7%) used liquid and the formulation was not reported in two (7%) of the studies.

Probiotic duration: The probiotics were typically administered as an adjunct for the same duration as the standard eradication therapy, but some RCT continued the probiotic/control intervention for an additional week. The most frequent duration of probiotic was for 2 wk (16 arms, 57%), while five (18%) gave probiotics for only one week and four (14%) gave probiotics for three weeks. Two treatment arms gave probiotics for 10 d (7%) and one (4%) gave for 20 d. All trials reported duration of probiotic given (Table 4).

Length of follow-up: In most trials, participants were followed and tested for *H. pylori* presence 4-8 wk after the intervention treatments were discontinued. Of the 28 treatment arms, 21 (75%) had 1-7 wk of follow-up and four (14%) had longer follow-up times, while three (11%) did not report any follow-up times (Table 4).

Efficacy of adjunct probiotics for H. pylori eradication

Of the 28 treatment arms, 26 (93%) reported *H. pylori* eradication rates in their paper. A low amount of heterogeneity was found when all strains were pooled together (I^2 =25%, P = 0.12), thus a fixed effects model was used for this outcome. The overall pooled RR indicated that probiotics, in general, were effective for *H. pylori* eradication (pRR = 1.10, 95%CI: 1.06-1.14) with a number-needed-to-treat (NNT) of 14. However, as recommended by the literature[24,79], the efficacy should be assessed separately by probiotic strain, as shown by the forest plot (Figure 3). This figure shows that only *S. boulardii* I-745 (n = 10 treatment arms, pRR = 1.11, 95%CI: 1.07-1.16) was significantly effective as an adjunct for *H. pylori* eradication. None of the pooled RR from the other five strains (*C. butyricum* 588, *L. rhamnosus GG*, *L. acidophilus* Lb, *L. reuteri* 55730 or *L. casei* DG significantly improved *H. pylori* eradiation rates with standard therapy. Deletion of the trial with the unknown strain of *L. acidophilus* did not significantly affect the pooled RR estimates.

Sub-group analysis: Results from the meta-regression analysis for the adjunctive use of probiotics for *H. pylori* eradication did not find significant differences in associations between the study population (adult

Table 4 Description of the interventions and *Helicobacter pylori* eradication rates *n* (%)

Probiotic strain	Daily dose (cfu/d)	Form	Duration treatment (wk)	Follow-up post-treatment (wk)	*H. pylori* eradication probiotic	*H. pylori* eradication in controls	Ref.
S. boulardii I-745	1×10^{10}	Sachet	2	5-7	17 (81)	16 (80)	Cremonini *et al*[53]
S. boulardii I-745	1×10^{10}	Capsule	2	4	nr	nr	Duman *et al*[54]
S. *boulardii* I-745	2×10^{10}	Sachet	2	6	44 (71)	37 (60)	Cindoruk *et al*[55]
S. boulardii I-745	1×10^{10}	Capsule	4	4-6	45 (93.3)	34 (80.9)	Hurduc *et al*[56]
S. boulardii I-745	2×10^{10}	Capsule	4	4	264 (80)[a]	237 (71.6)	Song *et al*[57]
S. boulardii I-745 + MPA	2×10^{10}	Capsule	4	4	271 (82.1)[b]	237 (71.6)	Song *et al*[57]
S. boulardii I-745	5×10^{9}	Capsule	2	5	71 (72)	82 (86)[a]	Ozdil *et al*[58]
S. boulardii I-745	5×10^{9}	Sachet	2	52	42 (84)[a]	32 (64)	Chu *et al*[59]
S. boulardii I-745	1×10^{10}	Capsule	2	8	70 (87.5)	65 (81)	Zojaji *et al*[60]
S. boulardii I-745	6×10^{6}	Capsule	2	6	30 (83.4)[a]	20 (58.8)	Kyriakos *et al*[61]
S. boulardii I-745	1×10^{10}	Capsule	2	4	102 (85)[c]	91 (75.8)	Zhao *et al*[62]
Clost. butyricum 588	1×10^{7}	Tablet	1	4	44 (94)	44 (88)	Guo *et al*[63]
Clost. butyricum 588	3×10^{7}	Tablet	2	6	17 (94)	13 (76)	Shimbo *et al*[64]
Clost. butyricum 588 (low dose)	6×10^{7}	Tablet	1	0	7 (100)	6 (87)	Imase *et al*[65]
Clost. butyricum 588 (high dose)	1.2×10^{8}	Tablet	1	0	4 (80)	6 (87)	Imase *et al*[65]
L. rhamnosus GG	1.2×10^{10}	Sachet	2	6	48 (80)	46 (76.6)	Armuzzi *et al*[66]
L. rhamnosus GG	1.2×10^{10}	Sachet	2	6	25 (83)	24 (80)	Armuzzi *et al*[67]
L. rhamnosus GG	1.2×10^{10}	Sachet	2	5-7	16 (76)	16 (80)	Cremonini *et al*[53]
L. rhamnosus GG	2×10^{9}	Capsule	1	6	23 (69)	22 (68)	Szajewska *et al*[68]
L. rhamnosus GG	1.2×10^{10}	Liquid	2	0	nr	nr	Padilla Ruiz *et al*[69]
L. acidophilus Lb	1.5×10^{10}	Capsule	1.4	6	52 (87)[a]	42 (70)	Canducci *et al*[70]
L. acidophilus Lb	2×10^{10}	Capsule	2	4-6	30 (64)	26 (70)	De Francesco *et al*[71]
L. acidophilus nr	2×10^{8}	nr	2	4-8	19 (83)	21 (95.5)	Yeom *et al*[72]
L. reuteri 55730	1×10^{8}	Tablet	2.9	8	17 (85)	16 (80)	Lionetti *et al*[73]
L. reuteri 55730	2×10^{8}	Tablet	1	4-6	9 (53)	10 (62)	Scaccianoce *et al*[74]
L. reuteri 55730	3×10^{8}	Liquid	2	6	36 (80)[a]	27 (60)	Ojetti *et al*[75]
L. casei DG	1.6×10^{10}	Capsule	1.4	4	33 (94.3)	30 (85.7)	Tursi *et al*[76]
L. casei DG	2×10^{10}	nr	4	6	22 (73)	21 (70)	Giovannone *et al*[77]

[a]$P < 0.05$, [b]$P < 0.01$, [c]Trend, $0.05 \leqslant P < 1.0$. This strain is now designated: Saccharomyces boulardii CNCM I-745, nr: Not reported; *S. boulardii*: *Saccharomcyes boulardii; L. rhamnosus: Lactobacillus rhamnosus.*

versus pediatric, $P = 0.76$), baseline disease state (asymptomatic carriage versus symptoms, $P = 0.17$), daily dose of probiotic (above or below 10^9 cfu/d, $P = 0.26$), or study quality ($P = 0.11$). Only probiotic strain group showed significance, confirming the validity of analyzing efficacy by strain type. Sub-group analysis for duration probiotic given and by type of *H. pylori* eradication therapy was not possible, as most trials used similar durations and types of eradication therapy.

Efficacy of adjunct probiotics for prevention of any adverse events

Of the 28 treatment arms, 18 (64%) planned *a priori* to document any adverse events that might occur during the intervention and follow-up period (if done), while 10 (36%) did not document total adverse events during their trials (Table 5). Overall, the pooled RR showed a protective effect (pRR = 0.54, 95%CI: 0.42-0.70, NNT = 8), and as significant heterogeneity was found (I^2 = 56%, $P = 0.003$), random effects models were used for this outcome. The forest plot (Figure 4) shows that only *S. boulardii* I-745 ($n = 7$ treatment arms, pRR = 0.42, 95%CI: 0.28-0.62) significantly reduced the incidence of adverse events associated with standard *H. pylori* eradication therapies. *L. acidophilus* Lb and *L. rhamnosus* GG had no significant protective effect for adverse events and the other three strains of probiotics only had a single treatment arm evaluating adverse events.

Efficacy of adjunct probiotics for the prevention of antibiotic associated diarrhea

Of the 28 treatment arms, 20 (71%) planned *a priori* to document AAD during the intervention and follow-up period (if done), while eight (29%) did not document AAD outcomes (Table 5). Overall, the pooled RR showed a protective effect (pRR = 0.43, 95%CI: 0.35-0.53, NNT = 10), and as significant heterogeneity was not found ($I^2 = 0$, $P = 0.88$), fixed effects models were used to summarize AAD trials. The forest plot (Figure 5) shows that only *S. boulardii* I-745 (n = 9 treatment arms, pRR = 0.47, 95%CI: 0.37-0.60) and *L. rhamnosus* GG ($n = 5$ treatment arms, pRR = 0.29, 95%CI: 0.17-0.48) significantly reduced the incidence of AAD associated with *H. pylori* eradication therapy. The pooled RR from *C. butyricum* 588 and *L. reuteri* 55730 did not find a significant protective effect on AAD. Two strains (*L. acidophilus* Lb and *L. casei* DG) could not be assessed with pooled RRs due to insufficient trials with AAD outcome data.

Publication bias

A funnel plot analysis (Figure 6) provides no compelling

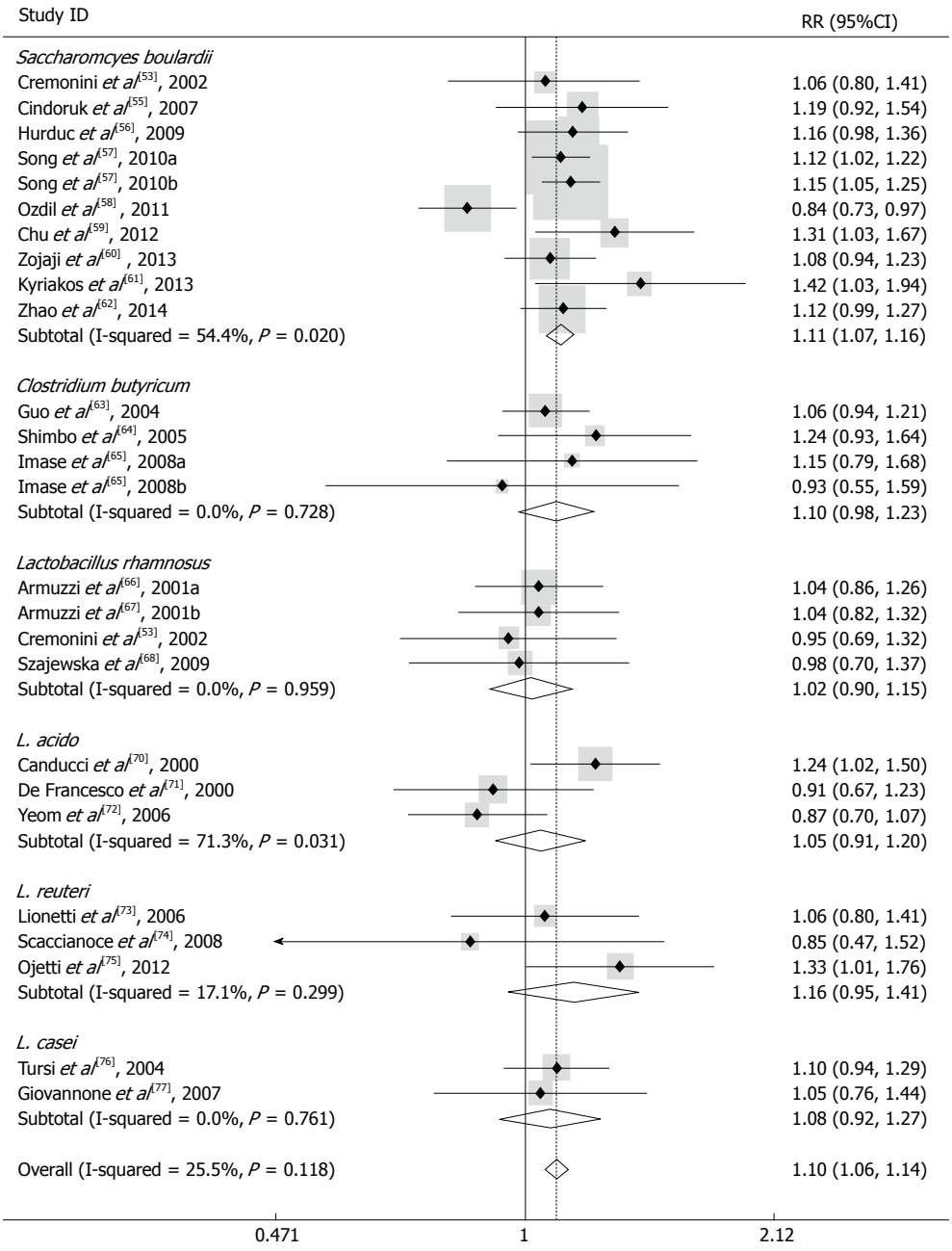


Figure 3 Forest plot of *Helicobacter pylori* eradication by probiotic strain.

indication of publication bias for trials evaluating *H. pylori* eradication outcomes, showing general symmetry of the funnel for the relationship between risk ratio and standard error. The funnel plot shows a lack of published small sized trials with an improved eradication rate. However, Egger's regression test for small study effects (P = 0.71) and Begg's rank test (P = 0.37) fail to suggest significant publication bias. No significant publication bias was found for the RCT assessing the prevention of all adverse reactions (Egger's regression P = 0.42 and Begg's rank P = 0.74). Potential publication bias may be present in RCTs assessing AAD (Egger's regression P = 0.003 and Begg's rank P = 0.025), as there were few outliers noted for small study sizes (Figure 7).

Quality of studies

Of the 25 RCTs, 3 (12%) were rated as high quality studies, 18 (72%) moderate quality and 4 (16%) were low quality trials. The concordance from the reviewers was acceptable (kappa = 0.62, P < 0.001) and any disagreements typically involved only 1-2 of the 33 items in the data extraction form. All disagreements were resolved. As shown in Figure 8, most trials had high quality study design (60%), but only 16% included sample size calculations, 76% failed to indicate "randomized controlled trial" in the title and only 48% described how participants were recruited. There were a low number of trials with selection bias, as all were randomized, but only 40% described the method of randomization used. There was a high

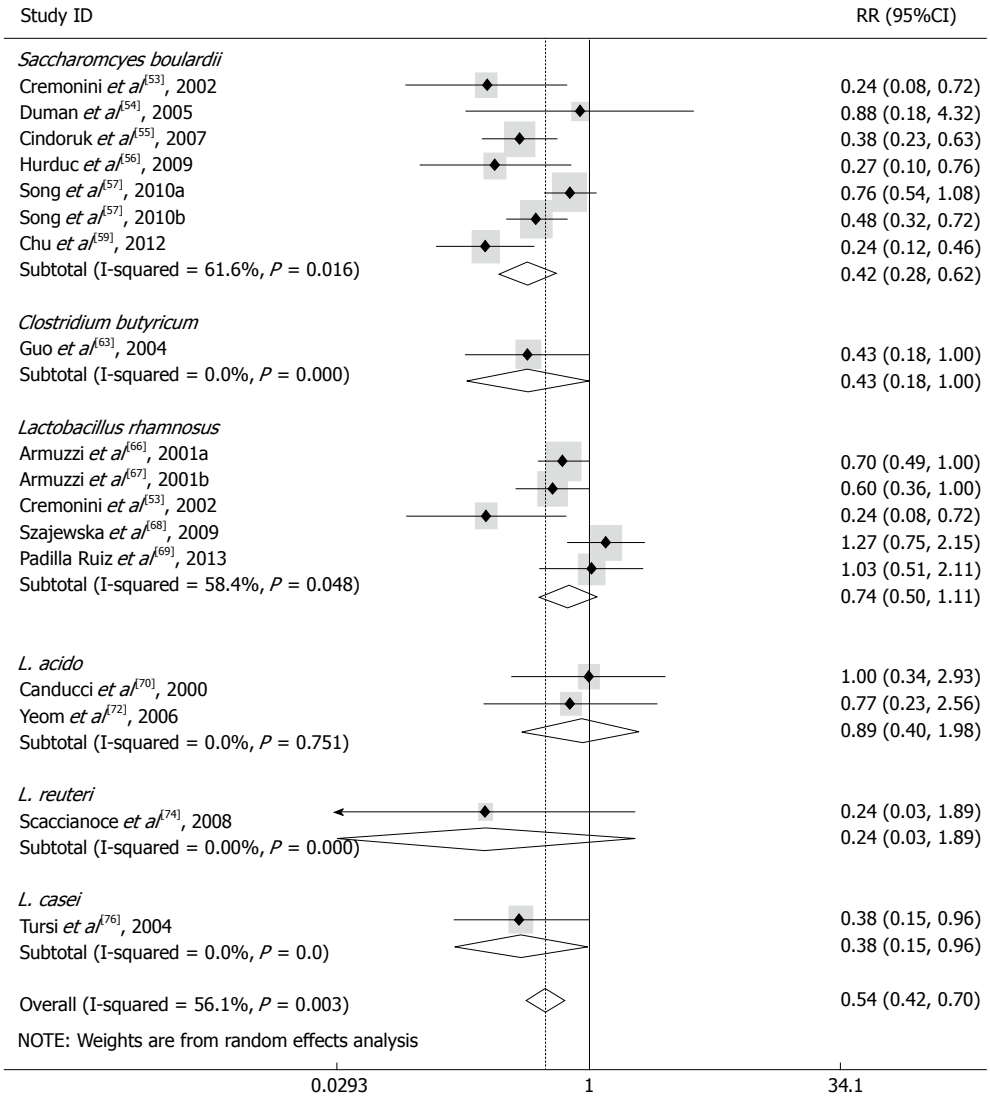


Figure 4 Forest plot of any adverse events by probiotic strain.

degree of detection bias due to the frequent used of open study designs (only 40% were double-blinded) and only 24% described the method of treatment concealment. Most (80%) of the trials reported their attrition rates, but only 65% provided the reasons for attrition by treatment groups. Reporting bias of the outcomes was generally high-moderate quality, but only 44% provided a consort figure describing study flow and only 56% provided a comparison of the two treatment groups at baseline. Other sources of bias were typically of poor quality due to the lack of trial registration or funding source descriptions. In the discussion section of the papers, although 84% compared their results to other studies, only 36% discussed limitations and few (8%) discussed generalizability of their results.

GRADE criteria

For the *H. pylori* eradication, we recommend the following adjunct probiotic strains: *S. boulardii* CNCM I-745 (high quality and strong strength). For the prevention of adverse events associated with standard *H. pylori* eradication therapy, we recommend *S. boulardii* CNCM I-745 (high quality and strong strength). For the prevention of antibiotic-associated diarrhea associated with standard *H. pylori* eradication therapy, we recommend the following adjunct probiotic strains: *S. boulardii* CNCM I-745 (high quality and strong strength) and *L. rhamnosus GG* (strong quality and strong strength). All other strains require additional multiple randomized, controlled trials before a recommendation can be provided.

DISCUSSION

Our meta-analyses found only one probiotic strain significantly improved *H. pylori* eradication rates: *S. boulardii* CNCM I-745 (pRR = 1.11, 95%CI: 1.07-1.15). Only one probiotic strain (*S. boulardii* CNCM I-745) significantly prevented any adverse events (pRR = 0.42, 95%CI: 0.28-0.62). Two probiotic strains significantly reduced antibiotic-associated diarrhea, *S. boulardii* CNCM I-745 and *L. rhamnosus* GG (pRR = 0.47, 95%CI: 0.37-0.60 and pRR = 0.29, 95%CI:

Table 5 Prevention of adverse events associated with *Helicobacter pylori* eradication therapy in 28 treatment arms with adjunct probiotics *n* (%)

Probiotic strain	Any adverse events in probiotic	Any adverse events in controls	Antibiotic associated diarrhea in probiotic	Antibiotic associated diarrhea in controls	Ref.
S. boulardii I-745	3 (14)[b]	12 (60)	1 (5)[a]	6 (30)	Cremonini *et al*[53]
S. boulardii I-745	3 (1.5)	3 (1.7)	14 (6.9)[b]	28 (15.6)	Duman *et al*[54]
S. boulardii I-745	14 (23)[b]	37 (60)	9 (14.5)[a]	19 (30.6)	Cindoruk *et al*[55]
S. boulardii I-745	4 (8)[a]	13 (31)	nr	nr	Hurduc *et al*[56]
S. boulardii I-745	48 (14)	63 (19)	9 (3.3)[a]	20 (6)	Song *et al*[57]
S. boulardii I-745 + MPA	30 (9)[b]	63 (19)	11 (3)	20 (6)	Song *et al*[57]
S. boulardii I-745	nr	nr	nr	nr	Ozdil *et al*[58]
S. boulardii I-745	8 (16)[b]	34 (68)	3 (6)	8 (16)	Chu *et al*[59]
S. boulardii I-745	nr	nr	10 (12.5)[a]	21 (26)	Zojaji *et al*[60]
S. boulardii I-745	nr	nr	1 (2.8)[a]	7 (20.6)	Kyriakos *et al*[61]
S. boulardii I-745	nr	nr	27 (22.5)[b]	47 (39.1)	Zhao *et al*[62]
Clost. butyricum 588	6 (12.8)[a]	15 (30)	nr	nr	Guo *et al*[63]
Clost. butyricum 588	nr	nr	1 (6)	2 (11.8)	Shimbo *et al*[64]
Clost. butyricum 588 (low dose)	nr	nr	1 (14)	3 (43)	Imase *et al*[65]
Clost. butyricum 588 (high dose)	nr	nr	0 (0)	3 (43)	Imase *et al*[65]
L. rhamnosus GG	26 (43)[a]	37 (62)	8 (13.2)[b]	29 (48.2)	Armuzzi *et al*[66]
L. rhamnosus GG	12 (40)[a]	20 (66.6)	1 (3.3)[b]	8 (26.6)	Armuzzi *et al*[67]
L. rhamnosus GG	3 (15)[b]	12 (60)	1 (5)	6 (30)	Cremonini *et al*[53]
L. rhamnosus GG	18 (51)	13 (41)	2 (6)	6 (20)	Szajewska *et al*[68]
L. rhamnosus GG	10 (34)	10 (33)	4 (13.8)	6 (20)	Padilla Ruiz *et al*[69]
L. acidophilus Lb	6 (10)	6 (10)	nr	nr	Canducci *et al*[70]
L. acidophilus Lb	nr	nr	nr	nr	De Francesco *et al*[71]
L. acidophilus nr	4 (15)	5 (19)	nr	nr	Yeom *et al*[72]
L. reuteri 55730	0 (0)	0 (0)	nr	nr	Lionetti *et al*[73]
L. reuteri 55730	1 (5.9)[c]	4 (26.7)	0 (0)[a]	2 (13)	Scaccianoce *et al*[74]
L. reuteri 55730	nr	nr	10 (22)[b]	26 (58)	Ojetti *et al*[75]
L. casei DG	5 (14.3)[a]	13 (37)	0 (0)	3 (8.6)	Tursi *et al*[76]
L. casei DG	nr	nr	nr	nr	Giovannone *et al*[77]

[a]$P < 0.05$, [b]$P < 0.01$, [c]Trend, $0.05 \leqslant P < 1.0$. This strain is now designated: Saccharomyces boulardii CNCM I-745. nr: Not reported in paper/abstract. Numbers in table given as frequency and percent (%). *S. boulardii*: *Saccharomcyes boulardii; L. rhamnosus: Lactobacillus rhamnosus.*

0.17-0.48, respectively). The most promising probiotics strains for *H. pylori* infections also have documented mechanisms of action directed against *H. pylori*. *S. boulardii* produces a neuraminidase that attacks sialic acid, an attachment receptor for *H. pylori*[18] and also induces a morphologic change from the spiral form to a coccoid form of *H. pylori*[80]. There is no direct evidence linking *L. rhamnosus* GG to specific anti-*H. pylori* actions. However, both *S. boulardii* CNCM I-745 and *L rhamnosus* GG have been shown to prevent AAD given for other infections[15,79,81-83].

Our findings are similar to other meta-analyses of probiotics for *H. pylori* infections, which differ by including fewer numbers of trials or did not examine all three outcomes (eradication, adverse reactions and AAD). Szajewska *et al*[84] pooled five randomized trials with *S boulardii* and found significantly better *H. pylori* eradication (pRR = 1.13, 95%CI: 1.05-1.21) and significantly less AAD (pRR = 0.47, 95%CI: 0.32-0.69). Our meta-analysis confirms the robustness of this efficacy from 10 RCTs showing a mild (9%) increase in mean *H. pylori* eradication rates from 73% in control arms to 82% in *S. boulardii* arms, and a reduced rate of AAD in *S. boulardii* arms compared to control arms (8.5% and 21%, respectively). We could not find any other meta-analyses that limited their review to one probiotic strain for *H. pylori* infections.

Tactics for limiting heterogeneity due to the differences of strain-specific probiotic efficacies can be done at the beginning (inclusion criteria only allowing one strain to be included) or post-literature harvesting (by performing sub-group analysis by strain type). Tong *et al*[85] reviewed 14 randomized trials from various probiotic strains and did a sub-group analysis by the type of probiotic and reported only one strain, *L. rhamnosus* GG, showed better *H. pylori* eradication rates odds ratio (OR) from four trials (pOR = 2.09, 95%CI: 1.28-3.4), although one of those trials was actually *L. casei*, not *L. rhamnosus*[85]. Zou *et al*[86] pooled eight trials for *H. pylori* eradication, but incorrectly combined different strains in their subgroup analyses. When Zou *et al*[86] presented data for adverse event rates, they reported five RCT identified as "*L. casei*", however the data presented was actually for eradication rates and three of the five studies used *L. rhamnosus* GG, while the two other studies used different *L. casei* strains (DN11400 and DG). One of the two pooled studies identified as "*L. acidophilus*" used a mixture of two different Lactobacilli strains[86]. Some meta-analyses did not separate out probiotic strains using sub-group analysis and only presented summary risk estimates combining many different probiotic

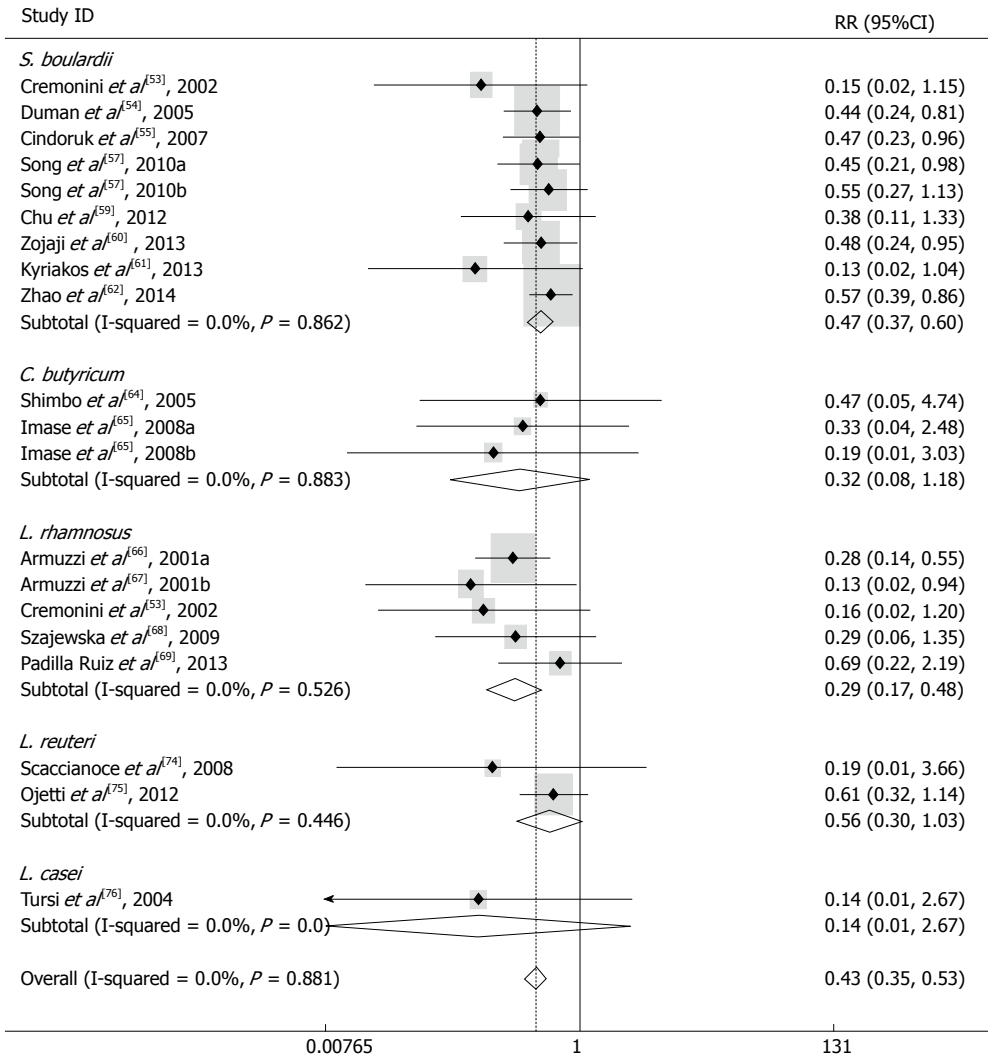


Figure 5 Forest plot of prevention of antibiotic-associated diarrhea by probiotic strain.

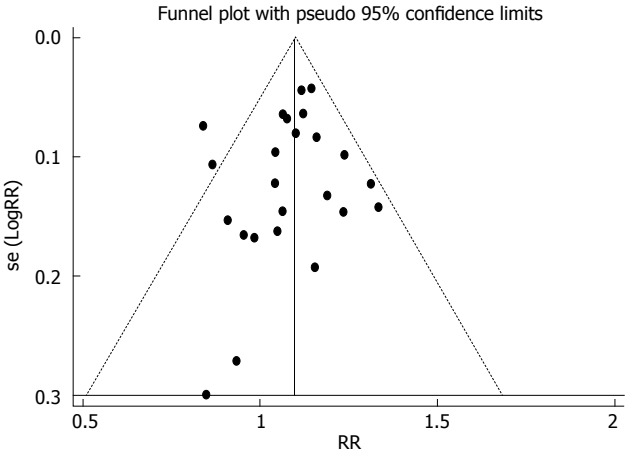


Figure 6 Funnel plot for publication bias assessment from for *Helicobacter pylori* eradication and probiotics.

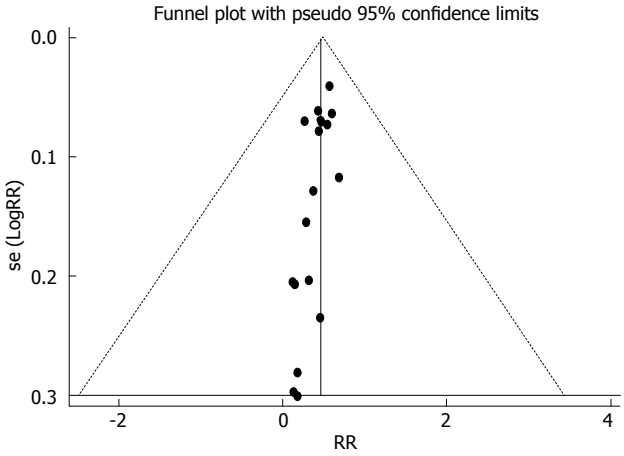


Figure 7 Funnel plot for publication bias assessment from for prevention of antibiotic associated diarrhea and probiotics.

strains[87-89]. Sachdeva *et al*[90] did not find an effect by probiotic strain in their meta-regression analysis. Wang *et al*[91] pooled 10 RCT using different mixtures containing Lactobacilli and/or Bifidobacterium and did a sub-group analysis on race, quality, symptoms, age and types of eradication therapy, but failed to analyze the strains of probiotics separately.

Other reviews and meta-analysis have also analyzed

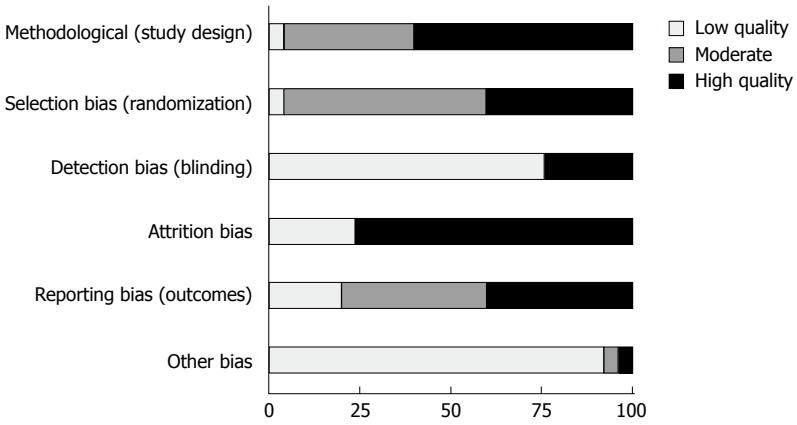


Figure 8 Frequency of study quality based on six different types of potential bias. High quality, low bias (76%-100% quality items within category present), moderate quality and moderate bias (51%-75% items present), low quality, high bias (0%-50% items present).

the effect of probiotics for the prevention of adverse events and AAD related to *H. pylori* eradication therapy, but typically have pooled different strains together into one group[85,87,89,91]. Zou *et al*[86] reported no significant effect of Lactobacilli probiotics on adverse events, but pooled together studies using *L. rhamnosus* GG (3 studies), *L. acidophilus*, *L. casei*, and *L. reuteri* (one study each) into the same group. As our meta-analysis shows a distinct strain specificity to both the efficacy of eradicating *H. pylori* and the prevention of adverse events (including AAD), future studies need to be aware that pooling similar probiotics by species is no longer appropriate and their outcomes need to analyzed by the same type of probiotic strain.

The quality of clinical trials in our analysis varied from a score of 0.32-0.89, which was not surprising as some of the trials were done before standardized randomized controlled trial guidelines were widely published and two trials with low quality scores were from meeting abstracts that never resulted in full article publications. The advantage of scoring trials on quality is the results allow an assessment of recommendations to improve future studies. Future trials would benefit from better study designs (use of placebos, study size calculations), more complete descriptions of their outcomes and discussion of limitations and generalizability.

A question that arises from discussions on how best to treat patients with *H. pylori* is whether probiotics alone are sufficient to treat these infections, or is adjunctive therapy with the standard antibiotic and PPI therapy more effective. The study by Gotteland *et al*[40] tested *S. boulardii* alone or heat-killed *L. acidophilus* Lb alone versus triple therapy *H. pylori* eradication therapies and found *S. boulardii* alone or *L. acidophilus* alone was significantly poorer (12% and 6% eradication, respectively) than triple therapy used alone (66%, $P < 0.05$), thus strengthening the position that probiotics are most effective when combined with antibiotic-PPI eradication therapy. Most other studies testing probiotics alone (without the standard eradication therapies) have failed to show a significant effect of the probiotic[41,42,44], while a few found significant improvement of eradication rates using just a probiotic[45,46], although one study treated patients with either only a PPI (omeprazole) or *L. reuteri*/PPI and did not use any antibiotics in the control group[46].

The results of the Maastricht IV/Florence Consensus, which involved 44 experts on *H. pylori,* reported the decreasing eradication rates of the triple therapy (only 70%) may be due to the development of resistance to clarithromycin and poor compliance due to adverse events associated with triple therapies[7]. This group found better eradication rates using either sequential treatments [5 d of PPI and amoxicillin followed by 5 d of PPI, clarithromycin and metronidazole (or tinidazole)] or quadruple therapy (PPI with two antibiotics and bismuth). This group also recommended extending the duration of therapy from 7 d to 10-14 d. While eradication rates may improve with these regimes, the incidence of adverse events remains high. At the time of the meeting (2010), they did not recommend the use of probiotics, citing the poor quality of the studies due to mixing different species and strains in published meta-analyses, but they did recommend further studies. In recent years, more probiotic trials have been done and this meta-analysis does present the outcomes separated by probiotic species and strain.

It was difficult to assess the most effective combination of probiotic strain and type of *H. pylori* eradication therapy, as most trials used a similar eradication therapy. In our review of 28 treatment arms, over 89% used triple therapy and the most common combination was amoxicillin, clarithromycin and omeprazole (36% of all triple therapies), followed by amoxicillin, clarithromycin and lansoprazole (18%). Eradication rates did not significantly differ by the type of eradication therapy and probiotic strain given, but the lack of variation and studies using the same eradication therapy and probiotic strain limited our analysis. It is also difficult to recommend the best daily dose and duration of a probiotic. Our subgroup analysis did not show a significant effect of daily dose, and doses used in trials with the same strain often had similar daily doses. Other meta-analyses that have investigated the effect of the dose and duration of the probiotic regime have not found a significant effect[88].

Most of the trials (89%) had sufficient follow-up

times (4-8 wk) to allow adverse events to occur, but 11% did not have any follow-up post-treatment. As only one trial followed patients for a prolonged time (one year), it is uncertain if the *H. pylori* eradication rates reported in the trials are transient or more permanent.

This systematic review has several strengths. We had specific outcomes selected *a* priori and the search strategy for this review was comprehensive including any relevant trials irrespective of language or publication status (*i.e.*, we included published data from meeting abstracts, obtained specific data from authors, and translated three non-English trials). Additional strengths of the review include its application of the GRADE criteria for each of the outcomes[31] and the rigorous evaluation of each of the subgroups (*i.e.*, same probiotic strain, probiotic dose, study population, and risk of bias) using the 33 criteria for assessing subgroup credibility[92]. The results of this meta-analysis may be generalizable to the global population, because we included a wide range of ages, countries and settings (inpatients and outpatients, adults and children were included). It should be noted however, that ethnicity and race data were not reported, nor were immunocompromised patients included in most of the trials, so the applicability of our results to these types of these populations is not known.

This review also has several limitations. While we did a more comprehensive search of the grey literature, we did not search all conference proceedings or dissertation abstracts. One of the main limitations for doing meta-analysis on probiotics is the limited number of probiotic strains that have data from multiple trials. Probiotic strain has been cited as the key indicator of efficacy for several diseases[23-25], but the limited number of trials on the same strain limits our ability draw robust conclusions on most of the strains used for all cited studies. We had to exclude 18 studies that only had one randomized controlled trial for a specific probiotic strain and, as a consequence, not all probiotic strains were included in this analysis. Another limitation is the changing designation of the probiotic strain over time. Older trials may refer to the same strain, but under a different strain type or the strain designation may not be provided in the published article. Other meta-analyses have grouped several strains of *L. casei* into one group (DG or DN114001 or Shirota), perhaps due to the lack of a current consensus on the taxonomy of these strains[93]. We did include one *L. acidophilus* study into our analysis, but it should be noted that the strain designation could not be determined retrospectively. This makes a systematic review challenging, as the authors must retrospectively find the matching strain designations as they change over time to include or exclude studies from specific probiotic strain groups.

Recommendations for future research include multiple randomized, controlled trials on the same probiotic strain, allowing confirmation of single clinical trial results. Improvements in the quality of study design should include complete description of the probiotic intervention (strain designation, daily dose, duration, source, *etc.*), use of treatment concealment (double blinding), calculating sample size *a* priori to power a sufficiently large study to detect significant results, use of intent-to-treat analysis to account for patient attrition effects, the collection of adverse event data and having sufficient follow-up time after the treatments are discontinued. In our meta-analysis, only four the trials had sufficient follow-up times (> 8 wk) to capture prolonged eradication of *H. pylori*. Future clinical trials need to incorporate sufficient follow-up times in their study protocols. None of the RCT in this meta-analysis reported any adverse events associated with probiotic use, which has been substantiated in other papers[94-96], but adverse event data should be collected and assessed for future studies.

In conclusion, our meta-analyses found only one strain of probiotic (*S. boulardii*, CNCM, I-745) is beneficial and safe in the eradication of *H. pylori* when combined with standard eradication therapy, and two strains of probiotics (*S. boulardii or L. rhamnosus* GG) decreased the adverse events of eradication therapy (including AAD), which may improve compliance in infected patients.

COMMENTS

Background

Helicobacter pylori (H. pylori) infections are a global problem and may lead to the development of a wide range of symptoms from dyspepsia to gastric cancer. The current therapy of multiple antibiotics and a proton pump inhibitor is associated with high frequencies of adverse events, which reduces compliance and increases treatment failure rates. The addition of probiotics to the standard treatments may assist in improving compliance, but the correct choice of probiotic strain is paramount.

Research frontiers

Over the years, many randomized controlled trials have been done to evaluate the efficacy of probiotics as adjunctive therapy for the eradication of *H. pylori* and/or development of adverse events, but previous reviews have been flawed or incomplete and may have inappropriately combined different types of probiotics into one group and thus could not achieve a comprehensive conclusion.

Innovations and breakthroughs

This comprehensive meta-analysis has used current guidelines for evaluating probiotic efficacy separately by the type of probiotic (only single strain probiotic trials grouped together) and evaluated each of three outcomes (*H. pylori* eradication, reducing any adverse events, reducing antibiotic-associated diarrhea) separately to determine which single probiotic strain may be efficious for each of the three outcomes. A total of 25 randomized controlled trials (with 28 treatment arms) of single strain probiotics were assessed. Of the six different probiotic strains evaluated, only two (*Saccharomyces boulardii* CNCM I-745 and *Lactobacillus rhamnosus* GG) were significantly associated with an improvement in at least one of the three outcomes.

Applications

These two probiotic strains can be used as adjunctive therapy to antibiotics used to treat *H. pylori* infections and may both improve compliance and reduce the development of adverse events, leading to better cure rates.

Terminology

Probiotics are living microbes (either fungal or bacterial), which when given at

appropriate doses, can affect the health status of the host.

Peer-review

The authors conducted a comprehensive literature review and data analysis on eradication of *H. pylori* by a single strain of probiotics. From literature collection to data analysis, it is all scientifically sound and the manuscript is well written.

REFERENCES

1 **Marshall BJ**, Warren JR. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* 1984; **1**: 1311-1315 [PMID: 6145023]

2 **Papamichael K,** Mantzaris GJ. Pathogenesis of Helicobacter pylori infection: colonization, virulence factors of the bacterium and immune and non-immune host response. *Hospital Chronicles* 2012; **7**: 32-37

3 **Malfertheiner P**, Link A, Selgrad M. Helicobacter pylori: perspectives and time trends. *Nat Rev Gastroenterol Hepatol* 2014; **11**: 628-638 [PMID: 25001975 DOI: 10.1038/nrgastro.2014.99]

4 **Gatta L**, Vakil N, Vaira D, Scarpignato C. Global eradication rates for Helicobacter pylori infection: systematic review and meta-analysis of sequential therapy. *BMJ* 2013; **347**: f4587 [PMID: 23926315 DOI: 10.1136/bmj.f4587]

5 **Liang X**, Xu X, Zheng Q, Zhang W, Sun Q, Liu W, Xiao S, Lu H. Efficacy of bismuth-containing quadruple therapies for clarithromycin-, metronidazole-, and fluoroquinolone-resistant Helicobacter pylori infections in a prospective study. *Clin Gastroenterol Hepatol* 2013; **11**: 802-807.e1 [PMID: 23376004 DOI: 10.1016/j.cgh.2013.01.008]

6 **Liou JM**, Chen CC, Chen MJ, Chen CC, Chang CY, Fang YJ, Lee JY, Hsu SJ, Luo JC, Chang WH, Hsu YC, Tseng CH, Tseng PH, Wang HP, Yang UC, Shun CT, Lin JT, Lee YC, Wu MS. Sequential versus triple therapy for the first-line treatment of Helicobacter pylori: a multicentre, open-label, randomised trial. *Lancet* 2013; **381**: 205-213 [PMID: 23158886 DOI: 10.1016/S0140-6736(12)61579-7]

7 **Malfertheiner P**, Megraud F, O'Morain CA, Atherton J, Axon AT, Bazzoli F, Gensini GF, Gisbert JP, Graham DY, Rokkas T, El-Omar EM, Kuipers EJ. Management of Helicobacter pylori infection--the Maastricht IV/ Florence Consensus Report. *Gut* 2012; **61**: 646-664 [PMID: 22491499]

8 **Francavilla R**, Zullo A, Vaira D. How often do you fail to take all of your medication? *J Pediatr Gastroenterol Nutr* 2012; **55**: 338 [PMID: 22411268 DOI: 10.1097/MPG.0b013e318253c9ff]

9 **Ierardi E**, Giorgio F, Losurdo G, Di Leo A, Principi M. How antibiotic resistances could change Helicobacter pylori treatment: A matter of geography? *World J Gastroenterol* 2013; **19**: 8168-8180 [PMID: 24363506 DOI: 10.3748/wjg.v19.i45.8168]

10 **Wermeille J**, Cunningham M, Dederding JP, Girard L, Baumann R, Zelger G, Buri P, Metry JM, Sitavanc R, Gallaz L, Merki H, Godin N. Failure of Helicobacter pylori eradication: is poor compliance the main cause? *Gastroenterol Clin Biol* 2002; **26**: 216-219 [PMID: 11981460]

11 **Megraud F**, Coenen S, Versporten A, Kist M, Lopez-Brea M, Hirschl AM, Andersen LP, Goossens H, Glupczynski Y. Helicobacter pylori resistance to antibiotics in Europe and its relationship to antibiotic consumption. *Gut* 2013; **62**: 34-42 [PMID: 22580412 DOI: 10.1136/gutjnl-2012-302254]

12 **Yang JC**, Lu CW, Lin CJ. Treatment of Helicobacter pylori infection: current status and future concepts. *World J Gastroenterol* 2014; **20**: 5283-5293 [PMID: 24833858 DOI: 10.3748/wjg.v20.i18.5283]

13 **Schwizer W**, Menne D, Schütze K, Vieth M, Goergens R, Malfertheiner P, Leodolter A, Fried M, Fox MR. The effect of Helicobacter pylori infection and eradication in patients with gastro-oesophageal reflux disease: A parallel-group, double-blind, placebo-controlled multicentre study. *United European Gastroenterol J* 2013; **1**: 226-235 [PMID: 24917966 DOI: 10.1177/2050640613484020]

14 **Ayala G**, Escobedo-Hinojosa WI, de la Cruz-Herrera CF, Romero I. Exploring alternative treatments for Helicobacter pylori infection. *World J Gastroenterol* 2014; **20**: 1450-1469 [PMID: 24587621 DOI: 10.3748/wjg.v20.i6.1450]

15 **McFarland LV**, Goh S. Preventing Pediatric Antibiotic-Associated Diarrhea and Clostridium difficile Infections with Probiotics: a meta-analysis. *World J Meta-Anal* 2013; **1**: 102-120 [DOI: 10.13105/wjma.v1.i3.02]

16 **Hongying F**, Xianbo W, Fang Y, Yang B, Beiguo L. Oral immunization with recombinant Lactobacillus acidophilus expressing the adhesin Hp0410 of Helicobacter pylori induces mucosal and systemic immune responses. *Clin Vaccine Immunol* 2014; **21**: 126-132 [PMID: 24285819 DOI: 10.1128/CVI.00434-13]

17 **Myllyluoma E**, Ahonen AM, Korpela R, Vapaatalo H, Kankuri E. Effects of multispecies probiotic combination on helicobacter pylori infection in vitro. *Clin Vaccine Immunol* 2008; **15**: 1472-1482 [PMID: 18579692]

18 **Sakarya S**, Gunay N. Saccharomyces boulardii expresses neuraminidase activity selective for α2,3-linked sialic acid that decreases Helicobacter pylori adhesion to host cells. *APMIS* 2014; **122**: 941-950 [PMID: 24628732 DOI: 10.1111/apm.12237]

19 **Yang YJ**, Sheu BS. Probiotics-containing yogurts suppress Helicobacter pylori load and modify immune response and intestinal microbiota in the Helicobacter pylori-infected children. *Helicobacter* 2012; **17**: 297-304 [PMID: 22759330 DOI: 10.1111/j.1523-5378.2012.00941.x]

20 **Fujimura S**, Watanabe A, Kimura K, Kaji M. Probiotic mechanism of Lactobacillus gasseri OLL2716 strain against Helicobacter pylori. *J Clin Microbiol* 2012; **50**: 1134-1136 [PMID: 22205802 DOI: 10.1128/JCM.06262-11]

21 **Plummer SF**, Garaiova I, Sarvotham T, Cottrell SL, Le Scouiller S, Weaver MA, Tang J, Dee P, Hunter J. Effects of probiotics on the composition of the intestinal microbiota following antibiotic therapy. *Int J Antimicrob Agents* 2005; **26**: 69-74 [PMID: 15967639]

22 **Myllyluoma E**, Ahlroos T, Veijola L, Rautelin H, Tynkkynen S, Korpela R. Effects of anti-Helicobacter pylori treatment and probiotic supplementation on intestinal microbiota. *Int J Antimicrob Agents* 2007; **29**: 66-72 [PMID: 17141481]

23 **Azaïs-Braesco V**, Bresson JL, Guarner F, Corthier G. Not all lactic acid bacteria are probiotics, ...but some are. *Br J Nutr* 2010; **103**: 1079-1081 [PMID: 20230653 DOI: 10.1017/S0007114510000723]

24 **Hill C**, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, Morelli L, Canani RB, Flint HJ, Salminen S, Calder PC, Sanders ME. Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol* 2014; **11**: 506-514 [PMID: 24912386 DOI: 10.1038/nrgastro.2014.66]

25 **Rijkers GT**, de Vos WM, Brummer RJ, Morelli L, Corthier G, Marteau P. Health benefits and health claims of probiotics: bridging science and marketing. *Br J Nutr* 2011; **106**: 1291-1296 [PMID: 21861940 DOI: 10.1017/S000711451100287X]

26 **Liberati A**, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009; **62**: e1-34 [PMID: 19631507 DOI: 10.1016/j.jclinepi.2009.06.006]

27 **Egger M**, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-634 [PMID: 9310563]

28 **Moher D**, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg* 2012; **10**: 28-55 [PMID: 22036893 DOI: 10.1016/j.ijsu.2011.10.001]

29 **Davis LM**, Martínez I, Walter J, Hutkins R. A dose dependent impact of prebiotic galactooligosaccharides on the intestinal microbiota of healthy adults. *Int J Food Microbiol* 2010; **144**: 285-292 [PMID: 21059476 DOI: 10.1016/j.ijfoodmicro.2010.10.007]

30 **Gibson GR**. Dietary modulation of the human gut microflora using prebiotics. *Br J Nutr* 1998; **80**: S209-S212 [PMID: 9924286]

31 **Guyatt GH**, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ. GRADE: an emerging consensus

on rating quality of evidence and strength of recommendations. *BMJ* 2008; **336**: 924-926 [PMID: 18436948 DOI: 10.1136/bmj.39489.470347.AD]

32 **Mustafa RA**, Santesso N, Brozek J, Akl EA, Walter SD, Norman G, Kulasegaram M, Christensen R, Guyatt GH, Falck-Ytter Y, Chang S, Murad MH, Vist GE, Lasserson T, Gartlehner G, Shukla V, Sun X, Whittington C, Post PN, Lang E, Thaler K, Kunnamo I, Alenius H, Meerpohl JJ, Alba AC, Nevis IF, Gentles S, Ethier MC, Carrasco-Labra A, Khatib R, Nesrallah G, Kroft J, Selk A, Brignardello-Petersen R, Schünemann HJ. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *J Clin Epidemiol* 2013; **66**: 736-742; quiz 742.e1-5 [PMID: 23623694 DOI: 10.1016/j.jclinepi.2013.02.004]

33 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120]

34 **Begg CB**, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994; **50**: 1088-1101 [PMID: 7786990]

35 **Peters JL**, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. *JAMA* 2006; **295**: 676-680 [PMID: 16467236]

36 **Sterne JA**, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001; **323**: 101-105 [PMID: 11451790]

37 **Boonyaritichaikij S**, Kuwabara K, Nagano J, Kobayashi K, Koga Y. Long-term administration of probiotics to asymptomatic pre-school children for either the eradication or the prevention of Helicobacter pylori infection. *Helicobacter* 2009; **14**: 202-207 [PMID: 19702850 DOI: 10.1111/j.1523-5378.2009.00675.x]

38 **Gawrońska A**, Dziechciarz P, Horvath A, Szajewska H. A randomized double-blind placebo-controlled trial of Lactobacillus GG for abdominal pain disorders in children. *Aliment Pharmacol Ther* 2007; **25**: 177-184 [PMID: 17229242]

39 **Francavilla R**, Lionetti E, Castellaneta SP, Magistà AM, Maurogiovanni G, Bucci N, De Canio A, Indrio F, Cavallo L, Ierardi E, Miniello VL. Inhibition of Helicobacter pylori infection in humans by Lactobacillus reuteri ATCC 55730 and effect on eradication therapy: a pilot study. *Helicobacter* 2008; **13**: 127-134 [PMID: 18321302 DOI: 10.1111/j.1523-5378.2008.00593.x]

40 **Gotteland M**, Poliak L, Cruchet S, Brunser O. Effect of regular ingestion of Saccharomyces boulardii plus inulin or Lactobacillus acidophilus LB in children colonized by Helicobacter pylori. *Acta Paediatr* 2005; **94**: 1747-1751 [PMID: 16421034]

41 **Miki K**, Urita Y, Ishikawa F, Iino T, Shibahara-Sone H, Akahoshi R, Mizusawa S, Nose A, Nozaki D, Hirano K, Nonaka C, Yokokura T. Effect of Bifidobacterium bifidum fermented milk on Helicobacter pylori and serum pepsinogen levels in humans. *J Dairy Sci* 2007; **90**: 2630-2640 [PMID: 17517703]

42 **Cats A**, Kuipers EJ, Bosschaert MA, Pot RG, Vandenbroucke-Grauls CM, Kusters JG. Effect of frequent consumption of a Lactobacillus casei-containing milk drink in Helicobacter pylori-colonized subjects. *Aliment Pharmacol Ther* 2003; **17**: 429-435 [PMID: 12562457]

43 **Takagi A,** Uemura N, Inoue K, et al. Effect of L. gasseri on dyspeptic symptoms in subjects with H. pylori infection. Meeting abstract (#Mo1869), presented by Digestive Disease Week. *Gastroenterol* 2013; **144**: S-679

44 **Pantoflickova D**, Corthésy-Theulaz I, Dorta G, Stolte M, Isler P, Rochat F, Enslen M, Blum AL. Favourable effect of regular intake of fermented milk containing Lactobacillus johnsonii on Helicobacter pylori associated gastritis. *Aliment Pharmacol Ther* 2003; **18**: 805-813 [PMID: 14535874]

45 **Gotteland M**, Andrews M, Toledo M, Muñoz L, Caceres P, Anziani A, Wittig E, Speisky H, Salazar G. Modulation of Helicobacter pylori colonization with cranberry juice and Lactobacillus johnsonii La1 in children. *Nutrition* 2008; **24**: 421-426 [PMID: 18343637 DOI: 10.1016/j.nut.2008.01.007]

46 **Saggioro A,** Caroli M, Girardi L, Chiozzini G, Pasini M. H. pylori eradication with Lactobacillus reuteri. A double-blind placebo-controlled study. *Dig Live Dis* 2005; **37** (Suppl 1): A88

47 **Nista EC**, Candelli M, Cremonini F, Cazzato IA, Zocco MA, Franceschi F, Cammarota G, Gasbarrini G, Gasbarrini A. Bacillus clausii therapy to reduce side-effects of anti-Helicobacter pylori treatment: randomized, double-blind, placebo controlled trial. *Aliment Pharmacol Ther* 2004; **20**: 1181-1188 [PMID: 15569121]

48 **Yaşar B**, Abut E, Kayadıbı H, Toros B, Sezıklı M, Akkan Z, Keskın Ö, Övünç Kurdaş O. Efficacy of probiotics in Helicobacter pylori eradication therapy. *Turk J Gastroenterol* 2010; **21**: 212-217 [PMID: 20931422]

49 **Dajani AI**, Hammour AMA, Yang DH, et al. Do probiotics improve eradication response to Helicobacter pylori on standard triple or sequential therapy? *Saudi J Gastroenterol* 2013; **19**: 113-120 [DOI: 10.4103/1319-3767.111953]

50 **Felley CP**, Corthésy-Theulaz I, Rivero JL, Sipponen P, Kaufmann M, Bauerfeind P, Wiesel PH, Brassart D, Pfeifer A, Blum AL, Michetti P. Favourable effect of an acidified milk (LC-1) on Helicobacter pylori gastritis in man. *Eur J Gastroenterol Hepatol* 2001; **13**: 25-29 [PMID: 11204805]

51 **Sýkora J**, Valecková K, Amlerová J, Siala K, Dedek P, Watkins S, Varvarovská J, Stozický F, Pazdiora P, Schwarz J. Effects of a specially designed fermented milk product containing probiotic Lactobacillus casei DN-114 001 and the eradication of H. pylori in children: a prospective randomized double-blind study. *J Clin Gastroenterol* 2005; **39**: 692-698 [PMID: 16082279]

52 **Sahagún-Flores JE**, López-Peña LS, de la Cruz-Ramírez Jaimes J, García-Bravo MS, Peregrina-Gómez R, de Alba-García JE. [Eradication of Helicobacter pylori: triple treatment scheme plus Lactobacillus vs. triple treatment alone]. *Cir Cir* 2007; **75**: 333-336 [PMID: 18158878]

53 **Cremonini F**, Di Caro S, Covino M, Armuzzi A, Gabrielli M, Santarelli L, Nista EC, Cammarota G, Gasbarrini G, Gasbarrini A. Effect of different probiotic preparations on anti-helicobacter pylori therapy-related side effects: a parallel group, triple blind, placebo-controlled study. *Am J Gastroenterol* 2002; **97**: 2744-2749 [PMID: 12425542]

54 **Duman DG**, Bor S, Ozütemiz O, Sahin T, Oğuz D, Iştan F, Vural T, Sandkci M, Işksal F, Simşek I, Soytürk M, Arslan S, Sivri B, Soykan I, Temizkan A, Beşşk F, Kaymakoğlu S, Kalayc C. Efficacy and safety of Saccharomyces boulardii in prevention of antibiotic-associated diarrhoea due to Helicobacterpylori eradication. *Eur J Gastroenterol Hepatol* 2005; **17**: 1357-1361 [PMID: 16292090]

55 **Cindoruk M**, Erkan G, Karakan T, Dursun A, Unal S. Efficacy and safety of Saccharomyces boulardii in the 14-day triple anti-Helicobacter pylori therapy: a prospective randomized placebo-controlled double-blind study. *Helicobacter* 2007; **12**: 309-316 [PMID: 17669103]

56 **Hurduc V**, Plesca D, Dragomir D, Sajin M, Vandenplas Y. A randomized, open trial evaluating the effect of Saccharomyces boulardii on the eradication rate of Helicobacter pylori infection in children. *Acta Paediatr* 2009; **98**: 127-131 [PMID: 18681892 DOI: 10.1111/j.1651-2227.2008.00977.x]

57 **Song MJ**, Park DI, Park JH, Kim HJ, Cho YK, Sohn CI, Jeon WK, Kim BI. The effect of probiotics and mucoprotective agents on PPI-based triple therapy for eradication of Helicobacter pylori. *Helicobacter* 2010; **15**: 206-213 [PMID: 20557362]

58 **Ozdil K**, Calhan T, Sahin A, Senates E, Kahraman R, Yüzbasioglu B, Demirdag H, Demirsoy H, Sökmen MH. Levofloxacin based sequential and triple therapy compared with standard plus probiotic combination for Helicobacter pylori eradication. *Hepatogastroenterology* 2011; **58**: 1148-1152 [PMID: 21937367 DOI: 10.5754/hge11075]

59 **Chu Y**, Zhu H, Zhou Y, Lv L, Huo J. Intervention study on Saccharomyces boulardii with proton pump inhibitor (PPI)-based triple therapy for Helicobacter pylori related peptic ulcer. *African J Pharmacy Pharmacol* 2012; **6**: 2900-2904 [DOI: 10.5897/AJIPP12.400]

60 **Zojaji H**, Ghobakhlou M, Rajabalinia H, Ataei E, Jahani Sherafat S, Moghimi-Dehkordi B, Bahreiny R. The efficacy and safety of adding the probiotic Saccharomyces boulardiito standard triple therapy for eradication of H.pylori: a randomized controlled trial. *Gastroen-*

terol Hepatol Bed Bench 2013; **6**: S99-S104 [PMID: 24834296]
61 **Kyriakos N,** Papamichael K, Roussos A, Theodoropoulos I, Karakoidas C, Smyrnidis A, Archavlis E, Lariou K, Mantzaris GJ. Lyophilized Form of Saccharomyces boulardii enhances the Helicobacter pylori eradication rates of omeprazole-triple therapy in patients with peptic ulcer disease or functional dyspepsia. *Hosp Chronicles* 2013; **8**: 127-133
62 **Zhao HM**, Ou-Yang HJ, Duan BP, Xu B, Chen ZY, Tang J, You JY. [Clinical effect of triple therapy combined with Saccharomyces boulardii in the treatment of Helicobacter pylori infection in children]. *Zhongguo Dangdai Erke Zazhi* 2014; **16**: 230-233 [PMID: 24661511]
63 **Guo JB**, Yang PF, Wang MT. [The application of clostridium to the eradication of Helicobacter pylori]. *Chin J Celiopathy* 2004; **3**: 163-165
64 **Shimbo I**, Yamaguchi T, Odaka T, Nakajima K, Koide A, Koyama H, Saisho H. Effect of Clostridium butyricum on fecal flora in Helicobacter pylori eradication therapy. *World J Gastroenterol* 2005; **11**: 7520-7524 [PMID: 16437727]
65 **Imase K**, Takahashi M, Tanaka A, Tokunaga K, Sugano H, Tanaka M, Ishida H, Kamiya S, Takahashi S. Efficacy of Clostridium butyricum preparation concomitantly with Helicobacter pylori eradication therapy in relation to changes in the intestinal microbiota. *Microbiol Immunol* 2008; **52**: 156-161 [PMID: 18402597]
66 **Armuzzi A**, Cremonini F, Ojetti V, Bartolozzi F, Canducci F, Candelli M, Santarelli L, Cammarota G, De Lorenzo A, Pola P, Gasbarrini G, Gasbarrini A. Effect of Lactobacillus GG supplementation on antibiotic-associated gastrointestinal side effects during Helicobacter pylori eradication therapy: a pilot study. *Digestion* 2001; **63**: 1-7 [PMID: 11173893]
67 **Armuzzi A**, Cremonini F, Bartolozzi F, Canducci F, Candelli M, Ojetti V, Cammarota G, Anti M, De Lorenzo A, Pola P, Gasbarrini G, Gasbarrini A. The effect of oral administration of Lactobacillus GG on antibiotic-associated gastrointestinal side-effects during Helicobacter pylori eradication therapy. *Aliment Pharmacol Ther* 2001; **15**: 163-169 [PMID: 11148433]
68 **Szajewska H**, Albrecht P, Topczewska-Cabanek A. Randomized, double-blind, placebo-controlled trial: effect of lactobacillus GG supplementation on Helicobacter pylori eradication rates and side effects during treatment in children. *J Pediatr Gastroenterol Nutr* 2009; **48**: 431-436 [PMID: 19330931]
69 **Padilla Ruiz M**, Fernández Aguiar ME, Arce Nuñez M, Polo Amorín R. [Lactobacillus rhamnosus GG supplementation to reduce side-effects of anti-Helicobacter pylori treatment]. *Rev Gastroenterol Peru* 2013; **33**: 121-130 [PMID: 23838939]
70 **Canducci F**, Armuzzi A, Cremonini F, Cammarota G, Bartolozzi F, Pola P, Gasbarrini G, Gasbarrini A. A lyophilized and inactivated culture of Lactobacillus acidophilus increases Helicobacter pylori eradication rates. *Aliment Pharmacol Ther* 2000; **14**: 1625-1629 [PMID: 11121911]
71 **De Francesco V**, Stoppino V, Sgarro C, Faleo D. Lactobacillus acidophilus administration added to omeprazole/amoxycillin-based double therapy in Helicobacter pylori eradication. *Dig Liver Dis* 2000; **32**: 746-747 [PMID: 11142590]
72 **Yeom H,** Shim K, Ryu K. Effect of Lactobacillus acidophilus on Helicobacter pylori treatment. Meeting Abstract M1271. Presented at Digestive Disease Week. Los Angeles, California, May 20-25, 2006
73 **Lionetti E**, Miniello VL, Castellaneta SP, Magistá AM, de Canio A, Maurogiovanni G, Ierardi E, Cavallo L, Francavilla R. Lactobacillus reuteri therapy to reduce side-effects during anti-Helicobacter pylori treatment in children: a randomized placebo controlled trial. *Aliment Pharmacol Ther* 2006; **24**: 1461-1468 [PMID: 17032283]
74 **Scaccianoce G**, Zullo A, Hassan C, Gentili F, Cristofari F, Cardinale V, Gigliotti F, Piglionica D, Morini S. Triple therapies plus different probiotics for Helicobacter pylori eradication. *Eur Rev Med Pharmacol Sci* 2008; **12**: 251-256 [PMID: 18727457]
75 **Ojetti V**, Bruno G, Ainora ME, Gigante G, Rizzo G, Roccarina D, Gasbarrini A. Impact of Lactobacillus reuteri Supplementation on Anti-Helicobacter pylori Levofloxacin-Based Second-Line Therapy. *Gastroenterol Res Pract* 2012; **2012**: 740381 [PMID: 22690211 DOI: 10.1155/2012/740381]
76 **Tursi A**, Brandimarte G, Giorgetti GM, Modeo ME. Effect of Lactobacillus casei supplementation on the effectiveness and tolerability of a new second-line 10-day quadruple therapy after failure of a first attempt to cure Helicobacter pylori infection. *Med Sci Monit* 2004; **10**: CR662-CR666 [PMID: 15567983]
77 **Giovannone M**, Barberani F, Boschetto S, Gigliozzi A, Tosoni M. Lactobacillus casei DG effectiveness on Helicobacter pylori eradication treatment-side effectis; a placebo-controlled, double-blind randomized pilot study. Meeting Abstract. Presented at Amer Gastroenterol Association Meeting, Washington DC, May 19-24, 2007
78 **Dinleyici EC**, Kara A, Ozen M, Vandenplas Y. Saccharomyces boulardii CNCM I-745 in different clinical conditions. *Expert Opin Biol Ther* 2014; **14**: 1593-1609 [PMID: 24995675]
79 **McFarland LV**. Deciphering meta-analytic results: a mini-review of probiotics for the prevention of paediatric antibiotic-associated diarrhoea and Clostridium difficile infections. *Benef Microbes* 2014; **6**: 1-6 [PMID: 24889895]
80 **Vandenplas Y**, Brunser O, Szajewska H. Saccharomyces boulardii in childhood. *Eur J Pediatr* 2009; **168**: 253-265 [PMID: 19096876 DOI: 10.1007/s00431-008-0879-7]
81 **McFarland LV**. Meta-analysis of probiotics for the prevention of antibiotic associated diarrhea and the treatment of Clostridium difficile disease. *Am J Gastroenterol* 2006; **101**: 812-822 [PMID: 16635227]
82 **Szajewska H**, Ruszczyński M, Radzikowski A. Probiotics in the prevention of antibiotic-associated diarrhea in children: a meta-analysis of randomized controlled trials. *J Pediatr* 2006; **149**: 367-372 [PMID: 16939749]
83 **Johnston BC**, Supina AL, Ospina M, Vohra S. Probiotics for the prevention of pediatric antibiotic-associated diarrhea. *Cochrane Database Syst Rev* 2007; **2**: CD004827 [PMID: 17443557]
84 **Szajewska H**, Horvath A, Piwowarczyk A. Meta-analysis: the effects of Saccharomyces boulardii supplementation on Helicobacter pylori eradication rates and side effects during treatment. *Aliment Pharmacol Ther* 2010; **32**: 1069-1079 [PMID: 21039671 DOI: 10.111/j.1365-2036.2010.04457.x]
85 **Tong JL**, Ran ZH, Shen J, Zhang CX, Xiao SD. Meta-analysis: the effect of supplementation with probiotics on eradication rates and adverse events during Helicobacter pylori eradication therapy. *Aliment Pharmacol Ther* 2007; **25**: 155-168 [PMID: 17229240]
86 **Zou J**, Dong J, Yu X. Meta-analysis: Lactobacillus containing quadruple therapy versus standard triple first-line therapy for Helicobacter pylori eradication. *Helicobacter* 2009; **14**: 97-107 [PMID: 19751434 DOI: 10.1111/j.1523-5378.2009.00716.x]
87 **Li S**, Huang XL, Sui JZ, Chen SY, Xie YT, Deng Y, Wang J, Xie L, Li TJ, He Y, Peng QL, Qin X, Zeng ZY. Meta-analysis of randomized controlled trials on the efficacy of probiotics in Helicobacter pylori eradication therapy in children. *Eur J Pediatr* 2014; **173**: 153-161 [PMID: 24323343 DOI: 10.1007/s00431-013-2220-3]
88 **Ritchie ML**, Romanuk TN. A meta-analysis of probiotic efficacy for gastrointestinal diseases. *PLoS One* 2012; **7**: e34938 [PMID: 22529959 DOI: 10.1371/journal.pone.0034938]
89 **Zheng X**, Lyu L, Mei Z. Lactobacillus-containing probiotic supplementation increases Helicobacter pylori eradication rate: evidence from a meta-analysis. *Rev Esp Enferm Dig* 2013; **105**: 445-453 [PMID: 24274441]
90 **Sachdeva A**, Rawat S, Nagpal J. Efficacy of fermented milk and whey proteins in Helicobacter pylori eradication: a review. *World J Gastroenterol* 2014; **20**: 724-737 [PMID: 24574746 DOI: 10.3748/wjg.v20.i3.724]
91 **Wang ZH**, Gao QY, Fang JY. Meta-analysis of the efficacy and safety of Lactobacillus-containing and Bifidobacterium-containing probiotic compound preparation in Helicobacter pylori eradication therapy. *J Clin Gastroenterol* 2013; **47**: 25-32 [PMID: 23090045 DOI: 10.1097/MCG.0b013e318266f6cf]
92 **Sun X**, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010; **340**: c117 [PMID: 20354011 DOI: 10.1136/bmj.c117]
93 **Sato H**, Torimura M, Kitahara M, Ohkuma M, Hotta Y, Tamura H. Characterization of the Lactobacillus casei group based on the

profiling of ribosomal proteins coded in S10-spc-alpha operons as observed by MALDI-TOF MS. *Syst Appl Microbiol* 2012; **35**: 447-454 [PMID: 23099260 DOI: 10.1016/j.syapm.2012.08.008]

94 **McFarland LV**. Systematic review and meta-analysis of Saccharomyces boulardii in adult patients. *World J Gastroenterol* 2010; **16**: 2202-2222 [PMID: 20458757]

95 **Hempel S**, Newberry S, Ruelaz A, Wang Z, Miles JN, Suttorp MJ, Johnsen B, Shanman R, Slusser W, Fu N, Smith A, Roth B, Polak J, Motala A, Perry T, Shekelle PG. Safety of probiotics used to reduce risk and prevent or treat disease. *Evid Rep Technol Assess* (Full Rep) 2011; **200**: 1-645 [PMID: 23126627]

96 **Salminen MK**, Tynkkynen S, Rautelin H, Saxelin M, Vaara M, Ruutu P, Sarna S, Valtonen V, Järvinen A. Lactobacillus bacteremia during a rapid increase in probiotic use of Lactobacillus rhamnosus GG in Finland. *Clin Infect Dis* 2002; **35**: 1155-1160 [PMID: 12410474]

P- Reviewer: Wang Y, Zhang Z **S- Editor**: Ji FF **L- Editor**: A
E- Editor: Lu YJ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i2.118

World J Meta-Anal 2015 April 26; 3(2): 118-124
ISSN 2308-3840 (online)

META-ANALYSIS

Addition of hip exercises to treatment of patellofemoral pain syndrome: A meta-analysis

Kimberly M Morelli, Maria Carrelli, Maria A Nunez, Caroline A Smith, Gordon L Warren

Kimberly M Morelli, Maria Carrelli, Maria A Nunez, Caroline A Smith, Gordon L Warren, Department of Physical Therapy, Georgia State University, Atlanta, GE 30302-4019, United States
Author contributions: All authors contributed to the design of the research study and interpretation of data; Morelli KM, Carrelli M, Nunez MA and Smith CA contributed to the acquisition of data as described in the methods; Morelli KM and Warren GL performed the data analyses and wrote the manuscript.
Conflict-of-interest: The authors have declared no conflict of interest related to the manuscript.
Data sharing: Technical appendix, statistical code, and dataset available from the corresponding author at gwarren@gsu.edu.

Correspondence to: Gordon L Warren, Professor, Department of Physical Therapy, Georgia State University, P.O. Box 4019, Atlanta, GE 30302-4019,
United States. gwarren@gsu.edu
Telephone: +1-404-4131255
Fax: +1-404-4131230
Received: November 5, 2014
Peer-review started: November 5, 2014
First decision: January 20, 2015
Revised: February 13, 2015
Accepted: March 30, 2015
Article in press: April 2, 2015
Published online: April 26, 2015

Abstract

AIM: To determine if the addition of hip-strengthening exercises decreases pain and improves function in patients with patellofemoral pain syndrome.

METHODS: The authors completed a systematic review searching eight databases (*i.e.*, PubMed, Cochrane, CINHAL, MEDLINE, SportsDiscus, EMBASE, APTA Hooked on Evidence, and PEDro). Two independent reviewers screened and excluded studies if they did not meet the following inclusion criteria: subjects had a primary diagnosis of patellofemoral pain syndrome (PFPS), intervention group included hip-strengthening exercises, control group included a traditional physical therapy intervention, study included outcome measures of pain and/or function, study used a randomized controlled trial design, PEDro score was $\geqslant$ 7, and study was published in a peer-reviewed journal. Primary outcome measures were subjective scales of pain and function. These measures were converted to standardized mean difference [effect size (ES)], and a random-effects model was used to calculate the overall ES.

RESULTS: Two hundred eighty-three studies were screened for inclusion in our meta-analysis. Nine studies were deemed suitable for data extraction and analysis. A total of 426 subjects were used in the nine studies. Overall, there was a significant positive effect of hip-strengthening exercises on measures of pain and function in subjects with PFPS (ES = 0.94, P = 0.00004). None of the individual studies had a negative ES, with study ES ranging from 0.35 to 2.59. Because of the high degree of between-study variance (I^2 = 76%; Q = 34.0, P < 0.001), subgroup meta-analyses and meta-regressions were performed. None of the potential moderator variables that were investigated (*e.g.*, outcome type, hip region targeted, duration of treatment) could explain a significant amount of the between-study variance in ES ($P \geqslant 0.23$).

CONCLUSION: Overall, the addition of hip-strengthening exercises to traditional physical therapy produced greater improvements in measures of pain and function.

Key words: Exercise therapy; Systematic review; Knee joint; Physical therapy modalities

Core tip: The most effective treatment to improve pain and function in patellofemoral pain syndrome is uncertain. We performed a systematic review and meta-analysis to determine if the addition of hip-strengthening exercises to traditional physical therapy interventions could effectively reduce pain and increase function in patients with patellofemoral pain syndrome. Our analysis indicates that the addition of hip-strengthening exercises provides a significant and relatively large additional reduction in pain and increase in function.

Morelli KM, Carrelli M, Nunez MA, Smith CA, Warren GL. Addition of hip exercises to treatment of patellofemoral pain syndrome: A meta-analysis. *World J Meta-Anal* 2015; 3(2): 118-124 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i2/118.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i2.118

INTRODUCTION

Patellofemoral pain syndrome (PFPS) is a prevalent lower-extremity disorder. PFPS can account for over 10% of physician office visits in an orthopedic setting, and account for 25%-40% of patients with knee pain and/or injury[1-4]. Women are twice as likely to be affected compared to men[3,5-7]. The etiology of PFPS has historically been attributed to abnormal tracking of the patella resulting from abnormal muscle forces, either weakness or tightness, and/or biomechanical factors (*e.g.,* Q angle, shallow trochlear groove) that alter the normal compressive and shear forces at the patellofemoral joint[1-8]. Lateral tracking of the patella can occur with an excessive Q angle at the knee, which is a measure of the angle of pull of the knee extensors in relation to the patellar tendon[2,3,8]. However, there is no consensus on PFPS's etiology.

Factors proximal to the patellofemoral joint are emerging as possible significant contributors to the cause of PFPS. There is a recent focus on the role of the hip abductor muscles in controlling the genu valgum angle at the knee during dynamic activity, finding that weakness of the hip abductors leads to an increased adduction/genu valgum moment with activity[8-10]. Weakness of the hip external rotators and extensors may also contribute to increased adduction and internal rotation of the lower leg with activity, thereby increasing biomechanical forces of shear and compression at the patellofemoral joint[10].

Traditional physical therapy interventions have focused on knee extensor strengthening, as well as bracing, taping, and modalities in treating patients with PFPS[2]. Often times, interventions focused strictly at the knee joint and knee extensors are not successful at decreasing a patient's pain complaint[2]. With the recent interest in the role of the proximal hip joint musculature contributing to PFPS, the objective of this study was to determine, utilizing a systematic review and meta-analysis, if the addition of hip-strengthening exercises to a traditional physical therapy intervention reduces pain and improves function in patients with PFPS more so than the traditional physical therapy intervention alone.

MATERIALS AND METHODS

Systematic review

We reviewed the research literature to identify studies that examined the effects of hip-strengthening exercises on pain and functional limitations in patients with PFPS. Our literature search began September 2013 and continued through October 2014. Databases including PubMed, Cochrane, CINHAL, MEDLINE, SportsDiscus, EMBASE, APTA Hooked on Evidence, and PEDro were searched electronically. The search terms included: "patellofemoral AND hip strength*" and MeSH terms (patellofemoral pain syndrome/rehabilitation AND hip) OR (patellofemoral pain syndrome/therapy AND hip).

Study inclusion and exclusion criteria: Two independent reviewers screened and excluded studies if they did not meet the following inclusion criteria: (1) study utilized subjects with a principal medical diagnosis of patellofemoral pain syndrome; (2) study included a treatment group performing hip-strengthening exercises in combination with or without a traditional physical therapy intervention; (3) study included a control group performing a traditional physical therapy intervention; (4) study named the muscles or muscle region targeted with exercises performed; (5) study measured pain or function as outcomes; (6) studies were randomized controlled trials and had a PEDro score greater than or equal to 7[11]; and (7) study was published in a peer-reviewed journal.

Selection of studies: Two hundred eighty-three studies were identified through the database searches and review of article reference lists. Of those, 126 studies were eliminated as duplicates among the different databases. Then, 135 studies were excluded on the basis of the title and/or review of the abstract. Twenty-two studies were fully evaluated via a careful review of the full text. On the basis of the inclusion and exclusion criteria, 13 studies were excluded leaving a total of nine studies to be included in the meta-analysis[12-20] (Figure 1).

Data extraction: For the meta-analysis, pain and function as reported by the numeric pain rating scale (NPRS), pain visual analog scale (VAS), Kujala anterior knee pain scale (AKPS), lower extremity functional scale (LEFS), and Womac pain rating scale data were extracted in the form of means, standard deviations, and sample sizes for the intervention (*i.e.,* group employing hip-strengthening exercises) and

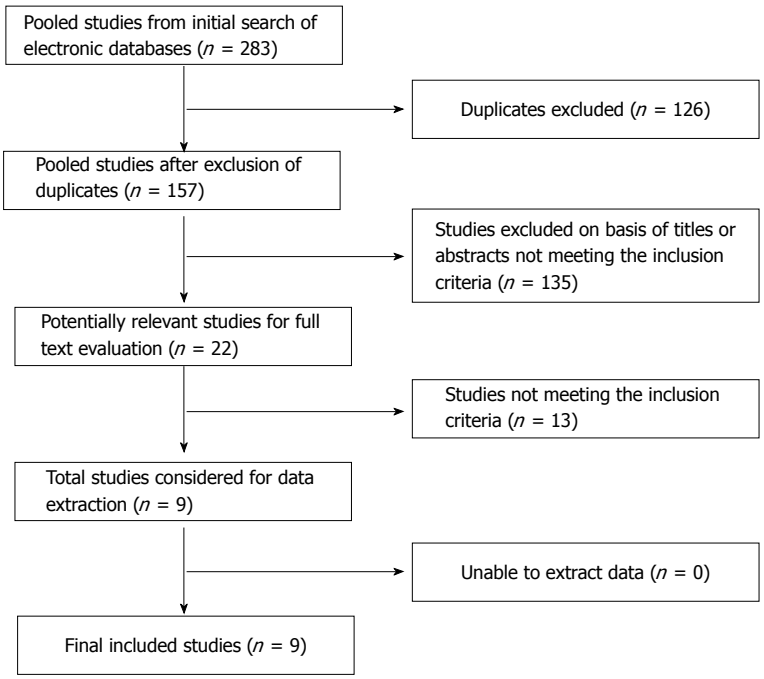


Figure 1 Flow diagram of the number of studies identified, the number excluded and the reasons for exclusions, and the final number of studies included in the systematic review and meta-analysis.

control groups. The number of sessions, region of hip targeted, total duration of treatment sessions, time for follow-up assessments, and subject descriptive measures were also extracted from the studies. The region of the hip targeted with exercise was extracted from the studies, and categorized as posterolateral if the exercises were isolated to the posterolateral hip muscles or general if the exercises involved the major thigh and muscles that cross the hip joint. The time for follow-up was extracted from the studies and categorized as immediate follow-up if outcome assessments were done when the intervention ceased or as long-term for any assessment performed 3 or more months after the intervention ceased. Exercise intensity was extracted for all studies but was not found to be usable because of the variability among studies in how intensity was expressed or because it varied within and between sessions in some studies but not in others.

Meta-analysis

The extracted pain and functional measures data were converted to a standard format, *i.e.*, standardized mean difference, which will be referred to as an effect size (ES) from this point on. Meta-analyses were run using the random-effects model that accounts for true between-study variation in effects as well as random error within each study. A random-effects model was employed for this analysis because the nine studies used dissimilar experimental designs and/or procedures[21]. Between-study variance was assessed using the Q value and I^2. Because substantial between-study variance was detected, we sought to determine the role that different experimental factors might have in explaining this variance. These factors can be treated as potential moderator variables. Meta-regressions (using a methods of moment model) or subgroup meta-analyses were used to probe the following potential moderator variables: region of hip targeted during the hip exercises, number of exercise sessions, duration in weeks of exercise intervention, control group type, outcome type, and time of follow-up. Subgroup meta-analyses and meta-regressions were used for analysis of categorical and continuous variables, respectively. In studies with more than one experimental factor level being evaluated (*e.g.,* a study using both pain and function outcomes in the subgroup meta-analysis evaluating the effect of outcome type), an ES was calculated for each level and was treated as if it originated from an independent study.

Meta-analyses were conducted using comprehensive meta-analysis software (version 2.2; Biostat Inc., Englewood, NJ). An α level of 0.05 was used in all analyses. Effect sizes of 0.2, 0.5, and 0.8 were considered to be small, moderate and large respectively[22]. The possible effect of publication bias on the meta-analysis was assessed by visual assessment of a funnel plot and using Duval and Tweedie's trim and fill correction.

RESULTS

Description of included studies

In total, nine studies were included in the meta-analysis examining the effect of hip-strengthening exercises on pain and function in persons with PFPS. The characteristics of these studies are summarized in Table 1. All nine studies were published in peer-reviewed journals and used a randomized controlled trial design. Subjects were randomly assigned to the two groups, *i.e.,* one receiving a traditional intervention and one receiving traditional intervention plus hip exercises. Therapy providers were not blinded

Table 1 Characteristics of the nine studies examining the effects of hip strengthening exercises on pain and function in patients with patellofemoral pain syndrome

Ref.	Subject information	Subject mean age (min-max)	Hip region targeted	Outcome measured	Time to follow-up (mo)	Number of exercise sessions	Exercise duration (wk)	PEDro quality score (0-11)
Dolak *et al*[12]	33 women	25.5 (16-35)	Posterolateral	Pain VAS and LEFS	0, 1[1], 2[1]	12	4	7
Fukuda *et al*[14]	41 women	25 (20-40)	Posterolateral	NPRS, Kujala AKPS, LEFS	0	12	4	9
Fukuda *et al*[13]	49 women	22.5 (20-40)	Posterolateral	NPRS, Kujala AKPS, LEFS	3, 6, 12	12	4	9
Herrington *et al*[15]	30 men	26.9 (18-35)	General	Pain VAS, Kujala AKPS	0	18	6	9
Ismail *et al*[16]	32 (9 men, 23 women)	21 (18-30)	Posterolateral	Pain VAS, Kujala AKPS	0	18	6	8
Khayambashi *et al*[17]	36 (18 men, 18 women)	27.8 (12-44)	Posterolateral	Pain VAS, Womac	0, 6	24	8	7
Nakagawa *et al*[18]	14 (4 men, 10 women)	23.6 (17-40)	Posterolateral	Pain VAS	0	30	6	10
van Linschoten *et al*[19]	131 (47 men, 84 women)	24 (14-40)	General	Pain VAS, Kujala AKPS	0, 9	84	12	7
Witvrouw *et al*[20]	60 (20 men, 40 women)	20.3 (14-33)	General	Pain VAS, Kujala AKPS	3, 60[1]	15	5	8

[1]Data for this time point not included in analysis due to lack of baseline data matched to subjects in that time to follow-up and/or a change in the intervention regimen that did not meet our criteria. VAS: Visual analog scale; NPRS: Numeric pain rating scale; LEFS: Lower extremity functional scale; AKPS: Anterior knee pain scale.

to which group they were assigned to. Assessors administering the outcome assessments (*i.e.,* Pain VAS, NPRS, LEFS, and Kujala AKPS) were blinded to the groups that the subjects were assigned to; however, subjects completed these questionnaires and were aware of the group they were assigned to. All but one study included measures of both pain and function; the one study included only a measure of pain[18]. Seven of the studies measured outcomes immediately after completing the intervention[12,14-19], while two studies did not make assessments of pain and function until at least 3 mo post intervention[13,20]. Dolak *et al*[12] made a follow-up assessment at 1 and 2 mo post intervention; however, this data was not included in the analysis because the exercise regimen changed after post-treatment and did not meet our inclusion criteria. Witvrouw *et al*[20] made a follow-up assessment at 5 years post intervention; however, these data were not included in the analysis because baseline measures were not available for the subjects who reported for the 5-year follow-up and subjects were inconsistent in adhering to their exercise regimen during this period. Six studies specified hip exercises as targeting the posterolateral musculature of the hip, such as hip abduction, hip lateral rotation and hip extension[12-14,16-18], while three studies' hip exercises were considered general to the hip musculature[15,19,20]. A total of 426 subjects were used in the nine studies. Subject gender in the studies was generally a mixture of men and women but one study used men only[15] and three studies used women only[12-14]. The duration of intervention varied among studies from 4 to 12 wk, with total number of treatment sessions ranging from 12 to 84.

Meta-analysis on pain and function outcomes

When combining all outcome types and times for follow-up, meta-analysis of all nine studies yielded a statistically significant and large effect size (ES = 0.94, P = 0.00004), indicating that patients with PFPS performing hip exercises in addition to traditional physical therapy interventions reported less pain and increased function than control subjects receiving traditional interventions (Figure 2). None of the individual studies had a negative ES, with the standardized mean difference ranging from 0.35 to 2.59. No one study was found to dominate the calculation of the overall ES. Fukuda *et al*[13] 2012 had the single largest effect on the overall ES but even if it was removed from the analysis, the overall ES was still moderate-to-large and statistically significant (ES = 0.67, P = 0.000001). Publication bias also did not appear to affect the overall ES. We did not observe any overt asymmetry in the funnel plot of standard error versus study ES. Furthermore, when the Duval and Tweedie's trim and fill adjustment was applied to correct for potential publication bias, no studies were trimmed and thus the procedure made no adjustment to the overall ES.

The two assessments of variation in ES among the studies indicate that the variation is both large (I^2 = 76%) and statistically significant (Q = 34.0, df = 8, P < 0.001). Because of this variability, subgroup meta-analyses and meta-regressions were used to probe possible roles for six experimental factors that might help to explain ES variation among the nine

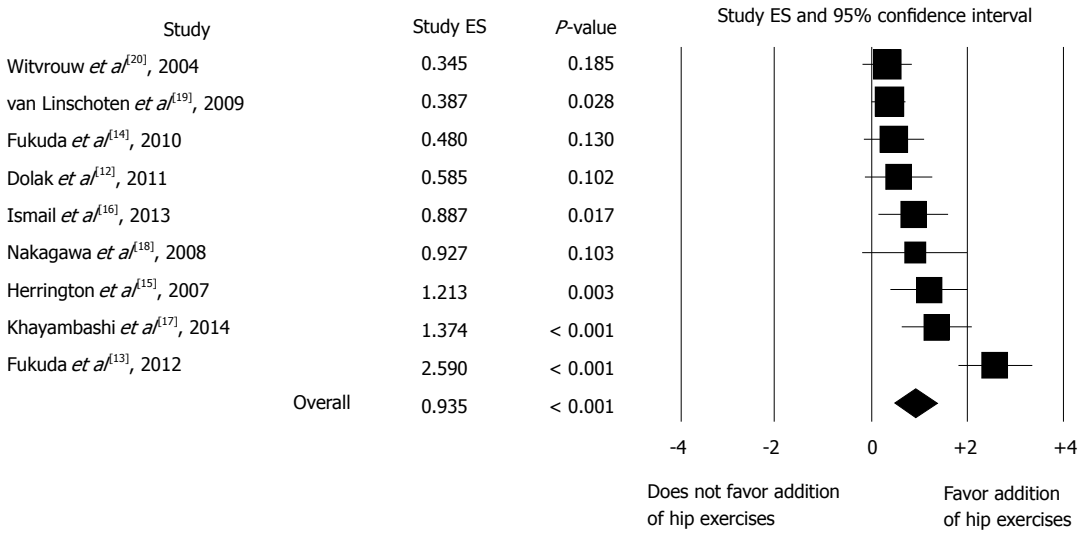


Figure 2 Forest plot of effect sizes from studies that assessed the effect of hip strengthening exercises on pain and function in patients with patellofemoral pain syndrome. A square represents the effect size for a given study with the size of the square being proportional to the weighting of that study in the meta-analysis. A horizontal line indicates the 95%CI for an effect. Studies are arranged from lowest to highest effect size. The diamond at the bottom represents the overall effect size calculated using a random-effects model. The width of the diamond represents the 95%CI for the overall effect size.

studies. Table 2 summarizes the findings of those analyses. None of the experimental factors were able to significantly account for any ES variation. Subgroup analysis of outcome type, time of follow-up, control group type, and hip region targeted with exercise indicated that these variables could not explain a significant amount of the between-study ES variation ($P \geqslant 0.23$). Using meta-regression, it was determined that exercise duration in number of weeks and the total number of treatment sessions also could not explain a significant amount of the between-study ES variation ($P \geqslant 0.38$).

DISCUSSION

The main finding of this study is that in persons diagnosed with patellofemoral pain syndrome the addition of hip-strengthening exercises to traditional physical therapy produced greater improvements in measures of pain and function than traditional therapy alone. Given this finding, developing a targeted program to strengthen both the hip and knee musculature as opposed to alternatives such as strengthening only the knee extensors may lead to fewer number of physical therapy and doctor visits and overall quicker recovery times. Interestingly, the number of exercise sessions and/or number of weeks of exercise intervention did not appear to affect the variation between studies in the effectiveness of the hip-strengthening exercises.

There are several potential limitations of our systematic review and meta-analysis, as well as some methodological concerns with the underlying studies themselves. One possible limitation of our systematic review was publication bias. Publication bias occurs when published research is systematically unrepresentative of the total population of studies[21]. Studies with non-significant and/or negative findings are less likely to be published, and this may influence the overall ES in a meta-analysis that is based largely on published studies. Publication bias was assessed in our review by examination of the funnel plot. Additionally, the Duval and Tweedie's trim and fill adjustment was applied but there was no correction to the overall ES. But because of the relatively few studies, the sensitivity of these analyses could be lacking. Furthermore, we did not rigorously examine the grey literature for unpublished studies.

A second potential limitation of our analysis was the inability to explain the substantial between-study variance in ES. Subgroup analyses and meta-regressions did not identify any experimental factors that could help explain this variance. Many of these analyses probably did not have adequate statistical power because of the limited number of studies in the review and because some subgroups had as few as three studies in them. We tried to assess the ability of gender to explain the between-study variance in ES but could not run a subgroup analysis on gender because there was only one study that used only male subjects. Another concern of the systematic review and meta-analysis is that the exercises performed in each study were categorized by the region of the thigh the exercises targeted (*i.e.*, knee extensors, general hip, posterolateral hip) *vs* listing each specific exercise performed. Thus, we were not able to assess how the performance of specific exercises might explain the between-study variance in ES and enable us to hypothesize a particular exercise to be more effective in reducing pain and improving function in patellofemoral pain syndrome. We also were not able to assess if the exercise intensity for the interventions might explain the between-study variance in ES.

A third potential limitation of our analysis is the

Table 2 Summary of subgroup meta-analyses and meta-regression analyses examining potential moderator variables that might explain the variation in effect size among studies

Moderator variable	Comparison (or slope for continuous variables)	*P* value
Outcome type	Function (n = 8, ES = 0.92) *vs* Pain (n = 9, ES = 0.95)	0.95
Time of follow-up	Immediate (n = 7, ES = 0.79) *vs* Long term (n = 4, ES = 1.111)	0.44
Control group type	Knee extensor strengthening only (n = 6, ES = 1.15) *vs* Knee extensor strengthening plus other (n = 3, ES = 0.56)	0.24
Hip region targeted with exercise	General hip (n = 3, ES = 0.60) *vs* Posterolateral hip (n = 6, ES = 1.13)	0.23
Number of exercise sessions	-0.009/session	0.38
Number of weeks of exercise	-0.066/wk	0.48

ES: Effect size.

inability to completely blind the subjects and therapy providers within the individual studies. All studies are randomized control trials with random assignment of subjects to groups. Subjects that have basic knowledge of anatomy and exercise would likely be aware of which group they were assigned to. While assessors administering the outcome assessment tools (*i.e.,* Pain VAS, NPRS, LEFS, and Kujala AKPS) were blinded to subject group assignment, the subjects themselves completed the outcome tools which consists of questionnaires. Whether an assessor is blinded or not should not affect how a subject completes these forms.

This study's findings provide justification for future research. All study ES including the overall ES were positive, suggesting that despite the large variation in experimental design among studies, the addition of hip strengthening to traditional physical therapy interventions is beneficial in reducing pain and function in patellofemoral pain syndrome when compared to traditional knee-focused interventions alone. Future research examining whether hip-strengthening exercises are equally effective in men and women is important to know, especially when considering that women are more frequently diagnosed with patellofemoral pain syndrome. It would also be helpful, with a larger number of studies, to be able to identify individual hip exercises that are more beneficial than others in decreasing pain and improving function.

COMMENTS

Background

Patellofemoral pain syndrome (PFPS), is a common disorder of the knee. There is no consensus on the etiology of PFPS, however there is an emerging focus on the contribution of proximal structures, *i.e.,* the hip, on PFPS. Traditional therapeutic exercises performed to address PFPS focus on strengthening the knee extensor muscles.

Research frontiers

Interventions targeting the more proximal segment, the hip, in treating PFPS are becoming more of a focus in rehabilitation than targeting the knee extensor muscles alone.

Innovations and breakthroughs

Previous systematic reviews that looked at PFPS only performed review of the literature. The present study included more high quality studies and performed a meta-analysis to quantitatively assess the effect of hip exercises on PFPS compared to traditional interventions.

Applications

The present study suggests that the addition of hip strengthening exercises to traditional therapy improves pain and function in patients with PFPS.

Terminology

Patellofemoral pain syndrome is a diagnosis characterized by anterior knee pain surrounding the patella.

Peer-review

The authors present a well written manuscript with a sound conclusion.

REFERENCES

1 **Bizzini M**, Childs JD, Piva SR, Delitto A. Systematic review of the quality of randomized controlled trials for patellofemoral pain syndrome. *J Orthop Sports Phys Ther* 2003; **33**: 4-20 [PMID: 12570282 DOI: 10.2519/jospt.2003.33.7.F4]

2 **Bolgla LA**, Boling MC. An update for the conservative management of patellofemoral pain syndrome: a systematic review of the literature from 2000 to 2010. *Int J Sports Phys Ther* 2011; **6**: 112-125 [PMID: 21713229]

3 **Petersen W**, Ellermann A, Gösele-Koppenburg A, Best R, Rembitzki IV, Brüggemann GP, Liebau C. Patellofemoral pain syndrome. *Knee Surg Sports Traumatol Arthrosc* 2014; **22**: 2264-2274 [PMID: 24221245 DOI: 10.1007/s00167-013-2759-6]

4 **Rixe JA**, Glick JE, Brady J, Olympia RP. A review of the management of patellofemoral pain syndrome. *Phys Sportsmed* 2013; **41**: 19-28 [PMID: 24113699 DOI: 10.3810/psm.2013.09.2023]

5 **Baldon Rde M**, Serrão FV, Scattone Silva R, Piva SR. Effects of functional stabilization training on pain, function, and lower extremity biomechanics in women with patellofemoral pain: a randomized clinical trial. *J Orthop Sports Phys Ther* 2014; **44**: 240-A8 [PMID: 24568258 DOI: 10.2519/jospt.2014.4940]

6 **Boling M**, Padua D, Marshall S, Guskiewicz K, Pyne S, Beutler A. Gender differences in the incidence and prevalence of patellofemoral pain syndrome. *Scand J Med Sci Sports* 2010; **20**: 725-730 [PMID: 19765240 DOI: 10.1111/j.1600-0838.2009.00996.x]

7 **Nakagawa TH**, Baldon Rde M, Muniz TB, Serrão FV. Relationship among eccentric hip and knee torques, symptom severity and functional capacity in females with patellofemoral pain syndrome. *Phys Ther Sport* 2011; **12**: 133-139 [PMID: 21802040 DOI: 10.1016/j.ptsp.2011.04.004]

8 **Powers CM**. The influence of altered lower-extremity kinematics on patellofemoral joint dysfunction: a theoretical perspective. *J Orthop Sports Phys Ther* 2003; **33**: 639-646 [PMID: 14669959 DOI: 10.2519/jospt.2003.33.11.639]

9 **Ferber R**, Kendall KD, Farr L. Changes in knee biomechanics after a hip-abductor strengthening protocol for runners with patellofemoral pain syndrome. *J Athl Train* 2011; **46**: 142-149 [PMID: 21391799 DOI: 10.4085/1062-6050-46.2.142]

10 **Peters JS**, Tyson NL. Proximal exercises are effective in treating patellofemoral pain syndrome: a systematic review. *Int J Sports Phys Ther* 2013; **8**: 689-700 [PMID: 24175148]

11 **Maher CG**, Sherrington C, Herbert RD, Moseley AM, Elkins M.

Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther* 2003; **83**: 713-721 [PMID: 12882612]

12 **Dolak KL**, Silkman C, Medina McKeon J, Hosey RG, Lattermann C, Uhl TL. Hip strengthening prior to functional exercises reduces pain sooner than quadriceps strengthening in females with patellofemoral pain syndrome: a randomized clinical trial. *J Orthop Sports Phys Ther* 2011; **41**: 560-570 [PMID: 21654093 DOI: 10.2519/jospt.2011.3499]

13 **Fukuda TY**, Melo WP, Zaffalon BM, Rossetto FM, Magalhães E, Bryk FF, Martin RL. Hip posterolateral musculature strengthening in sedentary women with patellofemoral pain syndrome: a randomized controlled clinical trial with 1-year follow-up. *J Orthop Sports Phys Ther* 2012; **42**: 823-830 [PMID: 22951491 DOI: 10.2519/jospt.2012.4184]

14 **Fukuda TY**, Rossetto FM, Magalhães E, Bryk FF, Lucareli PR, de Almeida Aparecida Carvalho N. Short-term effects of hip abductors and lateral rotators strengthening in females with patellofemoral pain syndrome: a randomized controlled clinical trial. *J Orthop Sports Phys Ther* 2010; **40**: 736-742 [PMID: 21041965 DOI: 10.2519/jospt.2010.3246]

15 **Herrington L**, Al-Sherhi A. A controlled trial of weight-bearing versus non-weight-bearing exercises for patellofemoral pain. *J Orthop Sports Phys Ther* 2007; **37**: 155-160 [PMID: 17469667 DOI: 10.2519/jospt.2007.2433]

16 **Ismail MM**, Gamaleldein MH, Hassa KA. Closed kinetic chain exercises with or without additional hip strengthening exercises in management of patellofemoral pain syndrome: a randomized controlled trial. *Eur J Phys Rehabil Med* 2013; **49**: 687-698 [PMID: 23820880]

17 **Khayambashi K**, Fallah A, Movahedi A, Bagwell J, Powers C. Posterolateral hip muscle strengthening versus quadriceps strengthening for patellofemoral pain: a comparative control trial. *Arch Phys Med Rehabil* 2014; **95**: 900-907 [PMID: 24440362 DOI: 10.1016/j.apmr.2013.12.022]

18 **Nakagawa TH**, Muniz TB, Baldon Rde M, Dias Maciel C, de Menezes Reiff RB, Serrão FV. The effect of additional strengthening of hip abductor and lateral rotator muscles in patellofemoral pain syndrome: a randomized controlled pilot study. *Clin Rehabil* 2008; **22**: 1051-1060 [PMID: 19052244 DOI: 10.1177/0269215508095357]

19 **van Linschoten R**, van Middelkoop M, Berger MY, Heintjes EM, Verhaar JA, Willemsen SP, Koes BW, Bierma-Zeinstra SM. Supervised exercise therapy versus usual care for patellofemoral pain syndrome: an open label randomised controlled trial. *BMJ* 2009; **339**: b4074 [PMID: 19843565 DOI: 10.1136/bmj.b4074]

20 **Witvrouw E**, Danneels L, Van Tiggelen D, Willems TM, Cambier D. Open versus closed kinetic chain exercises in patellofemoral pain: a 5-year prospective randomized study. *Am J Sports Med* 2004; **32**: 1122-1130 [PMID: 15262632 DOI: 10.1177/0363546503262187]

21 **Borenstein MHL**, Higgins J, Rothstein HR. Introduction to meta-analysis. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates, 1988

22 **Cohen J.** Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates, 1988

P- Reviewer: Lee SP, Mashreky SR, van der Veen HC **S- Editor**: Qi Y
L- Editor: A **E- Editor**: Lu YJ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i2.125

World J Meta-Anal 2015 April 26; 3(2): 125-132
ISSN 2308-3840 (online)

Red meat intake and the risk of endometrial cancer: Meta-analysis of observational studies

Woong Ju, NaNa Keum, Dong Hoon Lee, Yun Hwan Kim, Seung Cheol Kim, Eric L Ding, Eunyoung Cho

Woong Ju, Yun Hwan Kim, Seung Cheol Kim, Department of Obstetrics and Gynecology, Medical Research Institute, College of Medicine, Ewha Womans University, Seoul 158710, South Korea
NaNa Keum, Dong Hoon Lee, Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, United States
NaNa Keum, Eric L Ding, Department of Nutritron, Harvard School of Public Health, Boston, MA 02115, United States
Eunyoung Cho, Department of Dermatology, the Warren Alpert Medical School of Brown University, Providence, RI 02903, United States
Author contributions: Ju W and Cho E were responsible for the initial plan, study design, conducting the study, data interpretation and manuscript drafting; Ju W, Keum N and Ding EL were responsible for statistical analysis; Ju W, Keum N and Lee DH were responsible for data collection, and data extraction, and data interpretation; Kim YH and Kim SC were responsible for data interpretation and manuscript drafting; Cho E is the guarantor for this paper and has full responsibility for this study.
Supported by A grant of the Korean Health Technology R and D Project, Ministry of Health and Welfare, Republic of Korea, No. HI12C0050.
Conflict-of-interest: All authors have no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work, no other relationships or activities that could appear to have influenced the submitted work.
Data sharing: No additional data available.

Correspondence to: Eunyoung Cho, ScD, Department of Dermatology, the Warren Alpert Medical School of Brown University, 339 Eddy St, Providence, RI 02903, United States. eunyoung.cho@brown.edu
Telephone: +1-617-5252091
Fax: +1-617-5252008
Received: June 16, 2014
Peer-review started: June 16, 2014
First decision: August 7, 2014
Revised: November 17, 2014
Accepted: February 10, 2015
Article in press: February 12, 2015
Published online: April 26, 2015

Abstract

AIM: To evaluate whether red meat intake is related to the risk of endometrial cancer (EC) using meta-analysis.

METHODS: We searched PubMed, EMBASE, and the Cochrane Library up to June 2013, using common keywords related to red meat and EC. Case-control studies and cohort studies comparing the risk of endometrial cancer among categories by the amount of intake were included. Eleven case-control studies and five cohort studies met our criteria. We performed a conventional and a dose-response meta-analysis of case-control studies using the DerSimonian-Laird method for random-effects. For cohort studies we performed a conventional meta-analysis. Publication bias was evaluated using Egger's test.

RESULTS: In the meta-analysis of 11 case-control studies including 5419 cases and 12654 controls, higher red meat consumption was associated with an increased risk of EC [summary relative risk (SRR) = 1.43, 95%CI: 1.15-1.79; I^2 = 73.3% comparing extreme intake categories). In a dose-response analysis, for red meat intake of 100 g/d, SRR was 1.84 (95%CI: 1.64-2.05). In contrast, in the meta-analysis of five prospective studies including a total of 2549 cases among 247746 participants, no significant association between red meat intake and EC risk (SRR = 0.97, 95%CI: 0.85-1.11; I^2 = 4.9% comparing extreme intake categories) was observed.

CONCLUSION: Our meta-analysis found a significant

linear association between red meat intake and EC risk based on case-control studies but this was not confirmed in prospective studies.

Key words: Red meat; Endometrial cancer; Dose-response; Meta-analysis; Observational studies

Core tip: By conducting a dose-response meta-analysis, we found a significant linear association between red meat intake and endometrial cancer risk based on case-control studies. However this association was not confirmed in prospective studies. In our paper, we argue that those findings are attributable to methodological difference between retrospective case-control studies and prospective studies.

Ju W, Keum N, Lee DH, Kim YH, Kim SC, Ding EL, Cho E. Red meat intake and the risk of endometrial cancer: Meta-analysis of observational studies. *World J Meta-Anal* 2015; 3(2): 125-132 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i2/125.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i2.125

INTRODUCTION

Endometrial cancer (EC) is estimated to be the fourth most common cancer in females in the United States in 2013[1]. Risk factors for EC include obesity, postmenopausal hormone replacement therapy (HRT), type Ⅱ diabetes, tamoxifen use, and conditions related to unopposed estrogen such as chronic anovulation and estrogen-only HRT[2-4]. In particular, obesity measured by body-mass index (BMI) has been a well-established risk factor present in almost 50% of women with EC[5]. Recently, red meat intake has received an increasing attention as a potential risk factor for EC[6-8].

An harmful effect of red meat intake has been most studied with colorectal cancer (CRC)[9]. A meta-analysis of 26 cohort studies found an approximately 20% increased risk of colorectal cancer with higher red meat intake[10]. Given that CRC and EC share similar risk factors such as obesity, diabetes, and low physical activity, it has been hypothesized that high red meat intake may increase the risk of EC. This hypothesis is further supported by several mechanisms. Heterocyclic amines generated by overcooking or N-nitroso compounds from proteins have been suggested to act as carcinogens[11]. Iron component in red meat may increase the risk of EC by damaging DNA through oxidative stress[12].

While considerable observational studies have been conducted to examine the effect of red meat intake on EC risk, the epidemiologic relationship remains inconclusive. World Cancer Research Fund (WCRF) panel concluded that there was limited evidence suggesting red meat as a risk factor for EC[13]. While past meta-analysis suggested evidence for a significant inverse association (SRR = 1.59, 95%CI: 1.24-2.05; I^2 = 50.2%, comparing extreme intake categories) and quantified that 100 g/d intake was significantly associated with an approximately 60% increased risk of EC (SRR = 1.6, 95%CI: 1.26-2.03)[14], this meta-analysis included only one prospective study[15]. Several prospective studies have been published afterward[6,7,16,17]. Therefore, we aim to conduct an up-to date dose-response meta-analysis of red meat consumption and the risk of EC based on case-control studies and prospective cohort studies.

MATERIALS AND METHODS

Literature search

We searched PubMed, EMBASE, and the Cochrane Library up to June 2013, using common keywords related to red meat and EC. The keywords were combined as follows: "meat products" as a Medical Subject Headings (MeSH) term or "animal protein" or "red meat" or "meat" and "endometrial neoplasm" as a MeSH term or "uterine cancer" or "corpus cancer" or "endometrial carcinoma" or "uterine carcinoma" or "corpus carcinoma" We also reviewed the bibliographies of relevant articles to locate additional publications. The language of publication was restricted to English.

Selection criteria

We included observational studies that met all of the following inclusion criteria: studies with human subjects, measured outcomes with pathologic confirmation, RR(s) of red meat intake for EC, and statistical information sufficient enough to restore CI(s) for RR(s). If data were duplicated or shared in more than one study, the most comprehensive study with the greatest number of cases was included in the analysis.

Selection of relevant studies

Based on the pre-determined selection criteria, two of the authors (Ju W, Keum N) independently reviewed all studies retrieved from the databases and bibliographies. Two authors screened titles for initial selection and reviewed abstracts/tables of the initially selected articles to identify relevant studies. The reference lists of articles included in our analysis and studies included in the previous meta-analyses were also reviewed for additional papers. Inconsistency between researchers was resolved through discussion based on full articles or in consultation with the third author (Lee DH).

Data extraction

From each study, the following information was extracted: author, year of publication, study design, cohort name, country of study, study period, age range at baseline, types of exposure (red meat, all type of meat except fish), intake range (g/d, g/wk, servings/wk recent), the most fully adjusted measures of association [odds ratio

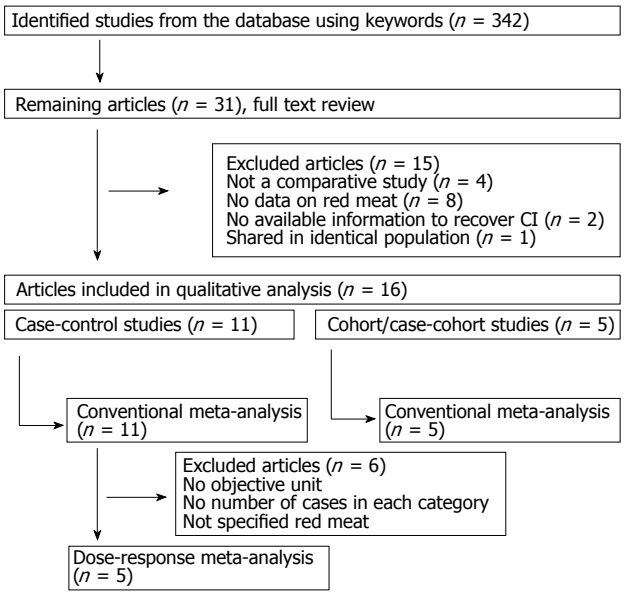


Figure 1 Flow diagram for identification of relevant studies.

(OR), rate ratio(RR), or hazard ratio (HR)], 95%CI, the number of cases, and total number of participants or person-time in each exposure categories. For the quality assessment of each study, exclusion criteria (hysterectomy, prior endometrial cancer), use of validated questionnaire, variables adjusted for confounding were extracted. For case-control studies, information on the types of controls (population *vs* hospital) was additionally extracted.

Statistical analysis

Both conventional and linear dose-response meta-analyses were performed. We assumed a linear dose-response relation in two points. First the previous studies[18,19] with CRC showed linear dose-response association up to 100 g/d intake. Second even in non-linearity model, over 100 g/d did not showed a reduced risk but a blunted slope of increasing risk, which means additional harmful effect of the same direction. In the conventional meta-analysis, random effects model was used to calculate the summary OR for the highest *vs* lowest intake and red meat and the 95%CI. Heterogeneity was assessed using Q test and I^2. Potential sources of heterogeneity were explored using meta-regression based on a priori selected variables. The quality of respective studies was evaluated by performing meta-regression in relation to proper definition of exclusion criteria, types of controls, use of validated questionnaire, type of exposure, adjustment for at least BMI, parity, and menopausal status. Potential publication bias was visually checked using funnel plot and statistically assessed with Egger's regression asymmetry test. Sensitivity analyses were performed by omitting each study at a time.

For the dose-response meta-analysis, a subset of studies included in the conventional meta-analysis was used if they satisfy the following criteria: availability of red meat intake in objectively quantifiable units; having at least three categories of red meat intake including the reference category; availability of number of cases, either number of participants or person-time, and 95%CI for each exposure category, Aggregate method assuming random effects model was used to calculate the SRR of EC associated with 100 g/d intake of red meat and 95%CI. For every study, the mean level of red meat intake in each category was assigned to the corresponding measure of association. In order to calculate the category-specific mean intake for the open-ended highest category, the length of the adjacent interval was assumed; for the open-ended lowest category, 0 g/d was set as a lower limit.

All statistical analyses were conducted using STATA 12 software package (StataCorp, College Station, TX) and based on 2-tailed α set at $P \leqslant 0.05$ for statistical significance. In dose-response meta-analysis we used "Generalized Least Squares" in STATA, which considers the correlation among exposure categories by approximating covariance with GL method.

RESULTS

Figure 1 shows how we identified relevant studies. Initial search identified a total of 342 articles, of which 311studies were excluded for not satisfying the pre-determined selection criteria. We reviewed the full texts of the remaining 31 articles and further excluded 15 articles for the following reasons: not a comparative study (n = 4); no data on red meat (n = 8); no information to recover CI (n = 2); shared in identical population (n = 1). Finally, a total of 16 studies (11 case-control studies, 5 prospective studies) were included in our meta-analyses[6,7,15,16,20-30]. and their characteristics are summarized in Tables 1 and 2, respectively.

Retrospective case-control studies

The 11 case-control studies included a total of 5419 cases and 12654 controls. The year of publication of the included studies ranged between 1993 and 2009. The countries where the studies were conducted were as follows: United States (n = 5), China (n = 2), Greece (n = 1), Italy (n = 1), Sweden (n = 1), and Switzerland (n = 1).

Both conventional and linear dose-response meta-analyses were performed.

Conventional meta-analysis: The SRR comparing the "highest" with the "lowest" categories of red meat intake was 1.43 (95%CI: 1.15-1.79), with considerable heterogeneity (I^2 = 73.3%, $P_{heterogeneity}$ < 0.001), which are shown in Figure 2. In this random-effects meta-analysis, z-value for the overall effect was 3.18 and P-value for test of effect size was 0.001.

None of a priori selected factors such as type of exposure (red meat *vs* all type of meat) and publication year, and BMI adjustment was a significant source of heterogeneity. Publication bias was not evident with Funnel plot showing a symmetric dispersion of studies

Table 1 Summary of case-control studies of red meat consumption and endometrial cancer

Ref.	Location	Study base, subjects	Nutrient	Measurement (unit)	Reference year	Adjustment factors	OR (95%CI)	Meta-analysis	
								Con-ventional	Dose-response
Shu *et al*[20]	China	Population; 268 cases, 278 controls	Red meat	1 liang (50 g)	10 yr prior to interview	Age, number of pregnancies, BMI, caloric intake	2.5	√	√
Potischman *et al*[21]	United States	Population; 399 cases, 296 controls	Red meat	Times/wk	Past few years	Age, BMI, estrogen use, oral contraceptive, number of births, current smoking, education, total calories	1.3 (0.8-2.4)	√	√
Levi *et al*[22]	Swiss	Hospital; 274 cases, 572 controls	Beef	Subjective score	Year before the occurrence of symptoms	Study center, age	2.26 (1.57-3.24)	√	
Goodman *et al*[23]	United States (Hawaii)	Population; 332 cases, 511 controls	Red meat	g	Year prior to diagnosis	Age, ethnicity, pregnancy history, oral contraceptive, diabetes, BMI, total calories	2 (1.1-3.7)	√	√
McCann *et al*[24]	United States	Population; 232 cases, 639 controls	Red meat	Times/mo	2 yr prior to interview	Age, education, BMI, diabetes, hypertension, smoking, age at menarche, parity, oral contraceptive, menopausal status, estrogen	0.8 (0.5-1.4)	√	√
Tavani *et al*[25]	Italy	Hospital; 750 cases, 4770 controls	Red meat	Portions/wk	2 yr preceding diagnosis	Age, year of recruitment, education, smoking, alcohol, fat, fruit, vegetables	1.5 (1.2-1.8)	√	√
Littman *et al*[26]	United States	Population; 679 cases, 944 controls	All meat	Servings/d	5 yr prior to diagnosis	Age, residence, total energy intake, unopposed estrogen, smoking, BMI	1 (0.75-1.4)	√	
Terry *et al*[27]	Sweden	Hospital; 709 cases, 2887 controls	All meat	Quartile	1 yr before diagnosis	Age, BMI, smoking, physical activity, diabetes, fatty fish consumption, total food consumption	1.3 (1.0-1.8)	√	
Dalvi *et al*[28]	United States	Population; 488 cases, 461 controls	Western diet	Quintile	1 yr preceding diagnosis	Age, race, age at menarche, oral contraceptive, parity, daily calorie intake, physical activity, menopause, hormone therapy, BMI	1.5 (0.77-3.0)	√	
Xu *et al*[29]	China	Population; 1204 cases, 1212 controls	Red meat	1 liang (50 g)	Past 5 yr	Age, menopause, diabetes, alcohol, BMI, physical activity, total energy intake, other meat	1.3 (1.0-1.8)	√	√
Petridou *et al*[30]	Greece	Hospital 84 cases 84 controls	All meat	Frequency/mo	1 yr preceding onset of disease	Education, BMI, pregnancy, total energy intake	0.78 (0.53-1.16)	√	

OR: Odds ratio; BMI: Body mass index.

(Figure 3A) and with Egger's test non-significant (P = 0.911, intercept: -0.20, 95%CI: -4.21-3.80).

Meta-regression for assessing the quality of individual studies showed that methodological components such as exclusion criteria, types of controls, validation of dietary questionnaire, and confounding adjustment did not significantly modify the relationship between red meat intake and EC.

Dose-response meta-analysis: Six out of the eleven studies were eligible for the dose-response meta-analysis, including a total of 3364 cases and 10916 controls. Figure 4 illustrates a significant linear dose-response relationship between red meat intake and EC risk. For each 100 g/d increase of red meat intake, SRR was 1.84 (95%CI: 1.64-2.05), with no significant evidence for heterogeneity (I^2 = 21.7%, $P_{heterogeneity}$ = 0.21). In this dose-response meta-analysis, z-value for the overall effect was 10.75 and P-value for test of effect size was less than 0.001.

Prospective observational studies

Four cohort studies and one case-cohort study were identified and included a total 549 cases among 247746 participants. Only conventional meta-analysis was conducted because the five studies did not provide all the necessary information needed for linear dose-response meta-analysis. The SRR of EC for the highest *vs* lowest category of red meat intake was 0.97 (95%CI: 0.85-1.11), with little heterogeneity (I^2 = 4.9%, $P_{heterogeneity}$ = 0.38) (Figure 5). In this random-effects meta-analysis, z-value for the overall effect was 0.44 and P-value for test of effect size was 0.66. No publication bias was indicated by Funnel plot inspection (Figure 3B) and Egger's test (P = 0.142, intercept = 1.61, 95%CI: -0.98-4.20).

Table 2 Summary of cohort, case-cohort studies of red meat consumption and endometrial cancer

Ref.	Location	Cohort	Study design	Population	Nutrient	Intake unit	Reference year	Adjustment factors	RR (95%CI)
Zheng *et al*[15]	United States	Iowa Women's Health Study	Cohort	23070 total 216 cases	Total meat	g/d	Baseline	Age, age at menopause, parity, hormone therapy, total energy intake	1.1
Kabat *et al*[16]	Canada	National Breast Screening Study	Cohort	426 cases 33722 non-cases	Red meat	g/d	Baseline	Age, BMI, menopause, parity, age at menarche, estrogen use, oral contraceptive, total calories, raw vegetable, alcohol intake, physical activity, education	0.86 (0.61-1.22)
van Lonkhuijzen *et al*[6]	Canada	Canadian Study of Diet, Lifestyle, and Health	Case-cohort	56837 total 221 cases 3697 non-cases	Red meat	g/d	Baseline	Age, BMI, age at menarche, number of live birhts, breastfeeding, oral contraceptive, exercise, average calorie, vegetable intake, postmenopausal status, hormone therapy	1.62 (0.86-3.08)
Genkinger *et al*[7]	Sweden	Swedish Mammography Cohort	Cohort	60895 total 720 cases	Red meat	g/wk	Baseline	Age, energy, BMI, parity, and education	1.06 (0.68-1.66)
Arem *et al*[17]	United States	NIH-AARP Diet and Health Study	Cohort	72796 total 966 cases	Red meat	g/1000 kcal	Baseline	Age, BMI, smoking, total energy intake, age at menarche, age at first child's birth, parity, age at menopause, hormone therapy, oral contraceptive, diabetes, physical activity	0.91 (0.77-1.08)

RR: Risk ratio.

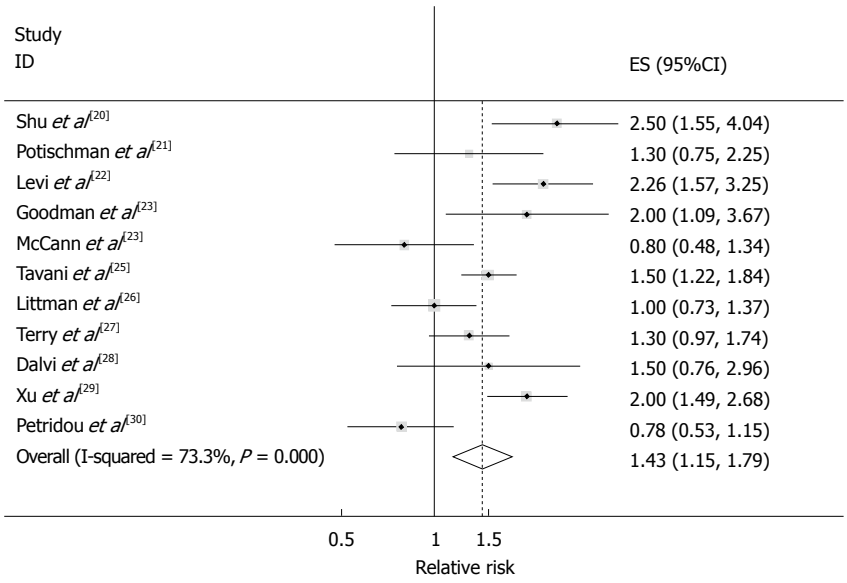


Figure 2 Random-effects meta-analysis of red meat intake and endometrial cancer risk in case-control studies, which shows I^2 = 73.3%, $P_{heterogeneity}$ < 0.000.

DISCUSSION

In this conventional and dose-response meta-analysis of observational studies, we found inconsistent results with retrospective case-control studies suggesting a significant increase in EC risk approximately by 84% associated with 100 g/d intake of red meat while with prospective observational studies indicating no such association.

Conventional meta-analysis which dichotomizes continuous exposures as highest *vs* lowest categories and collapses intake categories regardless of intake level ignores absolute intake and is not optimal to elucidate a dose-response relationship between dietary intake and disease outcomes. This approach may be particularly problematic in populations with wide intake range where cutoffs of intake categories are substantially different. A highest dosage in one study could be a reference dosage in another study, which means that the SRR from conventional meta-analysis

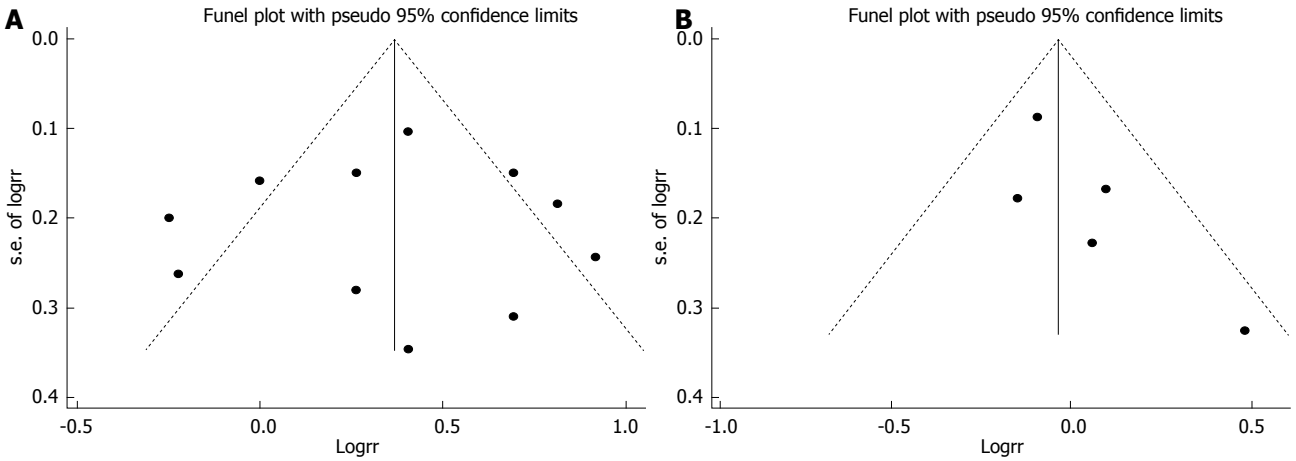


Figure 3 Funnel plot of eleven case-control studies (A) and five cohort studies (B) which were included in conventional meta-analysis.

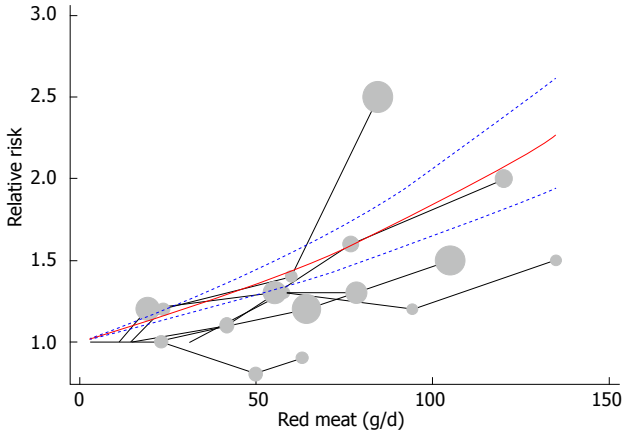


Figure 4 Dose-response spaghetti plot of red meat intake and endometrial cancer risk in case-control studies with pooled OR (solid line) and 95%CI (dotted lines).

might not give proper weights proportional to the dose-related risk. Dose-response meta-analysis has its practical importance than conventional meta-analysis of comparing extreme intake groups by providing summary estimates per absolute amount of intake, which can be easily incorporated in cancer prevention strategy and dietary policy.

Our findings based on 11 retrospective case-control studies are consistent with results from the previous meta-analysis of 7 case-control studies. The SRRs of 1.43 (95%CI: 1.15-1.79) for the highest *vs* lowest category of red meat intake and of 1.84 (95%CI: 1.64-2.05) per 100 g/d increment in the intake were similar to corresponding SRRs of 1.59 (95%CI: 1.24-2.05) and 1.60 (95%CI: 1.26-2.03) in the previous meta-analysis[14].

Despite those cumulative evidences arising from case-control studies suggests positive associations, most prospective studies have not supported an increased risk of EC associated with red meat intake. In 2000, Trichopoulou *et al*[31] summarized the nutritional etiology of various forms of cancer in their review, but did not find an evidence for a positive relationship between red meat intake and EC. In 2007, the Panel of WCRF concluded that although evidence for harmful effects of red meat and processed meat on EC risk was stronger than it had been in the mid-1990s, overall evidence remained suggestive, at most[13]. The most recent meta-analysis performed by Bandera *et al*[14] did not reported a pooled RR for cohort studies because they included only one cohort study by Zheng *et al*[15], which found no significant association (OR, 1.10, 95%CI: 0.79-1.52). Our updated meta-analysis of five prospective studies still suggests that red meat intake was not significantly associated with the risk of EC (SRR, 0.97, 95%CI: 0.85-1.11; I^2 = 4.9%). The discrepant results between retrospective case-control and prospective observational studies can be partially explained by several issues related to the measurement of red meat intake. First, retrospective *vs* prospective nature of measurement is an important consideration. In retrospective case-control studies, red meat intake was assessed after diagnosis of EC and thus, participants' knowledge about disease status could lead to differential measurement error. For instance, since cases are more sensitive to their dietary intake than controls in general, it is entirely possible that cases over-report their red meat intake, which could lead to the observed positive association between red meat intake and EC risk. In contrast, in prospective studies, red meat intake was assessed prior to the diagnosis of EC and thus, measurement errors are likely to be random with respect to disease status. Random measurement error of dichotomous exposure mostly attenuates a measure of association toward the null and thus, could partially account for the null association observed in our meta-analysis of cohort studies.

Second, difference in reference year for exposure measurement relates to differential assumption regarding etiologic window of red meat intake in affecting EC risk, which could lead to inconsistent results. In case-control

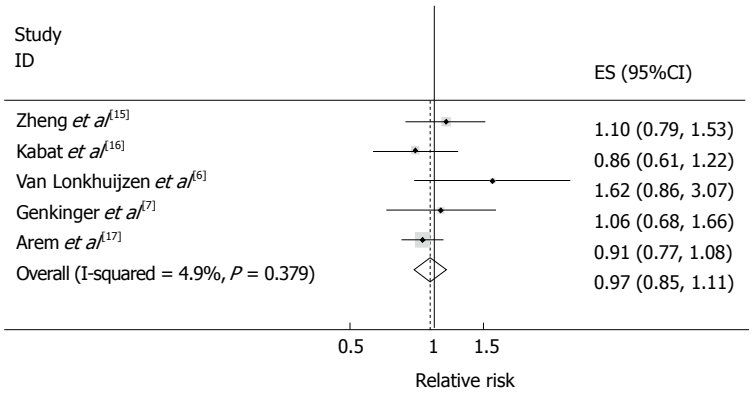


Figure 5 Random-effects meta-analysis of red meat intake and endometrial cancer risk in prospective studies.

studies, participants were asked to recall red meat intake during 1-5 years before the assessment. This inherently assumes that recent red meat intake is relevant to current EC risk. In cohort studies, baseline assessment of red meat intake is usually assumed to represent a long-term diet and participants were followed-up for 7 to 21 years. Thus, long-term red meat intake was assumed to modulate EC risk. Thus, it is possible that case-control studies and prospective observational studies addressed different questions regarding the red meat intake-EC relationship and thus, reached different conclusions.

Our study has several limitations. First we only investigated a role of red meat intake. Animal derived fat or processed meat has also been reported as risk factors of EC[32-35]. We focused on red meat in the context of consumers' intuition contrasting red meat from white meat or fish as people usually classify red meat as one of representative category when they shop at a market or order at a restaurant. Second we often used rather arbitrary intake as a representative intake of corresponding category when the mean or median intakes were not provided in dose-response meta-analysis. Since the representative dosage should not be missed in each range for dose-response meta-analysis, such an extrapolation can be accepted as technically inevitable.

Nonetheless this study has strength in that it provides updated evidence regarding the relationship between red meat intake and EC risk by incorporating recently published prospective studies.

In summary, our meta-analysis found a significant linear association between red meat intake and EC risk based on case-control studies but this was not confirmed in prospective studies. This discrepancy seems to be attributable to the differences in robustness against biases and reference year of assessment of red meat intake between the retrospective and prospective studies. When the implication of the current study is addressed, however, it should be considered that the quality of evidence from cohort studies be higher because it is more likely to represent the real world situation.

COMMENTS

Background

The incidence of endometrial cancer (EC) is increasing as the life styles become westernized globally. EC is estimated to be the fourth most common cancer in females in the. The association between red meat intake and the risk of EC is currently unclear.

Research frontiers

While past meta-analysis suggested evidence for a significant inverse association and quantified that 100 g/d intake was significantly associated with an approximately 60% increased risk of EC, this meta-analysis included only one prospective study, which could not be sufficient at this time because several prospective studies have been published afterward.

Innovations and breakthroughs

The aim of the current study was to conduct an up-to date dose-response meta-analysis of red meat consumption and the risk of EC based on case-control studies and prospective cohort studies.

Applications

This meta-analysis found a significant linear association between red meat intake and EC risk based on case-control studies but this was not confirmed in prospective studies. More results from prospective studies with long-term follow up are in need to confirm the association between red meat intake and the risk of EC.

Terminology

EC is a carcinoma originated from the inner mucous membrane of mammalian uterus, which is also referred as uterine cancer or corpus cancer.

Peer-review

This review article is well written and will contribute to the clinical practice of the readers.

REFERENCES

1 **Siegel R**, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin* 2013; **63**: 11-30 [PMID: 23335087 DOI: 10.3322/caac.21166]

2 **Zeleniuch-Jacquotte A**, Akhmedkhanov A, Kato I, Koenig KL, Shore RE, Kim MY, Levitz M, Mittal KR, Raju U, Banerjee S, Toniolo P. Postmenopausal endogenous oestrogens and risk of endometrial cancer: results of a prospective study. *Br J Cancer* 2001; **84**: 975-981 [PMID: 11286480 DOI: 10.1054/bjoc.2001.1704]

3 **Kaaks R**, Lukanova A, Kurzer MS. Obesity, endogenous hormones, and endometrial cancer risk: a synthetic review. *Cancer Epidemiol Biomarkers Prev* 2002; **11**: 1531-1543 [PMID: 12496040]

4 **Calle EE**, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *N Engl J Med* 2003; **348**: 1625-1638 [PMID: 12711737 DOI: 10.1056/NEJMoa021423]

5 **Parslov M**, Lidegaard O, Klintorp S, Pedersen B, Jønsson L, Eriksen PS, Ottesen B. Risk factors among young women with endometrial cancer: a Danish case-control study. *Am J Obstet Gynecol* 2000; **182**: 23-29 [PMID: 10649152]

6 **van Lonkhuijzen L**, Kirsh VA, Kreiger N, Rohan TE. Endometrial cancer and meat consumption: a case-cohort study. *Eur J Cancer Prev* 2011; **20**: 334-339 [PMID: 21422932 DOI: 10.1097/CEJ.0b013e328344747c]

7 **Genkinger JM**, Friberg E, Goldbohm RA, Wolk A. Long-term dietary heme iron and red meat intake in relation to endometrial cancer risk. *Am J Clin Nutr* 2012; **96**: 848-854 [PMID: 22952183]

8 **Lanou AJ**, Svenson B. Reduced cancer risk in vegetarians: an analysis of recent reports. *Cancer Manag Res* 2010; **3**: 1-8 [PMID: 21407994 DOI: 10.2147/CMR.S6910]

9 **Norat T**, Riboli E. Meat consumption and colorectal cancer: a review of epidemiologic evidence. *Nutr Rev* 2001; **59**: 37-47 [PMID: 11310774]

10 **Huxley RR**, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. *Int J Cancer* 2009; **125**: 171-180 [PMID: 19350627 DOI: 10.1002/ijc.24343]

11 **Larsson SC**, Orsini N, Wolk A. Processed meat consumption and stomach cancer risk: a meta-analysis. *J Natl Cancer Inst* 2006; **98**: 1078-1087 [PMID: 16882945 DOI: 10.1093/jnci/djj301]

12 **Kabat GC**, Rohan TE. Does excess iron play a role in breast carcinogenesis? An unresolved hypothesis. *Cancer Causes Control* 2007; **18**: 1047-1053 [PMID: 17823849 DOI: 10.1007/s10552-007-9058-9]

13 **Marmot M,** Atinmo T, Byers T, Chen J, Hirohata T, Jackson A, James W, Kolonel L, Kumanyika S, Leitzmann C. Food, nutrition, physical activity, and the prevention of cancer: a global perspective, the American Institute for Cancer Research. Available from: URL: http: //health-equity.pitt.edu/868/

14 **Bandera EV**, Kushi LH, Moore DF, Gifkins DM, McCullough ML. Consumption of animal foods and endometrial cancer risk: a systematic literature review and meta-analysis. *Cancer Causes Control* 2007; **18**: 967-988 [PMID: 17638104 DOI: 10.1007/s10552-007-9038-0]

15 **Zheng W**, Kushi LH, Potter JD, Sellers TA, Doyle TJ, Bostick RM, Folsom AR. Dietary intake of energy and animal foods and endometrial cancer incidence. The Iowa women's health study. *Am J Epidemiol* 1995; **142**: 388-394 [PMID: 7625403]

16 **Kabat GC**, Miller AB, Jain M, Rohan TE. Dietary iron and haem iron intake and risk of endometrial cancer: a prospective cohort study. *Br J Cancer* 2008; **98**: 194-198 [PMID: 18059399 DOI: 10.1038/sj.bjc.6604110]

17 **Arem H**, Gunter MJ, Cross AJ, Hollenbeck AR, Sinha R. A prospective investigation of fish, meat and cooking-related carcinogens with endometrial cancer incidence. *Br J Cancer* 2013; **109**: 756-760 [PMID: 23695021 DOI: 10.1038/bjc.2013.252]

18 **Chan DS**, Lau R, Aune D, Vieira R, Greenwood DC, Kampman E, Norat T. Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. *PLoS One* 2011; **6**: e20456 [PMID: 21674008 DOI: 10.1371/journal.pone.0020456]

19 **Norat T**, Lukanova A, Ferrari P, Riboli E. Meat consumption and colorectal cancer risk: dose-response meta-analysis of epidemiological studies. *Int J Cancer* 2002; **98**: 241-256 [PMID: 11857415]

20 **Shu XO**, Zheng W, Potischman N, Brinton LA, Hatch MC, Gao YT, Fraumeni JF. A population-based case-control study of dietary factors and endometrial cancer in Shanghai, People's Republic of China. *Am J Epidemiol* 1993; **137**: 155-165 [PMID: 8452119]

21 **Potischman N**, Swanson CA, Brinton LA, McAdams M, Barrett RJ, Berman ML, Mortel R, Twiggs LB, Wilbanks GD, Hoover RN. Dietary associations in a case-control study of endometrial cancer. *Cancer Causes Control* 1993; **4**: 239-250 [PMID: 8318640]

22 **Levi F**, Franceschi S, Negri E, La Vecchia C. Dietary factors and the risk of endometrial cancer. *Cancer* 1993; **71**: 3575-3581 [PMID: 8490907]

23 **Goodman MT**, Hankin JH, Wilkens LR, Lyu LC, McDuffie K, Liu LQ, Kolonel LN. Diet, body size, physical activity, and the risk of endometrial cancer. *Cancer Res* 1997; **57**: 5077-5085 [PMID: 9371506]

24 **McCann SE**, Freudenheim JL, Marshall JR, Brasure JR, Swanson MK, Graham S. Diet in the epidemiology of endometrial cancer in western New York (United States). *Cancer Causes Control* 2000; **11**: 965-974 [PMID: 11142531 DOI: 10.1023/a: 1026551309873]

25 **Tavani A**, La Vecchia C, Gallus S, Lagiou P, Trichopoulos D, Levi F, Negri E. Red meat intake and cancer risk: a study in Italy. *Int J Cancer* 2000; **86**: 425-428 [PMID: 10760833]

26 **Littman AJ**, Beresford SA, White E. The association of dietary fat and plant foods with endometrial cancer (United States). *Cancer Causes Control* 2001; **12**: 691-702 [PMID: 11562109 DOI: 10.1023/a: 1011292003586]

27 **Terry P**, Vainio H, Wolk A, Weiderpass E. Dietary factors in relation to endometrial cancer: a nationwide case-control study in Sweden. *Nutr Cancer* 2002; **42**: 25-32 [PMID: 12235647 DOI: 10.1207/s15327914nc421_4]

28 **Dalvi TB**, Canchola AJ, Horn-Ross PL. Dietary patterns, Mediterranean diet, and endometrial cancer risk. *Cancer Causes Control* 2007; **18**: 957-966 [PMID: 17638105 DOI: 10.1007/s10552-007-9037-1]

29 **Xu WH**, Dai Q, Xiang YB, Zhao GM, Zheng W, Gao YT, Ruan ZX, Cheng JR, Shu XO. Animal food intake and cooking methods in relation to endometrial cancer risk in Shanghai. *Br J Cancer* 2006; **95**: 1586-1592 [PMID: 17060930 DOI: 10.1038/sj.bjc.6603458]

30 **Petridou E**, Kedikoglou S, Koukoulomatis P, Dessypris N, Trichopoulos D. Diet in relation to endometrial cancer risk: a case-control study in Greece. *Nutr Cancer* 2002; **44**: 16-22 [PMID: 12672637 DOI: 10.1207/s15327914nc441_3]

31 **Trichopoulou A**, Lagiou P, Kuper H, Trichopoulos D. Cancer and Mediterranean dietary traditions. *Cancer Epidemiol Biomarkers Prev* 2000; **9**: 869-873 [PMID: 11008902]

32 **Cui X**, Rosner B, Willett WC, Hankinson SE. Dietary fat, fiber, and carbohydrate intake in relation to risk of endometrial cancer. *Cancer Epidemiol Biomarkers Prev* 2011; **20**: 978-989 [PMID: 21393567 DOI: 10.1158/1055-9965.epi-10-1089]

33 **McTiernan A**, Irwin M, Vongruenigen V. Weight, physical activity, diet, and prognosis in breast and gynecologic cancers. *J Clin Oncol* 2010; **28**: 4074-4080 [PMID: 20644095 DOI: 10.1200/jco.2010.27.9752]

34 **Prentice RL**, Thomson CA, Caan B, Hubbell FA, Anderson GL, Beresford SA, Pettinger M, Lane DS, Lessin L, Yasmeen S, Singh B, Khandekar J, Shikany JM, Satterfield S, Chlebowski RT. Low-fat dietary pattern and cancer incidence in the Women's Health Initiative Dietary Modification Randomized Controlled Trial. *J Natl Cancer Inst* 2007; **99**: 1534-1543 [PMID: 17925539 DOI: 10.1093/jnci/djm159]

35 **Daniel CR**, Cross AJ, Graubard BI, Hollenbeck AR, Park Y, Sinha R. Prospective investigation of poultry and fish intake in relation to cancer risk. *Cancer Prev Res* (Phila) 2011; **4**: 1903-1911 [PMID: 21803982]

P- Reviewer: Itamochi H, Trkulja V, Yokoyama Y, Zhang YJ
S- Editor: Ji FF **L- Editor**: A **E- Editor**: Lu YJ

Published by **Baishideng Publishing Group Inc**
8226 Regency Drive, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

ISSN 2308-3840 (online)

World Journal of *Meta-Analysis*

World J Meta-Anal 2015 June 26; 3(3): 133-187

Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

A peer-reviewed, online, open-access journal of meta-analysis

Editorial Board

2013-2018

The *World Journal of Meta-Analysis* Editorial Board consists of 400 members, representing a team of worldwide experts in clinical meta-analysis research. They are from 41 countries, including Argentina (2), Australia (3), Austria (1), Belgium (5), Brazil (11), Canada (16), Chile (3), China (121), Croatia (1), Egypt (1), Finland (4), France (2), Germany (9), Greece (11), Hungary (2), India (12), Iran (2), Ireland (1), Israel (2), Italy (41), Japan (5), Lithuania (1), Malaysia (1), Netherlands (8), New Zealand (1), Norway (1), Peru (1), Poland (4), Portugal (6), Romania (1), Saudi Arabia (4), Singapore (3), South Africa (1), South Korea (7), Spain (8), Sri Lanka (2), Sudan (1), Switzerland (2), Thailand (4), Turkey (3), United Kingdom (22), and United States (64).

Alessandro Cucchetti, *Bologna*
Gianfranco Damiani, *Rome*
Fabrizio D'Ascenzo, *Turin*
Massimo Del Fabbro, *Milano*
Valeria Fadda, *Florence*
Alessandro Fancellu, *Sassari*
Giuseppe Ferrante, *Milan*
Virginia Festa, *Rome*
Francesco Fiorica, *Ferrara*
Guglielmo Giraldi, *Rome*
Jenny Guidi, *Bologna*
Lorenzo Loffredo, *Rome*
Andrea Messori, *Firenze*
Eliano Pio Navarese, *Bydgoszcz*
Stefano Omboni, *Solbiate Arno*
Alvisa Palese, *Udine*
Stefano Palomba, *Reggio Emilia Stefano*
Carlo Perricone, *Rome*
Mario Petretta, *Naples*
Alessandro Pezzini, *Brescia*
Gianluca Pontone, *Milan*
Paolo Emilio Puddu, *Rome*
Andrea Rognoni, *Novara*
Giuseppe Scalabrino, *Milan*
Città della Salute e della Scienza, *Turin*
Fabrizio Sgolastra, *L'Aquila*
Maria Lucia Specchia, *Rome*
Stefano Trastulli, *Terni*
Fabio Tine, *Palermo*
Nereo Vettoretto, *Chiari*
Alberto Vianello, *Perugia*
Luigi Zorcolo, *Cagliari*

Japan

Nguyen Tien Huy, *Nagasaki*
Hiroharu Kamioka, *Tokyo*
Koji Kawakami, *Kyoto*
Keitaro Matsuo, *Nagoya*
Kazushi Okamoto, *Nagoya*

Lithuania

Edmundas Kadusevicius, *Kaunas*

Malaysia

SP Pani, *Ipoh*

Netherlands

Michel van den Bekerom, *Amsterdam*
Dimitra Dodou, *Delft*
D Haverkamp, *Amsterdam*
Vassilios Koussoulas, *Drachten*
BJ Polder, *Nijmegen*
Theo Stijnen, *Leiden*
RNM Weijers, *Amsterdam*
Joost de Winter, *Delft*

New Zealand

Shaofeng Li, *Auckland*

Norway

Eivind Berge, *Oslo*

Peru

Rafael Bolaños Díaz, *Lima*

Poland

Maciej Banach, *Lodz*
Krzysztof Jonderko, *Sosnowiec*
Jolanta Lissowska, *Warsaw*
Maciej Plaszewski, *Warsaw*

Portugal

Daniel Caldeira, *Lisboa*
J Costa, *Lisbon*
Ana Miguel, *Coimbra*
Manuel Morgado, *Covilh*
Bárbara Peleteiro, *Porto*
Rui Torres, *Paredes*

Romania

Fratila Ovidiu, *Oradea*

Saudi Arabia

Hazem M Al-Mandeel, *Riyadh*
Ezzeldin M Ibrahim, *Jeddah*
Mutahir A Tunio, *Riyadh*
Hayfaa A Wahabi, *Riyadh*

Singapore

Nikos LD Chatzisarantis, *Singapore*
Edwin Choon Wyn Lim, *Singapore*
Roger Ho, *Singapore*

South Africa

Alaine Umubyeyi Nyaruhirira, *Pretoria*

South Korea

Jung-Hee Kim, *Cheonan*
Hyangsook Lee, *Seoul*
Myeong Soo Lee, *Daejeon*
Chi-Un Pae, *Bucheon*
Yong Hyun Park, *Yangsan*
Jae Hong Seo, *Seoul*
Yong Sang Song, *Seoul*

Spain

Pablo Avanzas, *Oviedo*
Joan Cid, *Barcelona*
Joaquin de Haro, *Madrid*
Joan Guardia-Olmos, *Barcelona*
Nabil Halaihlel, *Zaragoza*
JA Monge-Argilés, *Alicante*
Raul Moreno, *Madrid*
Inés Velasco, *Aracena*

Sri Lanka

Ranil Jayawardena, *Colombo*
Priyanga Ranasinghe, *Colombo*

Switzerland

Jay P Singh, *Zurich*
Giorgio Treglia, *Bellinzona*

Sudan

Samir MH Shaheen, *Khartoum*

Thailand

Chuenjid Kongkaew, *Phitsanulok*
Manop Pithukpakorn, *Bangkok*
Piyamitr Sritara, *Bangkok*
Surasak Saokaew, *Phayao*

Turkey

Nese Demirturk, *Afyonkarahisar*
Nilüfer Ozabaci, *Eskisehir*
Ilke Sipahi, *Istanbul*

United Kingdom

Omar M Aboumarzouk, *Wales*
Abeer Al-Namankany, *London*
Ernest A Azzopardi, *Cardif*
Umberto Benedetto, *Cambridge*
Joanne Brooke, *London*
Noriko Cable, *London*
David Chan, *Cardiff*
YC Cheong, *Southampton*
Andrew Currie, *Harrow*
G Nabi, *Dundee*

Lesley A Anderson, *Belfast*
Valentina Gallo, *London*
Gianpiero Gravante, *Leicester*
Peter N Lee, *Surrey*
Igho Onakpoya, *Oxford*
Ashish Pradhan, *Huntingdon*
Evridiki Patelarou, *London*
Jian-Qing Shi, *Newcastle*
Surendra P Singh, *Wolverhampton*
Natalie Taylor, *Leeds*
Yousef Shahin, *Hull*
Zheng Ye, *Cambridge*

United States

L Joseph Su, *Rockville*
Olusola Adesope, *Pullman*
Mike Allen, *Milwaukee*
Bhupinder Anand, *Bellaire*
Stephen Aronoff, *Philadelphia*
KoKo Aung, *San Antonio*
William L Baker, *Farmington*
Moritz C Wyler von Ballmoos, *Milwaukee*
Matthew L Bechtold, *Columbia*
Atul Bhardwaj, *Hershey*
Somjot S Brar, *Los Angeles*
Hui Cai, *Nashville*
Subhash Chandra, *Towson*
Wen-Pin Chang, *Omaha*
Yong Chen, *North Wales*
Myunghan Choi, *Phoenix*
John Coverdale, *Houston*
Prakash C Deedwania, *Fresno*
Eugene Demidenko, *Hanover*
Hong-Wen Deng, *New Orleans*
Eric M Deshaies, *Syracuse*
Tao Fan, *Whitehouse Station*
Shinga Feresu, *Bloomington*
Janvier Gasana, *Miami*
Kaveh Hajifathalian, *Boston*
Mohammad Obaidul Hoque, *Baltimore*
Larissa R Brunner Huber, *Charlotte*
Imran H Iftikhar, *Columbia*
Vijayvel Jayaprakash, *Buffalo*
Xuezhi Jiang, *Weat Reading*
Shuo Jiao, *Seattle*
Evelyn Johnson, *Boise*
Le Kang, *Silver Spring*
SR Kapadia, *Cleveland*
Lior Katz, *Houston*
Daniel M Laskin, *Richmond*
Yu Liang, *Foster City*
Paul Ellis Marik, *Norfolk*
Lynne V McFarland, *Seattle*
Marcovalerio Melis, *New York*
Brian J Miller, *Augusta*
Pavlos Msaouel, *New York*
Joshua E Muscat, *Hershey*
Chee Yuan Ng, *Loma Linda*
Nghi C Nguyen, *Saint Louis*
Brandi S Niemeier, *Whitewater*
Thomas D Parsons, *Denton*
Nidal Abi Rafeh, *New Orleans*
Praveen Roy, *Marshfield*
Ali Salavati, *Philadelphia*
Ankur Sethi, *Chicago*
Tatyana A Shamliyan, *Minneapoli*
Qian Shi, *Rochester*
Zhongjie Shi, *Philadelphia*
Param Puneet Singh, *Chicago*
KV Slavin, *Chicago*
Ali El Solh, *Buffalo*
Jielin Sun, *Winston-Salem*
Richard G Trohman, *Chicago*
Laurah Turner, *Cincinnati*
Sheila Wilhelm, *Detroit*
Alex K Wong, *Los Angeles*
Xiaohui Xu, *Gainesville*
Lu Yin, *Nashville*

WJMA World Journal of Meta-Analysis

Contents

Bimonthly Volume 3 Number 3 June 26, 2015

Contents

World Journal of Meta-Analysis
Volume 3 Number 3 June 26, 2015

NAME OF JOURNAL
World Journal of Meta-Analysis

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Bimonthly

EDITOR-IN-CHIEF
Giuseppe Biondi-Zoccai, MD, Assistant Professor, Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
World Journal of Meta-Analysis
Room 903, Building D, Ocean International Center, No. 62 Dongsihuan Zhonglu, Chaoyang District, Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
8226 Regency Drive,
Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
June 26, 2015

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i3.133

World J Meta-Anal 2015 June 26; 3(3): 133-138
ISSN 2308-3840 (online)

EDITORIAL

Evolving role of salvage reirradiation: Is global harmonization required before treatment guidelines can be developed?

Natalie Logie, C Suzanne Drodge, Oleksandr Boychak, Alysa Fairchild

Natalie Logie, Alysa Fairchild, Department of Radiation Oncology, Cross Cancer Institute, Edmonton, Alberta T6G 1Z2, Canada
C Suzanne Drodge, Dr H Bliss Murphy Cancer Centre, St John's, NL A1B 3V6, Canada
Oleksandr Boychak, UPMC Whitfield Cancer Centre, Butlers-town North, Waterford, Ireland

Author contributions: All authors contributed to this work.

Conflict-of-interest: The authors have no conflicts of interest to declare.

Correspondence to: Dr. Alysa Fairchild, MD, FRCPC, Department of Radiation Oncology, Cross Cancer Institute, 11560 University Avenue, Edmonton, Alberta T6G 1Z2, Canada. alysa@ualberta.ca
Telephone: +1-780-4328516
Fax: +1-780-4328380

Received: January 28, 2015
Peer-review started: January 28, 2015
First decision: March 6, 2015
Revised: March 31, 2015
Accepted: April 16, 2015
Article in press: April 20, 2015
Published online: June 26, 2015

Abstract

Up to 90% of patients initially treated with curative-intent radiotherapy (RT) will experience locoregional failure. Historically, reirradiation (ReRT) was offered purely with palliative intent, if considered at all, due to concerns surrounding toxicity, tolerance of normal tissues, and choice of appropriate dose schedule. With technological advancements in RT delivery, coupled with longer survival in many malignancies secondary to improvements in systemic therapy, a small subset of patients presenting with localized recurrence is increasingly being offered salvage ReRT. However, this is largely on an ad hoc basis, guided mainly by small retrospective, single-institution reports. The patient population retreated, RT modality, dose received, degree of attrition and follow-up are extremely variable. The opportunity presently exists to apply lessons learned from the harmonization of the research efforts within the bone metastases community to the salvage ReRT situation: the adoption of common endpoints, minimum features to be incorporated into clinical trial design, and methods of data analysis and reporting. The ReRT data available must be harmonized so that valid, clinically applicable conclusions can be drawn. Collaboration in the form of an international registry of prospectively collected outcomes of patients reirradiated for cure for a variety of tumour sites would further support the evolution of Radiation Oncology towards personalized medicine, and away from the current "one-dose-fits-all" approach.

Key words: Reirradiation; Salvage; Treatment planning; Toxicity; Registry; Dose; Radiotherapy

Core tip: Given the heterogeneity of the available reirradiation evidence, an international registry would provide a foundation on which to base consensus recommendations regarding many of the outstanding

questions surrounding patient selection and treatment planning. Inter-centre collaboration will be required to build a critical mass of data sufficient for robust statistical analysis; however, in order to achieve this, global harmonization is needed. Standardized nomenclature would facilitate consistent coding of treated volumes, doses, toxicity rates, and quality of life outcomes. A registry would also assist in determining the feasibility of both phase Ⅱ prospective studies and meta-analysis of currently available data.

Logie N, Drodge CS, Boychak O, Fairchild A. Evolving role of salvage reirradiation: Is global harmonization required before treatment guidelines can be developed? *World J Meta-Anal* 2015; 3(3): 133-138 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i3/133.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i3.133

INTRODUCTION

Depending on the type and stage of cancer at first presentation, up to 90% of patients initially treated with curative-intent radiotherapy (RT) will experience locoregional failure[1]. For example, in breast cancer, despite local radiation, locoregional recurrences occur in up to 14% at 18 years[2], and after RT for non-melanoma skin cancer, in-field recurrence has been reported in up to 16%[3]. Pelvic recurrence occurs in 20%-40% of patients after radical radiation or surgery for gynecologic cancer[4]. In lung cancer, approximately one-third of those treated with radical chemoRT will develop a locoregional recurrence within five years[5,6]. Likewise, locoregional failure is the dominant pattern of failure after radical chemoRT for both head and neck cancer[7] and glioblastoma multiforme, with the latter recurring more than 90% of the time despite optimal up-front treatment[8].

At the time of local recurrence, treatment options may include resection, systemic therapy, laser or radiofrequency ablation, cryotherapy, hyperthermia, or photodynamic therapy. However, these options are not universally available; each has different and often stringent eligibility criteria; strength of supporting evidence varies; and in some, proof of long-term efficacy is lacking. Reirradiation (ReRT) with repeat conventional external beam RT, highly conformal RT such as stereotactic body RT (SBRT) (Table 1), proton therapy, heavy ions or brachytherapy may also be considerations in those experiencing recurrence who have exhausted or are not eligible for other forms of therapy.

RERT: THE CASE FOR HARMONIZATION

Historically, the use of ReRT has been limited by concerns surrounding toxicity, tumour radioresistance, and lack of robust evidence[1,9,10]. The complexity of delivering RT a second time to the same volume has been exacerbated by a dearth of individual radiation oncologist experience, a lack of confidence in the ability to reproduce the previous treatment's dosimetric parameters, a scarcity of adequate data on recovery of normal organs after radiation injury, and the absence of guidelines supporting approaches to optimal RT planning. In a 2008 Canadian national survey, the majority of respondents reported a lack of departmental guidelines and "enthusiasm" for instituting ReRT[1]. Controversy surrounds the choice of appropriate prescription in the context of the initial dose and field arrangement, and the best combination of steps to limit further damage to normal structures which have already received maximum or near-tolerance doses. Consequently, repeat RT in past was primarily done with palliative intent[11]. This is echoed by results of the 2008 survey, in which only 32% of respondents would offer ReRT for salvage but 99% would institute ReRT if quality of life could be improved[1].

The situation where both RT courses are delivered with palliative intent has been extensively studied in the setting of bone metastases. However, it required significant international effort over more than a decade to bring the Radiation Oncology community to the point of being able to answer even the most fundamental question of optimal ReRT dose. Prior to 2002, differences in endpoint definition and measurement, timing of follow-up, and interval to retreatment, for example, plagued cross-trial comparisons[12]. An update of the International Bone Metastasis Consensus Working Party recommendations in 2012 again encouraged investigators to adopt a common set of endpoints, described minimum features which should be incorporated into the design of future trials, and suggested methods of data analysis and reporting[13]. Together with the results of multiple meta-analyses[14-19], the steady evolution towards consensus has culminated in the recent publication of a phase Ⅲ randomized controlled trial. This has finally provided level Ⅰ evidence supporting a specific approach for treatment planning and dosing for external beam ReRT for bone metastases[20].

Given technological advancements in diagnostic imaging and RT delivery, coupled with longer survival in many malignancies secondary to improvements in systemic therapy, a small subset of patients presenting with localized recurrence is increasingly being offered ReRT for salvage (*i.e.*, with curative intent). At present, this is on an ad hoc basis, guided by data mainly from retrospective single-institution series which commonly span twenty years or more. Conclusions are limited by small patient numbers, attrition, heterogeneous baseline characteristics, and the presence of selection and referral bias. Descriptions of the patient population retreated, RT modality and dose received, endpoints reported and follow-up are extremely variable. Consequently, whether ReRT is offered, and how it is implemented, remains highly dependent on the specific radiation oncologist and may be limited by resource availability[9]. The opportunity presently exists to apply lessons learned from the harmonization of the research efforts within the

Table 1 Selected results of reirradiation: Both courses external beam radiotherapy unless otherwise specified

Site of ReRT	Symptom overall response rate	Symptom response duration	Overall radiologic response rate	Radiologic response duration	Overall survival	Toxicity	% not completing ReRT	ReRT-related death
Head and Neck[24]	NR	NR	NR	NR	44% at 1 yr	23% grade 3+ late at 1 yr	13%	6.7%
Thoracic[21]	Average 69.2%	0.5-5 mo	55%-77% (0-11% CR; 7%-44% PR)	NR	9%-59% at 1 yr	Esophagitis 17.2% Pneumonitis 12.3% Skin 4.1% Fracture 0.5% Myelopathy 0.5%	4.5%	1.6%
Breast[25]	100% (56% PR; 44% CR)	"For a long time of the patients' lifetime in the majority"	NR	NR	61% at 1 yr	No grade 3-4 acute or late toxicity	NR	NR
[1]Pancreas[26]	57% at 1-2 mo	NR	"Tumour stabilization but… not…reduction in tumour size"	NR	Med surv after ReRT 8.8 mo (95%CI: 1.2-16.4 mo)	28% acute grade 2 toxicity (fatigue, abdominal pain, anorexia, nausea, diarrhea) No acute grade 3+ toxicity 6% grade 3 late toxicity	0%	NR
[1,3]Cervix[27]	71% achieved ⩾ 50% reduction from baseline at 1-2 mo	NR	35% CR, 30% PR, 17% SD, 17% PD at 4 mo	NR	43% at 2 yr	35% mild acute toxicity 13% late grade 4 toxicity (all rectovaginal fistulae requiring colostomy)	NR	NR
[1]Abdomen/pelvis[28]	95% - pain 75% - bleeding	NR	100%	NR	52% at 1 yr	0% grade 3-5 acute or late toxicity Acute 22% grade 1-2 pain 14% grade 1-2 skin reaction 8% grade 1-2 diarrhea 15% grade 1-2 nausea 4% grade 2 vomiting 4% grade 1 dysuria 4% grade 1 dysphagia Late 4% grade 2 pain 4% grade 2 skin reaction 4% grade 1 diarrhea 15% grade 1-2 dysuria 19% grade 1-2 nerve complaints 11% grade 1-2 limb dysfunction	NR	NR
Bone metastases[18,19]	58%-68% (16%-28% CR; 28%-50% PR)	1-9.7 mo	NR	NR	Median 3-6 mo	30% (nausea, vomiting, fatigue, diarrhea)	NR	NR
Bone metastases[20]	45%-51% of per protocol patients at 2 mo[2] (11%-14% CR; 31%-40% PR)	NR	NR	NR	NR	Acute[2] skin 14%-24% Anorexia 56%-66% Vomiting 13%-23% Diarrhea 23%-31% Late[2] Fracture 5%-7%	NR	0%

									Spinal cord compression 1%-2% Myelopathy 0%

[1]EBRT followed by SBRT; [2]Depending on dose; [3]7/23 patients in this series did not have EBRT up front but results not reported separately. CR: Complete response; EBRT: External beam radiotherapy; Med surv: Median survival; NR: Not reported; PD: Progressive disease; PR: Partial response; ReRT: Reirradiation; SBRT: Stereotactic body radiotherapy; SD: Stable disease.

bone metastases community to the salvage ReRT situation.

In future publications, eligibility for retreatment should be defined prospectively; this may be symptom or radiologic progression or both. Baseline characteristics such as current symptom burden (and methodology of measurement), performance status, and previous treatment modalities should be documented. Controversy exists as to whether a favourable response to initial RT over a long disease-free interval should be required before considering ReRT. Information on toxicity experienced after first RT should be reviewed. Comprehensive restaging and pathologic confirmation is encouraged as outcomes after ReRT for a new primary will differ from those expected after treatment for in-field recurrence.

Initial and ReRT techniques, energies, field sizes, calculation algorithms, prescription points, doses, planning techniques, and volumes have varied significantly as can be expected from differing treatment indications, intents, geographic locations, and years[21]. Many past studies did not include all RT details, with the lack of information often due to treatment planning software changes and evolution of RT delivery techniques[21]. When reported, total dose over both courses was often the arithmetic cumulative dose, which does not take into account dose per fraction or overall treatment time. In comparison, biologically equivalent dose (BED) and equivalent dose in 2 Gy fractions (EQD2) provide the ability to compare different dose fractionation schedules. Data sufficient to calculate BED or EQD2 are not found in most studies, so conclusions which can be drawn at present regarding ReRT schedules are limited.

The rationale for ReRT dosing and cumulative allowed organ at risk tolerance doses should be stated, as should the radiobiological justification for minimum interval between RT courses. Prospective data on utilization of and outcomes after highly conformal techniques such as SBRT after conventional RT are urgently needed, including cost-effectiveness, as these approaches are steadily migrating into the clinical setting. While in theory, these technologies should allow optimal tumour localization and therefore normal tissue sparing, they also deposit extensive low dose wash resulting in higher integral doses. The methods of constructing a composite plan (*i.e.*, rigid *vs* deformable registration) and the resulting dosimetric parameters should be available and cumulative tumour and normal tissue BEDs reported.

Once such additional volumetric data are available (*e.g.*, median degree of overlap of 50% or 90% isodose lines), correlations can be explored with outcomes such as symptom response, progression and especially toxicity. Further understanding of organ tolerance to ReRT is essential, as traditional recommendations based on the Emami[22] or QANTEC[23] guidelines may not be entirely generalizable to commonly used intensity-modulated and arc-based techniques. Construction of a prognostic score including demographic, disease and treatment-factors which render a patient likely to respond, and/or unlikely to complete a second course of RT, which can be easily applied in clinic is urgently needed.

Follow-up intervals as measured from a common starting point, endpoints assessed and investigations performed should be guided by standard practice for up-front curative-intent RT in the specific primary site, and patients should be monitored long-term by their radiation oncologist for outcomes and side effects[21]. Symptom improvement and progression rates and duration must be reported, notwithstanding that measurement of these can be confounded by progressive disease and comorbidities. The use of a validated patient-reported quality of life scale prior to ReRT and at regular follow-up intervals should be strongly considered. There is little data currently available on the important parameter of duration of symptom control in relation to overall survival which would be illustrative for patients during consent discussions.

Given the heterogeneity within the population of patients reirradiated for cure, an international registry would provide a foundation on which to base consensus recommendations regarding many of the outstanding questions regarding patient selection and treatment planning. Inter-centre collaboration will be required to build a critical mass of data sufficient for robust statistical analysis; however, in order to achieve this, global harmonization is needed. Standardized nomenclature would facilitate consistent coding of treated volumes, BEDs, toxicity measurement, systemic therapy use, quality of life outcomes, and duration of follow-up. Parameters such as the minimum recommended interval between courses for different indications and sites, along with guidelines around tolerance doses for critical organs at risk could be derived. Even the definition of ReRT could be conclusively addressed, given the lack of clarity at present due to the increasing sequential use of different RT modalities. A registry

would also assist in determining the feasibility of development of phase Ⅱ prospective studies and meta-analysis of currently available data.

CONCLUSION

Given the evolving technological climate and number of patients who are being considered for salvage ReRT, the data available must be harmonized so that valid conclusions can be available for translation to the clinic. In order to properly consent patients, physicians require information about the potential benefits as well as the potential risks in relation to other available treatment modalities. International collaboration in the form of a registry of prospectively collected data on patients reirradiated for cure for a variety of tumour sites would further support the evolution of Radiation Oncology towards personalized medicine, and away from the current "one-dose-fits-all" approach.

REFERENCES

1 **Joseph KJ**, Al-Mandhari Z, Pervez N, Parliament M, Wu J, Ghosh S, Tai P, Lian J, Levin W. Reirradiation after radical radiation therapy: a survey of patterns of practice among Canadian radiation oncologists. *Int J Radiat Oncol Biol Phys* 2008; **72**: 1523-1529 [PMID: 18501531 DOI: 10.1016/j.ijrobp.2008.03.048]

2 **Richards GM**, Tomé WA, Robins HI, Stewart JA, Welsh JS, Mahler PA, Howard SP. Pulsed reduced dose-rate radiotherapy: a novel locoregional retreatment strategy for breast cancer recurrence in the previously irradiated chest wall, axilla, or supraclavicular region. *Breast Cancer Res Treat* 2009; **114**: 307-313 [PMID: 18389365 DOI: 10.1007/s10549-008-9995-3]

3 **Khan L**, Breen D, Zhang L, Balogh J, Czarnota G, Lee J, Tsao MN, Barnes EA. Predictors of recurrence after radiotherapy for non-melanoma skin cancer. *Curr Oncol* 2014; **21**: e326-e329 [PMID: 24764714 DOI: 10.3747/co.21.1727]

4 **Morgia M**, Walsh L, Milosevic M, Levin W, Fyles A. Gynaecological Malignancies. In: C Nieder, Langendijk J, editors. Re-Irradiation: New Frontiers, 2011: 171-181

5 **Aupérin A**, Le Péchoux C, Rolland E, Curran WJ, Furuse K, Fournel P, Belderbos J, Clamon G, Ulutin HC, Paulus R, Yamanaka T, Bozonnat MC, Uitterhoeve A, Wang X, Stewart L, Arriagada R, Burdett S, Pignon JP. Meta-analysis of concomitant versus sequential radiochemotherapy in locally advanced non-small-cell lung cancer. *J Clin Oncol* 2010; **28**: 2181-2190 [PMID: 20351327 DOI: 10.1200/JCO.2009.26.2543]

6 **Turrisi AT**, Kim K, Blum R, Sause WT, Livingston RB, Komaki R, Wagner H, Aisner S, Johnson DH. Twice-daily compared with once-daily thoracic radiotherapy in limited small-cell lung cancer treated concurrently with cisplatin and etoposide. *N Engl J Med* 1999; **340**: 265-271 [PMID: 9920950 DOI: 10.1056/NEJM199901283400403]

7 **Pignon JP**, le Maître A, Maillard E, Bourhis J. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. *Radiother Oncol* 2009; **92**: 4-14 [PMID: 19446902 DOI: 10.1016/j.radonc.2009.04.014]

8 **Easaw JC**, Mason WP, Perry J, Laperrière N, Eisenstat DD, Del Maestro R, Bélanger K, Fulton D, Macdonald D. Canadian recommendations for the treatment of recurrent or progressive glioblastoma multiforme. *Curr Oncol* 2011; **18**: e126-e136 [PMID: 21655151 DOI: 10.3747/co.v18i3.755]

9 **Joseph K**, Tai P, Wu J, Barnes E, Levin W. Workshop report: A practical approach and general principles of re-irradiation for in-field cancer recurrence. *Clin Oncol (R Coll Radiol)* 2010; **22**: 885-889 [PMID: 20888198 DOI: 10.1016/j.clon.2010.08.009]

10 **Poltinnikov IM**, Fallon K, Xiao Y, Reiff JE, Curran WJ, Werner-Wasik M. Combination of longitudinal and circumferential three-dimensional esophageal dose distribution predicts acute esophagitis in hypofractionated reirradiation of patients with non-small-cell lung cancer treated in stereotactic body frame. *Int J Radiat Oncol Biol Phys* 2005; **62**: 652-658 [PMID: 15936541 DOI: 10.1016/j.ijrobp.2004.10.030]

11 **Wu KL**, Jiang GL, Qian H, Wang LJ, Yang HJ, Fu XL, Zhao S. Three-dimensional conformal radiotherapy for locoregionally recurrent lung carcinoma after external beam irradiation: a prospective phase I-II clinical trial. *Int J Radiat Oncol Biol Phys* 2003; **57**: 1345-1350 [PMID: 14630272]

12 **Chow E**, Wu JS, Hoskin P, Coia LR, Bentzen SM, Blitzer PH. International consensus on palliative radiotherapy endpoints for future clinical trials in bone metastases. *Radiother Oncol* 2002; **64**: 275-280 [PMID: 12242115]

13 **Chow E**, Hoskin P, Mitera G, Zeng L, Lutz S, Roos D, Hahn C, van der Linden Y, Hartsell W, Kumar E. Update of the international consensus on palliative radiotherapy endpoints for future clinical trials in bone metastases. *Int J Radiat Oncol Biol Phys* 2012; **82**: 1730-1737 [PMID: 21489705 DOI: 10.1016/j.ijrobp.2011.02.008]

14 **Sze WM**, Shelley MD, Held I, Wilt TJ, Mason MD. Palliation of metastatic bone pain: single fraction versus multifraction radiotherapy--a systematic review of randomised trials. *Clin Oncol* (R Coll Radiol) 2003; **15**: 345-352 [PMID: 14524489 DOI: 10.1016/S0936-6555(03)00113-4]

15 **Wu JS**, Wong R, Johnston M, Bezjak A, Whelan T. Meta-analysis of dose-fractionation radiotherapy trials for the palliation of painful bone metastases. *Int J Radiat Oncol Biol Phys* 2003; **55**: 594-605 [PMID: 12573746 DOI: 10.1016/S0360-3016(02)04147-0]

16 **Chow E**, Harris K, Fan G, Tsao M, Sze WM. Palliative radiotherapy trials for bone metastases: a systematic review. *J Clin Oncol* 2007; **25**: 1423-1436 [PMID: 17416863 DOI: 10.1200/JCO.2006.09.5281]

17 **Chow E**, Zeng L, Salvo N, Dennis K, Tsao M, Lutz S. Update on the systematic review of palliative radiotherapy trials for bone metastases. *Clin Oncol (R Coll Radiol)* 2012; **24**: 112-124 [PMID: 22130630 DOI: 10.1016/j.clon.2011.11.004]

18 **Wong E**, Hoskin P, Bedard G, Poon M, Zeng L, Lam H, Vulpe H, Tsao M, Pulenzas N, Chow E. Re-irradiation for painful bone metastases - a systematic review. *Radiother Oncol* 2014; **110**: 61-70 [PMID: 24094630 DOI: 10.1016/j.radonc.2013.09.004]

19 **Huisman M**, van den Bosch MA, Wijlemans JW, van Vulpen M, van der Linden YM, Verkooijen HM. Effectiveness of reirradiation for painful bone metastases: a systematic review and meta-analysis. *Int J Radiat Oncol Biol Phys* 2012; **84**: 8-14 [PMID: 22300568 DOI: 10.1016/j.ijrobp.2011.10.080]

20 **Chow E**, van der Linden YM, Roos D, Hartsell WF, Hoskin P, Wu JS, Brundage MD, Nabid A, Tissing-Tan CJ, Oei B, Babington S, Demas WF, Wilson CF, Meyer RM, Chen BE, Wong RK. Single versus multiple fractions of repeat radiation for painful bone metastases: a randomised, controlled, non-inferiority trial. *Lancet Oncol* 2014; **15**: 164-171 [PMID: 24369114 DOI: 10.1016/S1470-2045(13)70556-4]

21 **Drodge S**, Ghosh S, Fairchild A. Thoracic reirradiation for lung cancer: A literature review and practical guide. *Ann Pall Med* 2014; **3**: 75-91 [DOI: 10.3978/j.issn.2224-5820.2014.03.04]

22 **Emami B**, Lyman J, Brown A, Coia L, Goitein M, Munzenrider JE, Shank B, Solin LJ, Wesson M. Tolerance of normal tissue to therapeutic irradiation. *Int J Radiat Oncol Biol Phys* 1991; **21**: 109-122 [PMID: 2032882 DOI: 10.1016/0360-3016(91)90171-Y]

23 **Marks LB**, Ten Haken RK, Martel MK. Guest editor's introduction to QUANTEC: a users guide. *Int J Radiat Oncol Biol Phys* 2010; **76**: S1-S2 [PMID: 20171501 DOI: 10.1016/j.ijrobp.2009.08.075]

24 **Duprez F**, Berwouts D, Madani I, Bonte K, Boterberg T, De Gersem W, Deron P, Huvenne W, De Neve W. High-dose reirradiation with intensity-modulated radiotherapy for recurrent head-and-neck cancer: disease control, survival and toxicity. *Radiother Oncol* 2014; **111**: 388-392 [PMID: 24998706 DOI: 10.1016/j.radonc.2014.04.018]

25 **Würschmidt F**, Dahle J, Petersen C, Wenzel C, Kretschmer M, Bastian C. Reirradiation of recurrent breast cancer with and without concurrent chemotherapy. *Radiat Oncol* 2008; **3**: 28 [PMID: 18801165 DOI: 10.1186/1748-717X-3-28]

26 **Wild AT**, Hiniker SM, Chang DT, Tran PT, Khashab MA, Limaye MR, Laheru DA, Le DT, Kumar R, Pai JS, Hargens B, Sharabi AB, Shin EJ, Zheng L, Pawlik TM, Wolfgang CL, Koong AC, Herman JM. Re-irradiation with stereotactic body radiation therapy as a novel treatment option for isolated local recurrence of pancreatic cancer after multimodality therapy: experience from two institutions. *J Gastrointest Oncol* 2013; **4**: 343-351 [PMID: 24294505 DOI: 10.3978/j.issn.2078-6891.2013.044]

27 **Seo Y**, Kim MS, Yoo HJ, Jang WI, Rhu SY, Choi SC, Kim MH, Kim BJ, Lee DH, Cho CK. Salvage stereotactic body radiotherapy for locally recurrent uterine cervix cancer at the pelvic sidewall: Feasibility and complication. *Asia Pac J Clin Oncol* 2014; Epub ahead of print [PMID: 24889550 DOI: 10.1111/ajco.12185]

28 **Abusaris H**, Hoogeman M, Nuyttens JJ. Re-irradiation: outcome, cumulative dose and toxicity in patients retreated with stereotactic radiotherapy in the abdominal or pelvic region. *Technol Cancer Res Treat* 2012; **11**: 591-597 [PMID: 22568625 DOI: 10.7785/tcrt.2012.500261]

P- Reviewer: Damin DC, Yokoyama Y **S- Editor**: Ji FF
L- Editor: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i3.139

World J Meta-Anal 2015 June 26; 3(3): 139-141
ISSN 2308-3840 (online)

MINIREVIEWS

Systematic reviews and meta-analyses: Why are they clinically significant?

Xing-Shun Qi, Zhi-Ping Yang, Ming Bai, Yong-Ji Wang

Xing-Shun Qi, Zhi-Ping Yang, Ming Bai, Evidence-Based Medicine Group, Xijing Hospital of Digestive Diseases, Fourth Military Medical University, Xi'an 710000, Shaanxi Province, China
Xing-Shun Qi, Department of Gastroenterology, General Hospital of Shenyang Military Area, Shenyang 110840, Liaoning Province, China
Ming Bai, Department of Nephrology, Xijing Hospital, Fourth Military Medical University, Xi'an 710000, Shaanxi Province, China
Yong-Ji Wang, Medical Department, 309th Hospital of Chinese People's Liberation Army, Beijing 100000, China
Yong-Ji Wang, Department of Health Statistics, Fourth Military Medical University, Xi'an 710000, Shaanxi Province, China

Author contributions: Qi XS conceived this work and drafted the manuscript; Yang ZP, Bai M and Wang YJ gave critical comments and revised the manuscript; all the authors have made an intellectual contribution to the manuscript and approved the submission.

Conflict-of-interest: All authors disclosed no conflicts of interest regarding this work.

Correspondence to: Dr. Xing-Shun Qi, Department of Gastroenterology, General Hospital of Shenyang Military Area, 83 WenHua Road, Shenhe District, Shenyang 110840, Liaoning Province, China. xingshunqi@126.com
Telephone: +86-24-28851113

Received: January 7, 2015
Peer-review started: January 8, 2015
First decision: February 7, 2015
Revised: February 24, 2015
Accepted: May 26, 2015
Article in press: May 27, 2015
Published online: June 26, 2015

Abstract

This review aims to clarify the clinical significance of systematic reviews and meta-analyses by illustrating several classical examples. Firstly, systematic reviews can provide the highest level of evidence for clinical decisions. Secondly, systematic reviews can propose unresolved issues and future directions. Thirdly, systematic reviews can avoid harm to the human body. Fourthly, systematic reviews can prevent a waste of resources. Generally speaking, clinical researchers should be encouraged to perform systematic reviews and meta-analyses.

Key words: Systematic reviews; Meta-analyses; China; Publication; Science citation index

Core tip: Systematic reviews and meta-analyses are very important for clinicians and investigators because they can provide the highest level of evidence for clinical decisions, propose unresolved issues and future directions, avoid harm to the human body and prevent a waste of resources.

Qi XS, Yang ZP, Bai M, Wang YJ. Systematic reviews and meta-analyses: Why are they clinically significant? *World J Meta-Anal* 2015; 3(3): 139-141 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i3/139.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i3.139

INTRODUCTION

In recent years, the number of systematic reviews and meta-analyses has been steadily on the rise. By searching the PubMed database, about 500 relevant papers were published around the world in 1994 but more than 6000 relevant papers were published

in 2009[1]. Currently, systematic reviews and meta-analyses are also very hot in China. According to the statistics produced by *Ding Xiang Yuan* reporters, China contributed over 1000 meta-analysis papers in 2012[2]. There was a 40-fold increase in the annual number of meta-analyses in the genomic era for China from 2003 to 2011[3].

Investigators who perform original research need lots of time and costs for collecting clinical data and/or doing the experiments. By comparison, meta-analysis authors spend less time and fewer costs on synthesizing previously published data into a new result. It is said that a doctor wrote dozens of meta-analyses in Science Citation Index (SCI) journals with an accumulated impact factor > 200 in one year[4]. Ironically, the spectrum of his or her meta-analyses was very wide, including breast diseases, colon cancer, orthopedics, etc. As a criticism of the fact, publishing a meta-analysis in SCI journals is often regarded as opportunistic behavior. Some experts working at famous institutions strongly discourage their students from doing meta-analyses[5]. Herein, we highlight the significance of meta-analyses to correct such a distortion and encourage more investigators to perform meta-analyses.

SYSTEMATIC REVIEWS CAN PROVIDE THE HIGHEST LEVEL OF EVIDENCE FOR CLINICAL DECISIONS

According to the system produced by the Oxford Centre for Evidence-Based Medicine (March 2009), evidence for therapy/prevention and etiology/harm studies is divided into five levels[6]. They include level 1 (randomized controlled trials), level 2 (cohort studies), level 3 (case-control studies), level 4 (case series) and level 5 (expert opinion). Level 1 is further classified into level 1a (systematic review of randomized controlled trials) and 1b (individual randomized controlled trials). Similarly, systematic reviews of cohort and case-control studies are also classified as levels 2a and 3a, respectively. In the updated system produced by the Oxford Centre for Evidence-Based Medicine (2011), evidence for treatment benefit studies is also divided into five levels[7]. Systematic reviews of randomized trials provide the top level of evidence. On the other hand, the number of citations potentially reflects the hierarchy of evidence. Meta-analyses can receive the largest number of citations, followed by randomized controlled trials, cohort or case-control studies, nonsystematic review articles, decision and cost-effectiveness analyses and case reports[8].

SYSTEMATIC REVIEWS CAN PROPOSE UNRESOLVED ISSUES AND FUTURE DIRECTIONS

Systematic reviews are indispensable before initiating new clinical research[9,10]. Since August 2005, the *LANCET* editors have required authors to summarize previously published findings and explain the impact of their findings on existing knowledge[11]. In this renowned journal, the guidelines for authors obviously propose how the authors of clinical trials should do an updated systematic review if a recent systematic review is unavailable[12].

This consideration is also appropriate for every clinical researcher. In 2011, we published a meta-analysis to explore the significance of screening for JAK2 V617F mutation in patients with Budd-Chiari syndrome[13]. The prevalence of JAK2 V617F mutation was 37% and positive JAK2 V617F mutation could predict the presence and development of myeloproliferative neoplasms in such patients[13]. However, most available studies were conducted in the West and only one study was conducted in Asia (India). Given the ethnical differences between China and the West and the absence of related data from China, further evaluation of the prevalence of JAK2 V617F mutation in Chinese patients is warranted. In 2012, we reported the results of a clinical study in which the prevalence of JAK2 V617F mutation in Chinese patients with Budd-Chiari syndrome was only 4.3%[14]. This finding suggested a difference in the etiological distribution of Budd-Chiari syndrome between China and the West. Thus, we further performed a large-scale observational study to more comprehensively analyze the thrombotic risk factors for Budd-Chiari syndrome in Chinese patients[15]. Except for JAK2 V617F mutation and myeloproliferative neoplasms, paroxysmal nocturnal hemoglobinuria, factor Ⅴ Leiden mutation and prothrombin G20210A mutation were rarely found in our patients. These results were immediately confirmed by other peers[16,17].

SYSTEMATIC REVIEWS CAN AVOID HARM TO THE HUMAN BODY

Gilbert *et al*[18] performed a systematic review of observational studies and recommendations from textbooks about the association between infant sleeping position and sudden infant death syndrome. In books on infant care, the recommendation regarding whether the infants should be on a back or front sleeping position was controversial before 1989 but only a back sleeping position was recommended after that. In the meta-analysis, 25 individual studies published between 1965 and 2004 were identified. Indeed, the cumulative meta-analysis of the first two published studies (the first study was published in 1965 and the second one was published in 1970) demonstrated that the front sleeping position led to a statistically significant increase in the incidence of sudden infant death syndrome (cumulative odds ratio = 2.93, 95%CI: 1.15-7.47). In other words, if a meta-analysis was performed soon after the first two papers were published, the debate regarding the sleeping position would have disappeared, thereby

preventing more than 10000 infant deaths in the United Kingdom and more than 50000 in Europe, the United States and Australasia.

SYSTEMATIC REVIEWS CAN PREVENT A WASTE OF RESOURCES

Lau *et al*[19] performed a meta-analysis of clinical trials to compare the benefit of intravenous streptokinase *vs* placebo or no therapy for acute myocardial infarction. In the meta-analysis, 33 individual studies published between 1959 and 1988 were identified. Indeed, in the cumulative meta-analysis of the first four published studies with 962 patients, the benefit of intravenous streptokinase for acute myocardial infarction became statistically significant (P = 0.023) but the 95%CI was relatively wide. In the cumulative meta-analysis of the first 15 published studies with 4314 patients, the benefit remained significant (P < 0.001) and the odds ratio became steadier with a narrower 95%CI. Accordingly, the 18 trials published since then were unnecessary. More importantly, the additional 32660 participants should not have been enrolled because the participants assigned to the placebo/no therapy group would not have received intravenous streptokinase.

Another similar example was a meta-analysis to evaluate the risk of lung cancer in never-smoking women exposed to passive smoking by spouses[20]. Taylor *et al*[20] identified a total of 51 studies between 1981 and 2006. In the cumulative meta-analysis of the first 10 studies published before 1986, the association of passive smoking and lung cancer was significant. In the cumulative meta-analysis of the first 20 studies published before 1989, the statistical significance became steadier. Thus, the subsequent 31 studies may have been wasteful.

CONCLUSION

The importance of systematic reviews and meta-analyses in the contemporary era of evidence-based medicine needs to be clearly recognized. Clinical researchers should be accustomed to publishing their own data after the related evidence is systematically reviewed.

REFERENCES

1 **Booth A**, Clarke M, Ghersi D, Moher D, Petticrew M, Stewart L. An international registry of systematic-review protocols. *Lancet* 2011; **377**: 108-109 [PMID: 20630580 DOI: 10.1016/S0140-6736(10)60903-8]

2 Meta-analysis Lun Wen Qu Wei Tong Ji Mian Mian Guan (Article in Chinese). Available from: URL: http://paper.dxy.cn/article/26164

3 **Ioannidis JP**, Chang CQ, Lam TK, Schully SD, Khoury MJ. The geometric increase in meta-analyses from China in the genomic era. *PLoS One* 2013; **8**: e65602 [PMID: 23776510 DOI: 10.1371/journal.pone.0065602]

4 Meta-analysis- Yi Xue Ke Yan De Guai Xiang. Available from: URL: http://news.dxy.cn/bbs/topic/21336219

5 **Yang ZP**, Ye XF, Fan DM. Meta-analysis is victim to Chinese academic and educational systems. *J Formos Med Assoc* 2013; **112**: 235-236 [PMID: 23660217 DOI: 10.1016/j.jfma.2012.09.019]

6 Oxford Centre for Evidence-based Medicine - Levels of Evidence (March 2009). Available from: URL: http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/

7 **Howick J**, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A, Moschetti I, Phillips B, Thornton H. The 2011 Oxford CEBM Evidence Levels of Evidence. Available from: URL: http://www.cebm.net/ocebm-levels-of-evidence/

8 **Patsopoulos NA**, Analatos AA, Ioannidis JP. Relative citation impact of various study designs in the health sciences. *JAMA* 2005; **293**: 2362-2366 [PMID: 15900006 DOI: 10.1001/jama.293.19.2362]

9 **Clarke M**. Doing new research? Don't forget the old. *PLoS Med* 2004; **1**: e35 [PMID: 15578106 DOI: 10.1371/journal.pmed.0010035]

10 **Clarke M**, Hopewell S, Chalmers I. Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. *Lancet* 2010; **376**: 20-21 [PMID: 20609983 DOI: 10.1016/S0140-6736(10)61045-8]

11 **Young C**, Horton R. Putting clinical trials into context. *Lancet* 2005; **366**: 107-108 [PMID: 16005318 DOI: 10.1016/S0140-6736(05)66846-8]

12 The Lancet: Information for Authors. Available from: URL: http://www.thelancet.com/lancet/information-for-authors

13 **Qi X**, Yang Z, Bai M, Shi X, Han G, Fan D. Meta-analysis: the significance of screening for JAK2V617F mutation in Budd-Chiari syndrome and portal venous system thrombosis. *Aliment Pharmacol Ther* 2011; **33**: 1087-1103 [PMID: 21395632 DOI: 10.1111/j.1365-2036.2011.04627.x]

14 **Qi X**, Zhang C, Han G, Zhang W, He C, Yin Z, Liu Z, Bai W, Li R, Bai M, Yang Z, Wu K, Fan D. Prevalence of the JAK2V617F mutation in Chinese patients with Budd-Chiari syndrome and portal vein thrombosis: a prospective study. *J Gastroenterol Hepatol* 2012; **27**: 1036-1043 [PMID: 22142461 DOI: 10.1111/j.1440-1746.2011.07040.x]

15 **Qi X**, Wu F, Ren W, He C, Yin Z, Niu J, Bai M, Yang Z, Wu K, Fan D, Han G. Thrombotic risk factors in Chinese Budd-Chiari syndrome patients. An observational study with a systematic review of the literature. *Thromb Haemost* 2013; **109**: 878-884 [PMID: 23447059 DOI: 10.1160/TH12-10-0784]

16 **Wang H**, Sun G, Zhang P, Zhang J, Gui E, Zu M, Jia E, Xu H, Xu L, Zhang J, Lu Z. JAK2 V617F mutation and 46/1 haplotype in Chinese Budd-Chiari syndrome patients. *J Gastroenterol Hepatol* 2014; **29**: 208-214 [PMID: 23980667 DOI: 10.1111/jgh.12379]

17 **Cheng D**, Xu H, Lu ZJ, Hua R, Qiu H, Du H, Xu X, Zhang J. Clinical features and etiology of Budd-Chiari syndrome in Chinese patients: a single-center study. *J Gastroenterol Hepatol* 2013; **28**: 1061-1067 [PMID: 23425079 DOI: 10.1111/jgh.12140]

18 **Gilbert R**, Salanti G, Harden M, See S. Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. *Int J Epidemiol* 2005; **34**: 874-887 [PMID: 15843394 DOI: 10.1093/ije/dyi088]

19 **Lau J**, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992; **327**: 248-254 [PMID: 1614465 DOI: 10.1056/NEJM199207233270406]

20 **Taylor R**, Najafi F, Dobson A. Meta-analysis of studies of passive smoking and lung cancer: effects of study type and continent. *Int J Epidemiol* 2007; **36**: 1048-1059 [PMID: 17690135 DOI: 10.1093/ije/dym158]

P- Reviewer: Gao C, Specchia ML, Tang Y **S- Editor**: Tian YL
L- Editor: Roemmele A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i3.142

World J Meta-Anal 2015 June 26; 3(3): 142-150
ISSN 2308-3840 (online)

Development of the Documentation and Appraisal Review Tool for systematic reviews

Rebecca L Diekemper, Belinda K Ireland, Liana R Merz

Rebecca L Diekemper, American College of Chest Physicians, CHEST, Glenview, IL 60026, United States
Belinda K Ireland, The EvidenceDoc, Pacific, MO 63069, United States
Liana R Merz, Center for Clinical Excellence, BJC HealthCare, Saint Louis, MO 63108, United States

Author contributions: Diekemper RL was the primary developer of the tool and she participated in the testing of the tool and drafting parts of the paper; Ireland BK came up with the concept of developing the tool and was a co-developer of the tool and participated in the testing of the tool and drafting parts of the paper; Merz LR was a co-developer of the tool and participated in the testing of the tool and drafting parts of the paper.

Conflict-of-interest: All of the authors report that they receive no financial compensation for DART. Diekemper RL uses DART for assessing the quality of systematic reviews used to inform guideline recommendations for CHEST guidelines. Due to her role as a developer of DART, the tool has been adopted by CHEST for use in guideline development. Ireland BK reports that as a consultant who frequently conducts systematic reviews and overviews of reviews, she is interested in an effective and efficient tool for evaluating the quality of systematic reviews. Merz LR has no conflicts of interest to disclose.
Data sharing: No additional data are available.

Correspondence to: Rebecca L Diekemper, MPH, American College of Chest Physicians, 2595 Patriot Blvd, Glenview, IL 60026, United States. rdiekemper@chestnet.org
Telephone: +1-314-5319325

Received: January 27, 2015
Peer-review started: February 5, 2015
First decision: March 6, 2015
Revised: April 2, 2015
Accepted: April 27, 2015
Article in press: April 29, 2015
Published online: June 26, 2015

Abstract

AIM: To develop a tool to more explicitly assess and document the quality of systematic reviews.

METHODS: We developed the Documentation and Appraisal Review Tool (DART) using epidemiologic principles of study design and the following resources: the modified Overview Quality Assessment Questionnaire (modified OQAQ), Assessment of Multiple Systematic Reviews (AMSTAR), the Cochrane Handbook, and the standards promoted by the Agency for Healthcare Research and Quality, and the Institutes of Medicine (IOM). We designed the DART tool to include the following: more detail to provide guidance and improve standardization of use, an approach to assess quality of systematic reviews addressing a variety of research designs, and additional space for recording notes to facilitate recall. DART underwent multiple rounds of testing with methodologists of varying levels of training and experience. Based on the results of six phases of pilot testing, we revised DART to improve performance, clarity and consistency. Pilot testing also included comparisons between DART, and the two most commonly used tools to evaluate the quality of systematic reviews, the modified OQAQ and AMSTAR.

RESULTS: Compared to AMSTAR and modified OQAQ, DART includes two unique questions and several questions covered by modified OQAQ or AMSTAR but not both. Modified OQAQ and DART had the highest reporting consistency. Four AMSTAR questions were unclear and elicited inconsistent responses. Identifying reviewer rationale was most difficult using the modified OQAQ tool, and easiest using DART. DART allows

for documentation of reviewer rationale, facilitating reconciliation between reviewers and documentation for future updates. DART also provides a comprehensive, systematic approach for reviewers with limited experience with systematic review methodology, to critically analyze systematic reviews. In addition, DART is the only one of the three tools to explicitly include quality review for biases specific to observational studies. This is now more widely recognized as important for assessing risk in order to generate recommendations that balance benefit to harm. The tool also includes the assessment of standards recommended by the March 2011 IOM Standards for Systematic Review.

CONCLUSION: This comprehensive tool improves upon existing tools for assessing the quality of systematic reviews and guides reviewers through critically analyzing a systematic review.

Key words: Quality assessment tool; Methodology; Healthcare research; Systematic review; Meta-analysis; Guidelines

Core tip: Systematic reviews and meta-analyses are commonly used to inform the recommendations presented in evidence-based clinical practice guidelines. The purpose of this study was to evaluate the Documentation and Appraisal Review Tool (DART) for its comprehensiveness, identify areas addressed by DART that were not addressed by two other validated tools [Overview Quality Assessment Questionnaire (OQAQ) and Assessment of Multiple Systematic Reviews (AMSTAR)], and to test its performance in eliciting consistent responses. We found that our tool was more comprehensive and included several questions not included in the other tools. We also found that DART elicited the most consistent responses when compared to OQAQ and AMSTAR.

Diekemper RL, Ireland BK, Merz LR. Development of the Documentation and Appraisal Review Tool for systematic reviews. *World J Meta-Anal* 2015; 3(3): 142-150 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i3/142.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i3.142

INTRODUCTION

Systematically collected and critically evaluated evidence forms the backbone of evidence-based clinical practice guidelines, hospital order sets, and quality measurement. Grant *et al*[1] define a systematic review as a systematic search, appraisal and synthesis of research evidence, often adhering to guidelines for conducting a review. Systematic reviews are the most comprehensive and valid method of collecting and synthesizing the published and unpublished record of clinical science, making them a preferred source of evidence and encouraging increased production. In 2010, Bastian *et al*[2] estimated 11 systematic reviews are published each day.

The consistent application of well-defined processes is essential to creating valid systematic reviews. These processes include (1) development of specific clinical question(s) using an analytic framework and standard format to articulate the question(s); (2) use of comprehensive and systematic methods to search for evidence; (3) unbiased process for selecting relevant research; (4) critical evaluation of the quality of included studies; (5) the extraction and synthesis of data from the included studies; and (6) the use of a pre-specified system to evaluate the body of evidence[3]. Even though these processes for sound systematic review are well described, and reporting checklists like Preferred Reporting Items for systematic reviews and meta-analyses[4] are available to authors to ensure a higher quality systematic review, the quality of published systematic reviews is not uniformly high. In 2002, Shea *et al*[5] evaluated the quality of Cochrane and other systematic reviews published in paper based journals, using the Oxman and Guyatt scale and the Sacks checklist. They found the average quality low for both types of reviews.

The Institute of Medicine (IOM) recognized that variation in the quality of systematic reviews still exists and convened a panel in 2010 to develop national standards for the design and implementation of systematic reviews. In 2011, the IOM panel released a list of 21 recommended standards for conducting systematic reviews[3]. If implemented properly and consistently, these standards could greatly reduce the variability and improve the overall quality of systematic reviews.

Currently, providers and policy makers wanting to incorporate the findings from existing systematic reviews into care decisions, protocols, and guidelines need assistance in evaluating the quality of systematic reviews. Several tools have been developed and evaluated and two have been validated for content[5,6]. We reviewed published user experience with these two, the modified Overview Quality Assessment Questionnaire (modified OQAQ)[5] and the Assessment of Multiple Systematic Reviews (AMSTAR)[6]. Most current users report implementation of AMSTAR because methods for evaluating systematic reviews have advanced since the development of OQAQ, however some also report modifying AMSTAR because it did not meet all their needs[7,8]. The Agency for Healthcare Research and Quality (AHRQ) recommends that its Evidence-based Practice Centers (EPCs) supplement the use of AMSTAR with additional considerations when incorporating existing systematic reviews into their reviews[8].

We examined both tools for use in evaluating systematic reviews of clinical interventions in a health system setting. Neither met all our needs (Table 1), and so we first set out to enhance one of the existing assessment tools. However, ultimately we determined the need to develop a comprehensive tool that improves

Table 1 Assessment of existing systematic review quality assessment tools

Need	Modified OQAQ	AMSTAR
Standardized quality assessment process across multiple reviewers with varying levels of experience	Insufficient detail to evaluate disputes	Confusing questions leading to inconsistent responses by same reviewer as well as between reviewers
Single tool to assess a variety of included research designs including randomized trials and observational studies	Insufficient detail on methods	Insufficient detail on methods
Detailed record of the review to facilitate updates of the evidence review	Insufficient detail for replication	Confusing questions leading to inconsistent responses by same reviewer and insufficient detail for replication
Training tool for junior epidemiologists and interns in systematic review methods	Insufficient detail on methods	Insufficient detail on methods

OQAQ: Overview Quality Assessment Questionnaire; AMSTAR: Assessment of Multiple Systematic Reviews.

upon existing tools for assessing the quality of systematic reviews and that guides reviewers through critically analyzing a systematic review. Here we describe the development of a tool designed to more explicitly document the quality assessment of systematic reviews: the Documentation and Appraisal Review Tool (DART) for Systematic Reviews (Table 2). To download the complete tool, please go to http://www.theevidencedoc.com.

MATERIALS AND METHODS

Design

DART was developed using epidemiologic principles of study design, the AMSTAR tool[6], and the Cochrane Handbook for Systematic Reviews of Interventions (version 4.2.6)[9] as guides. Once completed, we compared our tool to the validated systematic review tools, modified OQAQ and AMSTAR, and to tools developed by some of the AHRQ EPCs to ensure that the tool was as comprehensive as possible. All questions in the DART tool include the following: more detail to provide guidance and improve standardization of use, an approach to assess quality of systematic reviews addressing a variety of research designs, and additional space for recording notes to facilitate recall.

First round testing

An internal group of six methodologists then reviewed and pilot-tested the tool. The group was given systematic reviews of varying quality and asked to use the tool to critically analyze the reviews. The group met weekly for several weeks, testing a different systematic review with the tool each week. This exercise resulted in several revisions. By the end of phase Ⅱ, we determined that the tool was designed well enough to elicit consistent responses and agreement regarding the overall quality of the studies reviewed.

Comparison of test performance to validated tools

The second round of testing focused on the review of systematic reviews using DART in addition to the modified OQAQ and AMSTAR, two widely accepted, validated tools for assessing the quality of systematic reviews. The goal of this round of testing was to compare the performance of DART to the modified OQAQ and AMSTAR to determine if we met our design goals. Four internal reviewers with varying levels of training and experience, ranging from a student enrolled in a Masters of Public Health program to a faculty epidemiologist with over 30 years of experience used the three tools to independently assess the quality of several published systematic reviews. The reviewers then used a modified nominal group technique to brainstorm the strengths, weaknesses, and suggestions for improvement of DART. The reviewers also compared the performance of the three tools and identified variation in the responses to the quality assessment questions. The three tools were then mapped against each other to identify and characterize areas of overlap between the questions (Table 3), in order to determine if design goals for DART were met.

Refinement

After evaluating results from the content mapping and comparing performance and utility of DART for reviewers with different levels of experience, the tool was once again revised. A third round of pilot testing was performed using the revised tool to appraise the quality of different systematic reviews.

Comparison to IOM standards for systematic reviews

As a final review of our tool, we compared content to the March 2011 Standards for Systematic Reviews from the IOM to ensure that the tool included an evaluation component for each IOM standard[3].

Final testing

Final modification of the tool was completed in April 2011, followed by more rounds of internal pilot testing to evaluate consistency of responses for each question when the same reviewer appraised the systematic review at different points in time (intra-observer reliability) and when used by different reviewers (inter-observer reliability).

RESULTS

Assessing comparability of content of the three tools

In order to determine if we met our design goals, we

Table 2 Documentation and Appraisal Review Tool for systematic reviews

Title of Systematic Review:				
Author:				
Publication date:		Article tracking number:		
Reviewer:		Date completed:		
1 Did the authors develop the research question(s) and inclusion/exclusion criteria before conducting the review?				Use this space to document the rationale for your answer
a	It was clear the authors developed the research question(s) and inclusion criteria before conducting the review and that they stated the question(s) clearly	Yes		
b	Not described or cannot tell	No		
2 Did the authors describe the search methods used to find evidence (original research) on the primary question(s)?				Use this space to document the rationale for your answer
a	Key words and/or MESH terms were stated and where feasible the search strategy was provided	Yes		
b	Not described or cannot tell	No		
3 Was the search for the evidence reasonably comprehensive? Were the following included?				Use this space to document the rationale for your answer
a	Search included at least two electronic sources	Yes	No	
b	Authors chose the most applicable electronic databases (*e.g.*, CINAHL for nursing journals, EMBASE for pharmaceutical journals, and MEDLINE for general, comprehensive search) and only limited search by date when performing an update of a previous systematic review	Yes	No	
c	Search methods are likely to capture all relevant studies (*e.g.*, includes languages other than English; gray literature such as conference proceedings, dissertations, theses, clinical trials registries and other reports) and authors hand-searched journals or reference lists to identify published studies which were not electronically available	Yes	No	
4 Did the authors do the following when selecting studies for the review?				Use this space to document the rationale for your answer
a	Provide in the inclusion criteria: population, intervention, outcome and study design?	Yes	No	
b	State whether the selection criteria were applied independently by more than one person?	Yes	No	
c	State how disagreements were resolved during study selection?	Yes	No	
d	Provide a flowchart or descriptive summary of the included and excluded studies?	Yes	No	
e	Include all study designs appropriate for the research questions posed?	Yes	No	
5 Were the characteristics of the included studies provided? (in an aggregated form such as a table, data from the original studies were provided on the participants, interventions and outcomes)				Use this space to document the rationale for your answer
a	Yes			
b	Partially			
c	No			
6 Did the authors make any statements about assessing for publication bias?				Use this space to document the rationale for your answer
a	The authors did assess for publication bias and if publication bias was detected they stated how it was handled	Yes		
b	The authors did assess for publication bias but did not state how it was handled if it was detected	Partially		
c	Not described or cannot tell	No		
7 Did the authors do the following to assess the overall quality of the individual studies included in the review?				Use this space to document the rationale for your answer
a	Was the quality assessment specified with adequate detail to permit replication?	Yes	No	
b	Was the quality assessment conducted independently by more than one person?	Yes	No	
c	Did the authors state how disagreements were resolved during the quality assessment?	Yes	No	
8 Did the authors appropriately assess for quality by appropriately examining the following sources of bias in all of the included studies?				Use this space to document the rationale for your answer
All studies:				
a	Confounding (assessed comparability of study groups at start of study, was randomization successful?)	Yes	No	
b	Sufficient sample size (only applicable to studies that summarize their results in a qualitative manner; it's not a concern for pooled results)	Yes	No	
c	Outcome reporting bias (assessed for each outcome reported using a system such as the ORBIT classification system)	Yes	No	
d	Follow up (assessed for completeness and any differential loss to follow-up)	Yes	No	
For Randomized Controlled Trials only:				
e	Randomization	Yes	No	
f	Allocation concealment	Yes	No	
g	Blinding	Yes	No	

For Case-Control and Cohort Studies only:				
h	Selection bias	Yes	No	
i	Information bias--recall and completeness to follow-up	Yes	No	
For Quasi-Experimental Studies only:				
j	Differences between the first and second study measurement point - such as changes or improvements in other interventions, changes in measurement techniques or definitions, or aging of subjects	Yes	No	
k	Selection bias	Yes	No	
For Diagnostic Accuracy Studies only:				
l	Selection (spectrum) bias - were subjects selected to be representative of patients to whom the test will be applied in clinical practice, and to represent the broadest spectrum of disease?	Yes	No	
m	Verification bias - were all patients subjected to the same reference standard of diagnosis, and was it measured blindly and independently of the test?	Yes	No	
9 Did the authors use appropriate methods to extract data from the included studies?				Use this space to document the rationale for your answer
a	Were standard forms developed and piloted prior to the systematic review conduct?	Yes	No	
b	Did the authors ensure that data from the same study but that appeared in multiple publications were counted only once in the synthesis?	Yes	No	
c	Was data extraction performed by more than one person?	Yes	No	
10 Did the authors assess and account for heterogeneity (differences in participants, interventions, outcomes, trial design, quality or treatment effects) among the studies selected for the review?				Use this space to document the rationale for your answer
a	The authors stated the differences among the studies and how they accounted for those differences	Yes		
b	The authors stated the differences but not how they accounted for them	Partially		
c	Not described or cannot tell	No		
11 Did the authors describe the methods they used to combine/synthesize the results of the relevant studies (to reach a conclusion) and were the methods used appropriate for the review question(s)?				Use this space to document the rationale for your answer
a	Methods were reported clearly enough to allow for replication. The overview included some assessment of the qualitative and quantitative heterogeneity of the study results and the results were appropriately combined/synthesized. For meta-analyses, an accepted pooling method (*i.e.*, more than simple addition) was used. Or the authors state that the evidence is conflicting and that they can't combine/ synthesize the results	Yes		
b	The methods were reported clearly enough to allow for replication but they were not combined appropriately	Partially		
c	Not described or cannot tell	No		
12 Did the authors perform sensitivity analyses on any changes in protocol, assumptions, and study selection? (For example, using sensitivity analysis to compare results from fixed effects and random effects models)				Use this space to document the rationale for your answer
a	Sensitivity analyses were used when appropriate on all changes in a priori design	Yes		
b	Sensitivity analyses were only used on some changes in a priori design	Partially		
c	Not described or cannot tell	No		
13 Are the conclusions of the authors supported by the reported data with consideration of the overall quality of that data?				Use this space to document the rationale for your answer
a	The conclusions are supported by the reported data and reflect both the scientific quality of the studies and the risk of bias in the data obtained from those studies	Yes		
b	The authors failed to consider study quality and/or their conclusions were not supported by the data, or cannot tell	No		
14 Were conflicts of interest stated and were individuals excluded from the review if they reported substantial financial and intellectual COIs?				Use this space to document the rationale for your answer
a	COIs were reported for each team member and individuals were excluded if they had substantial COIs	Yes		
b	COIs were reported but it was not clear whether individuals were excluded based on their COIs	Partially		
c	COIs were not reported and individuals were not excluded based on their COIs	No		
15 On a scale of 1-10, how would you judge the overall quality of the paper?				
Rating	Overall Comments			
Good (8-10)				
Fair (5-7)				
Poor (< 5)				

COIs: Conflicts of interests.

Table 3 Comparison of Documentation and Appraisal Review Tool to modified Overview Quality Assessment Questionnaire and Assessment of Multiple Systematic Reviews

DART questions	Corresponding AMSTAR question(s)	Corresponding modified OQAQ question(s)
(1) Did the authors develop the research question(s) and inclusion/exclusion criteria before conducting the review?	(1) Was an "a priori" design provided?	Not addressed
(2) Did the authors describe the search methods used to find evidence (original research) on the primary question(s)?	(3) Was a comprehensive literature search performed?	(1) Were the search methods used to find evidence on the primary question stated?
(2a) Are key words and/or MESH terms stated?	(3) Was a comprehensive literature search performed?	Not addressed
(3) Was the search for the evidence reasonably comprehensive?	(3) Was a comprehensive literature search performed?	(2) Was the search for evidence reasonably comprehensive?
(3a) Does the search include at least 2 databases?	(3) Was a comprehensive literature search performed?	Not addressed
(3b) Did the authors choose the most applicable electronic databases and only limit the search by date when performing an update?	Not addressed	Not addressed
(3c) Are search methods likely to capture all relevant studies and did the authors hand-search journals or reference lists to identify published studies which were not electronically available?	(3) Was a comprehensive literature search performed?	Not addressed
	(4) Was the status of publication (*i.e.*, grey literature) used as an inclusion criterion?	
(4a) Did the authors provide in the inclusion criteria: Population, intervention, outcome, and study design, when selecting studies for the review?	Not addressed	Not addressed
(4b) Did the authors state whether the selection criteria were applied by more than one person?[1]	(2) Was there duplicate study selection and data extraction?[1]	Not addressed
(4c) Did the authors state how disagreements were resolved during study selection?[1]	(2) Was there duplicate study selection and data extraction?[1]	Not addressed
(4d) Did the authors provide a flowchart or descriptive summary of the included and excluded studies?	(5) Was a list of studies (included and excluded) provided?	Not addressed
(4e) Did the authors include all study designs appropriate for the research questions posed?	Not addressed	Not addressed
(5) Were the characteristics of the included studies provided? (in an aggregated form such as a table, data from the original studies were provided on the participants, interventions and outcomes)	(6) Were the characteristics of the included studies provided?	Not addressed
(6) Did the authors make any statements about assessing for publication bias?	(10) Was the likelihood of publication bias assessed?	Not addressed
(7a) Was the quality assessment specified with adequate detail to permit replication?	(7) Was the scientific quality of the included studies assessed and documented?	(5) Were the criteria used for assessing the validity of the included studies reported?
(7b) Was the quality assessment conducted independently by more than one person?	Not addressed	Not addressed
(7c) Did the authors state how disagreements were resolved during the quality assessment?	Not addressed	Not addressed
(8) Did the authors appropriately assess for quality by appropriately examining the following sources of bias in all of the included studies: confounding, sufficient sample size, outcome reporting bias, follow-up, randomization, allocation concealment, blinding, selection bias, information bias, verification bias, and differences between the first and second study measurement point?	(7) Was the scientific quality of the included studies assessed and documented? (partial match)	(6) Was the validity of all studies referred to in the text assessed using appropriate criteria? (partial match)
(9) Did the authors use appropriate methods to extract data from the included studies?	Not addressed	Not addressed
(9a) Were standard forms developed and piloted prior to the systematic review conduct?	Not addressed	Not addressed
(9b) Did the authors ensure that data from the same study that appeared in multiple publications were counted only once in the synthesis?	Not addressed	Not addressed
(9c) Was data extraction performed by more than one person?	(2) Was there duplicate study selection and data extraction?	Not addressed
(10) Did the authors assess and account for heterogeneity (differences in participants, interventions, outcomes, and trial design, quality or treatment effects) among the studies selected for the review?	(9) Were the methods used to combine the findings of studies appropriate?	(7) Were the methods used to combine the findings of the relevant studies reported?

		(8) Were the findings of the relevant studies combined appropriately?
(11) Did the authors describe the methods they used to combine/synthesize the results of the relevant studies (to reach a conclusion) and were the methods used appropriate for the review question(s)?	(9) Were the methods used to combine the findings of studies appropriate?	(7) Were the methods used to combine the findings of the relevant studies reported?
		(8) Were the findings of the relevant studies combined appropriately?
(12) Did the authors perform sensitivity analyses on any changes in protocol, assumptions, and study selection? (For example, using sensitivity analysis to compare results from fixed effects and random effects models)	Not addressed	Not addressed
(13) Are the conclusions of the authors supported by the reported data with consideration of the overall quality of that data?	(8) Was the scientific quality of the included studies used appropriately in formulating conclusions? (partial match)	(9) Were the conclusions made by the author(s) supported by the data reported? (partial match)
(14) Were conflicts of interest stated and were individuals excluded from the review if they reported substantial financial and intellectual COIs?	(11) Was the conflict of interest stated? (partial match)	Not addressed
(15) On a scale of 1-10, how would you judge the overall quality of the paper?	Not addressed	(10) Overall quality

[1]Separate questions in DART, but concepts not separated in AMSTAR. DART: Documentation and Appraisal Review Tool; OQAQ: Overview Quality Assessment Questionnaire; AMSTAR: Assessment of Multiple Systematic Reviews; COIs: Conflicts of interests.

mapped OQAQ and AMSTAR to DART and displayed the results in Table 3. Table 3 shows that our tool includes several questions that are unique and not included in the modified OQAQ or AMSTAR, with several other questions covered by one or the other but not both tools.

Assessing consistency of performance of the three tools

Throughout the iterations of development, testing and group discussion and review of performance, we learned that the modified OQAQ and DART consistently produced similar overall assessments of quality. However, during these discussions we had more difficulty remembering or locating reviewer rationale for the responses using the modified OQAQ tool. DART has sufficient space to record page and line details to facilitate recall. This was important when resolving disputes. We also discovered that the AMSTAR tool had questions that were confusing and difficult to implement consistently. They are the following: (1) Question 4: Was the status of publication (*i.e.,* grey literature) used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports from the systematic review, based on their publication status, language, *etc.* This question was confusing since it seemed to equate an accurate description of the extent of the search with the actual execution of a thorough search; (2) Question 5: Was a list of studies (included and excluded) provided? This question was interpreted as being too specific by requiring lists, and did not allow for a good flow chart; it seemed to require more detail than most journal space would allow; (3) Question 7: Was the scientific quality of the included studies assessed and documented? A priori methods of assessment should be provided [*e.g.,* for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria]; for other types of studies alternative items will be relevant. This question did not provide sufficient detail to execute consistently. We found it more useful to specify the most important sources of bias by study type for consistent reporting both within and across reviewers; and (4) Question 11: Was the conflict of interest stated? Potential sources of support should be clearly acknowledged in both the systematic review and the included studies. The answer to this question was always no. Systematic review authors often mention their personal sources of support, but we did not find an example where potential sources of support were provided for the included studies. This needs to either be two questions, or allow for partial scoring.

DART was the only one of the three tools to explicitly include quality review for biases specific to observational studies. Since the importance of including evidence from observational data is now more widely recognized, particularly for assessing risk in order to generate recommendations that balance benefit to harm, we believe it is important to include careful assessment of the potential for biased measurement unique to this design.

DISCUSSION

We are aware that a revision of the AMSTAR tool exists and is known as R-AMSTAR[7]. The primary goal for revising AMSTAR was to produce an overall quantitative estimate of the quality of the systematic review. The performance of R-AMSTAR has been compared to the original tool using systematic reviews from the field of assisted reproduction for subfertility[10]. In that comparison study, R-AMSTAR was noted to provide more guidance to the reviewer than AMSTAR, but was

more difficult to apply consistently. Popovich *et al*[10] reported that the R-AMSTAR criteria were difficult to apply because of subjectivity of some of the domains, especially domain 8. That question "Was the scientific quality of the included studies used appropriately in formulating conclusion?" provided four criteria, which Popovich *et al*[10] report as being difficult to distinguish. Their kappa statistics also showed poor inter-rater reliability for this domain.

We designed the DART quality assessment tool to address limitations we discovered when using the modified OQAQ and AMSTAR tools. The specific improvements are: (1) Space for enhanced recording detail to facilitate reconciliation between reviewers and provide detailed reference for use in future updates; (2) An evaluation of major biases relevant to observational study designs and the assessment of standards recommended by the March 2011 IOM Standards for Systematic Review[3]; (3) Additional detail and guidance for junior epidemiologists, clinicians and other members of the review panel with less experience in systematic review methods; and (4) Consistent overall quality assessment of systematic reviews using a qualitative ranking that categorizes studies as good, fair or poor at the end of a detailed assessment.

In order to facilitate the use of systematic reviews, the American College of Chest Physicians (CHEST) adopted DART to assess the quality of systematic reviews included in their evidence reviews. CHEST guideline authors used DART to assess the quality of systematic reviews and meta-analyses included in the "Diagnosis and Management of Lung Cancer: CHEST Evidence-Based Clinical Practice Guideline (3rd Edition)"[11], and subsequent guidelines. DART has been used for other CHEST guidelines and it is discussed in the article Methodologies for the Development of CHEST Guidelines and Expert Panel Reports[12].

This paper describes the development of DART for systematic reviews. The next step is to quantify the performance of components of the tool through validation testing, assessing inter-rater agreement scores. Based on our preliminary evaluation with the modified OQAQ and AMSTAR, intra-rater reliability should also be tested when assessing the same systematic review at a later point in time, since updated evidence reviews are essential to ensuring that the best current evidence informs clinical guidelines and policy. The ability to facilitate accurate recall of prior reviews will improve the efficiency of that process.

The authors now have considerable experience and familiarity with DART and can complete the assessment form quickly. It is therefore important to use an external validation process to test performance in persons with a wide variety of backgrounds and without prior experience with the tool in order to evaluate inter and intra rater consistencies in response and time for completion.

Well-executed systematic reviews now form the foundation of evidence-based clinical practice guidelines. Even though the IOM has developed rigorous standards for conducting systematic reviews, there is still wide variation in how they are conducted and reported. Given this variation and the new reliance on systematic reviews, comprehensive tools are needed to assess the quality of systematic reviews. By creating the DART for Systematic Reviews we attempted to fill this gap.

ACKNOWLEDGMENTS

We would like to thank the interns in the Center for Clinical Excellence at BJC HealthCare for assisting us with testing DART and giving us feedback on modifications to the tool.

COMMENTS

Background

Systematic reviews are the foundation for evidence-based guidelines. Rigorous standards exist, but there is wide variation in implementation, highlighting the need for a more comprehensive quality assessment tool for systematic reviews.

Research frontiers

As the publication of systematic reviews increases, variability in the quality still exists. Users of systematic reviews need a way to assess the quality of systematic reviews that includes all relevant study designs. Since the importance of including evidence from observational data is now more widely recognized, especially to assess potential for harm, a single tool is needed that includes careful assessment of the potential for biased measurement unique to this design as well as for randomized trials.

Innovations and breakthroughs

The authors designed the the Documentation and Appraisal Review Tool (DART) quality assessment tool to address limitations they discovered when using the modified Overview Quality Assessment Questionnaire and Assessment of Multiple Systematic Reviews tools. The specific improvements include: the ability to record rationale for each criteria; criteria for assessing observational studies and for assessing standards recommended by the Institute of Medicine in 2011; additional guidance to assist less experienced reviewers in assessing the quality of systematic reviews; and consistent overall quality assessment of systematic reviews using a qualitative ranking.

Applications

DART provides a comprehensive, systematic approach for reviewers with limited experience with systematic review methodology, to critically analyze systematic reviews. It also provides a complete record of judgments and decisions made during the assessment to assist reconciliation between reviewers during the current review and for use in future updates.

Terminology

The terminology used in this article reflects the vocabulary familiar to an audience using systematic reviews for decision-making.

Peer-review

The peer reviewers did not report having any concerns about the paper. Reviewer comments included the following: Systematic reviews are the foundation for evidence-based guidelines and are increasing. The article discusses the development of a comprehensive tool that improves upon existing tools for assessing the quality of systematic reviews and that guides reviewers through critically analyzing a systematic review. It has significance to appraise a systematic review.

REFERENCES

1 **Grant MJ**, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J* 2009; **26**: 91-108 [PMID: 19490148]

2 **Bastian H**, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 2010; **7**: e1000326 [PMID: 20877712]

3 **Institute of Medicine (US) Committee on Standards for**

Systematic Reviews of Comparative Effectiveness Research. Finding what works in health care: standards for systematic reviews. Eden J, Levit L, Berg A, Morton S, editors. Washington (DC): National Academies Press (US), 2011 [PMID: 24983062]

4 **Moher D**, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151**: 264-269, W64 [PMID: 19622511 DOI: 10.7326/0003-4819-151-4-200908180-00135]

5 **Shea B**, Moher D, Graham I, Pham B, Tugwell P. A comparison of the quality of Cochrane reviews and systematic reviews published in paper-based journals. *Eval Health Prof* 2002; **25**: 116-129 [PMID: 11868441]

6 **Shea BJ**, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007; **7**: 10 [PMID: 17302989]

7 **Kung J**, Chiappelli F, Cajulis OO, Avezova R, Kossan G, Chew L, Maida CA. From systematic reviews to clinical recommendations for evidence-based health care: validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. *Open Dent J* 2010; **4**: 84-91 [PMID: 21088686]

8 **White CM,** Ip S, McPheeters M, Carey TS, Chou R, Lohr KN, Robinson K, McDonald K, Whitlock E. Using existing systematic reviews to replace de novo processes in conducting comparative effectiveness reviews methods guide for effectiveness and comparative effectiveness reviews. Rockville (MD): Agency for Healthcare Research and Quality (US), 2008 [PMID: 21433402]

9 **Higgins JPT**, Green S, editors. Assessment of study quality. cochrane handbook for systematic reviews of interventions 4.2.6 [updated September 2006]. In: The Cochrane Library. Chichester, UK: John Wiley and Sons, Ltd, 2006: 384

10 **Popovich I**, Windsor B, Jordan V, Showell M, Shea B, Farquhar CM. Methodological quality of systematic reviews in subfertility: a comparison of two different approaches. *PLoS One* 2012; **7**: e50403 [PMID: 23300526 DOI: 10.1371/journal.pone.0050403]

11 **Lewis SZ**, Diekemper R, Addrizzo-Harris DJ. Methodology for development of guidelines for lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest* 2013; **143**: 41S-50S [PMID: 23649432]

12 **Lewis SZ**, Diekemper R, Ornelas J, Casey KR. Methodologies for the development of CHEST guidelines and expert panel reports. *Chest* 2014; **146**: 182-192 [PMID: 25010961 DOI: 10.1378/chest.14-0824]

P- Reviewer: Cid J, Gao C, Tang Y, Trohman RG **S- Editor**: Tian YL
L- Editor: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i3.151

World J Meta-Anal 2015 June 26; 3(3): 151-162
ISSN 2308-3840 (online)

Effectiveness of 7-valent pneumococcal conjugate vaccine: A meta-analysis of post-marketing studies

Chiara de Waure, Maria Lucia Specchia, Silvio Capizzi, Mufida Aljicevic, Milos Dujovic, Admir Malaj, Walter Ricciardi

Chiara de Waure, Maria Lucia Specchia, Silvio Capizzi, Walter Ricciardi, Institute of Public Health, Catholic University of the Sacred Heart, 00168 Rome, Italy
Mufida Aljicevic, Faculty of Medicine, University of Sarajevo, 71000 Sarajevo, Bosnia and Herzegovina
Milos Dujovic, Faculty of Medicine, University of Belgrade, 11000 Belgrade, Serbia
Admir Malaj, Faculty of Pharmacy, Medical University of Tirana, 1001 Tirana, Albania

Author contributions: All authors contributed to this work.

Conflict-of-interest: The authors have no conflicts of interest.
Data sharing: No additional data are available.

Correspondence to: Maria Lucia Specchia, MD, MPH, PhD, Institute of Public Health, Catholic University of the Sacred Heart, L.go F. Vito 1, 00168 Rome, Italy. marialucia.specchia@rm.unicatt.it
Telephone: +39-6-30154396
Fax: +39-6-35001522

Received: July 28, 2014
Peer-review started: July 30, 2014
First decision: December 17, 2014
Revised: April 22, 2015
Accepted: May 16, 2015
Article in press: May 18, 2015
Published online: June 26, 2015

Abstract

AIM: To investigate the 7-valent pneumococcal conjugate vaccine (PCV7) effectiveness.

METHODS: A systematic literature review of studies which evaluated the effectiveness of PCV7 vaccine was performed searching the keyword "heptavalent pneumococcal conjugate vaccine" in PubMed and Scopus until March 16, 2013. The selection of potential eligible articles was done by two researchers independently on the basis of abstract and title and only post-marketing studies were included in the systematic review. Data extraction was carried out by two researchers with respect to invasive pneumococcal diseases due to both all and vaccine serotypes in pre-vaccine and post-vaccine periods in children less than 5 years. Results of studies which were considered suitable for meta-analysis were combined by means of relative risk (RR) with 95%CI. Vaccine effectiveness was calculated as (1-RR) × 100. Heterogeneity was assessed by I^2 and a random effects model was used to combine data in the case of heterogeneity. RevMan 5 was used to pool data.

RESULTS: On the whole, 757 eligible papers were identified from the literature search in PubMed and Scopus. Of them, 62 were finally considered in the systematic review and 38 were included in the meta-analysis. In all post-marketing studies included in the systematic review the incidence of invasive pneumococcal diseases due to vaccine serotypes declined significantly with the exception of few studies showing stability or a slight, but not significant, increase. Furthermore most of studies highlighted also a reduction in the incidence of invasive pneumococcal diseases due to all serotypes. With regards to meta-analysis, a random effects model was used to combine data because of the high heterogeneity. Data combination showed that the effectiveness of PCV7 in reducing invasive pneumococcal diseases due to vaccine serotypes and to all serotypes was 84% (95%CI: 74%-90%) and 53% (95%CI: 46%-59%) respectively. These results are confirmatory with respect to the efficacy of PCV7 against invasive pneumococcal diseases

due to vaccine serotypes.

CONCLUSION: PCV7 implementation determines a significant decrease of invasive pneumococcal diseases.

Key words: Streptococcus pneumoniae; Pneumococcal infections; Pneumococcal vaccines; Treatment outcome; Meta-analysis

Core tip: This systematic review and meta-analysis was performed with the aim to collect data from post-marketing studies on 7-valent pneumococcal conjugate vaccine (PCV7) and to provide evidence about the impact of the vaccine in the real world. Eligible articles were identified through a search on PubMed and Scopus. The meta-analysis showed that PCV7 is able to reduce invasive pneumococcal diseases due to both vaccine serotypes and to all serotypes. The effectiveness was 84% (95%CI: 74%-90%) and 53% (95%CI: 46%-59%) respectively. These data may be taken into consideration in order to foresee the impact under real conditions of PCV13 which has replaced PCV7 from 2010 onwards.

de Waure C, Specchia ML, Capizzi S, Aljicevic M, Dujovic M, Malaj A, Ricciardi W. Effectiveness of 7-valent pneumococcal conjugate vaccine: A meta-analysis of post-marketing studies. *World J Meta-Anal* 2015; 3(3): 151-162 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i3/151.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i3.151

INTRODUCTION

Streptococcus pneumoniae (*S. pneumoniae*) is a leading cause of severe bacterial infectious disease and World Health Organization has estimated that this bacteria causes 1.4-1.6 million child deaths annually[1,2], in that around 11% of all deaths in children < 5 years[3]. More than 90 serotypes of *S. pneumoniae* exist. These strains may cause invasive pneumococcal disease (IPD). The highest incidence of IPD is seen in children < 2 years old. In order to prevent disease caused by *S. pneumoniae,* two types of vaccines, polysaccharide (PPV) and conjugate (PCV) exist, even though the PPV vaccine is ineffective in children < 2 years old[4].

The PCV vaccines consist of capsular PPVs bound to proteins which are highly immunogenic and enhance an immune response by recruiting type 2 helper T cells, which allows for immunoglobulin type switching and production of memory B cells. The main drawbacks of PCV vaccines are that they only provide protection against a subset of serotypes covered by the PPV vaccines[5-7]. In fact, PCV vaccines encompass the 7-valent vaccine (PCV7), the PCV10 and the PCV13. Currently, PCV13 is used in prevention campaigns. Its marketing authorization in the European Union goes back to December 2009[8]. PCV13 has replaced PCV7 from 2010 onward.

The PCV7, providing protection against serotypes 4, 6B, 9V, 14, 18C, 19F and 23F, was introduced into routine childhood immunization program in the United States in 2000 and was shown to reduce the incidence of IPD by all and vaccine-serotypes[9,10]. Notwithstanding, some studies have described significant rises in non-vaccine serotypes after the implementation of universal PCV7 programs[11-14]. Based on the favourable United States experience and the proof of vaccine efficacy[15] a number of countries have introduced PCV7[16]. Worldwide the vaccine has been provided with different schedules. In Europe both the 2 + 1 and 3 + 1 schedules have been used[16].

In the light of monitoring the health impact of technologies and policies, data from the real practice should be collected and analysed. Because of the recent introduction and implementation of PCV13, many data from real practice are only available for PCV7 even though evidence is being produced on PCV13 also[17-23]. Notwithstanding, this evidence should be considered early and is still scant in order to make a meta-analysis. Furthermore, it is mostly related to the transition period between the use of PCV7 and the introduction of PCV13 which took place from 2010 onward with different time schedules across countries. Based on this premises, the objective of this study was to perform a systematic review and a meta-analysis of post-marketing studies on the effectiveness of PCV7 in comparison with no vaccination in preventing IPD in children less than 5 years of age worldwide. The final aim was to provide evidence about PCV7 effectiveness under real conditions and to foresee the potential impact of PCV13 on the basis of results. The systematic review was performed according to PRISMA Statement published by Moher *et al*[24].

MATERIALS AND METHODS

Selection of articles

A literature search was conducted using PubMed and Scopus search engines. The following search strategy was used: "heptavalent pneumococcal PCV vaccine" (Substance Name) NOT ["Clinical Trial" (Publication Type) OR "Clinical Trials as Topic" (Mesh) OR "Controlled Clinical Trial" (Publication Type) OR "Clinical Trial, Phase Ⅳ" (Publication Type) OR "Clinical Trial, Phase Ⅲ" (Publication Type) OR "Clinical Trial, Phase Ⅱ" (Publication Type) OR "Clinical Trial, Phase Ⅰ" (Publication Type)]. The search covered the period up to March 16, 2013, without starting date, and was limited to English-language publications.

The selection of potential eligible articles was done by two researchers independently on the basis of title and abstract. Full text of eligible articles was collected for the final judgment on inclusion. Disagreements were solved through consensus or the consultation of a third researcher.

We defined a priori criteria for the inclusion of studies

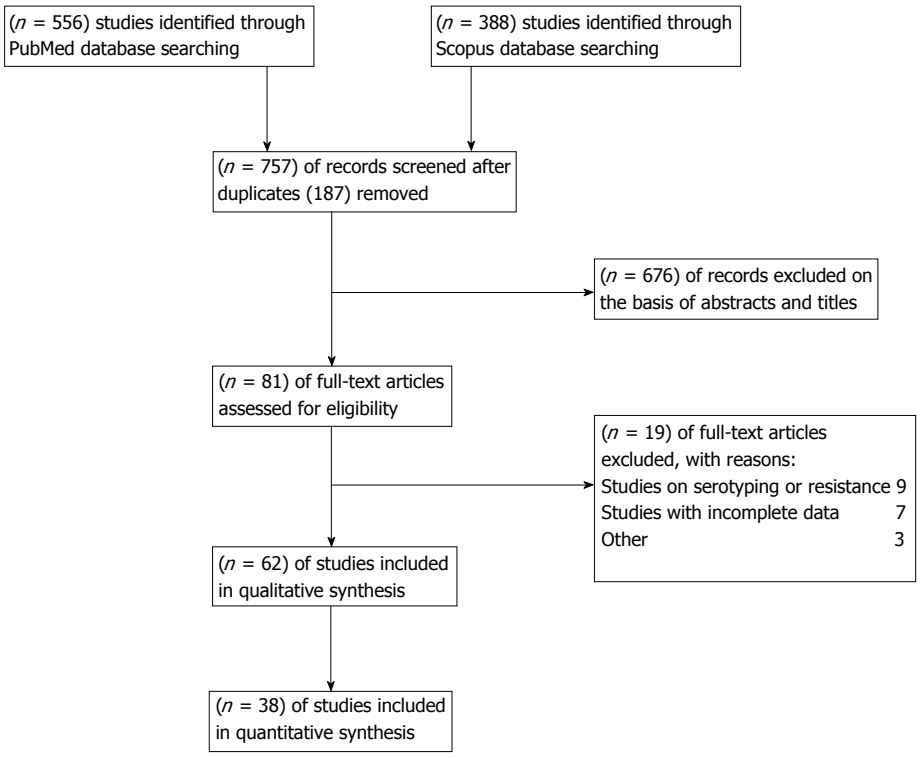


Figure 1 Flow-chart of studies selection.

in this meta-analysis, selecting studies dealing with the incidence of IPD in children less than 5 years of age in the period before and after the introduction of PCV7. Only articles releasing data on IPD incidence in pre- and post-vaccination periods were included in the quantitative assessment.

Data extraction

The following data were recorded from each study: first author, journal, published year, country, study population, IPD case definition, crude number or incidence of IPD before and after the introduction of PCV7. Data on IPD caused by all serotypes and due to vaccine serotypes, if available, were collected. Data extraction was performed by two researchers independently and disagreements were solved through consensus or the consultation of a third researcher.

Statistical analysis

Studies were included in the meta-analysis if they provided crude data or if it was possible to get them through computation.

The relative risk (RR) with 95%CI was used to combine data. Vaccine effectiveness was calculated as (1-RR) × 100. RevMan 5 was used to combine data and a fixed effects model was applied in the case of absence of heterogeneity (I^2 < 50%). On the other way around, a random effects model was used. Studies which were not considered in the meta-analysis were described qualitatively in Table 1. Finally, publication bias was assessed by means of funnel plots.

RESULTS

On the whole, 556 articles were yielded from PubMed and 388 from Scopus but 187 papers were shared by the two databases for a total of 757 papers. Of them, 62 were finally considered in the systematic review (Figure 1)[25-86]. Their characteristics and results are shown in Table 1.

With respect to meta-analysis, 38 articles provided data on IPD due to all serotypes while 22 allowed the collection of data on IPD due to vaccine serotypes. Data combination showed a vaccine effectiveness of 84% for IPD due to vaccine serotypes (RR = 0.16, 95%CI: 0.10%-0.26; I^2 = 95%, Figure 2) and 53% (RR = 0.47, 95%CI: 0.41-0.54; I^2 = 95%, Figure 3) for IPD related to all serotypes. Publication bias could not be excluded with respect to the assessment of effectiveness against IPD due to vaccine serotypes while may be excluded as regards IPD due to all serotypes (Figures 4 and 5).

DISCUSSION

Our study aimed to review and combine data of post-marketing studies on PCV7 worldwide.

The analysis and data combination allowed us to investigate the effectiveness of PCV7 and its impact in terms of public health. Results are indeed useful for supporting decision-makers in the field of vaccinations. In particular, findings of the meta-analysis showed that the effectiveness of PCV7 in reducing IPD due to vaccine serotypes is 84%. The effectiveness is estimated to be 53% with respect to IPD due to all serotypes.

Table 1 Summary of studies characteristics and results

Ref.	Country	Study period	Invasive pneumococcal disease definition	Main results
Albrich *et al*[25]	United States	1997-2004	Isolation of *S. pneumoniae* from sterile body fluid	1
Ampofo *et al*[26]	United States	1997-2010	Isolation of *S. pneumoniae* from sterile body fluid	The proportion of children younger than 2 yr with IPD decreased (54% *vs* 43% with respect to all serotypes and 56% *vs* 43% for vaccine serotypes), while the proportion of disease among children aged 2-4 slightly increased (27% *vs* 29% with respect to all and vaccine serotypes)
Aristegui *et al*[27]	Spain	1998-2003	Isolation of *S. pneumoniae* from sterile body fluid	1
Barricarte *et al*[28]	Spain	2001-2005	Isolation of *S. pneumoniae* from sterile body fluid	The overall effectiveness in reducing IPD was 31% (OR = 0.69, 95%CI: 0.37-1.27) and 88% (OR = 0.12, 95%CI: 0.02-0.91) for all serotypes and vaccine serotypes respectively
Benito-Fernández *et al*[29]	Spain	2000-2005	Isolation of *S. pneumoniae* from sterile body fluid	1
Ben-Shimol *et al*[30]	Israel	1989-2010	Isolation of *S. pneumoniae* from sterile body fluid	In 2009 and 2010, IPD incidence (due to vaccine serotypes) were 15.9 per 100000 and 5.4, per 100000 respectively (a 43% and 81% decrease compared to 2003-2007)
Bjornson *et al*[31]	Canada	2001-2005	Isolation of *S. pneumoniae* from sterile body fluid	1
Calbo *et al*[32]	Spain	1999-2004	Isolation of *S. pneumoniae* from sterile body fluid	The IPD incidence significantly decreased from 96.9 cases per 100000 person-years to 90.6 cases per 100000 person-years (7% reduction)
Carstairs *et al*[33]	United States	2000-2002	Isolation of *S. pneumoniae* from sterile body fluid	1
Casado-Flores *et al*[34]	Spain	2001-2006	Isolation of *S. pneumoniae* from sterile body fluid	1
CDC[35]	United States	1998-2005	Isolation of *S. pneumoniae* from sterile body fluid	1
De Serres *et al*[36]	United States	2001-2009	Isolation of *S. pneumoniae* from sterile body fluid	Effectiveness of PCV7 against IPD due to vaccine serotypes was 97% (95%CI: 92%-98%) among healthy children and 88% (95%CI: 78%-94%) among children with comorbid conditions. The incidence of IPD due to non-vaccine serotypes increased from 6.8 per 100000 (1998-1999) to 10.3 per 100000 in 2007 (51% increase)
De Wals *et al*[37]	Canada	2007-2010	Isolation of *S. pneumoniae* from sterile body fluid	A decrease in the frequency of IPD caused by vaccine serotypes was observed
Dias *et al*[38]	Portugal	1999-2004	Isolation of *S. pneumoniae* from sterile body fluid	1
Vestrheim *et al*[39]	Norway	2004-2008	Isolation of *S. pneumoniae* from sterile body fluid	1
Dubos *et al*[40]	France	2000-2005	Isolation of *S. pneumoniae* from sterile body fluid	A decrease of 82% (95%CI: 52%-95%) of cases was observed (from 8.9 cases per 100000 in 2001 to 1.8 per 100000 in 2005) in children < 2 yr
Fenoll *et al*[41]	Spain	1996-2001 2005-2006	Isolation of *S. pneumoniae* from sterile body fluid	A decrease of the incidence of IPD due to vaccine serotypes from 5.2 per 100000 in 1996-2001 to 2.4 per 100000 in 2005-2006 was observed
Flannery *et al*[42]	United States	1998-2002	Isolation of *S. pneumoniae* from sterile body fluid	1
Giele *et al*[43]	Australia	1996-2005	Isolation of *S. pneumoniae* from sterile body fluid	1
Schutze *et al*[44]	Arkansas	1998-2003	Isolation of *S. pneumoniae* from sterile body fluid	A decrease of IPD from 44.2 per 100000 person-years to 8.30 per 100000 person-years was observed in children < 2 yr
Guevara *et al*[45]	Spain	2001-2007	Isolation of *S. pneumoniae* from sterile body fluid	1
Haddy *et al*[46]	United States	1999-2002	Isolation of *S. pneumoniae* from sterile body fluid	1
Hanna *et al*[47]	Queensland	1999-2007	Isolation of *S. pneumoniae* from sterile body fluid	1
Hanquet *et al*[48]	Belgium	2002-2008	Isolation of *S. pneumoniae* from sterile body fluid	1
Harboe *et al*[49]	Denmark	2000-2008	Isolation of *S. pneumoniae* from sterile body fluid	In children < 2 yr, the overall incidence decreased from 54 to 23 cases per 100000 (IRR = 0.43, 95%CI: 0.29-0.62) and from 36.7 to 7.7 (IRR = 0.20, 95%CI: 0.09-0.38) for vaccine serotypes. A non-significant increase was observed in children aged 2-4 yr
Hennessy *et al*[50]	United States	1995-2003	Isolation of *S. pneumoniae* from sterile body fluid	1
CDC[51]	United States	1998-2003	Isolation of *S. pneumoniae* from sterile body fluid	A decrease of IPD due to vaccine serotypes from 80 cases per 100000 to 4.6 per 100000 was observed (decrease of 94% (95%CI: 92%-96%) from 1998-1999 to 2003
Hsu *et al*[52]	United States	1998-2005	Isolation of *S. pneumoniae* from sterile body fluid	1

Hsu *et al*[53]	United States	1990-1991 2001-2003	Isolation of *S. pneumoniae* from sterile body fluid	1
Hsu *et al*[54]	United States	2001-2007	Isolation of *S. pneumoniae* from sterile body fluid	IPD incidence was stable during the 6 yr period, although IPD due to vaccine serotypes decreased
Ingels *et al*[55]	Denmark	2000-2010	Isolation of *S. pneumoniae* from sterile body fluid	1
Wenger *et al*[56]	United States, Alaska	1986-2007	Isolation of *S. pneumoniae* from sterile body fluid	1
Johnson *et al*[57]	South Australia	2002-2009	Isolation of *S. pneumoniae* from sterile body fluid	1
Kellner *et al*[58]	Canada	1998-2007	Isolation of *S. pneumoniae* from sterile body fluid	1
Kyaw *et al*[59]	United States	1996-2004	Isolation of *S. pneumoniae* from sterile body fluid	1
Leal *et al*[60]	Alberta	1998-2010	Isolation of *S. pneumoniae* from sterile body fluid	1
Liao *et al*[61]	Taiwan	2000-2008	Isolation of *S. pneumoniae* from sterile body fluid	The overall incidence of IPD decreased by 33% (95%CI: 0%-72.2%)
Messina *et al*[62]	United States	1999-2001 2003-2005	Isolation of *S. pneumoniae* from sterile body fluid	1
Muñoz-Almagro *et al*[63]	Spain	1997-2006	Isolation of *S. pneumoniae* from sterile body fluid	1
Patrzalek *et al*[64]	Poland	2005-2010	Isolation of *S. pneumoniae* from sterile body fluid	1
Pérez *et al*[65]	Spain	1998-2008	Isolation of *S. pneumoniae* from sterile body fluid	1
Pérez-Trallero *et al*[66]	Spain	1996-2007	Isolation of *S. pneumoniae* from sterile body fluid	1
Pilishvili *et al*[67]	United States	1998-2007	Isolation of *S. pneumoniae* from sterile body fluid	1
Poehling *et al*[68]	United States	1997-2004	Isolation of *S. pneumoniae* from sterile body fluid	1
[69]	Canada	2002-2005	Isolation of *S. pneumoniae* from sterile body fluid	1
Ramani *et al*[70]	United States	1994-2001	Hospital discharges for IPD	A significant decrease was observed only for children aged < 1 yr (from 40 per 100000 to 23 per 100000 person years). All other age groups did not show a significant change in discharge rates for IPD
Rendi-Wagner *et al*[71]	Austria	2001-2007	Isolation of *S. pneumoniae* from sterile body fluid	1
Rodenburg *et al*[72]	Netherlands	2004-2008	Isolation of *S. pneumoniae* from sterile body fluid	1
Rückinger *et al*[73]	Germany	1997-2003 2007-2008	Isolation of *S. pneumoniae* from sterile body fluid	1
de Sevilla *et al*[74]	Spain	2007-2009	Isolation of *S. pneumoniae* from sterile body fluid	An increase of 44% of IPD (95%CI: 10%-89%) was shown
Shafinoori *et al*[75]	United States	1998-2004	Isolation of *S. pneumoniae* from sterile body fluid	A significant 68% and 70% decrease of IPD in children < 2 yr and aged 2 to 4 yr respectively was observed
Shah *et al*[76]	United States	1999-2003	Hospital discharges for IPD	A significant decrease from 12.03 per 100000 person-years in 1999 to 5.60 per 100000 person-years in 2003 was shown
Techasaensiri *et al*[77]	United States	1999-2008	Isolation of *S. pneumoniae* from sterile body fluid	The incidence of IPD significantly decreased in children < 2 yr
Tsai *et al*[78]	United States	1994-1999 2001-2004	Hospital discharges for pneumococcal meningitis	The average annualized rates of hospitalizations decreased from 7.7 per 100000 to 2.6 per 100000 in children < 2 yr and from 0.9 per 100000 to 0.5 per 100000 in children aged 2-4 (a reduction of 66%, 95%CI: 56.3%-73.5% and of 51.5%, 95%CI: 28.9%-66.9% respectively)
Tsigrelis *et al*[79]	United States	1995-2007	Isolation of *S. pneumoniae* from sterile body fluid	1
Tyrrell *et al*[80]	Canada	2000-2006	Isolation of *S. pneumoniae* from sterile body fluid	IPD due to vaccine serotypes decreased of 61% in children < 2 yr (from 96.7 per 100000 person-years to 25.8 per 100000 person-years) and of 57% in children from 2 to 4 yr (from 24.5 per 100000 person-years to 10.6 per 100000 person-years)
Van der Linden *et al*[81]	Germany	1997-2010	Isolation of *S. pneumoniae* from sterile body fluid	IPD incidence decreased from 2.4 per 100000 to 0.3 per 100000
Vestrheim *et al*[82]	Norway	2002-2007	Isolation of *S. pneumoniae* from sterile body fluid	1
Weatherholtz *et al*[83]	United States	1995-2006	Isolation of *S. pneumoniae* from sterile body fluid	Rates of IPD due to vaccine serotypes among children aged < 1 yr, 1-2 yr, and 2-5 yr decreased from 210, 263, and 51 cases per 100000 respectively in to 0 case per 100000
Whitney *et al*[84]	United States	1998-2001	Isolation of *S. pneumoniae* from sterile body fluid	1

Winters *et al*[85]	Canada	2002-2005	Isolation of *S. pneumoniae* from sterile body fluid	The incidence of IPD decreased from 54 per 100000 person-years to 16 per 100000 person-years (decrease of 70%). An even stronger decrease was observed in children < 1 yr, where the incidence decreased from 135 per 100000 to 15 per 100000 person-years (decrease of 89%)
Yildirim *et al*[86]	United States	2007-2010	Isolation of *S. pneumoniae* from sterile body fluid	IPD cases due to vaccine serotypes decreased

[1]Studies included in the meta-analysis. IPD: Invasive pneumococcal disease; CDC: Centers for Disease Control and Prevention; IRR: Incidence rate ratio; PCV7: 7-valent pneumococcal conjugate vaccine; *S. pneumoniae*: *Streptococcus pneumoniae*.

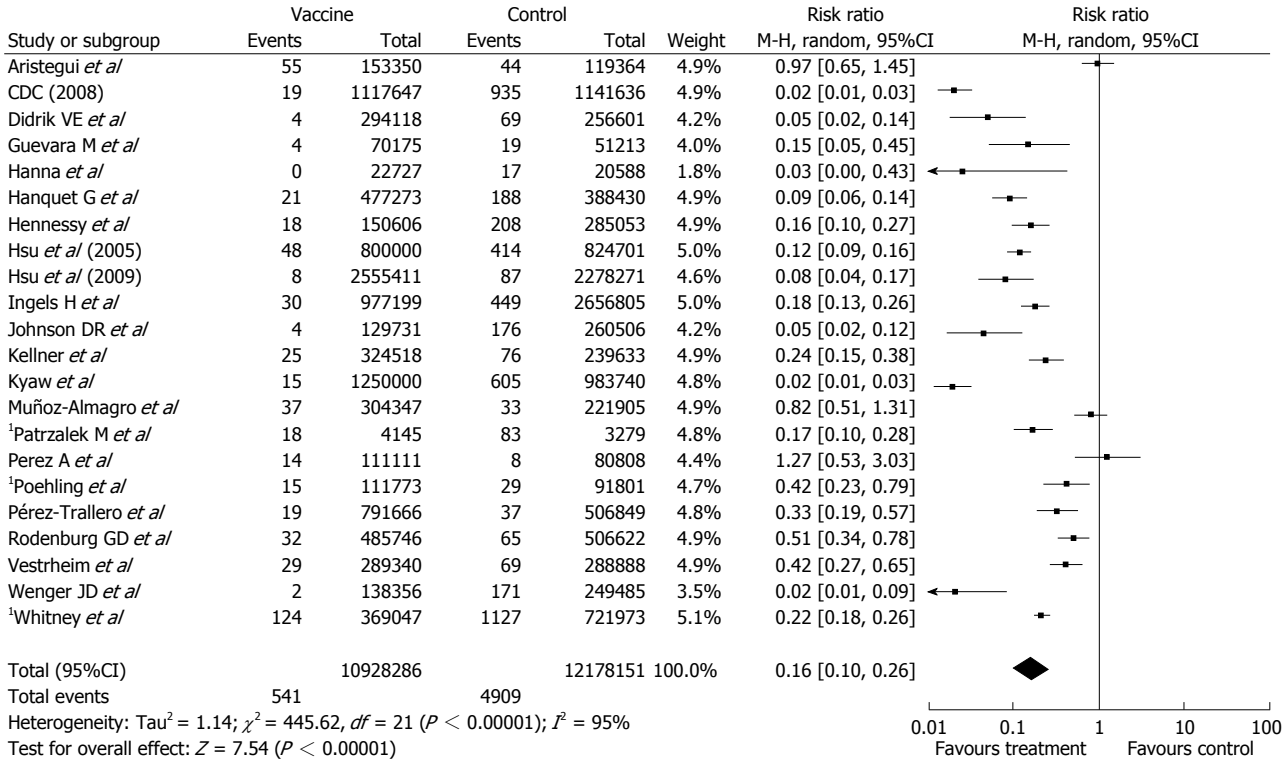


Figure 2 Data combination for invasive pneumococcal disease due to vaccine serotypes. [1]Data available not for the entire age group < 5 years. CDC: Centers for Disease Control and Prevention.

The results of our study are aligned with the evidence on the efficacy of PCV7 demonstrated in randomized clinical trials (RCT). In fact, a meta-analysis of RCT conducted by Pavia *et al*[15] showed an efficacy of 89% in preventing IPD due to vaccine serotypes, and of 63%-74% in preventing IPD due to all serotypes. Indeed, as IPD due to vaccine serotypes, effectiveness data have confirmed efficacy data. With this respect it is important to point out that the assessment of efficacy of interventions is critical in order to decide upon their adoption and is addressed through explanatory clinical trials[87]. Notwithstanding, the proof of efficacy is not always sufficient because it is also important to have evidence about how interventions work under more natural field conditions rather than in controlled clinical trials[87,88]. Indeed, overall effectiveness of interventions should be assessed by different study designs able to maximize external validity[87].

As far as PCV7 is concerned, all post-marketing studies showed that the incidence of IPD due to vaccine serotypes declined significantly after the implementation of vaccination, with the exception of few studies[36,27,49,63,65] showing a stability or a slight increase. As a consequence, the implementation of vaccination has definitively contributed in consistently preventing IPD in children up to 5 years of age with a strong impact on population health and costs due to hospitalizations[89,90]. In fact, a relevant reduction of IPD due to all serotypes was also shown by the meta-analysis even though, comparing with IPD due to vaccine serotypes, more studies highlighted a stability or an increase in the overall incidence of IPD[26,32,38,45,48,52,63,65,66,70-72,74,79]. In particular two studies[64,75] showed a significant increase although due to non-vaccine serotypes and in a context of low vaccination coverage. The increase in the incidence of non-vaccine serotypes is a well-known phenomenon which may be counteracted by the extension of serotypes coverage. In this view the availability and the implementation of PCV13 is useful in order to further reduce the incidence of IPD. In fact, the post-licensure assessment already carried out by Andrews *et al*[23] estimated that the effectiveness of at least 2 doses of

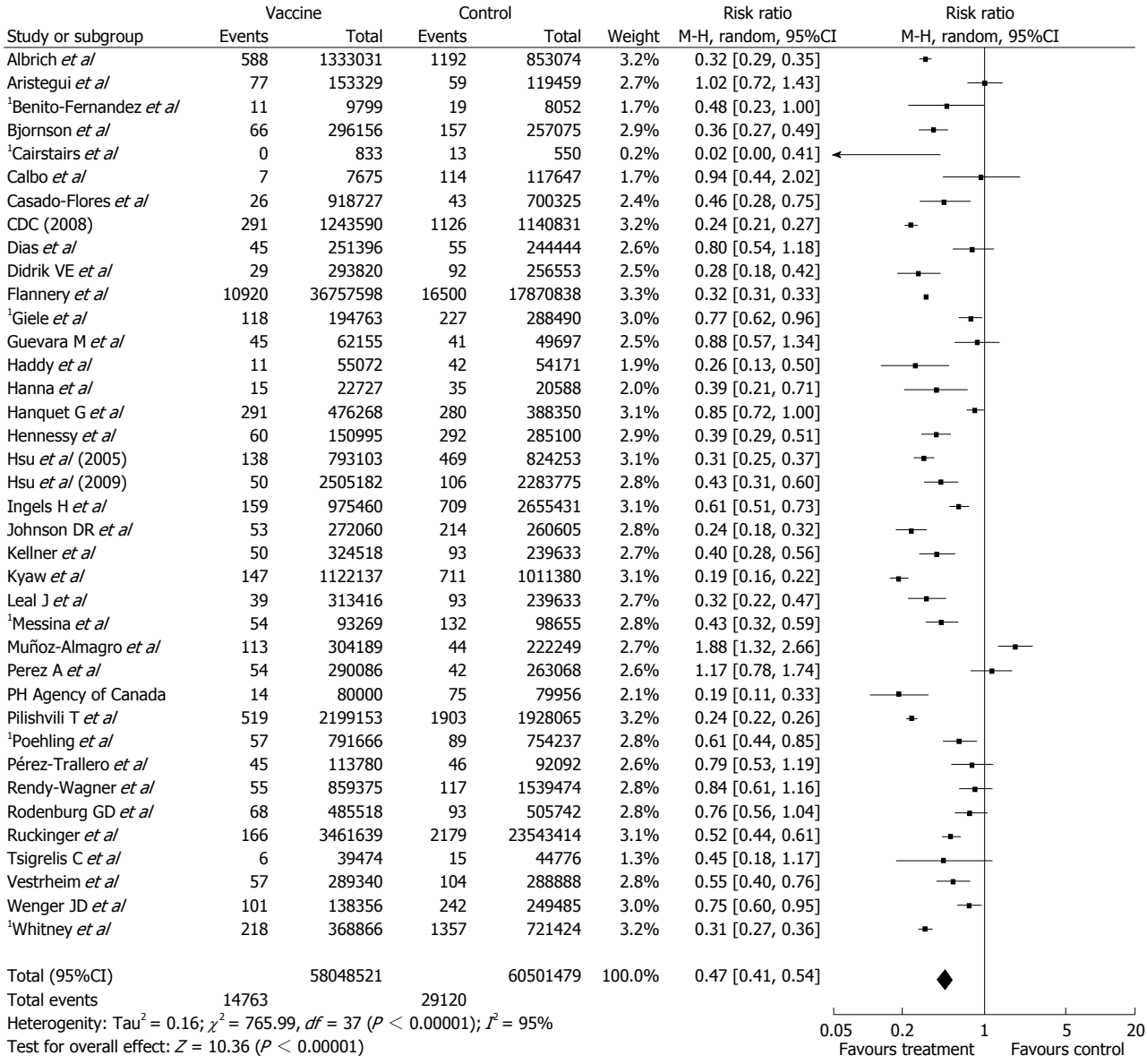


Figure 3 Data combination for invasive pneumococcal disease due to all serotypes. [1]Data available not for the entire age group < 5 years. CDC: Centers for Disease Control and Prevention.

PVC13 before 12 mo of age or of 1 dose from 12 mo onwards was 90% (95%CI: 34%-98%) against PCV7 serotypes. This result is aligned with data from our and Pavia *et al*[15] meta-analyses. Furthermore, PCV13 was shown to have an effectiveness of 73% (95%CI: 55%-84%) against the additional serotypes included in the vaccine[23]. PCV13 may indeed provide an added value in comparison to PCV7. In fact, already available population-based studies showed that IPD decreased of a percentage from 18% to 42% when PCV13 era is compared to PCV7 one[18,20,21]. The decline is more important in children less than 2 years of age in which the decrease in all IPD varies from 50% to 60%[18,20,21].

This study presents some limitations. The research was limited to only two specialized searching engines and, consequently, selection bias may be not excluded. Papers included in the review were heterogeneous with respect to countries and study design as also highlighted by the test of heterogeneity. Crude data were not obtainable from all the papers selected and only children < 5 years of age, independently by their health status, were considered in the analysis. Furthermore, neither a quality assessment nor stratified analyses in order to investigate heterogeneity were performed.

Strengths of this study are represented by the objective itself, because we focused on effectiveness instead of efficacy, and the large number of papers included in the analysis.

The consistent decrease of IPD due to vaccine serotypes after the PCV7 implementation is important as the new PCV13 is being implemented. In fact, it is expected that it will have the same effectiveness in preventing IPD due PCV7 vaccine serotypes and it will also have an important impact on cases due to new vaccine

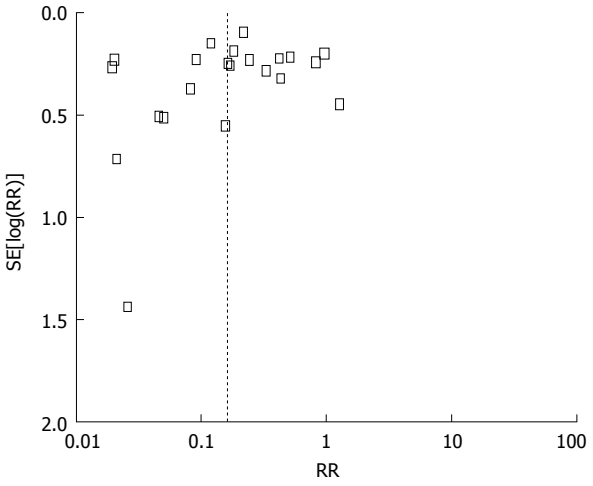


Figure 4 Funnel plot of studies on invasive pneumococcal disease due to vaccine serotypes. RR: Relative risk.

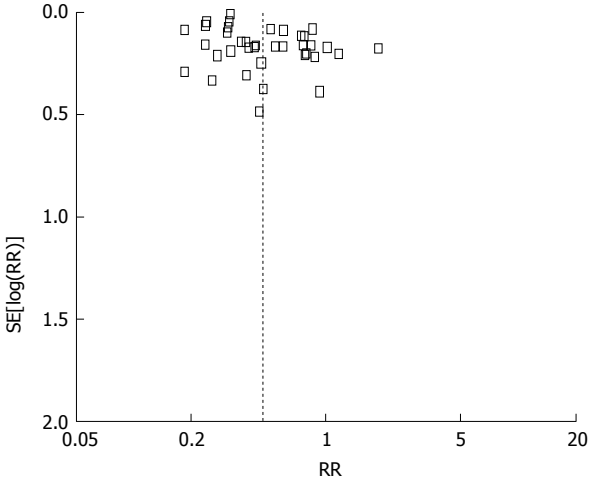


Figure 5 Funnel plot of studies on invasive pneumococcal disease due to all serotypes. RR: Relative risk.

serotypes[91,92].

COMMENTS

Background

Streptococcus pneumoniae (*S. pneumoniae*) is a leading cause of severe bacterial infectious disease, causing 1.4-1.6 million child deaths annually, in that around 11% of all deaths in children < 5 years. Two types of vaccines against *S. pneumoniae* exist, polysaccharide (PPV) and conjugate (PCV), even though the PPV vaccine is ineffective in children < 2 years old. PCV vaccines encompass the 7-valent vaccine (PCV7), the PCV10 and the PCV13. Currently, PCV13 is used in prevention campaigns. Its marketing authorization in the European Union goes back to December 2009 and it has replaced PCV7 from 2010 onward. Because of the recent introduction and implementation of PCV13, consistent data from real practice are only available for PCV7 and their assessment is of utmost importance in order to monitor the health impact of the vaccine.

Research frontiers

The monitoring of the overall health impact of technologies and policies is a key issue in medicine. It is mainly based on post-marketing studies on the effectiveness of interventions carried out through the collection and analysis of data from the real practice. In this context, the objective of the authors study was to perform a systematic review and a meta-analysis of post-marketing studies on the effectiveness of PCV7 worldwide.

Innovations and breakthroughs

The authors' findings showed that the effectiveness of PCV7 in reducing invasive pneumococcal disease (IPD) is 84% with respect to IPD due to vaccine serotypes and 53% with respect to IPD due to all serotypes. Concerning IPD due to vaccine serotypes, effectiveness data have confirmed efficacy data previously reported in a meta-analysis of randomized clinical trial conducted by Pavia *et al*. However, with this respect, it is important to emphasize that efficacy trials test the expected results of an intervention under ideal circumstances whereas effectiveness studies measure the beneficial effects under "real world" clinical settings. Indeed, the results of their meta-analysis represent and advance in the knowledge of PCV7 impact.

Applications

Given the consistent decrease of IPD due to vaccine serotypes after the PCV7 implementation, results of their systematic review and meta-analysis allow forecasting that the new PCV13, which is being implemented, will further decrease the number of IPD. In fact PCV13 effectiveness is expected to be the same as PCV7 in preventing IPD due to both PCV7 vaccine serotypes and new vaccine serotypes.

Terminology

IPD is defined as the isolation of *S. pneumoniae* from a sterile site/body fluid.

Peer-review

The authors performed an interesting and well-written meta-analysis on a highly relevant topic.

REFERENCES

1 **World Health Organization**. Acute respiratory infections. [Update 2009 Sept]. Available from: URL: http://apps.who.int/vaccine_research/diseases/ari/en/index3.html

2 **Rosen JB**, Thomas AR, Lexau CA, Reingold A, Hadler JL, Harrison LH, Bennett NM, Schaffner W, Farley MM, Beall BW, Moore MR; CDC Emerging Infections Program Network. Geographic variation in invasive pneumococcal disease following pneumococcal conjugate vaccine introduction in the United States. *Clin Infect Dis* 2011; **53**: 137-143 [PMID: 21690620 DOI: 10.1093/cid/cir326]

3 **Laitinen S**, Vaara M, Salo E. Pneumococcal serotype distribution in invasive paediatric disease in Southern Finland before the introduction of vaccine. *Scand J Infect Dis* 2010; **42**: 924-930 [PMID: 20735331 DOI: 10.3109/00365548.2010.510532]

4 **Isaacman DJ**, McIntosh ED, Reinert RR. Burden of invasive pneumococcal disease and serotype distribution among Streptococcus pneumoniae isolates in young children in Europe: impact of the 7-valent pneumococcal conjugate vaccine and considerations for future conjugate vaccines. *Int J Infect Dis* 2010; **14**: e197-e209 [PMID: 19700359 DOI: 10.1016/j.ijid.2009.05.010]

5 **Posfay-Barbe KM**, Wald ER. Pneumococcal vaccines: do they prevent infection and how? *Curr Opin Infect Dis* 2004; **17**: 177-184 [PMID: 15166818 DOI: 10.1097/00001432-200406000-00002]

6 **Stein KE**. Thymus-independent and thymus-dependent responses to polysaccharide antigens. *J Infect Dis* 1992; **165** Suppl 1: S49-S52 [PMID: 1588177 DOI: 10.1093/infdis/165-Supplement_1-S49]

7 **Westerink MA**, Schroeder HW, Nahm MH. Immune Responses to pneumococcal vaccines in children and adults: Rationale for age-specific vaccination. *Aging Dis* 2012; **3**: 51-67 [PMID: 22500271]

8 **Black S**, Shinefield H, Fireman B, Lewis E, Ray P, Hansen JR, Elvin L, Ensor KM, Hackell J, Siber G, Malinoski F, Madore D, Chang I, Kohberger R, Watson W, Austrian R, Edwards K. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. Northern California Kaiser Permanente Vaccine Study Center Group. *Pediatr Infect Dis J* 2000; **19**: 187-195 [PMID: 10749457 DOI: 10.1097/00006454-200003000-00003]

9 **Hendrickson DJ**, Blumberg DA, Joad JP, Jhawar S, McDonald RJ. Five-fold increase in pediatric parapneumonic empyema since introduction of pneumococcal conjugate vaccine. *Pediatr Infect Dis J* 2008; **27**: 1030-1032 [PMID: 18845981 DOI: 10.1097/INF.0b013e31817e5188]

10 **Barricarte A**, Castilla J, Gil-Setas A, Torroba L, Navarro-Alonso JA, Irisarri F, Arriazu M. Effectiveness of the 7-valent pneumococcal conjugate vaccine: a population-based case-control study. *Clin Infect Dis* 2007; **44**: 1436-1441 [PMID: 17479939]

11 **Chibuk TK**, Robinson JL, Hartfield DS. Pediatric complicated pneumonia and pneumococcal serotype replacement: trends in hospitalized children pre and post introduction of routine vaccination with Pneumococcal Conjugate Vaccine (PCV7). *Eur J Pediatr* 2010; **169**: 1123-1128 [PMID: 20383524 DOI: 10.1007/s00431-010-1195-6]

12 **Hicks LA**, Harrison LH, Flannery B, Hadler JL, Schaffner W, Craig AS, Jackson D, Thomas A, Beall B, Lynfield R, Reingold A, Farley MM, Whitney CG. Incidence of pneumococcal disease due to non-pneumococcal conjugate vaccine (PCV7) serotypes in the United States during the era of widespread PCV7 vaccination, 1998-2004. *J Infect Dis* 2007; **196**: 1346-1354 [PMID: 17922399 DOI: 10.1086/521626]

13 **Lepoutre A**, Varon E, Georges S, Gutmann L, Lévy-Bruhl D. Impact of infant pneumococcal vaccination on invasive pneumococcal diseases in France, 2001-2006. *Euro Surveill* 2008; **13**: pii: 18962 [PMID: 18761883]

14 **Moore MR**, Gertz RE, Woodbury RL, Barkocy-Gallagher GA, Schaffner W, Lexau C, Gershman K, Reingold A, Farley M, Harrison LH, Hadler JL, Bennett NM, Thomas AR, McGee L, Pilishvili T, Brueggemann AB, Whitney CG, Jorgensen JH, Beall B. Population snapshot of emergent Streptococcus pneumoniae serotype 19A in the United States, 2005. *J Infect Dis* 2008; **197**: 1016-1027 [PMID: 18419539 DOI: 10.1086/528996]

15 **Pavia M**, Bianco A, Nobile CG, Marinelli P, Angelillo IF. Efficacy of pneumococcal vaccination in children younger than 24 months: a meta-analysis. *Pediatrics* 2009; **123**: e1103-e1110 [PMID: 19482744 DOI: 10.1542/peds.2008-3422]

16 **Weil-Olivier C**, van der Linden M, de Schutter I, Dagan R, Mantovani L. Prevention of pneumococcal diseases in the post-seven valent vaccine era: a European perspective. *BMC Infect Dis* 2012; **12**: 207 [PMID: 22954038 DOI: 10.1186/1471-2334-12-207]

17 **Miller E**, Andrews NJ, Waight PA, Slack MP, George RC. Effectiveness of the new serotypes in the 13-valent pneumococcal conjugate vaccine. *Vaccine* 2011; **29**: 9127-9131 [PMID: 21983361 DOI: 10.1016/j.vaccine.2011.09.112]

18 **Kaplan SL**, Barson WJ, Lin PL, Romero JR, Bradley JS, Tan TQ, Hoffman JA, Givner LB, Mason EO. Early trends for invasive pneumococcal infections in children after the introduction of the 13-valent pneumococcal conjugate vaccine. *Pediatr Infect Dis J* 2013; **32**: 203-207 [PMID: 23558320 DOI: 10.1097/INF.0b013e318275614b]

19 **Martinelli D**, Pedalino B, Cappelli MG, Caputi G, Sallustio A, Fortunato F, Tafuri S, Cozza V, Germinario C, Chironna M, Prato R. Towards the 13-valent pneumococcal conjugate universal vaccination: effectiveness in the transition era between PCV7 and PCV13 in Italy, 2010-2013. *Hum Vaccin Immunother* 2014; **10**: 33-39 [PMID: 24096297 DOI: 10.4161/hv.26650]

20 **Steens A**, Bergsaker MA, Aaberge IS, Rønning K, Vestrheim DF. Prompt effect of replacing the 7-valent pneumococcal conjugate vaccine with the 13-valent vaccine on the epidemiology of invasive pneumococcal disease in Norway. *Vaccine* 2013; **31**: 6232-6238 [PMID: 24176490 DOI: 10.1016/j.vaccine.2013.10.032]

21 **Harboe ZB**, Dalby T, Weinberger DM, Benfield T, Mølbak K, Slotved HC, Suppli CH, Konradsen HB, Valentiner-Branth P. Impact of 13-valent pneumococcal conjugate vaccination in invasive pneumococcal disease incidence and mortality. *Clin Infect Dis* 2014; **59**: 1066-1073 [PMID: 25034421 DOI: 10.1093/cid/ciu524]

22 **Chacon-Cruz E**, Rivas-Landeros RM, Volker-Soberanes ML. Early trends in invasive pneumococcal disease in children following the introduction of 13-valent pneumococcal conjugate vaccine: results from eight years of active surveillance in a Mexican hospital. *Ther Adv Vaccines* 2014; **2**: 155-158 [PMID: 25364508 DOI: 10.1177/2051013614547199]

23 **Andrews NJ**, Waight PA, Burbidge P, Pearce E, Roalfe L, Zancolli M, Slack M, Ladhani SN, Miller E, Goldblatt D. Serotype-specific effectiveness and correlates of protection for the 13-valent pneumococcal conjugate vaccine: a postlicensure indirect cohort study. *Lancet Infect Dis* 2014; **14**: 839-846 [PMID: 25042756 DOI: 10.1016/S1473-3099(14)70822-9]

24 **Moher D**, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097 [PMID: 19621072 DOI: 10.1371/journal.pmed.1000097]

25 **Albrich WC**, Baughman W, Schmotzer B, Farley MM. Changing characteristics of invasive pneumococcal disease in Metropolitan Atlanta, Georgia, after introduction of a 7-valent pneumococcal conjugate vaccine. *Clin Infect Dis* 2007; **44**: 1569-1576 [PMID: 17516400]

26 **Ampofo K**, Pavia AT, Chris S, Hersh AL, Bender JM, Blaschke AJ, Weng HY, Korgenski KE, Daly J, Mason EO, Byington CL. The changing epidemiology of invasive pneumococcal disease at a tertiary children's hospital through the 7-valent pneumococcal conjugate vaccine era: a case for continuous surveillance. *Pediatr Infect Dis J* 2012; **31**: 228-234 [PMID: 22330164 DOI: 10.1097/INF.0b013e31823dcc72]

27 **Aristegui J**, Bernaola E, Pocheville I, García C, Arranz L, Durán G, Pérez L, Bastida M, Canduela C, Herranz Aguirre M, Garrote E, Fletcher MA, Pérez C. Reduction in pediatric invasive pneumococcal disease in the Basque Country and Navarre, Spain, after introduction of the heptavalent pneumococcal conjugate vaccine. *Eur J Clin Microbiol Infect Dis* 2007; **26**: 303-310 [PMID: 17457623 DOI: 10.1007/s10096-007-0294-4]

28 **Barricarte A**, Gil-Setas A, Torroba L, Castilla J, Petit A, Polo I, Arriazu M, Irisarri F, García Cenoz M. [Invasive pneumococcal disease in children younger than 5 years in Navarra, Spain (2000-2005). Impact of the conjugate vaccine]. *Med Clin* (Barc) 2007; **129**: 41-45 [PMID: 17588359 DOI: 10.1157/13106935]

29 **Benito-Fernández J**, Raso SM, Pocheville-Gurutzeta I, SánchezEtxaniz J, Azcunaga-Santibañez B, Capapé-Zache S. Pneumococcal bacteremia among infants with fever without known source before and after introduction of pneumococcal conjugate vaccine in the Basque Country of Spain. *Pediatr Infect Dis J* 2007; **26**: 667-671 [PMID: 17848875 DOI: 10.1097/INF.0b013e3180f610b3]

30 **Ben-Shimol S**, Greenberg D, Givon-Lavi N, Elias N, Glikman D, Rubinstein U, Dagan R; Israeli Bacteremia and Meningitis Active Surveillance Group. Rapid reduction in invasive pneumococcal disease after introduction of PCV7 into the National Immunization Plan in Israel. *Vaccine* 2012; **30**: 6600-6607 [PMID: 22939907 DOI: 10.1016/j.vaccine.2012.08.012]

31 **Bjornson G**, Scheifele DW, Bettinger J, Patrick DM, Gustafson L, Daly P, Tyrrell GJ. Effectiveness of pneumococcal conjugate vaccine in Greater Vancouver, Canada: 2004-2005. *Pediatr Infect Dis J* 2007; **26**: 540-542 [PMID: 17529875 DOI: 10.1097/INF.0b013e31803c56df]

32 **Calbo E**, Díaz A, Cañadell E, Fábrega J, Uriz S, Xercavins M, Morera MA, Cuchi E, Rodríguez-Carballeira M, Garau J; Spanish Pneumococcal Infection Study Network. Invasive pneumococcal disease among children in a health district of Barcelona: early impact of pneumococcal conjugate vaccine. *Clin Microbiol Infect* 2006; **12**: 867-872 [PMID: 16882291 DOI: 10.1111/j.1469-0691.2006.1502_1.x]

33 **Carstairs KL**, Tanen DA, Johnson AS, Kailes SB, Riffenburgh RH. Pneumococcal bacteremia in febrile infants presenting to the emergency department before and after the introduction of the heptavalent pneumococcal vaccine. *Ann Emerg Med* 2007; **49**: 772-777 [PMID: 17337092 DOI: 10.1016/j.annemergmed.2006.10.026]

34 **Casado-Flores J**, Rodrigo C, Arístegui J, Martínón JM, Fenoll A, Mendez C. Decline in pneumococcal meningitis in Spain after introduction of the heptavalent pneumococcal conjugate vaccine. *Pediatr Infect Dis J* 2008; **27**: 1020-1022 [PMID: 18845983 DOI: 10.1097/INF.0b013e31817bd2dc]

35 **Centers for Disease Control and Prevention (CDC).** Invasive

pneumococcal disease in children 5 years after conjugate vaccine introduction--eight states, 1998-2005. *MMWR Morb Mortal Wkly Rep* 2008; **57**: 144-148 [PMID: 18272956]

36 **De Serres G**, Pilishvili T, Link-Gelles R, Reingold A, Gershman K, Petit S, Farley MM, Harrison LH, Lynfield R, Bennett NM, Baumbach J, Thomas A, Schaffner W, Beall B, Whitney C, Moore M. Use of surveillance data to estimate the effectiveness of the 7-valent conjugate pneumococcal vaccine in children less than 5 years of age over a 9 year period. *Vaccine* 2012; **30**: 4067-4072 [PMID: 22525797 DOI: 10.1016/j.vaccine.2012.04.017]

37 **De Wals P**, Lefebvre B, Defay F, Deceuninck G, Boulianne N. Invasive pneumococcal diseases in birth cohorts vaccinated with PCV-7 and/or PHiD-CV in the province of Quebec, Canada. *Vaccine* 2012; **30**: 6416-6420 [PMID: 22921290 DOI: 10.1016/j.vaccine.2012.08.017]

38 **Dias R**, Caniça M. Invasive pneumococcal disease in Portugal prior to and after the introduction of pneumococcal heptavalent conjugate vaccine. *FEMS Immunol Med Microbiol* 2007; **51**: 35-42 [PMID: 17854472]

39 **Vestrheim DF**, Høiby EA, Bergsaker MR, Rønning K, Aaberge IS, Caugant DA. Indirect effect of conjugate pneumococcal vaccination in a 2+1 dose schedule. *Vaccine* 2010; **28**: 2214-2221 [PMID: 20056192 DOI: 10.1016/j.vaccine.2009.12.054]

40 **Dubos F**, Marechal I, Husson MO, Courouble C, Aurel M, Martinot A. Decline in pneumococcal meningitis after the introduction of the heptavalent-pneumococcal conjugate vaccine in northern France. *Arch Dis Child* 2007; **92**: 1009-1012 [PMID: 17626145]

41 **Fenoll A**, Granizo JJ, Aguilar L, Giménez MJ, Aragoneses-Fenoll L, Hanquet G, Casal J, Tarragó D. Temporal trends of invasive Streptococcus pneumoniae serotypes and antimicrobial resistance patterns in Spain from 1979 to 2007. *J Clin Microbiol* 2009; **47**: 1012-1020 [PMID: 19225097 DOI: 10.1128/JCM.01454-08]

42 **Flannery B**, Schrag S, Bennett NM, Lynfield R, Harrison LH, Reingold A, Cieslak PR, Hadler J, Farley MM, Facklam RR, Zell ER, Whitney CG. Impact of childhood vaccination on racial disparities in invasive Streptococcus pneumoniae infections. *JAMA* 2004; **291**: 2197-2203 [PMID: 15138241 DOI: 10.1001/jama.291.18.2197]

43 **Giele C**, Moore H, Bayley K, Harrison C, Murphy D, Rooney K, Keil AD, Lehmann D. Has the seven-valent pneumococcal conjugate vaccine had an impact on invasive pneumococcal disease in Western Australia? *Vaccine* 2007; **25**: 2379-2384 [PMID: 17064825 DOI: 10.1016/j.vaccine.2006.09.004]

44 **Schutze GE**, Tucker NC, Mason EO. Impact of the conjugate pneumococcal vaccine in arkansas. *Pediatr Infect Dis J* 2004; **23**: 1125-1129 [PMID: 15626950]

45 **Guevara M**, Barricarte A, Gil-Setas A, García-Irure JJ, Beristain X, Torroba L, Petit A, Polo Vigas ME, Aguinaga A, Castilla J. Changing epidemiology of invasive pneumococcal disease following increased coverage with the heptavalent conjugate vaccine in Navarre, Spain. *Clin Microbiol Infect* 2009; **15**: 1013-1019 [PMID: 19673968 DOI: 10.1111/j.1469-0691.2009.02904.x]

46 **Haddy RI**, Perry K, Chacko CE, Helton WB, Bowling MG, Looney SW, Buck GE. Comparison of incidence of invasive Streptococcus pneumoniae disease among children before and after introduction of conjugated pneumococcal vaccine. *Pediatr Infect Dis J* 2005; **24**: 320-323 [PMID: 15818291]

47 **Hanna JN**, Humphreys JL, Murphy DM. Invasive pneumococcal disease in Indigenous people in north Queensland: an update, 2005-2007. *Med J Aust* 2008; **189**: 43-46 [PMID: 18601643]

48 **Hanquet G**, Lernout T, Vergison A, Verhaegen J, Kissling E, Tuerlinckx D, Malfroot A, Swennen B, Sabbe M; Belgian IPD Scientific Committee. Impact of conjugate 7-valent vaccination in Belgium: addressing methodological challenges. *Vaccine* 2011; **29**: 2856-2864 [PMID: 21342667 DOI: 10.1016/j.vaccine.2011.02.016]

49 **Harboe ZB**, Valentiner-Branth P, Benfield TL, Christensen JJ, Andersen PH, Howitz M, Krogfelt KA, Lambertsen L, Konradsen HB. Early effectiveness of heptavalent conjugate pneumococcal vaccination on invasive pneumococcal disease after the introduction in the Danish Childhood Immunization Programme. *Vaccine* 2010; **28**: 2642-2647 [PMID: 20096392 DOI: 10.1016/j.vaccine.2010.01.017]

50 **Hennessy TW**, Singleton RJ, Bulkow LR, Bruden DL, Hurlburt DA, Parks D, Moore M, Parkinson AJ, Schuchat A, Butler JC. Impact of heptavalent pneumococcal conjugate vaccine on invasive disease, antimicrobial resistance and colonization in Alaska Natives: progress towards elimination of a health disparity. *Vaccine* 2005; **23**: 5464-5473 [PMID: 16188350]

51 **Centers for Disease Control and Prevention (CDC).** Direct and indirect effects of routine vaccination of children with 7-valent pneumococcal conjugate vaccine on incidence of invasive pneumococcal disease--United States, 1998-2003. *MMWR Morb Mortal Wkly Rep* 2005; **54**: 893-897 [PMID: 16163262]

52 **Hsu HE**, Shutt KA, Moore MR, Beall BW, Bennett NM, Craig AS, Farley MM, Jorgensen JH, Lexau CA, Petit S, Reingold A, Schaffner W, Thomas A, Whitney CG, Harrison LH. Effect of pneumococcal conjugate vaccine on pneumococcal meningitis. *N Engl J Med* 2009; **360**: 244-256 [PMID: 19144940 DOI: 10.1056/NEJMoa0800836]

53 **Hsu K**, Pelton S, Karumuri S, Heisey-Grove D, Klein J. Population-based surveillance for childhood invasive pneumococcal disease in the era of conjugate vaccine. *Pediatr Infect Dis J* 2005; **24**: 17-23 [PMID: 15665705]

54 **Hsu KK**, Shea KM, Stevenson AE, Pelton SI; Massachusetts Department of Public Health. Changing serotypes causing childhood invasive pneumococcal disease: Massachusetts, 2001-2007. *Pediatr Infect Dis J* 2010; **29**: 289-293 [PMID: 19935447 DOI: 10.1097/INF.0b013e3181c15471]

55 **Ingels H**, Rasmussen J, Andersen PH, Harboe ZB, Glismann S, Konradsen H, Hoffmann S, Valentiner-Branth P, Lambertsen L; Danish Pneumococcal Surveillance Collaboration Group 2009-2010. Impact of pneumococcal vaccination in Denmark during the first 3 years after PCV introduction in the childhood immunization programme. *Vaccine* 2012; **30**: 3944-3950 [PMID: 22504662 DOI: 10.1016/j.vaccine.2012.03.060]

56 **Wenger JD**, Zulz T, Bruden D, Singleton R, Bruce MG, Bulkow L, Parks D, Rudolph K, Hurlburt D, Ritter T, Klejka J, Hennessy T. Invasive pneumococcal disease in Alaskan children: impact of the seven-valent pneumococcal conjugate vaccine and the role of water supply. *Pediatr Infect Dis J* 2010; **29**: 251-256 [PMID: 19952861 DOI: 10.1097/INF.0b013e3181bdbed5]

57 **Johnson DR**, D'Onise K, Holland RA, Raupach JC, Koehler AP. Pneumococcal disease in South Australia: vaccine success but no time for complacency. *Vaccine* 2012; **30**: 2206-2211 [PMID: 22273663 DOI: 10.1016/j.vaccine.2011.12.119]

58 **Kellner JD**, Vanderkooi OG, MacDonald J, Church DL, Tyrrell GJ, Scheifele DW. Changing epidemiology of invasive pneumococcal disease in Canada, 1998-2007: update from the Calgary-area Streptococcus pneumoniae research (CASPER) study. *Clin Infect Dis* 2009; **49**: 205-212 [PMID: 19508165 DOI: 10.1086/599827]

59 **Kyaw MH**, Lynfield R, Schaffner W, Craig AS, Hadler J, Reingold A, Thomas AR, Harrison LH, Bennett NM, Farley MM, Facklam RR, Jorgensen JH, Besser J, Zell ER, Schuchat A, Whitney CG. Effect of introduction of the pneumococcal conjugate vaccine on drug-resistant Streptococcus pneumoniae. *N Engl J Med* 2006; **354**: 1455-1463 [PMID: 16598044]

60 **Leal J**, Vanderkooi OG, Church DL, Macdonald J, Tyrrell GJ, Kellner JD. Eradication of invasive pneumococcal disease due to the seven-valent pneumococcal conjugate vaccine serotypes in Calgary, Alberta. *Pediatr Infect Dis J* 2012; **31**: e169-e175 [PMID: 22673137]

61 **Liao WH**, Lin SH, Lai CC, Tan CK, Liao CH, Huang YT, Wang CY, Hsueh PR. Impact of pneumococcal vaccines on invasive pneumococcal disease in Taiwan. *Eur J Clin Microbiol Infect Dis* 2010; **29**: 489-492 [PMID: 20108017 DOI: 10.1007/s10096-010-0873-7]

62 **Messina AF**, Katz-Gaynor K, Barton T, Ahmad N, Ghaffar F, Rasko D, McCracken GH. Impact of the pneumococcal conjugate vaccine on serotype distribution and antimicrobial resistance of

invasive Streptococcus pneumoniae isolates in Dallas, TX, children from 1999 through 2005. *Pediatr Infect Dis J* 2007; **26**: 461-467 [PMID: 17529859]

63 **Muñoz-Almagro C**, Jordan I, Gene A, Latorre C, Garcia-Garcia JJ, Pallares R. Emergence of invasive pneumococcal disease caused by nonvaccine serotypes in the era of 7-valent conjugate vaccine. *Clin Infect Dis* 2008; **46**: 174-182 [PMID: 18171247 DOI: 10.1086/524660]

64 **Patrzalek M**, Gorynski P, Albrecht P. Indirect population impact of universal PCV7 vaccination of children in a 2 + 1 schedule on the incidence of pneumonia morbidity in Kielce, Poland. *Eur J Clin Microbiol Infect Dis* 2012; **31**: 3023-3028 [PMID: 22895889 DOI: 10.1007/s10096-012-1656-0]

65 **Pérez A**, Giménez M, Sala P, Sierra M, Esteve A, Rodrigo C. Increase in invasive nonvaccine pneumococcal serotypes at two hospitals in Barcelona: was replacement disease to blame? *Acta Paediatr* 2011; **100**: 1572-1575 [PMID: 21623903 DOI: 10.1111/j.1651-2227.2011.02365.x]

66 **Pérez-Trallero E**, Marimon JM, Ercibengoa M, Vicente D, Pérez-Yarza EG. Invasive Streptococcus pneumoniae infections in children and older adults in the north of Spain before and after the introduction of the heptavalent pneumococcal conjugate vaccine. *Eur J Clin Microbiol Infect Dis* 2009; **28**: 731-738 [PMID: 19153783 DOI: 10.1007/s10096-008-0693-1]

67 **Pilishvili T**, Lexau C, Farley MM, Hadler J, Harrison LH, Bennett NM, Reingold A, Thomas A, Schaffner W, Craig AS, Smith PJ, Beall BW, Whitney CG, Moore MR. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. *J Infect Dis* 2010; **201**: 32-41 [PMID: 19947881 DOI: 10.1086/648593]

68 **Poehling KA**, Talbot TR, Griffin MR, Craig AS, Whitney CG, Zell E, Lexau CA, Thomas AR, Harrison LH, Reingold AL, Hadler JL, Farley MM, Anderson BJ, Schaffner W. Invasive pneumococcal disease among infants before and after introduction of pneumococcal conjugate vaccine. *JAMA* 2006; **295**: 1668-1674 [PMID: 16609088 DOI: 10.1001/jama.295.14.1668]

69 Incidence of invasive pneumococcal disease after introduction of the Universal Infant Immunization Program, British Columbia (2002-2005). *Can Commun Dis Rep* 2006; **32**: 157-161 [PMID: 16869067]

70 **Ramani RR**, Hall WN, Boulton M, Johnson DR, Zhu BP. Impact of PCV7 on invasive pneumococcal disease among children younger than 5 years: a population-based study. *Am J Public Health* 2004; **94**: 958-959 [PMID: 15249298 DOI: 10.2105/AJPH.94.6.958]

71 **Rendi-Wagner P**, Paulke-Korinek M, Kundi M, Burgmann H, Georgopoulos A, Vécsei A, Kollaritsch H. National paediatric immunization program of high risk groups: no effect on the incidence of invasive pneumococcal diseases. *Vaccine* 2009; **27**: 3963-3968 [PMID: 19393711 DOI: 10.1016/j.vaccine.2009.04.044]

72 **Rodenburg GD**, de Greeff SC, Jansen AG, de Melker HE, Schouls LM, Hak E, Spanjaard L, Sanders EA, van der Ende A. Effects of pneumococcal conjugate vaccine 2 years after its introduction, the Netherlands. *Emerg Infect Dis* 2010; **16**: 816-823 [PMID: 20409372 DOI: 10.3201/eid1605.091223]

73 **Rückinger S**, van der Linden M, Reinert RR, von Kries R, Burckhardt F, Siedler A. Reduction in the incidence of invasive pneumococcal disease after general vaccination with 7-valent pneumococcal conjugate vaccine in Germany. *Vaccine* 2009; **27**: 4136-4141 [PMID: 19406190 DOI: 10.1016/j.vaccine.2009.04.057]

74 **de Sevilla MF**, García-García JJ, Esteva C, Moraga F, Hernández S, Selva L, Coll F, Ciruela P, Planes AM, Codina G, Salleras L, Jordan I, Domínguez A, Muñoz-Almagro C. Clinical presentation of invasive pneumococcal disease in Spain in the era of heptavalent conjugate vaccine. *Pediatr Infect Dis J* 2012; **31**: 124-128 [PMID: 22173137 DOI: 10.1097/INF.0b013e318241d09e]

75 **Shafinoori S**, Ginocchio CC, Greenberg AJ, Yeoman E, Cheddie M, Rubin LG. Impact of pneumococcal conjugate vaccine and the severity of winter influenza-like illnesses on invasive pneumococcal infections in children and adults. *Pediatr Infect Dis J* 2005; **24**: 10-16 [PMID: 15665704]

76 **Shah SS**, Ratner AJ. Trends in invasive pneumococcal disease-associated hospitalizations. *Clin Infect Dis* 2006; **42**: e1-e5 [PMID: 16323082]

77 **Techasaensiri C**, Messina AF, Katz K, Ahmad N, Huang R, McCracken GH. Epidemiology and evolution of invasive pneumococcal disease caused by multidrug resistant serotypes of 19A in the 8 years after implementation of pneumococcal conjugate vaccine immunization in Dallas, Texas. *Pediatr Infect Dis J* 2010; **29**: 294-300 [PMID: 19949357 DOI: 10.1097/INF.0b013e3181c2a229]

78 **Tsai CJ**, Griffin MR, Nuorti JP, Grijalva CG. Changing epidemiology of pneumococcal meningitis after the introduction of pneumococcal conjugate vaccine in the United States. *Clin Infect Dis* 2008; **46**: 1664-1672 [PMID: 18433334 DOI: 10.1086/587897]

79 **Tsigrelis C**, Tleyjeh IM, Huskins WC, Lahr BD, Nyre LM, Virk A, Baddour LM. Incidence of invasive pneumococcal disease among children after introduction of a 7-valent pneumococcal conjugate vaccine: a population-based study in Olmsted County, Minnesota. *Mayo Clin Proc* 2009; **84**: 871-875 [PMID: 19797776 DOI: 10.1016/S0025-6196(11)60504-1]

80 **Tyrrell GJ**, Lovgren M, Chui N, Minion J, Garg S, Kellner JD, Marrie TJ. Serotypes and antimicrobial susceptibilities of invasive Streptococcus pneumoniae pre- and post-seven valent pneumococcal conjugate vaccine introduction in Alberta, Canada, 2000-2006. *Vaccine* 2009; **27**: 3553-3560 [PMID: 19464534 DOI: 10.1016/j.vaccine.2009.03.063]

81 **van der Linden M**, Weiß S, Falkenhorst G, Siedler A, Imöhl M, von Kries R. Four years of universal pneumococcal conjugate infant vaccination in Germany: impact on incidence of invasive pneumococcal disease and serotype distribution in children. *Vaccine* 2012; **30**: 5880-5885 [PMID: 22771186 DOI: 10.1016/j.vaccine.2012.06.068]

82 **Vestrheim DF**, Løvoll O, Aaberge IS, Caugant DA, Høiby EA, Bakke H, Bergsaker MR. Effectiveness of a 2+1 dose schedule pneumococcal conjugate vaccination programme on invasive pneumococcal disease among children in Norway. *Vaccine* 2008; **26**: 3277-3281 [PMID: 18456376 DOI: 10.1016/j.vaccine.2008.03.087]

83 **Weatherholtz R**, Millar EV, Moulton LH, Reid R, Rudolph K, Santosham M, O'Brien KL. Invasive pneumococcal disease a decade after pneumococcal conjugate vaccine use in an American Indian population at high risk for disease. *Clin Infect Dis* 2010; **50**: 1238-1246 [PMID: 20367225 DOI: 10.1086/651680]

84 **Whitney CG**, Farley MM, Hadler J, Harrison LH, Bennett NM, Lynfield R, Reingold A, Cieslak PR, Pilishvili T, Jackson D, Facklam RR, Jorgensen JH, Schuchat A. Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *N Engl J Med* 2003; **348**: 1737-1746 [PMID: 12724479 DOI: 10.1056/NEJMoa022823]

85 **Winters M**, Patrick DM, Marra F, Buxton J, Chong M, Isaac-Renton JL, Shaw C, Tyrrell GJ, Lovgren M, Paulus S. Epidemiology of invasive pneumococcal disease in BC during the introduction of conjugated pneumococcal vaccine. *Can J Public Health* 2008; **99**: 57-61 [PMID: 18435393]

86 **Yildirim I**, Stevenson A, Hsu KK, Pelton SI. Evolving picture of invasive pneumococcal disease in massachusetts children: a comparison of disease in 2007-2009 with earlier periods. *Pediatr Infect Dis J* 2012; **31**: 1016-1021 [PMID: 22673142 DOI: 10.1097/INF.0b013e3182615615]

87 **Godwin M**, Ruhland L, Casson I, MacDonald S, Delva D, Birtwhistle R, Lam M, Seguin R. Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. *BMC Med Res Methodol* 2003; **3**: 28 [PMID: 14690550]

88 **Weinberg GA**, Szilagyi PG. Vaccine epidemiology: efficacy, effectiveness, and the translational research roadmap. *J Infect Dis* 2010; **201**: 1607-1610 [PMID: 20402594 DOI: 10.1086/652404]

89 **Lee KK**, Rinaldi F, Chan MK, Chan ST, So TM, Hon EK, Lee VW. Economic evaluation of universal infant vaccination with 7vPCV in Hong Kong. *Value Health* 2009; **12** Suppl 3: S42-S48 [PMID: 20586981 DOI: 10.1111/j.1524-4733.2009.00626.x]

90 **Marchetti M**, Colombo GL. Cost-effectiveness of universal pneumococcal vaccination for infants in Italy. *Vaccine* 2005; **23**: 4565-4576 [PMID: 15992969]

91 **Feikin DR**, Kagucia EW, Loo JD, Link-Gelles R, Puhan MA, Cherian T, Levine OS, Whitney CG, O'Brien KL, Moore MR; Serotype Replacement Study Group. Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. *PLoS Med* 2013; **10**: e1001517 [PMID: 24086113 DOI: 10.1371/journal.pmed.1001517]

92 **Hanquet G**, Kissling E, Fenoll A, George R, Lepoutre A, Lernout T, Tarragó D, Varon E, Verhaegen J. Pneumococcal serotypes in children in 4 European countries. *Emerg Infect Dis* 2010; **16**: 1428-1439 [PMID: 20735928 DOI: 10.3201/eid1609.100102]

P- Reviewer: Sadoghi P, Trohman RG **S- Editor**: Gong XM
L- Editor: A **E- Editor**: Jiao XK

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i3.163

World J Meta-Anal 2015 June 26; 3(3): 163-180
ISSN 2308-3840 (online)

META-ANALYSIS

Is the traditional Chinese medicine helpful for patients with hematologic malignant diseases? A meta-analysis of randomized controlled trials

Cheng-Liang Qian, Fei Yan, Yan-Zhi Song, Dong Li, Ke-Zhou Dong, Yi-Min Zhu

Cheng-Liang Qian, Department of Traditional Chinese Medicine, Nanjing BenQ Medical Center, Nanjing Medical University, Nanjing 021000, Jiangsu Province, China
Fei Yan, Department of Medical Oncology, Jiangsu Cancer Hospital, Nanjing 021000, Jiangsu Province, China
Yan-Zhi Song, Dong Li, Department of Hematology, Nanjing BenQ Medical Center, Nanjing Medical University, Nanjing 021000, Jiangsu Province, China
Ke-Zhou Dong, Yi-Min Zhu, Department of Respiration, the 2st Jiangsu Province Hospital of TCM, Nanjing University of Chinese Medicine, Nanjing 021000, Jiangsu Province, China

Author contributions: Song YZ conceived and designed the study, searched and selected trials for inclusion, assessed methodological quality of included trials, extracted data, performed the statistical analysis and wrote the review; Qian CL searched trials, selected trials for inclusion, assessed methodological quality of included trials and extracted data; Yan F searched and selected trials for inclusion and wrote the review; Li D, Dong KZ and Zhu YM wrote and revised the review.

Supported by The Six Peak Talent Program of Jiangsu Province, No. 2009-47-D.

Conflict-of-interest: There is no conflict of interest reported.
Data sharing: No.

Correspondence to: Dr. Yan-Zhi Song, Department of Hematology, Nanjing BenQ Medical Center, Nanjing Medical University, 71# Hexi Street, Jianye District, Nanjing 021000, Jiangsu Province, China. yandgics@126.com
Telephone: +86-25-52238800
Fax: +86-25-52238800

Received: October 19, 2014
Peer-review started: October 21, 2014
First decision: December 26, 2014
Revised: April 13, 2015
Accepted: May 5, 2015
Article in press: May 6, 2015
Published online: June 26, 2015

Abstract

AIM: To evaluate the efficacy of traditional Chinese medicine (TCM) for the treatment of hematologic malignant diseases.

METHODS: We searched the Cochrane CENTRAL, PubMed, Embase, Web of Science, AMED, CNKI, Wanfang Platform; China Sinomed and the clinical trial registry web sites and Googlescholar electronically up to June 19th, 2014 and hand searched related publications. Only randomized controlled trials (RCTs) researching on whether TCM as the adjuvant treatment improved the effect for hematologic malignant diseases were included. Two reviewers extracted data and evaluated the studies independently. Pooled risk ratios (RR) were calculated as outcome measures. Our primary outcomes were the overall response (OR) rate.

RESULTS: We retrieved 13143 references and included 11 RCTs involved 891 participants after screening. Because the non-significant heterogeneity we used the fixed effect model to combine data and TCM had a significantly higher OR and CR (complete response) rates than the control [RR = 1.17, 95%CI: (1.10, 1.25), $P < 0.00001$; RR = 1.24, 95%CI: (1.11, 1.37), $P < 0.0001$, respectively]. Only three studies included in the survival rate analysis. We combined them with random effects model and there was no significant difference between the TCM and control arms. Because

of the low heterogeneity we used the fixed effect model to combine the non-hematologic adverse effects (AEs) data. Our results showed that TCM significantly decreased non-hematologic AEs rates we researched, the gastrointestinal reaction [RR = 0.50, 95%CI: (0.37, 0.68), $P < 0.0001$], liver and/or kidney injury [RR = 0.37, 95%CI: (0.26, 0.53), $P < 0.00001$] and heart injury [RR = 0.24, 95%CI: (0.09, 0.68), $P = 0.007$]. Additionally, TCM had a trend to decrease the infection rate [RR = 0.16, (0.02, 1.12), $P = 0.07$], but not statistically significantly.

CONCLUSION: TCM increases OR and CR rates for hematologic malignances and reduces treatment associated serious non-hematologic AEs. Therefore, TCM should be included in the treatment of hematologic malignances.

Key words: Hematologic malignant disease; Leukemia; Lymphoma; Chinese medicine

Core tip: We pooled all the studies complied to our inclusion criteria that were retrieved by extensively searching the related databases, journals and websites. Our result suggested that adding traditional Chinese medicine (TCM) increased overall response and complete response rates for malignant hematologic diseases treatment. Although it was based on the evidence of low level of GRADE quality, our result demonstrated that TCM reduced treatment associated serious non-hematologic adverse effects (AEs). Furthermore, considering the rare AEs and drugs interactions, TCM should be included in the hematologic malignances treatment, at least for adult acute leukemia.

Qian CL, Yan F, Song YZ, Li D, Dong KZ, Zhu YM. Is the traditional Chinese medicine helpful for patients with hematologic malignant diseases? A meta-analysis of randomized controlled trials. *World J Meta-Anal* 2015; 3(3): 163-180 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i3/163.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i3.163

INTRODUCTION

The incidence and mortality of malignant tumors have increased greatly in recent years[1]. Albeit the treatment methods of malignant diseases progress quickly the general prognosis of this kind of diseases is poor[1]. Whereby the hematologic malignancies have a particular high-grade malignancy and are systemic diseases that are common to involve multiple systems and organs. Hence, the systemic chemotherapy with western medicine becomes the standard treatment of these kind of diseases[2]. However, the same as other malignant diseases, even in nowadays, the response and survival rates are still not ideal[3,4]. As well as the chemotherapy always causes serious adverse effects (AEs), such as Ⅲ-Ⅳ grade bone marrow suppression, serious nausea and vomiting, hepatic and renal dysfunction and heart injury etc. Attempts to improve therapy by intensifying the number of chemotherapeutic agents or their doses lead only to increase side effects[5]. Even the targeted molecular therapy developed in recent years also causes obvious side effects. For example, the rituximab increases the response rate and survival time for B cell lymphoma[6,7], alternatively, it will obviously suppress the bodies' normal immune response to pathogens for a long period of more than one year. As a result of it, patients who received it are sensitive to infection and sometimes it is fatal[8,9]. And furthermore, in most conditions, these new medicines need to be administered with chemotherapy together not to mention the tumor cells will become resistant to the therapy after treated for a period[10].

On the other hand, many studies reported that adding traditional Chinese medicine (TCM) into the malignant diseases treatment strategy not only increased the response rate but also significantly lowered the treatment associated AEs rate[11-14]. There are a variety of herbs being used in different combinations and forms, such as oral administration and intravenous injection for hematologic malignancies yet. Many randomized controlled studies have shown that TCM as the adjuvant agent improved the malignant hematologic diseases response and reduced the AEs associated with chemotherapy[15]. But most of the published studies were small sample sized and the results were not consensus. So we wrote the meta-analysis to evaluate the efficacy of TCM for the treatment of hematologic malignant diseases.

MATERIALS AND METHODS

This meta-analysis followed the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) Statement issued in 2009 (Table 1).

Inclusion criteria

We only included randomized controlled trials (RCTs) that researched on whether TCM as the adjuvant treatment improved the effect for malignant hematologic diseases. There was no age, sex, race, complicated diseases or language limits of the study. Our primary outcomes were the overall response (OR) rate calculated by summating the complete response (CR), partial response and stable disease rates. The survival and serious AEs rates and the change of quality of life were our secondary outcomes. The diagnosis must be confirmed by pathological sections or bone marrow smears.

Since some TCMs for acute promyelocytic leukemia treatment, such as the compound Huang Dai Tablets, have been administered as the primary maintenance treatment, not the adjuvant treatment and their active ingredients has been recognized as Tetraarsenic tetra-

Table 1 The Preferred Reporting Items for Systematic Review and Meta-Analysis checklist

Section/topic	*n*	Checklist item	Reported on page
Title			
Title	1	Identify the report as a systematic review, meta-analysis, or both	1
Abstract			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number	2
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to PICOS	3-4
Methods			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (*e.g.*, Web address), and, if available, provide registration information including registration number	
Eligibility criteria	6	Specify study characteristics (*e.g.*, PICOS, length of follow-up) and report characteristics (*e.g.*, years considered, language, publication status) used as criteria for eligibility, giving rationale	5
Information sources	7	Describe all information sources (*e.g.*, databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	5-6, Table 2
Study selection	9	State the process for selecting studies (*i.e.*, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	6-7
Data collection process	10	Describe method of data extraction from reports (*e.g.*, piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	7
Data items	11	List and define all variables for which data were sought (*e.g.*, PICOS, funding sources) and any assumptions and simplifications made	7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis	6-7
Summary measures	13	State the principal summary measures (*e.g.*, risk ratio, difference in means)	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (*e.g.*, I^2) for each meta-analysis	7

PICOS: Participants, interventions, comparisons, outcomes, and study design.

sulfide we did not include these studies. The efficacy and safety of this kind of TCM is the focus of our next study.

Searching method

YS and CQ searched the following databases independently, the Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library, PubMed, Embase, Web of Science, Allied and Alternative Medicine (AMED), Googlescholar, China National Knowledge Infrastructure (CNKI), Wanfang Data Knowledge Service Platform; China biomedical literature service system (Sinomed); and the well-known clinical trial registry sites (http://www.clinicaltrial.gov/; http://apps.who.int/trialsearch/). The electronic search was up to June 19th, 2014. The detailed searching strategy for PubMed was recorded in Table 2.

We specified three searching themes: First, we searched TCM related words, we used the terms "complementary medicine", and the free words "tradition or tradition* or china or chinese or herb or herbal or complement* or tcm or 'zhong yi' or chm or ethno* or folk or home or indigenous or primitive or materia* or nosod* or east or eastern or orient or oriental or Asian or Korea* or Tibet* or herbaceous or plant or plants or botan* or kampo or mongol* or phytogenic or phytotherapy or alternative"; Second, we searched hematologic diseases related words, we used the terms "leukemia" or "lymphoma" or "multiple myeloma", and the free words "hemotolog* or anemia or thrombocytopen* or pancytope* or 'bone marrow'or transplant or 'stem cell' or 'leukemia or lymphoma' or cancer or dysplas* or malignant or hyperplas* or hypoplas* or myelom* or Hodgkin or non-Hodgkin or blast or blasts or 'progression free survival (PFS)' or 'disease free survival (DFS)' or 'overall survival (OS)' or OS or PFS or DFS or chemotherapy or (chemical treatment) or radiotherapy or irradia* or oncolog* or monoclon*"; and third, we used the Cochrane highly sensitive search filters to retrieve randomized trials in Medline and Embase[16].

We also hand searched other journals that might publish relative clinical trials, PubMed related articles, reference lists of retrieved articles. Considering there might be some ongoing studies which did not register in the clinical trial registry sites and some finished studies which did not published, we contacted some researchers, relative manufacturers and specialists for further information of unpublished trials. Our study did not set limits of ages, sexes, races, published languages and

Table 2 The PubMed searching strategy

(1) "Complementary therapies" (Mesh)
(2) Tradition or tradition* OR china or chinese OR herb or herbal OR complement* or tcm or "zhong yi" or chm or ethno* or folk or home or indigenous or primitive or materia* or nosod* or east or eastern or orient or oriental or asian or Korea* or Tibet* or herbaceous or plant or plants or botan* or kampo or mongol* or phytogenic or phytotherapy or alternative
(3) Medicine or medicinal or medical or remed* or therapy or therapies or therapeutic or therapeutics or therapist or treat or treatment or drug or drugs
(4) (2) and (3)
(5) (1) or (4)
(6) Leukemia or lymphoma or "multiple myeloma" (mesh)
(7) Hemotolog* or anemia or thrombocytopen* or pancytope* or "bone marrow" or transplant or "stem cell"
(8) Leukemia OR lymphoma OR cancer OR dysplas* OR malignant OR hyperplas* OR hypoplas* or myelom* or Hodgkin or non-hodgkin or blast or blasts or "progression free survival" or "disease free survival" or "overall survival" or OS or PFS or DFS or chemotherapy or (chemical treatment) or radiotherapy or irradia* or oncolog* or monoclon*
(9) (7) and (8)
(10) (6) or (9)
(11) (((((Randomized controlled trial [Publication type]) OR controlled clinical trial [Publication type]) OR (randomized or placebo[Title/Abstract])) OR drug therapy [MeSH Subheading]) OR (randomly or groups or trial [Title/Abstract])) OR rct
(12) Animals [mh] NOT humans [mh]
(13) (11) not (12)
(14) (5) and (10) and (13)
(15) (Cancer or carcinoma or sarcoma)[ti]
(16) Carcinoma[mesh] or sarcoma[mesh]
(17) (14) not (15) or (16)

Table 3 Characteristics of Dian Rong 2009 study

Methods	A randomized double blind placebo controlled Ⅰ multicenter study
Participants	Refractory acute leukemia patients
Interventions	TCM group: Combine Chinese interventions with standard chemotherapy of western medicine Control group: Standard chemotherapy with western medicine
Outcomes	The primary outcome: the response rate

TCM: Traditional Chinese medicine.

regions.

Data extraction, evaluation and analysis

YS and CQ extracted data from the retrieved studies. Then they independently used the Cochrane Collaboration tool for assessing risk of bias[17] to assess the quality of the trials (Tables 3-24). The tool comprised of seven specific domains (named sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting and other issues). We only included studies in "low risk" of bias in the randomization sequence generation and did not show high risks in any other domains. We used the funnel plot to detect the publication bias. If it was symmetrical we considered there was no publication bias, or else, we considered there was publication bias. If there was some disagreement between the two authors, they would resolve it by discussion.

We analyzed the included data with the Revman software (Version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012). We used the relative risk (RR) to evaluate the outcomes. If there was not significant heterogeneity between the included studies (detected by the P value of the χ^2 test over 0.10 and $I^2 \leqslant 25\%$) we used the Mantel-Haenszel fixed effect model to analyze data. If there was significant heterogeneity (detected by the P value was less than 0.10 and/or $I^2 \geqslant 50\%$) we detected if there was clinical heterogeneity. In the condition of absence of clinical heterogeneity we pooled data with random effects model. If $P \geqslant 0.10$ and $25\% \leqslant I^2 \leqslant 50\%$, we decided to choose the fixed effect or random effects models to combine data by discussion. Considering there might be clinical heterogeneity between different diseases we performed subgroup analyses (studies were divided into four subgroups: the adult acute leukemia, chronic myelogenous leukemia, lymphoma and pediatric acute myeloid leukemia subgroups). We also used sensitivity analyses to assess the association of the quality of included studies and the clinical characteristics. A two-sided P value of less than 0.05 was considered as a significant difference. We also used the GRADE grid to evaluate the quality of evidence on the primary outcome.

Statistical analysis

Technical appendix, statistical code, and dataset available from the corresponding author at yandgics@126.com. The article was reviewed by the statistician Xiaoxiao Wang. In her opinion, the RR rate was suitable, the heterogeneity of the included articles was effectively detected and the appropriate pooling methods (the random effects model or fixed effect model) was chosen for the systematic review. He also supported using the funnel plot to detect the publication bias.

RESULTS

We searched 13143 references in total. There were 367 papers retained after we examined the titles and abstracts. We excluded 347 references in the further assessment with the reason of that focused on the solid

Table 4 Risk assessment of Dian Rong 2009 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: Central randomized Comment: Probably done. Several studies published by this research group reported reliable randomization method
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Quote: A double-blind and placebo controlled Comment: Probably done. Several studies published by this research group reported reliable method to warrant the double blindness
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: Mortality and survival time are objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

Table 5 Characteristics of Xiu Mei 1997 study

Methods	A randomized controlled study
Participants	Non-Hodgkin lymphoma patients
Interventions	TCM group: Standard chemotherapy + traditional Chinese medicine Control group: Standard chemotherapy
Outcomes	The primary outcome: The overall response rate

TCM: Traditional Chinese medicine.

tumors but not the hematologic malignancies, were not real RCTs, did not report the primary outcome of our study or clearly described the randomization methods, and had other reasons that did not conform our inclusion criteria[18-20]. Finally all the reviewers agreed 11 studies[15,21-35] involved 891 participants should be included for meta-analysis (Figure 1).

Characteristics of included trials

The 11 included studies all compared the OR rate between the addition of TCM or not in the treatment of hematologic malignant diseases, such as acute leukemia, lymphoma, *etc.* Seven studies[15,21,24-31,33,34] researched the effect of adding TCM to the standard treatment for adult acute leukemia patients. Among the 7 studies, Mao Sheng 2007, Chuan Xin 2013, Ji Hong 2011, Rui Rong 2004 and Wen Jiang 2010[26,27,29,30,32,34] focused on acute myeloid leukemia. [Wang, 2007 #18; Wang, 2007 #9; Xu, 2010 #11; Zhu, 2011 #21]. The rest two studies did not restrict the type of the acute leukemia (lymphoblastic or non-lymphoblastic). Only one study Chuan Xin 2013[32] focused on pediatric acute myeloid leukemia patients while no study focused on pediatric acute lymphoblastic leukemia patients. Also only one study Xiu Mei 1997[35] focused on lymphoma patients. Two studies (sHai Yan 2007 and sWei Hong 2013)[22,23] focused on chronic myelogenous leukemia patients. But the basic treatment of the two studies were the hydroxyurea and/or a-interferon treatment but not the tyrosine kinase inhibitors which was the standard treatment recently[36]. Hence we did a sensitivity analysis of excluding the two studies. There was not study included was about the multiple myeloma (MM) or myelodysplastic syndrome (MDS). Only the study Dian Rong 2009[15] published one article in English, all of the rest studies were published in Chinese. Only one study reported the quality of life hence we did not analyze this outcome. There was not significant difference in the demographic characteristics of the two treatment groups in the 11 included studies (Table 25).

Quality of included trials

Five studies (Dian Rong 2009; Ying Fei 2005; Su Juan 2005; Mao Sheng 2007; Rui Rong 2004)[15,21,24-29,31,33] were multi-center double-blind RCT studies. The rest six studies[22,23,30,32,34,35] were single center studies and did not use the blind method. All of the included studies were not large sampled with the largest sample size (Xiu Mei 1997)[35] was 167 and the smallest sample size was 18 (sHai Yan 2007)[22]. All of the included studies did not use the intention to treat strategy to analyze results. There was no other factors influenced the quality of included studies. The funnel plot of the primary outcome was symmetric (Figure 2 and Tables 3-24, 26).

Efficacy analysis

Studies in both the OR and CR meta-analyses did not show significant heterogeneity so we combined data with the fixed effect model. The efficacy analyses showed the TCM arm had a significantly higher OR rate than the control arm (RR = 1.17 with a 95%CI: (1.10, 1.25), $P < 0.00001$) (Figure 3). The higher response rate was also statistically significant in the sensitivity analysis of excluding the two chronic myelogenous studies (RR = 1.17, 95%CI: (1.09, 1.26), $P < 0.00001$). As for the CR rate, the TCM arm was significantly higher than the control group as well [RR = 1.24, 95%CI: (1.11, 1.37), $P < 0.0001$] (Figure 4). And also it was

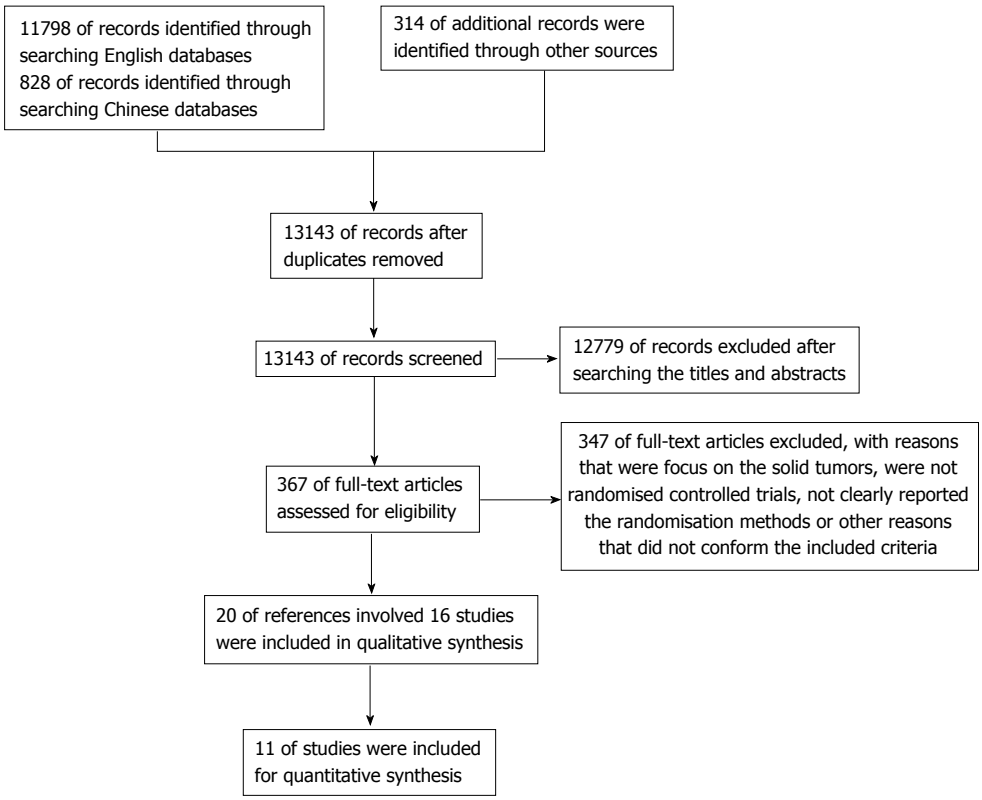


Figure 1 Study selection.

Table 6 Risk assessment of Xiu Mei 1997 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: The random sequence produced by rolling the dice Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: Mortality and survival is an objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

Table 7 Characteristics of Ji Hong 2011 study

Methods	A randomized controlled study
Participants	Initial treat old AML patients
Interventions	TCM group: HAG + TCM Control group: HAG
Outcomes	The primary outcome: The overall response rate

TCM: Traditional Chinese medicine; AML: Acute myeloid leukemia; HAG: Homoharringtonine + cytoarabine + granulocye colony stimulating factor.

not changed in the sensitivity analysis that excluded the two chronic myelogenous leukemia studies [RR = 1.21, 95%CI: (1.08, 1.35), P = 0.0007]. However, the Summary of findings (SoF) table showed the quality of the evidence was low (Table 27). There were three studies[15,21,25-27,35] reported the survival rate. The pooled results of the three studies did not show significant difference between the TCM arm and the control arm [RR = 1.22, 95%CI: (0.77, 1.94), P = 0.40] (Figure 5). Studies included in this analysis reported the survival rate of different period and the heterogeneity was significant. As a result of it, we used the random effects model to pool data.

Table 8 Risk assessment of Ji Hong 2011 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: Use the random number table to get the allocation sequence Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: The response rate is an objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: 7 participants in 53 randomized lost to follow-up Comment
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

Table 9 Characteristics of Ying Fei 2005 study

Methods	A multicenter double-blinded randomized controlled study
Participants	Initial treat leukemia patients
Interventions	TCM group: standard chemotherapy + Shen Qi Fu Zheng Ye Control group: Standard chemotherapy
Outcomes	The primary outcome: The overall response rate

TCM: Traditional Chinese medicine.

Serious AEs analysis

Our study demonstrated the TCM arm had a significantly less non-hematologic serious AEs rates in the gastrointestinal reaction [RR = 0.50, 95%CI: (0.37, 0.68), $P < 0.0001$], liver and/or kidney injury [RR = 0.37, 95%CI: (0.26, 0.53), $P < 0.00001$] and heart injury [RR = 0.24, 95%CI: (0.09, 0.68), $P = 0.007$] analyses (Figure 6). Additionally, the TCM showed a trend of reducing the infection rate [RR = 0.16, 95%CI: (0.02, 1.12), $P = 0.07$] but it was not statistically significant (Figure 7). The rates of Ⅲ-Ⅳ grade agranulocytosis and thrombocytosis were not different between adding TCM in the treatment method and not adding it [RR = 0.52, 95%CI: (0.14, 1.84), $P = 0.31$; RR = 0.52, 95%CI: (0.14, 1.91), $P = 0.33$, respectively] (Figure 7). Most of the included studies did not report the myelosuppression recovery time. So we did not analyze this outcome. In the non-hematologic serious AEs analyses, studies were pooled with the fixed effect model while in the hematologic AEs analyses, studies were pooled with the random effects model because of the significant heterogeneity. Because there were only two to three studies included in the serious AEs meta-analyses, we did not perform subgroup analysis to detect the clinical heterogeneity.

DISCUSSION

Oncologists begin to pay attention to the effect of TCM for the malignant diseases treatment in nowadays.

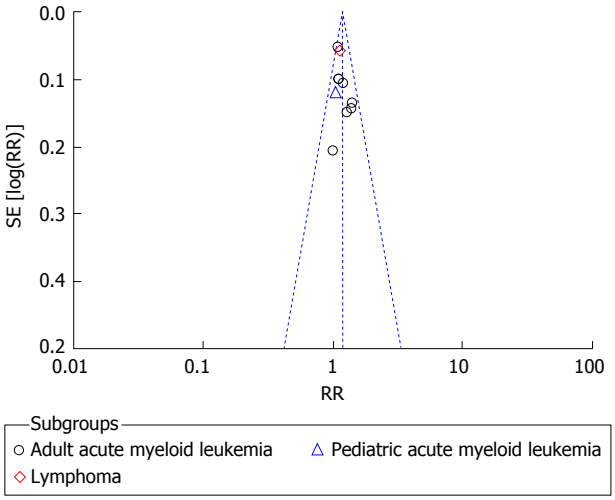


Figure 2 Funnel plot of the overall response meta-analysis. RR: Risk ratios.

Several meta-analyses revealed that TCM could improve response rate for some kinds of solid tumors[11-13]. There were also several RCTs showed that some TCM could increase the OR rate and decrease the AEs rate for hematologic malignancies. But the results published were not consistent[15,34]. At the same time there is not large sample sized RCT reported. As is generally accepted, meta-analysis attempts to identify all studies that would meet the eligibility criteria, subjectively assess the validity of the findings of the included studies and systematically present and synthesize the characteristics and findings of the included studies[37]. Therefore, it increases the sample size and reports a more reliable result. In The Oxford 2011 Levels of Evidence Table, meta-analysis of RCT has become the highest level of evidence[38]. In consequence, meta-analysis is a good method to evaluate the efficacy and safety of TCM for hematologic malignancies.

Response rate of TCM

Our results showed that TCM significantly increased the OR and CR rates. Although the GRADE SoF tables (Table

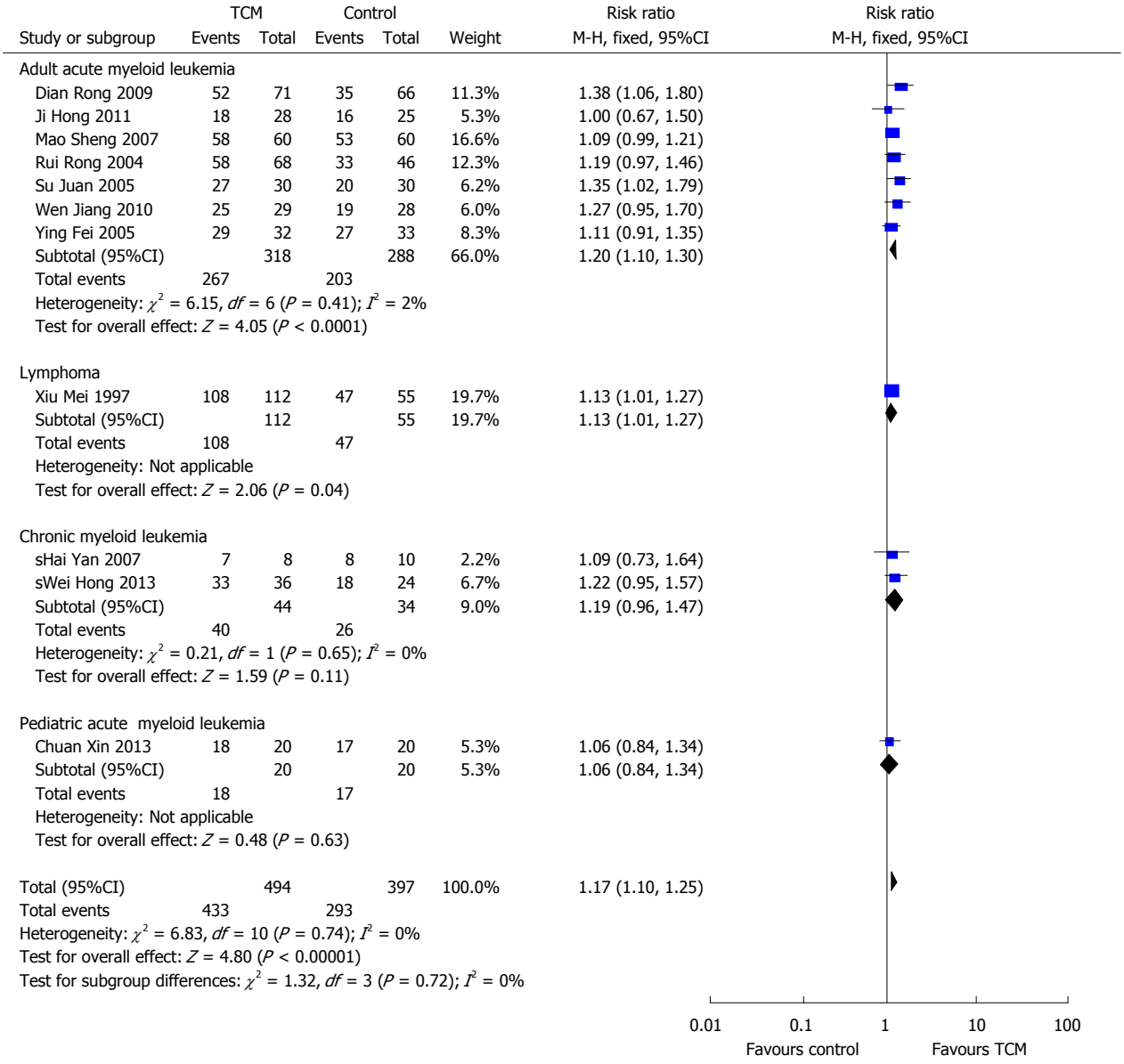


Figure 3 Overall response meta-analysis. TCM: Traditional Chinese medicine.

Table 10 Risk assessment of Ying Fei 2005 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: Generate randomization sequence by drawing lots Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: The response rate is an objective parameter subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

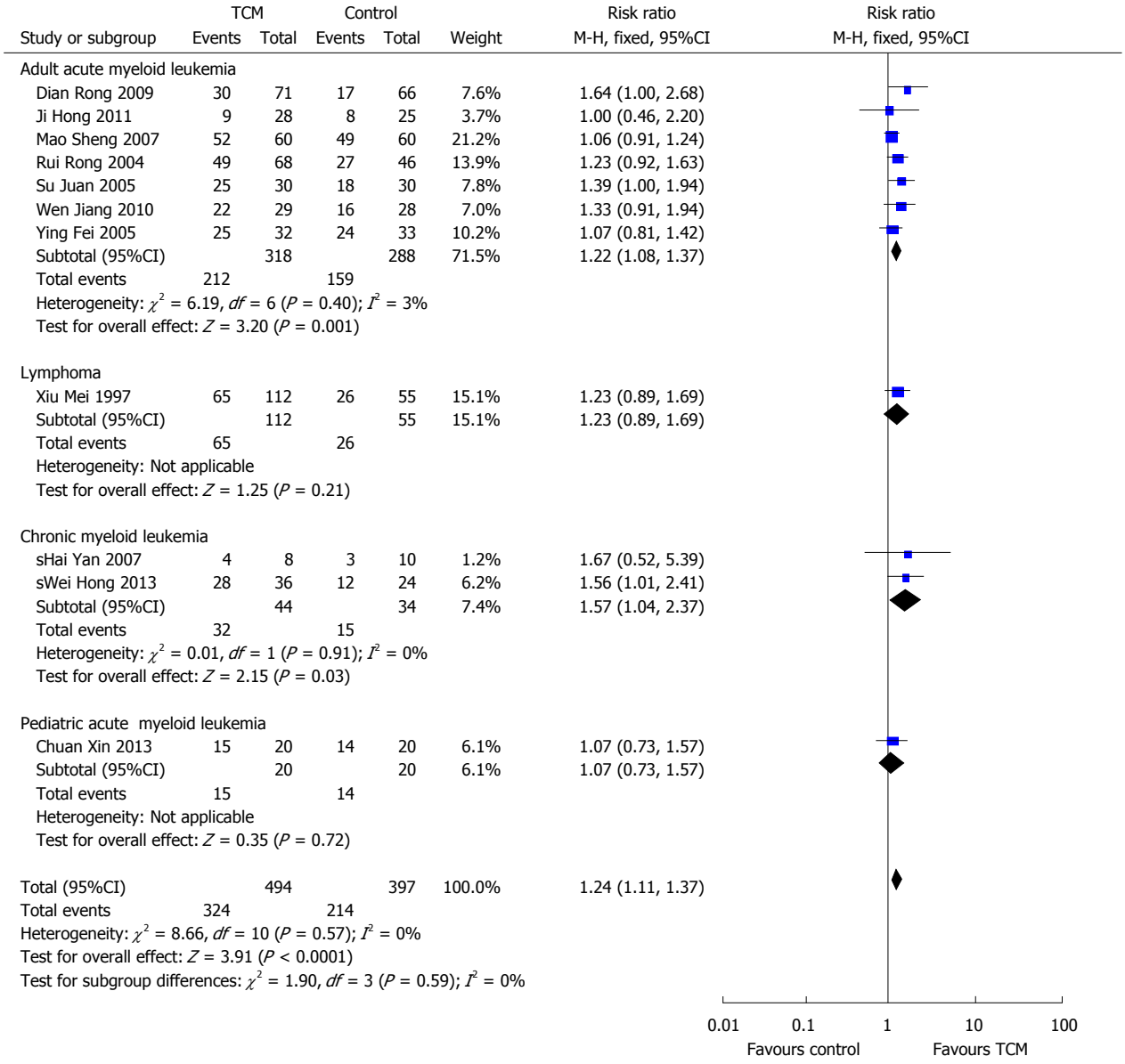


Figure 4 Complete response meta-analysis. TCM: Traditional Chinese medicine.

Table 11 Characteristics of Wen Jiang 2010 study

Methods	A randomized placebo controlled study
Participants	Initial treat acute leukemia patients
Interventions	TCM group: Standard chemotherapy + Shen Qi Qing Re Ke Li
	Control group: Standard chemotherapy
Outcomes	The primary outcome: The overall response rate

TCM: Traditional Chinese medicine.

27) showed the evidence quality of the two meta-analyses was low and the recommendation strength was weak (data not show), the TCM causes little side effects and it is economical. Furthermore, even though we included studies of different diseases there was not significant heterogeneity in the meta-analyses. So we could pooled data with the fixed effect model which made the result more reliable. Subsequently, it is suggested that TCM, as an adjuvant treatment method, can improve the efficacy of hematologic malignant diseases treatment.

However, there were two studies included in the chronic myelogenous leukemia subgroup prescribed the hydroxyurea or interferon as the fundamental treatment rather than the tyrosine kinase inhibitors which should be the first choice[36] nowadays. We excluded the two studies in the sensitivity analyses and then we got the same result that the TCM arm had significantly higher response rates (both OR and CR) than the control arm. The results of the sensitivity analyses strengthened the evidence that the response rate could be increased by adding TCM for hematologic malignancies. But there was only one study included in the pediatric acute myeloid leukemia and lymphoma subgroups and no studies on MM and MDS. As it was shown in the efficacy forest, the better effect of the TCM was mainly contributed by

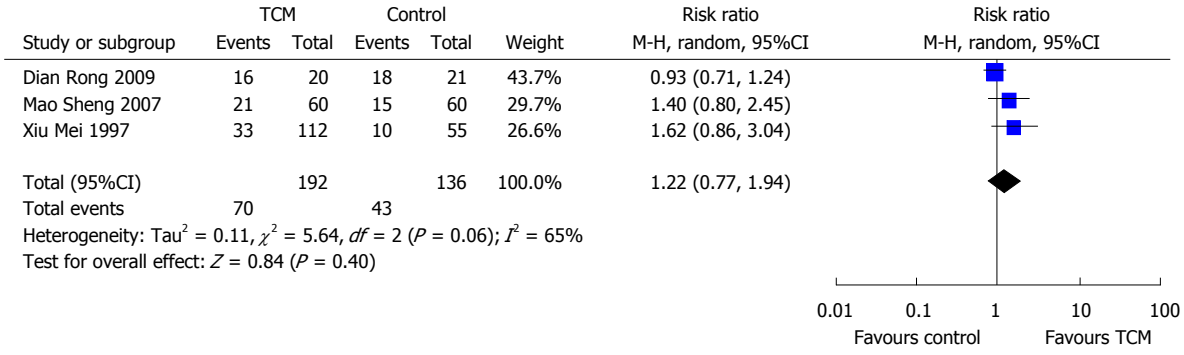


Figure 5 Survival rate meta-analysis. TCM: Traditional Chinese medicine.

Table 12 Risk assessment of Wen Jiang 2010 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: Use the random number table to get the allocation sequence Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: The response rate is an objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

Table 13 Characteristics of Su Juan 2005 study

Methods	A multicenter randomized controlled study
Participants	Acute leukemia
Interventions	TCM group: Standard chemotherapy + TCM Qing Re Jie Du Kang Bai Fang Control group: Standard chemotherapy
Outcomes	The primary outcome: The overall response rate

TCM: Traditional Chinese medicine.

Table 14 Risk assessment of Su Juan 2005 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: Use the random number table to get the allocation sequence Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: The response rate is an objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

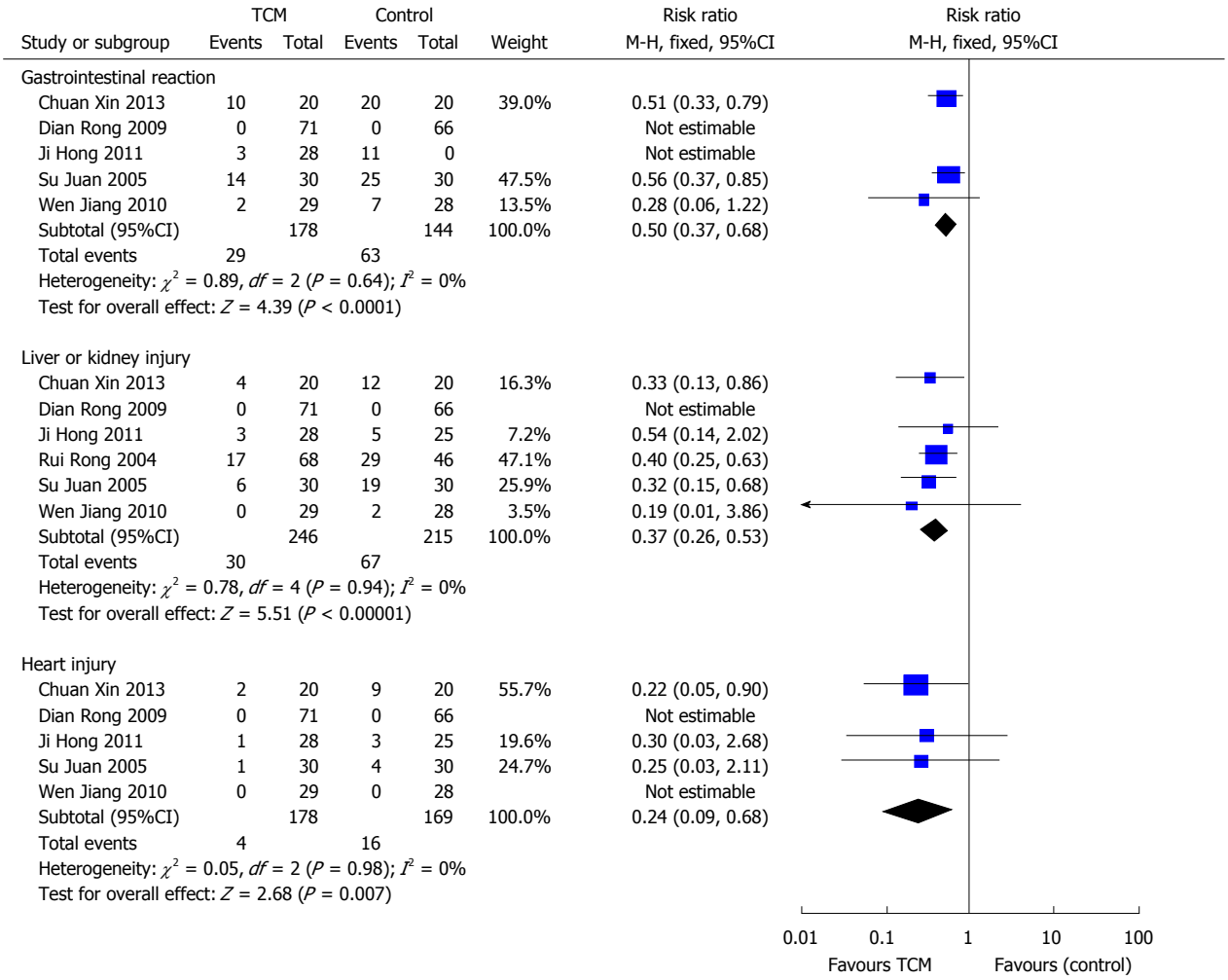


Figure 6 Non-hematologic serious adverse effects meta-analysis. TCM: Traditional Chinese medicine.

Table 15 Characteristics of Mao Sheng 2007 study

Methods	A multicenter double-blinded randomized placebo controlled study
Participants	Acute myeloid leukemia patients with micro residual disease
Interventions	TCM group: Standard chemotherapy + Yi Qi Jie Du Huo Xue Fang Control group: Standard chemotherapy
Outcomes	The primary outcome: The overall response rate

TCM: Traditional Chinese medicine.

the adult acute leukemia subgroup. For this reason we concluded TCM can be used as the adjuvant treatment for acute myeloid leukemia and there was in lack of studies on other hematologic malignant diseases, including chronic myelogenous leukemia.

Survival rate of TCM

There were only three studies with significantly heterogeneity involved 328 participants included in the survival rate meta-analysis. We did not show the difference between adding the TCM or not for treatment of malignant hematologic diseases. The result might because the small number of included studies was not enough to show a statistical significance or the addition of TCM can not change the survival rate. We need more high quality studies to clarify the problem. As a result of it, the data included was not enough to draw a conclusion of besides increasing the response rates, whether the addition of TCM can further improve patients survival rate.

Serious AEs rate of TCM

It is well known in the solid tumors treatment, TCM can decrease the AEs of chemotherapy[39], our results also showed that TCM significantly decreased the serious non-hematologic AEs and had a trend to reduce the serious infection rate. The result enhanced the role of TCM for hematologic malignant diseases treatment. Decreasing the serious non-hematologic AEs makes the chemotherapy safer and improves patients' tolerance

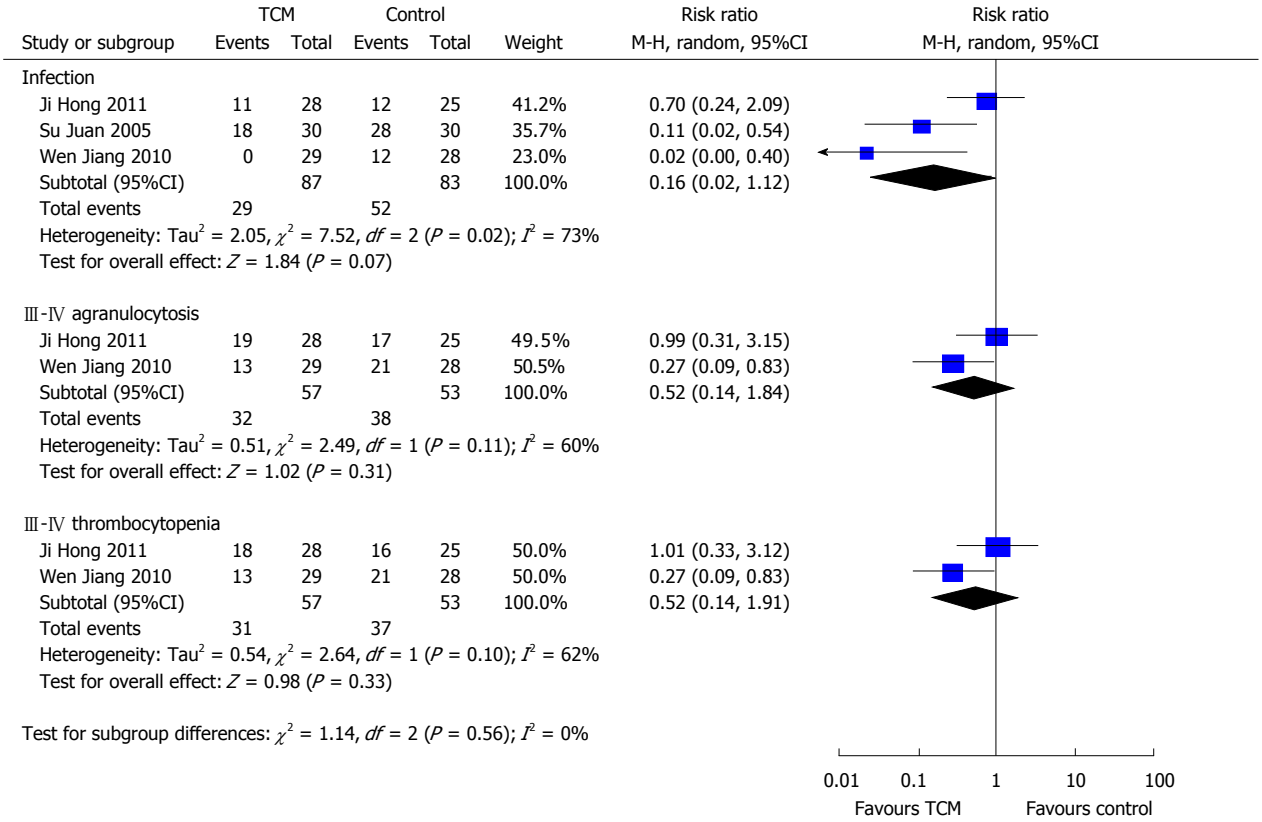


Figure 7 Hematologic serious adverse effects meta-analysis. TCM: Traditional Chinese medicine.

Table 16 Risk assessment of Mao Sheng 2007 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: Use the random number table to get the allocation sequence Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: The response rate is an objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

Table 17 Characteristics of Rui Rong 2004 study

Methods	A multicenter double-blinded randomized placebo controlled study
Participants	Acute myeloid leukemia
Interventions	TCM group: Standard chemotherapy + Yi Qi Yang Yin Qing Re Fa Control group: Standard chemotherapy
Outcomes	The primary outcome: The overall response rate

TCM: Traditional Chinese medicine.

and adherence. This point is especially important for hematologic malignant diseases because most of such patients do not have the opportunity of surgical operation and rely on chemotherapeutic treatment. Additionally, the chemotherapy usually has better effect for hematologic malignant diseases than solid tumors. Infection is the most common cause of death among patients with acute leukemia accounting for up to 75% of mortality[40]. In our study, we showed a trend of reducing infection rate but it was not statistically significant. Since the three included studies all showed better effect of TCM and two were statistically significant we inferred the reason might be there were not enough studies included. More data was

Table 18 Risk assessment of Rui Rong 2004 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: Use the random number table to get the allocation sequence Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: The response rate is an objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

Table 19 Characteristics of Chuan Xin 2013 study

Methods	A randomized controlled study
Participants	Child acute myeloid leukemia patients
Interventions	TCM group: Standard chemotherapy + traditional Chinese medicine Control group: Standard chemotherapy
Outcomes	The primary outcome: The overall response rate

TCM: Traditional Chinese medicine.

Table 20 Risk assessment of Chuan Xin 2013 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: Use the random number table to get the allocation sequence Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: The response rate is an objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

Table 21 Characteristics of sWei Hong 2013 study

Methods	A randomized controlled study
Participants	Chronic myeloid leukemia patients
Interventions	TCM group: A-interferon or hydroxyurea + TCM Control group: A-interferon or hydroxyurea
Outcomes	The primary outcome: The response rate

TCM: Traditional Chinese medicine.

needed to confirm whether it was the truth. There were only two studies included in the serious hematologic AEs meta-analyses and we were in need of more studies to clarify this question.

Comparison with other studies

Our study result was consistent with several meta-analyses on the solid tumors[11-14]. In the studies, the authors showed that the Chinese herbal medicine (CHM) can increase the response and survival more than one year rates. Among the diseases studied, the non small cell lung cancer (NSCLC) is also sensitive to chemotherapeutic agents that is something like the hematologic malignancies. Our study also showed that TCM increased the response rate but failed to show that TCM increased the survival rate. This might be because there were not enough participants involved in our meta-analysis or the different clinical features of the diseases we researched. In the NSCLC study, authors

Table 22 Risk assessment of sWei Hong 2013 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: The random number table was used to generate the allocation sequence Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: Mortality and survival is an objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

Table 23 Characteristics of sHai Yan 2007 study

Methods	A randomized controlled study
Participants	Chronic myeloid leukemia patients
Interventions	Traditional Chinese medicine group: Hydroxyurea + traditional Chinese medicine Control group: Hydroxyurea
Outcomes	The primary outcome: The response rate

Table 24 Risk assessment of sHai Yan 2007 study

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: The random number table was used to generate the allocation sequence Comment: Probably done
Allocation concealment (selection bias)	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of participants and personnel (performance bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: Not mentioned Comment: Mortality and survival is an objective parameter. Subjective judgement can not influent the result
Incomplete outcome data (attrition bias) All outcomes	Unclear	Quote: Not mentioned Comment: Unclear
Selective reporting (reporting bias)	Low risk	The primary outcome listed in the method section are all reported Comment: Probably done
Other bias	Unclear	The study did not use the intention to treat strategy to analyze the result

demonstrated the CHM decreased the morbidity of serious agranulocytosis and thrombocytosis which was not revealed in our study. As well, this might be caused by the lack of studies included or the different clinical features of the diseases. The consistency of our study with other studies strengthened our results.

Limitation of the meta-analysis

We have tried our best to make our research more reliable but we still have some limitation. First, none of included studies were performed out of China and all of the included studies except one were published in Chinese. As the funnel plot was symmetric, the publication bias was unavoidable. Second, six of the included studies were small sample sized and did not mention any blindness methods that had the risk of compromising concealment allocation[41]. Third, except the acute leukemia subgroup, there were rare studies of other hematologic malignant diseases included in the meta-analyses. Thus the efficacy result mainly reflected the efficacy of TCM for acute leukemia. According to our result, it was not clear whether the TCM usage had the same efficacy for other hematologic malignant diseases. Finally, all of the included studies were not large sample sized. Only 5 studies used the central randomization method. As a result of it, the quality of evidence of our study was compromised and the GRADE recommendation level was low. Because of these limitations the reliability might be influenced and the

Table 25 Characteristics of included studies

Studies	Age	Sex (male:female)	Race	Disease	No. of participants (TCM:control)	Intervention		Published language
						TCM	Control	
Dian Rong 2009[15,21,25,31,33]	TCM 39.52 ± 18.87 Control 37.94 ± 18.55	TCM 50:21 Control 39:27	Chinese	Acute leukemia	71:66	Compound Zhe Bei granule + standard chemotherapy	Placebo + standard chemotherapy	English
Mao Sheng 2007[26,27]	TCM 35.63 ± 6.46 Control 36.57 ± 7.38	TCM 33:27 Control 31:29	Chinese	Acute myeloid leukemia	60:60	Yi Qi Jie Du Huo Xue decoction + standard western mechicine	Standard western mechicine	Chinese
sHai Yan 2007[22]	TCM 18-65 Control 19-63	TCM 5:3 Control 7:3	Chinese	Chronic myelogenous leukemia	8:10	Qu Du Hua Yu decoction + hydroxyurea	Hydroxyurea	Chinese
sWei Hong 2013[23]	TCM 25-60 Control 25-65	TCM 22:14 Control 17:7	Chinese	Chronic myelogenous leukemia	22:17	TCM + interferon-α	Interferon-α	Chinese
Chuan Xin 2013[32]	TCM 4.30 ± 1.81 Control 4.95 ± 2.04	TCM 12:8 Control 10:10	Chinese	Pediatric acute myeloid leukemia	20:20	TCM + standard chemotherapy	Standard chemotherapy	Chinese
Ji Hong 2011[34]	TCM 60-71 Control 61-72	TCM 16:16 Control 15:13	Chinese	Elderly acute myeloid leukemia	32:28	TCM + HAG chemotherapy	HAG chemotherapy	Chinese
Rui Rong 2004[29]	TCM 12-78 Control 11-76	TCM 40:28 Control 27:19	Chinese	Acute myeloid leukemia	68:46	TCM + standard chemotherapy	Standard chemotherapy	Chinese
Su Juan 2005[24]	TCM 32.5 ± 12.45 Control 31.53 ± 12.41	TCM 16:14 Control 17:13	Chinese	Acute leukemia	30:30	Qing Re Jie Du kang Bai decoction + standard chemotherapy	Standard chemotherapy	Chinese
Wen Jiang 2010[30]	TCM 47-78 Control 46-79	TCM 17:12 Control 15:13	Chinese	Acute myeloid leukemia	29:28	Shen Qi Qing Re Ke Li + HAG chemotherapy	HAG chemotherapy	Chinese
Xiu Mei 1997[35]	TCM 6-73 Control 6-71	TCM 72:40 Control 36:19	Chinese	Non-Hodgkin lymphoma	112:55	TCM + standard chemotherapy	Standard chemotherapy	Chinese
Ying Fei 2005[28]	TCM 13-72 Control 15-71	TCM 22:10 Control 25:8	Chinese	Acute leukemia	32:33	Shen Qi Fu Zheng injection + standard chemotherapy	Standard chemotherapy	Chinese

TCM: Traditional Chinese medicine; HAG: Homoharringtonine + cytoarabine + granulocye colony stimulating factor.

Table 26 Quality assessment of included studies

Studies	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias) All outcomes	Blinding of outcome assessment (detection bias) All outcomes	Incomplete outcome data (attrition bias) All outcomes	Selective reporting (reporting bias)	Other bias
Dian Rong 2009[15,21,25,31,33]	Low risk	Unclear	Low risk	Low risk	Unclear	Low risk	Unclear
Mao Sheng 2007[26,27]	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	Unclear
sHai Yan 2007[22]	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	Unclear
sWei Hong 2013[23]	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	Unclear
Chuan Xin 2013[32]	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	Unclear
Ji Hong 2011[34]	Low risk	Unclear	Unclear	Low risk	Low risk	Low risk	Unclear
Rui Rong 2004[29]	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	Unclear
Su Juan 2005[24]	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	Unclear
Wen Jiang 2010[30]	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	Unclear
Xiu Mei 1997[35]	Low risk	Unclear	Unclear	Unclear	Unclear	Low risk	Unclear
Ying Fei 2005[28]	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	Unclear

Table 27 Summary of findings of the overall response and complete response outcomes

Outcomes	Illustrative comparative risks[1] (95%CI)		Relative effect (95%CI)	No. of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Control	Corresponding risk Overall response rate				
Overall response rate	Study population		RR = 1.14	974	++--	
	761 per 1000	867 per 1000 (784-959)	(1.03-1.26)	(12 studies)	Low	
	Moderate					
	775 per 1000	883 per 1000 (798-976)				
Complete response rate	Study population		RR = 1.21	974	++--	
	579 per 1000	701 per 1000 (579-846)	(1-1.46)	(12 studies)	Low[2,3]	
	Moderate					
	579 per 1000	701 per 1000 (579-845)				
Overall response rate for malignant hematologic disease Patient or population: Patients with malignant hematologic disease Settings: Intervention: Overall response rate						

[1]The basis for the assumed risk (*e.g.*, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95%CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95%CI); [2]Not all studies included were high quality randomized controlled trial; [3]Most studies showed better effect when traditional Chinese medicine was added while some studies did not show statistically significant better effect. GRADE Working Group grades of evidence. High quality: Further research is very unlikely to change our confidence in the estimate of effect; Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate; Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate; Very low quality: We are very uncertain about the estimate.

results should be interpreted with caution. As there were some limitations, we extensively searched the related databases, publications and websites, strictly screened and evaluated retrieved articles and analyzed the pooled data. Our study assessed the evidence available recently so it is still significant for evaluating the role of TCM for hematologic malignancies.

Because TCM causes little AEs, has little interaction with other drugs or treatment methods it can be safely prescribed in most of the malignant diseases treatment. It is especially popular among the complementary and alternative medicine usage in the palliative care of cancer patients[42]. But recently, it plays more important role in the tumor treatment. Our meta-analysis demonstrated that TCM not only had the advantage of reducing the chemotherapy associated serious non-hematologic AEs and had a trend to reduce the serious infection rate, but also significantly increased the response rate. Our result suggests TCM is helpful for hematologic malignant diseases treatment. Although we failed to show a better survival rate of TCM compared with control, we believed to recommend adding TCM to the hematologic malignancies treatment as an adjuvant therapy is reasonable, at least for adult acute leukemia.

Conclusion and implications for research

TCM increases the OR and CR rate for acute leukemia treatment and reduced the treatment associated serious non-hematologic AEs. Therefore, we recommend including TCM in the hematologic malignancies treatment, at least for adult acute leukemia treatment.

Except adult acute leukemia, we need more high quality studies on other hematologic malignant diseases, pediatric patients and in other regions apart from China. We are also in need of studies of TCM on the survival, infection and hematologic AEs rates for hematologic malignancies treatment.

COMMENTS

Background

Albeit as the standard treatment, the chemotherapy always causes serious adverse effects (AEs) and its efficacy is still not satisfactory. Recently, many studies showed that traditional Chinese medicine (TCM) can improve the effect of the standard treatment and reduce the AEs.

Research frontiers

In recent years, more and more researchers begin to pay attention to the effect of TCM for malignant diseases. Many studies showed that TCM can increase the efficacy of the standard treatment and decrease the AEs.

Innovations and breakthroughs

Although there were many clinical studies published on the TCM for hematologic malignances, as far as we know, there was no systematic review published on this issue. As far as we know, the authors first summarized the evidence now available on it with systematic review and demonstrated a subjective result. The result confirmed the effectiveness of TCM for hematologic malignances and could be used in the clinical practice.

Applications

The result showed that TCM can increase the overall response and complete response rates. In addition, TCM also reduced the non-hematologic serious AEs. The authors consider TCM should be used for hematologic malignances treatment.

Peer-review

The manuscript is quite interesting.

REFERENCES

1 **Epplein M**, Bostick RM, Mu L, Ogino S, Braithwaite D, Kanetsky PA. Challenges and opportunities in international molecular cancer prevention research: An ASPO Molecular Epidemiology and the Environment and International Cancer Prevention Interest Groups Report. *Cancer Epidemiol Biomarkers Prev* 2014; **23**: 2613-2617 [PMID: 25277796 DOI: 10.1158/1055-9965.epi-14-0848]

2 **Ramdass B**, Chowdhary A, Koka PS. Hematological malignancies: disease pathophysiology of leukemic stem cells. *J Stem Cells* 2013; **8**: 151-187 [PMID: 24699024]

3 **Avigan D**, Hari P, Battiwalla M, Bishop MR, Giralt SA, Hardy

NM, Kröger N, Wayne AS, Hsu KC. Proceedings from the National Cancer Institute's Second International Workshop on the Biology, Prevention, and Treatment of Relapse after Hematopoietic Stem Cell Transplantation: part II. Autologous Transplantation-novel agents and immunomodulatory strategies. *Biol Blood Marrow Transplant* 2013; **19**: 1661-1669 [PMID: 24018393 DOI: 10.1016/j.bbmt.2013.08.011]

4 **de Lima M**, Porter DL, Battiwalla M, Bishop MR, Giralt SA, Hardy NM, Kröger N, Wayne AS, Schmid C. Proceedings from the National Cancer Institute's Second International Workshop on the Biology, Prevention, and Treatment of Relapse After Hematopoietic Stem Cell Transplantation: part III. Prevention and treatment of relapse after allogeneic transplantation. *Biol Blood Marrow Transplant* 2014; **20**: 4-13 [PMID: 24018392 DOI: 10.1016/j.bbmt.2013.08.012]

5 **Fisher RI**, Gaynor ER, Dahlberg S, Oken MM, Grogan TM, Mize EM, Glick JH, Coltman CA, Miller TP. Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. *N Engl J Med* 1993; **328**: 1002-1006 [PMID: 7680764 DOI: 10.1056/nejm199304083281404]

6 **Cai Q**, Westin J, Fu K, Desai M, Zhang L, Huang H, Jiang W, Liang R, Qian Z, Champlin RE, Wang M. Accelerated therapeutic progress in diffuse large B cell lymphoma. *Ann Hematol* 2014; **93**: 541-556 [PMID: 24375125 DOI: 10.1007/s00277-013-1979-7]

7 **Ujjani C**, Cheson BD. The optimal management of follicular lymphoma: an evolving field. *Drugs* 2013; **73**: 1395-1403 [PMID: 23884816 DOI: 10.1007/s40265-013-0092-5]

8 **Furst DE**, Keystone EC, Braun J, Breedveld FC, Burmester GR, De Benedetti F, Dörner T, Emery P, Fleischmann R, Gibofsky A, Kalden JR, Kavanaugh A, Kirkham B, Mease P, Sieper J, Singer NG, Smolen JS, Van Riel PL, Weisman MH, Winthrop K. Updated consensus statement on biological agents for the treatment of rheumatic diseases, 2011. *Ann Rheum Dis* 2012; **71** Suppl 2: i2-i45 [PMID: 22460137 DOI: 10.1136/annrheumdis-2011-201036]

9 **Tavazzi E**, Ferrante P, Khalili K. Progressive multifocal leukoencephalopathy: an unexpected complication of modern therapeutic monoclonal antibody therapies. *Clin Microbiol Infect* 2011; **17**: 1776-1780 [PMID: 22082208 DOI: 10.1111/j.1469-0691.2011.03653.x]

10 **Stolz C**, Schuler M. Molecular mechanisms of resistance to Rituximab and pharmacologic strategies for its circumvention. *Leuk Lymphoma* 2009; **50**: 873-885 [PMID: 19373595 DOI: 10.1080/10428190902878471]

11 **Cho WC**, Chen HY. Transcatheter arterial chemoembolization combined with or without Chinese herbal therapy for hepatocellular carcinoma: meta-analysis. *Expert Opin Investig Drugs* 2009; **18**: 617-635 [PMID: 19388879 DOI: 10.1517/13543780902855308]

12 **Cho WC**, Chen HY. Clinical efficacy of traditional Chinese medicine as a concomitant therapy for nasopharyngeal carcinoma: a systematic review and meta-analysis. *Cancer Invest* 2009; **27**: 334-344 [PMID: 19212827 DOI: 10.1080/07357900802392683]

13 **Li SG**, Chen HY, Ou-Yang CS, Wang XX, Yang ZJ, Tong Y, Cho WC. The efficacy of Chinese herbal medicine as an adjunctive therapy for advanced non-small cell lung cancer: a systematic review and meta-analysis. *PLoS One* 2013; **8**: e57604 [PMID: 23469033 DOI: 10.1371/journal.pone.0057604]

14 **Zhong LL**, Chen HY, Cho WC, Meng XM, Tong Y. The efficacy of Chinese herbal medicine as an adjunctive therapy for colorectal cancer: a systematic review and meta-analysis. *Complement Ther Med* 2012; **20**: 240-252 [PMID: 22579437 DOI: 10.1016/j.ctim.2012.02.004]

15 **Lu DR**, Li DY, Chen XY, Ye PZ, Tian SD. Clinical research of compound zhebei granules for increasing the therapeutic effect of chemotherapy in refractory acute leukemia patients. *J Tradit Chin Med* 2009; **29**: 190-194 [PMID: 19894383]

16 **Lefebvre C**, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, green S, editors. Cochrane handbook for systematic reviews of interventions version 5.1.0 (updated march 2011). The Cochrane Collaboration, 2011. Available from: URL: http://www.cochrane-handbook.org

17 **Higgins JPT**, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 510 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: URL: http://www.cochrane-handbook.org

18 **Seely D**, Wu P, Fritz H, Kennedy DA, Tsui T, Seely AJ, Mills E. Melatonin as adjuvant cancer care with and without chemotherapy: a systematic review and meta-analysis of randomized trials. *Integr Cancer Ther* 2012; **11**: 293-303 [PMID: 22019490 DOI: 10.1177/1534735411425484]

19 **Zhu XY**, Zhang XZ, Zhong XY. [Effect of shenqi fuzheng injection for hemopoietic and immune function reconstruction in patients with hematologic malignancies undergoing chemotherapy]. *Zhongguo Zhongxiyi Jiehe Zazhi* 2010; **30**: 205-207 [PMID: 20462054]

20 **Ni H**. Clinical research of integrated chinese and western medicine for multiple myeloma. *Anhui Yixue Zazhi* 2006: 915-916

21 **Huang S**. The clinical study of compound granule of thunberg fritillary bulb for improving the survival of refractory acute leukemia patients. Beijing University Of Chinese Medicine, 2011. Available from: URL: http://xb.bucm.edu.cn/

22 **Li H**, Wang Q. Clinical research of qu du hua yu formula for treating chronic myeloid leukemia. *Liaoning Chuantong Yixue Zazhi* 2007; **34**: 169-170 [DOI: 10.3969/j.issn.1000-1719.2007.02.027]

23 **Pei W**. Curative effect research of combine traditional chinese and western medicine treatment of chronic myelogenous leukemia. *Zhongguo Zhongyao Zazhi* 2013: 935-936

24 **Peng S**. Clinical study of qingrejiedukangbai decoction combined with chemotherapy with western medicine treating acute leukemia with hyperactivity of virulent heat-evil. Hunan University of Chinese Medicine, 2005. Available from: URL: http://www.hnctcm.edu.cn/xueshuqikan/hunanzyydxxb/

25 **Tian S**. Clinical research of compound zhe bei mu granule assit chemotherapy for treating refractory acute leukemia. Beijing University of Chinese Medicine, 2006. Available from: URL: http://xb.bucm.edu.cn/

26 **Wang M**, Lang L, Zhao X, Di H, Li Z, Yang S, Hou W, Yan J. Clinical research of yi qi jie du huo xue chinese medicine combined with chemotherapy for treating the micro residual disease of adult aml. China Practical Medicine, 2007: 101-102. Available from: URL: http://c.wanfangdata.com.cn/Periodical-zglcsyyx.aspx

27 **Wang M**, Yang S, Hou W, Lang L, Yan J, Zhao X, Li Z. Clinical research of the yi qi jie du huo xue chinese medicine combined with chemotherapy for treating adult aml. Proceedings of the The 8th National conference of integrated Chinese and Western Medicine Hemotology, 2007: 5

28 **Wei YF**, Wang SY, Ren LL. [Efficacy of shenqi fuzheng injection combined with chemotherapy in treatment of acute leukemia and its effect on T-lymphocyte subsets, serum IFN-gamma, IL-10 and IL-2]. *Zhongguo Zhongxiyi Jiehe Zazhi* 2005; **25**: 303-306 [PMID: 15892271]

29 **Xu RR**, Cao F, Liu ZX. [Clinical observation on treatment of acute myelocytic leukemia by supplementing qi, nourishing yin and clearing heat principle]. *Zhongguo Zhongxiyi Jiehe Zazhi* 2004; **24**: 411-414 [PMID: 15199624]

30 **Xu W**, Yang S, Di H, Li Q, Qiao Zi, Jiang Q, Wang J, Liu X, Huo Y, Jia X, Zhao P, Ma Y. Clinical research of shen qi qing re ke li combined with hag chemotherapy for treatment of acute myeloid leukemia. Proceedings of the National Conference of Integrated Chinese and Western Medicine Hemotology, 2010: 4

31 **Ye F**. Clinical research of zhe bei mu granule reversing the multi-resistance of acute leukemia. Beijing University of Chinese Medicine, 2006. Available from: URL: http://xb.bucm.edu.cn/

32 **Zhang C**, Zou X, Li Y. Clinial research of 20 cases of pediatric acute myeloid leukmia treated withtraditional chinese medicine combined with western medicine chemotherapy. Traditional Chinese Medicinal Research, 2013: 32-33. Available from: URL: http:// www.cqvip.com/QK/96073X/index.asp

33 **Zhang Y**. Clinical research of compound zhe bei mu granules assit chemotherapy to improve the effecacy of acute leukemia treatment.

Beijing University of Chinese Medicine, 2007. Available from: URL: http://xb.bucm.edu.cn/

34 **Zhu J**. Clinical research of hag chemotherapy combined with chinese medicine for elderly acute myeloid leukemia. Modern Journal of Integrated Traditional Chinese and Western Medicine, 2011: 4637-4638. Available from: URL: http://xdjh.chinajournal.net.cn/WKC/WebPublication/index.aspx

35 **Guo XM**, Li JX, Yang XF. [Clinical observation on 112 cases with non-Hodgkin's lymphoma treated by Chinese herbs combined with chemotherapy]. *Zhongguo Zhongxiyi Jiehe Zazhi* 1997; **17**: 325-327 [PMID: 9863121]

36 **NCCN.org. Chronic myelogenous leukemia**. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines). 2015: Version 1. Available from: URL: http://www.nccn.org/professionals/physician_gls/pdf/cml.pdf

37 **Green S**, Higgins JPT, Alderson P, Clarke M, Mulrow CD, Oxman AD. Chapter 1: Introduction. In: Higgins JPT, green S (editors), cochrane handbook for systematic reviews of interventions version 5.1.0 (updated march 2011). The Cochrane Collaboration, 2011. Available from: URL: http://www.cochrane-handbook.org

38 **OCEBM Levels of Evidence Working Group**. The oxford 2011 levels of evidence. Oxford Centre for Evidence-Based Medicine. 2011

39 **Ling CQ**, Yue XQ, Ling C. Three advantages of using traditional Chinese medicine to prevent and treat tumor. *J Integr Med* 2014; **12**: 331-335 [PMID: 25074882 DOI: 10.1016/S2095-4964(14)60038-8]

40 **EJ B**. Infectious complications in patients receiving cytotoxic therapy for acute leukemia: History, background and approaches to management. In: Wingard JR, bowden RA, editors. Management of Infection in Oncology Patients. London: Martin Dunitz, 2003: 71-104

41 **Hills RK**, Gray R, Wheatley K. Balancing treatment allocations by clinician or center in randomized trials allows unacceptable levels of treatment prediction. *J Evid Based Med* 2009; **2**: 196-204 [DOI: 10.1111/j.1756-5391.2009.01023.x]

42 **Hyodo I**, Amano N, Eguchi K, Narabayashi M, Imanishi J, Hirai M, Nakano T, Takashima S. Nationwide survey on complementary and alternative medicine in cancer patients in Japan. *J Clin Oncol* 2005; **23**: 2645-2654 [PMID: 15728227]

P- Reviewer: Alshehabi Z, Romero MR **S- Editor**: Ji FF
L- Editor: A **E- Editor**: Liu SQ

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i3.181

World J Meta-Anal 2015 June 26; 3(3): 181-187
ISSN 2308-3840 (online)

META-ANALYSIS

Sleep-associated movement disorders and the risk of cardiovascular disease: A systematic review and meta-analysis

Zhen Fang, Yao-Wu Liu, Li-Yan Zhao, Yan Xu, Feng-Xiang Zhang

Zhen Fang, Yao-Wu Liu, Li-Yan Zhao, Yan Xu, Feng-Xiang Zhang, Department of Cardiology, the First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, Jiangsu Province, China

Author contributions: Fang Z, Liu YW and Zhang FX designed the review; Fang Z, Zhao LY and Xu Y collected the data; Fang Z and Liu YW analyzed the data and wrote the paper.

Supported by The National Natural Science Foundation of China, Nos. 81470456 and 81170160; The priority Academic Program Development of Jiangsu Higher Education Institutions.

Conflict-of-interest: No conflict of interest.
Data sharing: No additional data are available.

Correspondence to: Feng-Xiang Zhang, MD, PhD, Department of Cardiology, the First Affiliated Hospital of Nanjing Medical University, Guangzhou Road 300, Nanjing 210029, Jiangsu Province, China. njzfx6@njmu.edu.cn
Telephone: +86-25-68136056
Fax: +86-25-83717168

Received: January 25, 2015
Peer-review started: January 27, 2015
First decision: February 7, 2015
Revised: April 9, 2015
Accepted: April 16, 2015
Article in press: April 20, 2015
Published online: June 26, 2015

Abstract

AIM: To investigate whether an association exists between sleep-associated movement disorders and cardiovascular disease (CVD).

METHODS: Several studies have observed the relationship of sleep-associated movement disorders such as restless legs syndrome (RLS) and periodic limb movements during sleep with CVD, but the results were still contradictory. We performed an extensive literature search on PubMed, Medline and Web of Science published from inception to December 2014. Additional studies were manually searched from bibliographies of retrieved studies. Meta-analyses were conducted with Stata version 12.0 (Stata Corp, College Station, Texas). Pooled odds ratios (ORs) and 95%CIs were calculated to assess the strength of association using the random effects model. Sensitivity and subgroup analyses were performed to explore the underlying sources of heterogeneity. The publication bias was detected using Egger's test and Begg's test.

RESULTS: A total of 781 unique citations were indentified from electronic databases and 13 articles in English were finally selected. Among these studies, nine are cohort studies; two are case-control studies; and two are cross-sectional studies. The results showed that the summary OR of CVD associated with sleep-associated movement was 1.51 (95%CI: 1.29-1.77) in a random-effects model. There was significant heterogeneity between individual studies (P for heterogeneity = 0.005, I^2 = 57.6%). Further analysis revealed that a large-scale cohort study may account for this heterogeneity. A significant association was also found between RLS and CVD (OR = 1.54, 95%CI: 1.24-1.92). In a fixed-effects model, we determined a significant relationship between sleep-associated

movement disorders and coronary artery disease (CAD) (OR = 1.34, 95%CI: 1.16-1.54; *P* for heterogeneity = 0.210; I^2 = 30.0%). Our meta-analysis suggests that sleep-associated movement disorders are associated with prevalence of CVD and CAD.

CONCLUSION: This finding indicates that sleep-associated movement disorders may prove to be predictive of underlying CVD.

Key words: Sleep-associated movement disorders; Restless legs syndrome; Cardiovascular disease; Meta-analysis; Periodic limb movements during sleep

Core tip: We conducted a meta-analysis of 13 relevant studies to investigate the association between sleep-associated movement disorders and cardiovascular disease (CVD). The present study suggested that sleep-associated movement disorders are associated with prevalence of CVD. This finding indicates that sleep-associated movement disorders may prove to be predictive of underlying CVD.

Fang Z, Liu YW, Zhao LY, Xu Y, Zhang FX. Sleep-associated movement disorders and the risk of cardiovascular disease: A systematic review and meta-analysis. *World J Meta-Anal* 2015; 3(3): 181-187 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i3/181.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i3.181

INTRODUCTION

Sleep-associated movement disorders are a group of movement disorders which occur during sleep in relation to episodes of arousal and sleep disorder. They are characterized by the persistence of muscle tone or the emergence of motor activity. Among of them, restless legs syndrome (RLS) and periodic limb movements during sleep (PLMS) are the two most common disorders encountered in adult. RLS affects approximately 5%-10% of the general population and up to 80% of RLS patients may have PLMS[1,2]. RLS and PLMS can result in similar clinical problems due to sleep disruption[3]. Recently, several studies indicate that untreated RLS with PLMS may contribute partly to secondary causes of uncontrolled hypertension and cardiovascular disease (CVD), while some studies demonstrated negative results[4,5]. Therefore, the objective of the present study was to provide a systematic review and meta-analysis of the available evidence on the association between sleep-associated movement disorders and CVD in general populations.

MATERIALS AND METHODS

This meta-analysis was based on the guidelines of the Meta-analysis of Observational Studies in Epidemiology Group[6].

Data sources and search strategy

We performed a literature search of PubMed, Medline and Web of Science using key words of "periodic limb movements", "RLS", "heart disease", "CVD", "coronary artery disease (CAD)" and "sleep-associated movement disorders" published from inception to December 2014. Additional studies were manually searched from references of related studies or reviews and the language was limited in English. Review articles, abstracts, correspondence, conference proceedings and book chapters were excluded, and only one instance of the study found in multiple journals was included.

Inclusion and exclusion criteria

Prospective cohort, case-control, and cross-sectional studies based in general populations that assessed the association of sleep-associated movement disorders with CVD were eligible for this systematic review. Exclusion criteria were as follows: (1) duplicated studies; (2) no controls; and (3) no detail risk estimates and 95%CIs. We included only published full-text that assessed sleep-associated movement disorders and CVD, or that provided sufficient data to calculate risk estimates of CVD associated with sleep-associated movement disorders. Unpublished reports, abstracts, comments, reviews, case report or editorials were not considered in this review. CVD in our investigation were defined as CAD, heart failure (HF) and stroke, not including hypertension.

Data extraction

Two reviewers independently extracted eligible data by screening the titles and abstracts of the search results and evaluating the remaining full-text articles. Disagreements were discussed till consensus was achieved. The following data were extracted: the first authors' name, publication year, country where the study was conducted, study type, RLS or PLMS, number of samples, crude or adjusted risk estimates and 95%CIs. Different study types were divided into prospective cohort, case-control, and cross-sectional studies.

Statistical analysis

Summary odds ratios (ORs) and 95%CIs were used to measure the association strength between sleep-associated movement disorders and CVD risk. Cochran's Q statistic and the I^2 statistic were used to quantify between-study heterogeneity. The heterogeneity was considered as significant with a conservative *P* value of 0.10 and a value of I^2 exceeding 56%. We pooled ORs, relative risks and hazard ratios (HRs) with the random-effects model when a significant heterogeneity exists, otherwise, with the fixed-effect model[7]. We also performed subgroup analyses to explore the underlying confounding factor. Sensitivity analyses were carried out to test the reliability of results. We checked for funnel

Table 1 Characteristics of the eligible studies included in the meta-analysis

Ref.	Year	Country	Study type	Total	Source of patients	CVD-OR (95%CI)	CAD-OR (95%CI)
Hanly *et al*[10]	1996	Canada	Cohort	32	PLMS	8.73 (0.94-81.49)	-
Ulfberg *et al*[11]	2001	Sweden	Case-control	4000	RLS	2.50 (1.40-4.30)	-
Ohayon *et al*[12]	2002	5 European countries	Cross-sectional	18980	PLMS/RLS	1.47 (1.12-1.81)	-
Winkelman *et al*[13]	2006	United States	Cohort	2821	RLS	2.07 (1.31-3.27)	
Elwood *et al*[14]	2006	United Kingdom	Cohort	1871	RLS	1.38 (1.06-1.81)	1.24 (0.89-1.74)
Winkelman *et al*[15]	2008	United States	Cross-sectional	3433	RLS	2.07 (1.43-3.00)	2.05 (1.38-3.04)
Walters *et al*[16]	2010	United States	Cohort	267	RLS	2.46 (0.97-6.28)	-
Koo *et al*[17]	2011	United States	Cohort	2911	PLMS	1.28 (1.08-1.51)	1.23 (1.01-1.50)
Li *et al*[18]	2012	United States	Cohort	70977	RLS	1.46 (0.97-2.18)	1.46 (0.97-2.18)
Winter *et al*[19]	2012	United States	Cohort	48938	RLS	1.06 (0.90-1.26)	-
Lindner *et al*[20]	2012	Hungary	Cohort	150	PLMS	1.85 (0.46-7.51)	1.15 (0.35-3.81)
Mirza *et al*[21]	2013	United States	Case-control	584	PLMS	1.62 (1.14-2.30)	-
Szentkirályi *et al*[22]	2013	German	Cohort	4308	RLS	0.94 (0.42-2.10)	0.53 (0.12-2.27)

RLS: Restless legs syndrome; PLMS: Periodic limb movements during sleep; CVD: Cardiovascular disease; CAD: Coronary artery disease.

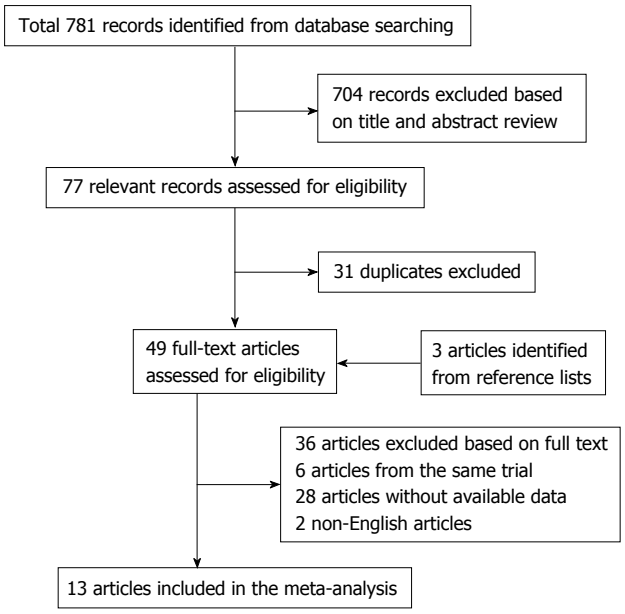


Figure 1 Flow diagram of the study selection process.

plot asymmetry, Begg's test and Egger's test to assess potential publication bias, and the significant *P* value was < 0.05[8,9]. The "trim and fill" procedure was utilized to further evaluate the possible effect of publication bias in the present meta-analysis[7]. All analyses were calculated with Stata version 12.0 (Stata Corp, College Station, Texas).

RESULTS

Characteristics of eligible studies

A total of 781 unique citations were identified: 279 from PubMed, 283 from Medline and 219 from Web of Science. The flow of study identification was shown in Figure 1[10-22]. Table 1 shows characteristics of eligible studies and the effect of sleep-associated movement disorders on the risk for CVD and CAD. Among these studies, nine are cohort studies; two are case-control studies; and two are cross-sectional studies. All participants were investigated from either European countries or United States. The sample sources of cases in nine studies were RLS patients and five were PLMS patients, including one study investigating both PLS and PLMS patients. The risk estimates and 95%CIs of most studies were extracted directly from original articles except for those of seven studies were recalculated by merging raw data[12-14,17,19,20,22].

Associations of sleep-associated movement disorders with CVD and CAD

Several studies indicated that sleep-associated movement disorders were associated with a significant increased risk for CVD; while others showed inconsistent findings (Figure 2). In a random-effects model, the summary OR of CVD associated with sleep-associated movement was 1.51 (95%CI: 1.29-1.77), with the evidence of heterogeneity (*P* for heterogeneity = 0.005, I^2 = 57.6%) (Figure 2). In subgroup analysis by study type, the summary OR was 1.36 for nine cohort studies (95%CI: 1.14-1.62; *P* for heterogeneity = 0.055; I^2 = 47.5%) (Figure 2). Figure 3 listed that a significant association was also found between RLS and CVD (OR = 1.54, 95%CI: 1.24-1.92). In a fixed-effects model, we determined a significant association of sleep-associated movement disorders with CAD (OR = 1.34, 95%CI: 1.16-1.54; *P* for heterogeneity = 0.210; I^2 = 30.0%) (Figure 4).

Sensitive analysis and publication bias evaluation

Sensitive analysis was performed by sequentially excluding each study to test the stability of the results in the present meta-analysis. After removing a study performed by Winter *et al*[19] which allowed the assessment of incident CVD cases, we found no significantly heterogeneity existed between overall studies (*P* = 0.112, I^2 = 34.8%). In addition, there was no significantly influence on the pooled OR of the CVD risk (OR = 1.49, 95%CI: 1.35-1.64). Therefore, the different study design may be a possible origin of heterogeneity. Then we conducted the funnel plot and

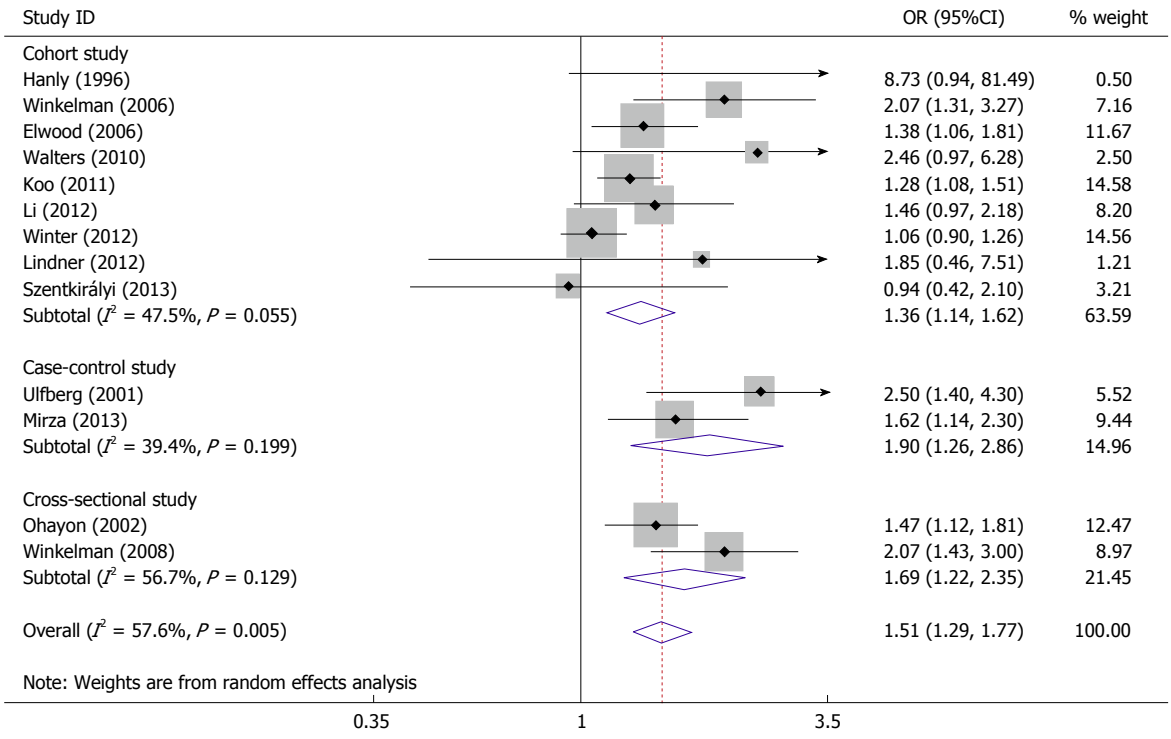


Figure 2 Forest plot (random effects model) of overall cardiovascular disease risk associated with sleep-associated movement disorders.

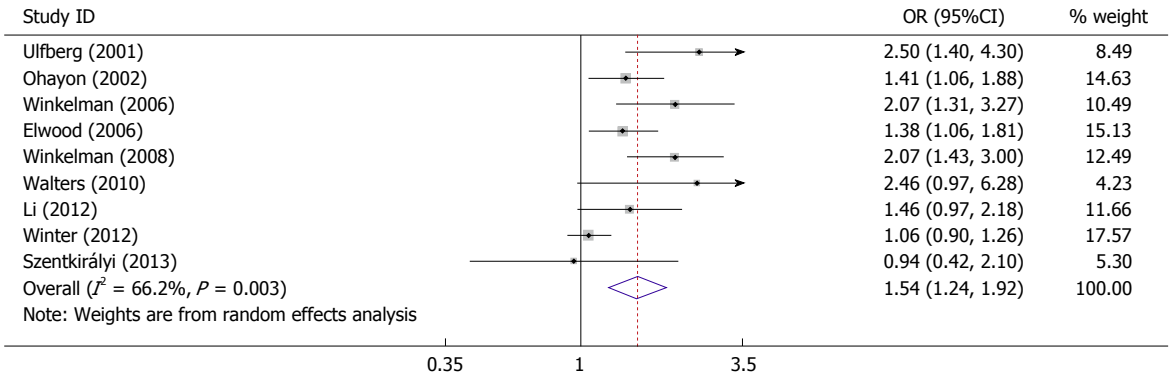


Figure 3 Forest plot (fixed effects model) of overall cardiovascular disease risk associated with restless legs syndrome.

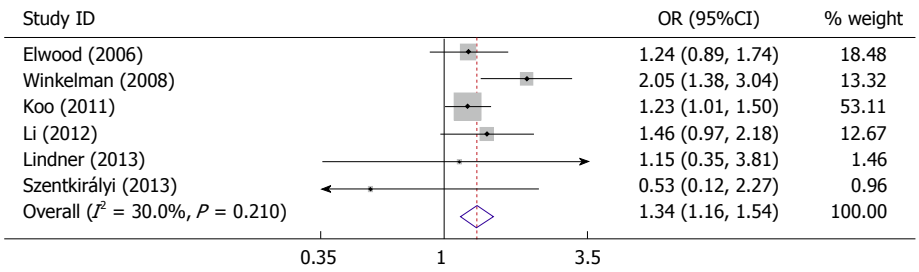


Figure 4 Forest plot (random effects model) of overall coronary artery disease risk associated with sleep-associated movement disorders.

Egger's test to assess the publication bias of literatures. Visual assessment of the Begg funnel plot revealed asymmetry (Figure 5A). This indicates the potential publication bias, although the Begg's test showed no statistically significance (Z = 1.53, P = 0.127). In order to identify and correct for funnel plot asymmetry arising from publication bias, we continued the analysis using the trim and fill method. The other four hypothetical studies were filled to produce a symmetrical funnel plot (Figure 5B). After that, the meta-analysis still showed a statistically significant association between sleep-associated movement disorders and CVD (OR = 1.39,

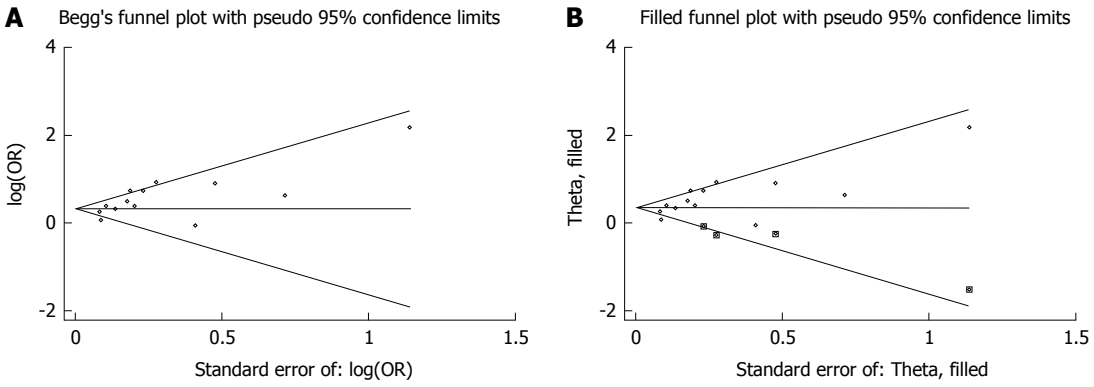


Figure 5 Funnel plots without and with trim and fill. A: Funnel plot without trim and fill; B: Funnel plot with trim and fill.

95%CI: 1.19-1.63).

DISCUSSION

Plenty of evidences have revealed screening, identification, and treatment of sleep disorders were important among patients with CVD. Several studies showed RLS were associated with hypertension and heart disease, because RLS may contribute to a high cardiovascular burden[1,11,12]. In 2001, Ulfberg *et al*[11] found an association of RLS with both self-reported hypertension and heart problems in 4000 Swedish men aged 18 to 64 years (hypertension: OR = 1.5, 95%CI: 0.9-2.4; heart problems: OR = 2.5, 95%CI: 1.4-4.3). Ohayon *et al*[12] reported heart disease made a significant independent contribution to RLS (OR = 1.41, 95%CI: 1.06-1.88). In a cohort study, Elwood *et al*[14] identified RLS is associated with a significant increase in ischaemic heart disease events among 1871 men in South Wales, United Kingdom during the following 10 years (OR = 1.24, 95%CI: 0.89-1.71). In the Wisconsin Sleep Cohort study of 2006, Winkelman *et al*[15] observed a dose-related association between RLS symptoms and CVD (Frequent: OR = 1.61, 95%CI: 0.82-3.13; Daily: OR = 2.58, 95%CI: 1.38-4.84). Moreover, Winkelman *et al*[15] also demonstrated the association of RLS with CVD and CAD in a large cross-sectional observational community-based study of 1559 men and 1874 women (CAD: OR = 2.05, 95%CI = 1.38-3.04; CVD: OR = 2.07, 95%CI: 1.43-3.00) for subjects with RLS compared to those without RLS, and the associations were stronger in those with RLS more frequent or severer symptoms[15]. Li *et al*[18] performed a large-scale prospective study to examine whether RLS was associated with an increased risk of CAD in women of the Nurses' Health Study (HR = 1.46, 95%CI: 0.97-2.18). The fact suggests that CVD could be result from the long-term impact of RLS or RLS-associated conditions. Nevertheless, a study from Walters *et al*[16] showed that there was no statistically difference in the prevalence of CVDs or risk factors between RLS patients and controls, which may be caused by the limited sample size. Another two large prospective cohort studies (Women's Health Study and Physicians' Health Study, United States) also did not support that RLS is a marker of increased risk of vascular disease. The discrepancy between these two results and those of previous studies may be explained by the prospective cohort study, which was designed to assess incident CVD cases[19].

Ninety-nine percent of PLMS are related to greater heart rate response, which result in sympathetic activation as a cause of cardiovascular complications[1,23,24]. In 1996, Hanly *et al*[10] for first time found an association between congestive HF and increased prevalence of PLMS. Furthermore, a cross-sectional study was performed in the five European countries, identifying CVD certainly associated with PLMS (OR = 1.61, 95%CI: 1.09-2.39). A study published in 2011 from Koo *et al*[17] supported PLMS frequency may be a predictive factor of incident CVD. In a recent study by Mirza *et al*[23], periodic limb movement index > 35/h were found to confer a high risk for HF (OR = 1.62; 95%CI: 1.14-2.30).

To clarify the controversial results of previous studies regarding the association of sleep-associated movement disorders with CVD, we performed this meta-analysis. Our analysis suggested that sleep-associated movement might play an important role in the development of heart disease, particular in prevalence of CAD. As different study design of the previous works might contribute to discrepancies between previous reports, thus we conducted subgroup analysis by study types which suggested the association was only to be weaker but still significant in cohort studies. In addition, our results also provided a stronger evidence for the significant relationship between RLS and CVD. However, the exact mechanism of the effect of sleep-associated movement disorders on cardiovascular system remains unclear. The most accepted hypothesis is these disorders may result from sustained adrenergic surges caused by sympathetic nervous system activation, which predispose to persistent elevated blood pressure as well as increased left ventricular afterload and heart rate. Another possible explanation is that sleep-associated movement disorders interrupt sleep which raises heart risk[25].

Some limitations of our meta-analysis should be considered. First, the results of the present meta-analysis remain cautious due to heterogeneity across

studies. Second, the risk estimate of each study included was not adjusted by the same covariable related to risk of CVD. Third, the asymmetry shape of the funnel plot suggested the possibility of publication bias, even the trim and fill sensitivity analysis has been used to test the stability of the results. Fourth, all sample sources are of European or United States descent, which lead to lacking data from other ethnicity backgrounds.

In conclusion, the current meta-analysis suggests that sleep-associated movement disorders are associated with prevalence of CVD, which may be predictive of CVD. This finding may settle the controversy among previous investigations. However, further well-designed and mechanistic work should undertake to confirm this association.

COMMENTS

Background

The burden of cardiovascular disease (CVD) is increasing globally, especially in developing countries such as China. CVD has been the first leading cause of mortality in China. It has been known that unhealthy life style is the most common induced factor of CVD which can also lead to other disease like diabetes, obesity and so on. Therefore, the prevention of CVD, which consumes less, is more important than treatment in developing countries. Sleep-associated movement disorders is a group of symptoms that easily been ignore by the public and some limited studies seem to indicate they may be also the underlying cause of CVD, although this association is not been well established.

Research frontiers

Over the recent 2 decades, many studies attempted to understand the associations between sleep-associated movement disorders and CVD. However, it is difficult to obtain an inconsistent conclusion about the association from the previous studies.

Innovations and breakthroughs

From this meta-analysis, sleep-associated movement disorders may increase the risk of CVD by approximately 51%. Significant associations also showed in subgroup analyses of nine cohort studies. And sleep-associated movement disorders may be predictively used in the prevention of coronary artery disease in the future based the current investigation.

Applications

Sleep-associated movement disorders appear to be either directly or indirectly associated with the risk of CVD. An exploration of the mechanism for this association may help us decrease the prevalence of CVD.

Terminology

Sympathetic nervous system is a web of nerves and neurons spreading excitement to each organ of body. Left ventricular afterload is the encountering resistance when the myocardial of left ventricular contracts.

Peer-review

Well written and concise meta-analysis.

REFERENCES

1 **Schaffernocker T**, Ho J, Hayes D. Sleep-associated movement disorders and heart failure. *Heart Fail Rev* 2009; **14**: 165-170 [PMID: 19051011 DOI: 10.1007/s10741-008-9118-6]

2 **Montplaisir J**, Boucher S, Poirier G, Lavigne G, Lapierre O, Lespérance P. Clinical, polysomnographic, and genetic characteristics of restless legs syndrome: a study of 133 patients diagnosed with new standard criteria. *Mov Disord* 1997; **12**: 61-65 [PMID: 8990055 DOI: 10.1002/mds.870120111]

3 **Ancoli-Israel S**, Kripke DF, Klauber MR, Mason WJ, Fell R, Kaplan O. Periodic limb movements in sleep in community-dwelling elderly. *Sleep* 1991; **14**: 496-500 [PMID: 1798881]

4 **Ferini-Strambi L**, Walters AS, Sica D. The relationship among restless legs syndrome (Willis-Ekbom Disease), hypertension, cardiovascular disease, and cerebrovascular disease. *J Neurol* 2014; **261**: 1051-1068 [PMID: 23963470 DOI: 10.1007/s00415-013-7065-1]

5 **Walters AS**, Rye DB. Review of the relationship of restless legs syndrome and periodic limb movements in sleep to hypertension, heart disease, and stroke. *Sleep* 2009; **32**: 589-597 [PMID: 19480225]

6 **Stroup DF**, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008-2012 [PMID: 10789670 DOI: 10.1001/jama.283.15.2008]

7 **Higgins JP**, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539-1558 [PMID: 12111919 DOI: 10.1002/sim.1186]

8 **Egger M**, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-634 [PMID: 9310563 DOI: 10.1136/bmj.315.7109.629]

9 **Begg CB**, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994; **50**: 1088-1101 [PMID: 7786990]

10 **Hanly PJ**, Zuberi-Khokhar N. Periodic limb movements during sleep in patients with congestive heart failure. *Chest* 1996; **109**: 1497-1502 [PMID: 8769500 DOI: 10.1378/chest.109.6.1497]

11 **Ulfberg J**, Nyström B, Carter N, Edling C. Prevalence of restless legs syndrome among men aged 18 to 64 years: an association with somatic disease and neuropsychiatric symptoms. *Mov Disord* 2001; **16**: 1159-1163 [PMID: 11748753 DOI: 10.1002/mds.1209]

12 **Ohayon MM**, Roth T. Prevalence of restless legs syndrome and periodic limb movement disorder in the general population. *J Psychosom Res* 2002; **53**: 547-554 [PMID: 12127170 DOI: 10.1016/S0022-3999(02)00443-9]

13 **Winkelman JW**, Finn L, Young T. Prevalence and correlates of restless legs syndrome symptoms in the Wisconsin Sleep Cohort. *Sleep Med* 2006; **7**: 545-552 [PMID: 16740407 DOI: 10.1016/j.sleep.2006.01.004]

14 **Elwood P**, Hack M, Pickering J, Hughes J, Gallacher J. Sleep disturbance, stroke, and heart disease events: evidence from the Caerphilly cohort. *J Epidemiol Community Health* 2006; **60**: 69-73 [PMID: 16361457 DOI: 10.1136/jech.2005.039057]

15 **Winkelman JW**, Shahar E, Sharief I, Gottlieb DJ. Association of restless legs syndrome and cardiovascular disease in the Sleep Heart Health Study. *Neurology* 2008; **70**: 35-42 [PMID: 18166705 DOI: 10.1212/01.wnl.0000287072.93277.c9]

16 **Walters AS**, Moussouttas M, Siddiqui F, Silveira DC, Fuentes K, Wang L, Berger K. Prevalence of stroke in Restless Legs Syndrome: Initial Results Point to the Need for More Sophisticated Studies. *Open Neurol J* 2010; **4**: 73-77 [PMID: 20721325 DOI: 10.2174/1874205X01004010073]

17 **Koo BB**, Blackwell T, Ancoli-Israel S, Stone KL, Stefanick ML, Redline S. Association of incident cardiovascular disease with periodic limb movements during sleep in older men: outcomes of sleep disorders in older men (MrOS) study. *Circulation* 2011; **124**: 1223-1231 [PMID: 21859975 DOI: 10.1161/CIRCULATIONAHA.111.038968]

18 **Li Y**, Walters AS, Chiuve SE, Rimm EB, Winkelman JW, Gao X. Prospective study of restless legs syndrome and coronary heart disease among women. *Circulation* 2012; **126**: 1689-1694 [PMID: 22967852 DOI: 10.1161/CIRCULATIONAHA.112.112698]

19 **Winter AC**, Schürks M, Glynn RJ, Buring JE, Gaziano JM, Berger K, Kurth T. Restless legs syndrome and risk of incident cardiovascular disease in women and men: prospective cohort study. *BMJ Open* 2012; **2**: e000866 [PMID: 22447047 DOI: 10.1136/bmjopen-2012-000866]

20 **Lindner A**, Fornadi K, Lazar AS, Czira ME, Dunai A, Zoller R, Veber O, Szentkiralyi A, Kiss Z, Toronyi E, Mucsi I, Novak M, Molnar MZ. Periodic limb movements in sleep are associated with stroke and cardiovascular risk factors in patients with renal failure. *J Sleep Res* 2012; **21**: 297-307 [PMID: 21917047 DOI: 10.1111/

j.1365-2869.2011.00956]

21 **Mirza M**, Shen WK, Sofi A, Jahangir A, Mori N, Tajik AJ, Jahangir A. Frequent periodic leg movement during sleep is associated with left ventricular hypertrophy and adverse cardiovascular outcomes. *J Am Soc Echocardiogr* 2013; **26**: 783-790 [PMID: 23622883 DOI: 10.1016/j.echo.2013.03.018]

22 **Szentkirályi A**, Völzke H, Hoffmann W, Happe S, Berger K. A time sequence analysis of the relationship between cardiovascular risk factors, vascular diseases and restless legs syndrome in the general population. *J Sleep Res* 2013; **22**: 434-442 [PMID: 23374090 DOI: 10.1111/jsr.12040]

23 **Sforza E**, Nicolas A, Lavigne G, Gosselin A, Petit D, Montplaisir J. EEG and cardiac activation during periodic leg movements in sleep: support for a hierarchy of arousal responses. *Neurology* 1999; **52**: 786-791 [PMID: 10078729]

24 **Yang CK**, Jordan AS, White DP, Winkelman JW. Heart rate response to respiratory events with or without leg movements. *Sleep* 2006; **29**: 553-556 [PMID: 16676789]

25 **Nannapaneni S**, Ramar K. Periodic limb movements during sleep and their effect on the cardiovascular system: is there a final answer? *Sleep Med* 2014; **15**: 379-384 [PMID: 24656911 DOI: 10.1016/j.sleep.2013.12.014]

P- Reviewer: Lazzeri C, Lin GM **S- Editor**: Ji FF
L- Editor: A **E- Editor**: Liu SQ

Published by **Baishideng Publishing Group Inc**
8226 Regency Drive, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

ISSN 2308-3840 (online)

World Journal of *Meta-Analysis*

World J Meta-Anal 2015 August 26; 3(4): 188-205

Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

A peer-reviewed, online, open-access journal of meta-analysis

Editorial Board

2013-2018

The *World Journal of Meta-Analysis* Editorial Board consists of 400 members, representing a team of worldwide experts in clinical meta-analysis research. They are from 41 countries, including Argentina (2), Australia (3), Austria (1), Belgium (5), Brazil (11), Canada (16), Chile (3), China (121), Croatia (1), Egypt (1), Finland (4), France (2), Germany (9), Greece (11), Hungary (2), India (12), Iran (2), Ireland (1), Israel (2), Italy (41), Japan (5), Lithuania (1), Malaysia (1), Netherlands (8), New Zealand (1), Norway (1), Peru (1), Poland (4), Portugal (6), Romania (1), Saudi Arabia (4), Singapore (3), South Africa (1), South Korea (7), Spain (8), Sri Lanka (2), Sudan (1), Switzerland (2), Thailand (4), Turkey (3), United Kingdom (22), and United States (64).

Alessandro Cucchetti, *Bologna*
Gianfranco Damiani, *Rome*
Fabrizio D'Ascenzo, *Turin*
Massimo Del Fabbro, *Milano*
Valeria Fadda, *Florence*
Alessandro Fancellu, *Sassari*
Giuseppe Ferrante, *Milan*
Virginia Festa, *Rome*
Francesco Fiorica, *Ferrara*
Guglielmo Giraldi, *Rome*
Jenny Guidi, *Bologna*
Lorenzo Loffredo, *Rome*
Andrea Messori, *Firenze*
Eliano Pio Navarese, *Bydgoszcz*
Stefano Omboni, *Solbiate Arno*
Alvisa Palese, *Udine*
Stefano Palomba, *Reggio Emilia Stefano*
Carlo Perricone, *Rome*
Mario Petretta, *Naples*
Alessandro Pezzini, *Brescia*
Gianluca Pontone, *Milan*
Paolo Emilio Puddu, *Rome*
Andrea Rognoni, *Novara*
Giuseppe Scalabrino, *Milan*
Città della Salute e della Scienza, *Turin*
Fabrizio Sgolastra, *L'Aquila*
Maria Lucia Specchia, *Rome*
Stefano Trastulli, *Terni*
Fabio Tine, *Palermo*
Nereo Vettoretto, *Chiari*
Alberto Vianello, *Perugia*
Luigi Zorcolo, *Cagliari*

Japan

Nguyen Tien Huy, *Nagasaki*
Hiroharu Kamioka, *Tokyo*
Koji Kawakami, *Kyoto*
Keitaro Matsuo, *Nagoya*
Kazushi Okamoto, *Nagoya*

Lithuania

Edmundas Kadusevicius, *Kaunas*

Malaysia

SP Pani, *Ipoh*

Netherlands

Michel van den Bekerom, *Amsterdam*
Dimitra Dodou, *Delft*
D Haverkamp, *Amsterdam*
Vassilios Koussoulas, *Drachten*
BJ Polder, *Nijmegen*
Theo Stijnen, *Leiden*
RNM Weijers, *Amsterdam*
Joost de Winter, *Delft*

New Zealand

Shaofeng Li, *Auckland*

Norway

Eivind Berge, *Oslo*

Peru

Rafael Bolaños Díaz, *Lima*

Poland

Maciej Banach, *Lodz*
Krzysztof Jonderko, *Sosnowiec*
Jolanta Lissowska, *Warsaw*
Maciej Plaszewski, *Warsaw*

Portugal

Daniel Caldeira, *Lisboa*
J Costa, *Lisbon*
Ana Miguel, *Coimbra*
Manuel Morgado, *Covilh*
Bárbara Peleteiro, *Porto*
Rui Torres, *Paredes*

Romania

Fratila Ovidiu, *Oradea*

Saudi Arabia

Hazem M Al-Mandeel, *Riyadh*
Ezzeldin M Ibrahim, *Jeddah*
Mutahir A Tunio, *Riyadh*
Hayfaa A Wahabi, *Riyadh*

Singapore

Nikos LD Chatzisarantis, *Singapore*
Edwin Choon Wyn Lim, *Singapore*
Roger Ho, *Singapore*

South Africa

Alaine Umubyeyi Nyaruhirira, *Pretoria*

South Korea

Jung-Hee Kim, *Cheonan*
Hyangsook Lee, *Seoul*
Myeong Soo Lee, *Daejeon*
Chi-Un Pae, *Bucheon*
Yong Hyun Park, *Yangsan*
Jae Hong Seo, *Seoul*
Yong Sang Song, *Seoul*

Spain

Pablo Avanzas, *Oviedo*
Joan Cid, *Barcelona*
Joaquin de Haro, *Madrid*
Joan Guardia-Olmos, *Barcelona*
Nabil Halaihlel, *Zaragoza*
JA Monge-Argilés, *Alicante*
Raul Moreno, *Madrid*
Inés Velasco, *Aracena*

Sri Lanka

Ranil Jayawardena, *Colombo*
Priyanga Ranasinghe, *Colombo*

Switzerland

Jay P Singh, *Zurich*
Giorgio Treglia, *Bellinzona*

Sudan

Samir MH Shaheen, *Khartoum*

Thailand

Chuenjid Kongkaew, *Phitsanulok*
Manop Pithukpakorn, *Bangkok*
Piyamitr Sritara, *Bangkok*
Surasak Saokaew, *Phayao*

Turkey

Nese Demirturk, *Afyonkarahisar*
Nilüfer Ozabaci, *Eskisehir*
Ilke Sipahi, *Istanbul*

United Kingdom

Omar M Aboumarzouk, *Wales*
Abeer Al-Namankany, *London*
Ernest A Azzopardi, *Cardif*
Umberto Benedetto, *Cambridge*
Joanne Brooke, *London*
Noriko Cable, *London*
David Chan, *Cardiff*
YC Cheong, *Southampton*
Andrew Currie, *Harrow*
G Nabi, *Dundee*

Lesley A Anderson, *Belfast*
Valentina Gallo, *London*
Gianpiero Gravante, *Leicester*
Peter N Lee, *Surrey*
Igho Onakpoya, *Oxford*
Ashish Pradhan, *Huntingdon*
Evridiki Patelarou, *London*
Jian-Qing Shi, *Newcastle*
Surendra P Singh, *Wolverhampton*
Natalie Taylor, *Leeds*
Yousef Shahin, *Hull*
Zheng Ye, *Cambridge*

United States

L Joseph Su, *Rockville*
Olusola Adesope, *Pullman*
Mike Allen, *Milwaukee*
Bhupinder Anand, *Bellaire*
Stephen Aronoff, *Philadelphia*
KoKo Aung, *San Antonio*
William L Baker, *Farmington*
Moritz C Wyler von Ballmoos, *Milwaukee*
Matthew L Bechtold, *Columbia*
Atul Bhardwaj, *Hershey*
Somjot S Brar, *Los Angeles*
Hui Cai, *Nashville*
Subhash Chandra, *Towson*
Wen-Pin Chang, *Omaha*
Yong Chen, *North Wales*
Myunghan Choi, *Phoenix*
John Coverdale, *Houston*
Prakash C Deedwania, *Fresno*
Eugene Demidenko, *Hanover*
Hong-Wen Deng, *New Orleans*
Eric M Deshaies, *Syracuse*
Tao Fan, *Whitehouse Station*
Shinga Feresu, *Bloomington*
Janvier Gasana, *Miami*
Kaveh Hajifathalian, *Boston*
Mohammad Obaidul Hoque, *Baltimore*
Larissa R Brunner Huber, *Charlotte*
Imran H Iftikhar, *Columbia*
Vijayvel Jayaprakash, *Buffalo*
Xuezhi Jiang, *Weat Reading*
Shuo Jiao, *Seattle*
Evelyn Johnson, *Boise*
Le Kang, *Silver Spring*
SR Kapadia, *Cleveland*
Lior Katz, *Houston*
Daniel M Laskin, *Richmond*
Yu Liang, *Foster City*
Paul Ellis Marik, *Norfolk*
Lynne V McFarland, *Seattle*
Marcovalerio Melis, *New York*
Brian J Miller, *Augusta*
Pavlos Msaouel, *New York*
Joshua E Muscat, *Hershey*
Chee Yuan Ng, *Loma Linda*
Nghi C Nguyen, *Saint Louis*
Brandi S Niemeier, *Whitewater*
Thomas D Parsons, *Denton*
Nidal Abi Rafeh, *New Orleans*
Praveen Roy, *Marshfield*
Ali Salavati, *Philadelphia*
Ankur Sethi, *Chicago*
Tatyana A Shamliyan, *Minneapoli*
Qian Shi, *Rochester*
Zhongjie Shi, *Philadelphia*
Param Puneet Singh, *Chicago*
KV Slavin, *Chicago*
Ali El Solh, *Buffalo*
Jielin Sun, *Winston-Salem*
Richard G Trohman, *Chicago*
Laurah Turner, *Cincinnati*
Sheila Wilhelm, *Detroit*
Alex K Wong, *Los Angeles*
Xiaohui Xu, *Gainesville*
Lu Yin, *Nashville*

WJMA World Journal of Meta-Analysis

Contents

Bimonthly Volume 3 Number 4 August 26, 2015

EDITORIAL

SYSTEMATIC REVIEWS

World Journal of Meta-Analysis
Volume 3 Number 4 August 26, 2015

Contents

NAME OF JOURNAL
World Journal of Meta-Analysis

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Bimonthly

EDITOR-IN-CHIEF
Giuseppe Biondi-Zoccai, MD, Assistant Professor, Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
World Journal of Meta-Analysis
Room 903, Building D, Ocean International Center, No. 62 Dongsihuan Zhonglu, Chaoyang District, Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
8226 Regency Drive,
Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
August 26, 2015

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i4.188

World J Meta-Anal 2015 August 26; 3(4): 188-192
ISSN 2308-3840 (online)

EDITORIAL

Application of meta-analysis to specific research fields: Lessons learned

Lynne V McFarland

Lynne V McFarland, Department of Medicinal Chemistry, University of Washington, Seattle, WA 98108-1597, United States

Lynne V McFarland, VA Puget Sound Healthcare System, Seattle, WA 98108-1597, United States

Author contributions: McFarland LV contributed solely to this work.

Conflict-of-interest statement: The author has no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work, not other relationships or activities that could appear to have influenced the submitted work.

Correspondence to: Lynne V McFarland, PhD, VA Puget Sound Healthcare System, 1660 S. Columbian Way, Seattle, WA 98108-1597, United States. lynne.mcfarland@va.gov
Telephone: +1-206-2771780
Fax: +1-206-7642935

Received: May 5, 2015
Peer-review started: May 7, 2015
First decision: June 3, 2015
Revised: June 15, 2015
Accepted: July 7, 2015
Article in press: July 9, 2015
Published online: August 26, 2015

Abstract

Scientific research is challenged to translate findings from multiple, often conflicting, clinical trials into a simple answer of whether a treatment works or not. The public and healthcare providers alike frequently voice their frustrations when the media reports a treatment working on one day, but seemingly the next day reports a study refuting the previous one. Meta-analyses are being used more commonly by researchers to convey an understandable summary of scientific studies to the general public and healthcare providers. As time goes by, we have learned how to improve meta-analytic techniques to reflect more valid results and when it is appropriate to pool or not to pool results from different studies. Retrospective reviews often don't acknowledge this learning curve and may fail to recommend the most current valid guidelines. This editorial presents an example of how the current use of meta-analysis has shifted in one field (the therapeutic effects of probiotics) and recommendations on how to correctly interpret the results of such an analysis.

Key words: Meta-analysis; Study designs; Probiotics; Sensitivity analysis; Meta-regression

Core tip: As meta-analyses are used more frequently and their findings reach a wider scope of people, it is the responsibility of researchers to use current guidelines and appropriately apply their findings to form valid conclusions. As researchers gain experience with this technique, we need to recognize that our methods may change over time. Meta-analysis remains a valuable tool for examining controversies arising from conflicting studies.

McFarland LV. Application of meta-analysis to specific research fields: Lessons learned. *World J Meta-Anal* 2015; 3(4): 188-192 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i4/188.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i4.188

INTRODUCTION

The lag-time between a scientific discovery and when it is translated into changes in public behavior and beliefs can be protracted. We have never been so challenged in our ability to sort through the onrush of volumes of scientific data to provide timely clinical advice to the public and healthcare providers. The public and healthcare providers face the challenge of obtaining accurate clinical recommendations when the amount of information available is vast and often discordant. In addition, as meta-analyses are used more often, issues in proper application of methods and interpretation of the results need to be re-visited, for example, when it is appropriate to pool studies.

Public awareness and belief in scientific study results

We researchers work hard to present new findings for clinical therapies, but how much does the general population really understand? Patients are often confused by information overload and the ambiguity of clinical results, while at the same time suffer from low health literacy[1]. Health literacy measures the ability of the public to understand healthcare providers, as well as reading and comprehending the vast amount of health information. What is alarming is that only 25%-50% of the American and European general populations are scientifically literate, making it difficult for them to decipher complex scientific articles on new medical treatments[2,3]. A global survey across 57 nations found 56% have "hardly any or only some confidence" in the scientific community[4]. Improving public confidence and awareness of scientific findings is paramount. Science is not a static world. Shifting through studies analyzing similar treatments for the same disease can be frustrating, even for scientists, as newer studies may refute established beliefs and the findings of other studies that seem similar in study design and enrolled participant populations.

Systematic reviews were first developed to provide the public and healthcare providers with a simple summary from multiple studies. The value of systematic reviews is that they allow a detailed examination of many studies, in which researchers try to explain the heterogeneity of the study outcomes, which may to be to a variety of sources ranging from the type of study population, different study designs, differences in study quality and differences in dose and duration of the investigational intervention. But to the disappointment of readers and clinicians, systematic reviews often do not provide a simple answer to the question: "Does this treatment work?"

Meta-analysis was developed to allow combining results from different studies in order to obtain a single pooled estimate of effect or risk from various exposures (behaviors, medications, or interventions) associated with a specific outcome. Meta-analytic analysis involves discerning where heterogeneity exists, identifying the sources from which it may spring and modeling the data to obtain a single pooled estimate of effect. The advantages of meta-analysis include: achieving a higher statistical power due to the greater numbers involved in the pooled population, ability to examine sources of variability over multiple studies while adjusting for common measures shared by studies, and providing one pooled outcome across a variety of studies.

If meta-analysis can provide an answer to "Does this treatment work?", why is there still confusion in the literature? The problem relates to limitations inherent in the meta-analytic methods and the often too rapid acceptance of conclusions reached by the researchers. If a meta-analysis has pooled dissimilar types of treatments together and the pooled risk estimate shows significant efficacy, people often jump to the conclusion that any of the treatments included in the meta-analysis must be effective. This isn't necessarily so.

Meta-analyses can be misleading if different types of treatments or different types of outcomes are grouped together. The pooled outcomes are also highly dependent upon how the search of the literature is conducted: including or excluding meeting abstracts, non-English trials, years covered, inclusion/exclusion criteria narrow or wide, how incomplete data is handled (exclude or attempt to contact authors of original papers) and different forms of bias (publication bias, study quality, study size). There is a distinct learning curve as both the field of meta-analysis matures and researchers using this technique gain experience.

LEARNING CURVE

History

Meta-analysis developed in the field of astronomy in the 1700s, with the application to the clinical field in 1907[5]. The use of meta-analysis was largely ignored by the clinical field until the 1980s, when the number of publications increased exponentially (Figure 1). Expert panels were convened to reach a consensus on how meta-analysis should be conducted and reported. The guidelines published as PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analysis, were developed in 1999 and are continually updated, yet the translation of meta-analytic results continues to challenge the general public and scientists alike[6-8].

Meta-analysis is used to assess a wide array of interventions and diseases ranging from the phases of the moon and lunacy[9] to the effects of wine[10,11] and chocolate on cardiovascular disease[12] and the effectiveness of probiotics for many diseases [for example, antibiotic associated diarrhea (AAD), allergies, acute pediatric diarrhea, *etc.*][13-15].

To demonstrate how we have learned to properly perform and interpret meta-analytic results, let's focus on the field of probiotics for the prevention and treatment of various diseases.

Examples of inappropriate pooling

In the early phase of scientific enquiry for a new

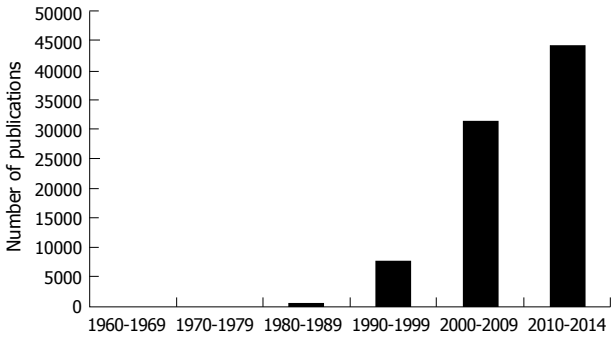


Figure 1 Number of publications from PubMed (January 1960-October 2014) citing Meta-analysis by decade. A total of 83786 citations.

treatment, there is a tendency to combine different types of interventions or different outcomes to generate a pooled estimate of risk, especially when there is a paucity of available clinical trials published in the literature. This was seen in the field of probiotic randomized clinical trials (RCT).

Some of the earliest meta-analyses of probiotics combined different probiotic strains and species, and reported one pooled estimate of effect. An early meta-analysis done in 2002 assessing probiotics for the prevention of AAD reported a significant protective effect (pooled OR = 0.39, 95%CI: 0.25-0.62) for four RCT of "yeast" probiotics[16]. Fortunately, the only available trials used the same type of yeast (*Saccharomyces boulardii*). However in this same meta-analysis, D'Souza *et al*[16] also concluded that all "non-yeast" probiotics were effective for AAD, based on a pooled OR from five RCT, which combined the results from four different types of probiotics (*Enterococcus faecium*, *Lactobacillus rhamnosus* GG and three trials with two different mixtures of probiotics). Subsequent meta-analyses did not support the efficacy for three of the types of probiotics and only *L. rhamnosus* GG was found to be effective for AAD in later studies. Another example is from a meta-analysis by Van Niel *et al*[17], who reported a single pooled relative risk (RR) estimate for three different strains and a mixture of two strains of Lactobacilli for the treatment of pediatric diarrhea, and concluded that "Lactobacilli was effective for treating pediatric diarrhea". Again, when these different strains were subsequently analyzed separately, some were effective, but some were not. This meta-analysis did not analyze the different species separately. Another meta-analysis by Deshpande *et al*[18] done in 2010 including 10 different types, either a single strain or mixtures, of probiotics concluded from their pooled RR that a "combination of Lactobacillus and at least one Bifidobacterium species" was effective for pediatric necrotizing enterocolitis. Unfortunately, there only three of the ten trials used a mixture of Lactobacilli and Bifidobacteria species, but none of the mixtures contained the same strains of bacteria. Even recent publications suffer from inappropriately concluding "any of the tested probiotic strains are effective" once as a significant pooled OR is found. King *et al*[19] pooled the results from 20 RCT using a variety of probiotics (eight RCT using five different Lactobacilli species and 12 trials with different mixtures of probiotic types) and concluded "probiotics reduce the duration of respiratory illness". They did observe significant heterogeneity in the analysis and, although they did sub-group analysis by degree of bias and type of patient population (adult *vs* pediatric), they failed to do a separate analysis by type of probiotic strain. Khalesi *et al*[20] did a meta-analysis with nine RCT using a variety of different probiotic formulations (four RCT used yogurts, two RCT used milk, one each used capsules, drink or cheese) with nine different probiotic species and concluded "consuming probiotics may improve blood pressure". Although the researchers did do separate analysis by single *vs* mixtures of probiotics, dose and duration of probiotic treatment and baseline blood pressure, they did not do a separate analysis by the type of probiotic strain. Perhaps the best example of inappropriate use of meta-analysis is seen in the review done by Hempel *et al*[21] published in JAMA in 2012. The authors concluded that "the pooled RR from 63 RCT indicated a statistically significant association of probiotic administration with reduction in AAD". However, eight studies were not on AAD, the pooled studies were a mix of treatment and prevention study designs, and the pooled probiotics included 13 different mixtures and 8 different species of single probiotics. The erroneous assumption that once a pooled estimate shows significant efficacy means that any of the treatments included in the analysis are independently effective continues to be a major source of misinterpretation using meta-analysis techniques.

Solutions for "to pool, or not to pool?"

We have learned over time that the efficacy of probiotics is species and strain specific. Thus, it became apparent that we need to separately analyze different probiotics and not lump them all together. It is now recommended to not pool studies using different probiotic strains and to either exclude non-identical interventions or use separate sub-group analyses for each group with significant heterogeneity or if the disease or intervention is known *a priori* to have different effects[22,23]. Although the use of random-effects models may help to statistically control for some of influences of heterogeneity due to study size or bias, other sources of heterogeneity need to be examined too.

Limit to one type of intervention: One of the earliest probiotic meta-analysis limited inclusion to six RCT using the same mixture of probiotics (*Enterococcus faecium* and *Strept. thermophilus*) and found a significant reduction in cholesterol in their pooled outcome measure[24]. As research into probiotics continues, more studies are focusing on including the same type of probiotic strain, allowing them to pool only those probiotics similar in genetic make-up, mechanism-of-action and strain efficacy.

Table 1 Recommendations for performing a valid meta-analysis

Begin with a systematic review of the literature, carefully reviewing original trial publications
Use a standardize data extraction form and independent reviewers to achieve reliable data for the analysis. Use a third reviewer to resolve conflicts
Use current guidelines for standardized reporting of meta-analysis results and incorporate a statistician into the analysis team
If the exposure/intervention/medication has been shown to have different effects within categories, narrow the groups until similar un-confounded groups are formed (for example, by strain of probiotics not by species or by genus, or by type of drug, not by class)
Use caution when using the overall pooled estimate of risk for concluding any type of exposure results in a similar risk of outcome

Sub-group analysis: Another strategy to examine heterogeneity is sub-group analysis. These sub-groups should be defined a priori in the protocol and several methods may be used, including exclusion sensitivity analysis, stratifying on sub-groups, or use of meta-regression[25,26]. As probiotic meta-analysis research progresses, researchers are separating the analysis by distinct subgroups, including species or strain of probiotic, dose and duration of intervention given, type of enrolled population (adult *vs* pediatric) or by study quality. Szajewska *et al*[27] did a meta-analysis of six RCT for the prevention of pediatric AAD, which showed an overall pooled protective RR (RR = 0.44, 95%CI: 0.25-0.77) and used sub-group analysis to show only one strain (*L. rhamnosus* GG) was independently effective. None of the other three types of probiotics had more than one RCT. The scarcity of published articles can limit the applicability of meta-analyses when sub-group analysis by probiotic species results in only one RCT per species. In these cases, we must wait until multiple trials are available before we can use meta-analysis to determine if a particular probiotic strain is effective or not.

Currently, most researchers are now incorporating current guidelines and performing sub-group analysis separately for each type of probiotic and reporting these pooled estimates separately[13,15]. The meta-analysis of Johnston *et al*[28] illustrates what happens when meta-analysis is used when the literature can provide only a limited number of trials for the disease being studied. He included six RCTs, but only two trials used the same type of probiotic (*L. rhamnosus* GG) and although he presented pooled RR for the other trials, this is inappropriate, as there was only one trial each for the other probiotic strains. When I look back at my 2006 meta-analysis of different probiotics for the prevention of AAD, I appropriately reported one pooled RR for six RCT using *S. boulardii* and another pooled RR for the six RCT using *L. rhamnosus* GG[13]. But if I were to report these same results today, I would not have combined the remaining single probiotics and the mixtures in a pooled estimate, rather I would have reported separate estimates of efficacy for each probiotic strain sub-group.

Another consideration is whether to pool the type of disease being treated. One meta-analysis limited inclusion to the same type of probiotic (*S. boulardii*), but included 27 RCT on a variety of diseases, ranging from AAD to traveler's diarrhea[29]. Each type of disease was assessed separately with a systematic review and a meta-analysis was appropriately done only for the disease indication that had sufficient numbers of trials (10 RCT for preventing AAD).

Inflammatory bowel disease encompasses ulcerative colitis, Crohn's disease and pouchitis, which have different mechanisms and treatment outcomes. In a recent meta-analysis of 23 RCT for IBD, researchers correctly separated the three different disease profiles and additionally did sub-group analyses by the type of probiotic[14]. Only one type of probiotic mixture (VSL#3) was effective for treating both ulcerative colitis and pouchitis, but not for Crohn's disease.

This learning curve for meta-analysis is not limited to just the field of probiotics. A similar learning curve to pool and not pool types of interventions is seen in the progression of studies investigating the role of red *vs* white wine *vs* beer for cardiovascular health. An early meta-analysis did not separate the different types of alcohol and pooled results from wine, beer and liquors[10]. A later meta-analysis did separate wine from beer and found a dose-effect on reduced cardiovascular risk from wine, but not beer[30]. Further research determined the main effect on cardiovascular health was due to the resveratrol (from red wine) and later meta-analyses focused on just red wine[11]. Similarly, Hooper *et al*[12] recognized that not all chocolate is alike, as they separated out white *vs* dark chocolate and cocoa into separate sub-groups in their meta-analysis of chocolate on cardiovascular health.

CONCLUSION

Now, we recognize that different types of interventions (for example, different probiotic species or different types of medications) should not be pooled in a meta-analysis, nor should different disease outcomes be pooled[15,22]. Yet, retrospective reviews often fail to acknowledge the learning curve as both the public and scientific researchers learn how best to perform these types of complex analyses. Table 1 shows my recommendations for the proper conduct of meta-analyses. Meta-analysis has its place, but we must resist the urge to jump to an easy conclusion that any of the treatments included in the analysis work once a pooled estimate shows efficacy. Systematic reviews may offer a more transparent analysis of multiple studies for clinicians and the public, but may be more cumbersome to interpret.

REFERENCES

1 **Gebele C**, Tscheulin DK, Lindenmeier J, Drevs F, Seemann AK. Applying the concept of consumer confusion to healthcare: development and validation of a patient confusion model. *Health Serv Manage Res* 2014; **27**: 10-21 [PMID: 25595013 DOI: 10.1177/0951484814546959]

2 **Shaw TC**. Uncovering health literacy: Developing a remotely administered questionnaire for determining health literacy levels in health disparate populations. *J Hosp Adm* 2014; **3**: 140-156 [PMID: 25126152 DOI: 10.5430/jha.v3n4p149]

3 **Darbyshire JL**, Holman RR, Price HC. Presenting the results of clinical trials to participants. *Clin Med* 2009; **9**: 415-416 [PMID: 19886097 DOI: 10.7861/clinmedicine.9-5-415]

4 **National Data Program for the Sciences (NORC)**. 2012 General Social Survey. [Accessed 2014 Nov 12]. Available from: URL: http://www3.norc.org/GSS/

5 **O'Rourke K**. An historical perspective on meta-analysis: dealing quantitatively with varying study results. *J R Soc Med* 2007; **100**: 579-582 [PMID: 18065712 DOI: 10.1258/jrsm.100.12.579]

6 **Moher D**, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; **354**: 1896-1900 [PMID: 10584742 DOI: 10.1016/S0140-6736(99)04149-5]

7 **Liberati A**, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009; **62**: e1-e34 [PMID: 19631507 DOI: 10.1016/j.jclinepi.2009.06.006]

8 **Moher D**, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015; **4**: 1 [PMID: 25554246 DOI: 10.1186/2046-4053-4-1]

9 **Rotton J**, Kelly IW. Much ado about the full moon: a meta-analysis of lunar-lunacy research. *Psychol Bull* 1985; **97**: 286-306 [PMID: 3885282 DOI: 10.1037/0033-2909.97.2.286]

10 **Rimm EB**, Williams P, Fosher K, Criqui M, Stampfer MJ. Moderate alcohol intake and lower risk of coronary heart disease: meta-analysis of effects on lipids and haemostatic factors. *BMJ* 1999; **319**: 1523-1528 [PMID: 10591709 DOI: 10.1136/bmj.319.7224.1523]

11 **Sahebkar A**. Effects of resveratrol supplementation on plasma lipids: a systematic review and meta-analysis of randomized controlled trials. *Nutr Rev* 2013; **71**: 822-835 [PMID: 24111838 DOI: 10.1111/nure.12081]

12 **Hooper L**, Kay C, Abdelhamid A, Kroon PA, Cohn JS, Rimm EB, Cassidy A. Effects of chocolate, cocoa, and flavan-3-ols on cardiovascular health: a systematic review and meta-analysis of randomized trials. *Am J Clin Nutr* 2012; **95**: 740-751 [PMID: 22301923 DOI: 10.3945/ajcn.111.023457]

13 **McFarland LV**. Meta-analysis of probiotics for the prevention of antibiotic associated diarrhea and the treatment of Clostridium difficile disease. *Am J Gastroenterol* 2006; **101**: 812-822 [PMID: 16635227 DOI: 10.1111/j.1572-0241.2006.00465.x]

14 **Shen J**, Zuo ZX, Mao AP. Effect of probiotics on inducing remission and maintaining therapy in ulcerative colitis, Crohn's disease, and pouchitis: meta-analysis of randomized controlled trials. *Inflamm Bowel Dis* 2014; **20**: 21-35 [PMID: 24280877 DOI: 10.1097/01.MIB.0000437495.30052.be]

15 **Szajewska H**. Pooling data on different probiotics is not appropriate to assess the efficacy of probiotics. *Eur J Pediatr* 2014; **173**: 975 [PMID: 24849615 DOI: 10.1007/s00431-014-2340-4]

16 **D'Souza AL**, Rajkumar C, Cooke J, Bulpitt CJ. Probiotics in prevention of antibiotic associated diarrhoea: meta-analysis. *BMJ* 2002; **324**: 1361 [PMID: 12052801 DOI: 10.1136/bmj.324.7350.1361]

17 **Van Niel CW**, Feudtner C, Garrison MM, Christakis DA. Lactobacillus therapy for acute infectious diarrhea in children: a meta-analysis. *Pediatrics* 2002; **109**: 678-684 [PMID: 11927715 DOI: 10.1542/peds.109.4.678]

18 **Deshpande G**, Rao S, Patole S, Bulsara M. Updated meta-analysis of probiotics for preventing necrotizing enterocolitis in preterm neonates. *Pediatrics* 2010; **125**: 921-930 [PMID: 20403939 DOI: 10.1542/peds.2009-1301]

19 **King S**, Glanville J, Sanders ME, Fitzgerald A, Varley D. Effectiveness of probiotics on the duration of illness in healthy children and adults who develop common acute respiratory infectious conditions: a systematic review and meta-analysis. *Br J Nutr* 2014; **112**: 41-54 [PMID: 24780623 DOI: 10.1017/S0007114514000075]

20 **Khalesi S**, Sun J, Buys N, Jayasinghe R. Effect of probiotics on blood pressure: a systematic review and meta-analysis of randomized, controlled trials. *Hypertension* 2014; **64**: 897-903 [PMID: 25047574 DOI: 10.1161/hypertensionaha.03469]

21 **Hempel S**, Newberry SJ, Maher AR, Wang Z, Miles JN, Shanman R, Johnsen B, Shekelle PG. Probiotics for the prevention and treatment of antibiotic-associated diarrhea: a systematic review and meta-analysis. *JAMA* 2012; **307**: 1959-1969 [PMID: 22570464 DOI: 10.1001/jama.2012.3507]

22 **McFarland LV**. Deciphering meta-analytic results: a mini-review of probiotics for the prevention of paediatric antibiotic-associated diarrhoea and Clostridium difficile infections. *Benef Microbes* 2015; **6**: 189-194 [PMID: 24889895 DOI: 10.3920/BM2014.0034]

23 **Guarner F**, Khan AG, Garisch J, Eliakim R, Gangl A, Thomson A, Krabshuis J, Lemair T, Kaufmann P, de Paula JA, Fedorak R, Shanahan F, Sanders ME, Szajewska H, Ramakrishna BS, Karakan T, Kim N. World Gastroenterology Organisation Global Guidelines: probiotics and prebiotics October 2011. *J Clin Gastroenterol* 2012; **46**: 468-481 [PMID: 22688142 DOI: 10.1097/MCG.0b013e3182549092]

24 **Agerholm-Larsen L**, Bell ML, Grunwald GK, Astrup A. The effect of a probiotic milk product on plasma cholesterol: a meta-analysis of short-term intervention studies. *Eur J Clin Nutr* 2000; **54**: 856-860 [PMID: 11114681 DOI: 10.1038/sj.ejcn.1601104]

25 **Higgins JPT**, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration, 2011

26 **Sun X**, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ* 2010; **340**: c117 [PMID: 20354011 DOI: 10.1136/bmj.c117]

27 **Szajewska H**, Ruszczyński M, Radzikowski A. Probiotics in the prevention of antibiotic-associated diarrhea in children: a meta-analysis of randomized controlled trials. *J Pediatr* 2006; **149**: 367-372 [PMID: 16939749 DOI: 10.1016/j.jpeds.2006.04.053]

28 **Johnston BC**, Supina AL, Vohra S. Probiotics for pediatric antibiotic-associated diarrhea: a meta-analysis of randomized placebo-controlled trials. *CMAJ* 2006; **175**: 377-383 [PMID: 16908901 DOI: 10.1503/cmaj.051603]

29 **McFarland LV**. Systematic review and meta-analysis of Saccharomyces boulardii in adult patients. *World J Gastroenterol* 2010; **16**: 2202-2222 [PMID: 20458757 DOI: 10.3748/wjg.v16.i18.2202]

30 **Di Castelnuovo A**, Rotondo S, Iacoviello L, Donati MB, De Gaetano G. Meta-analysis of wine and beer consumption in relation to vascular risk. *Circulation* 2002; **105**: 2836-2844 [PMID: 12070110 DOI: 10.1161/01.CIR.0000018653.19696.01]

P- Reviewer: Bashashati M, Lehert P, Sasaki K **S- Editor**: Tian YL
L- Editor: A **E- Editor**: Wu HL

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i4.193

World J Meta-Anal 2015 August 26; 3(4): 193-205
ISSN 2308-3840 (online)

s

SYSTEMATIC REVIEWS

Brassiere wearing and breast cancer risk: A systematic review and meta-analysis

Winnie KW So, Dorothy NS Chan, Yan Lou, Kai-Chow Choi, Carmen WH Chan, Kristina Shin, Ava Kwong, Diana TF Lee

Winnie KW So, Dorothy NS Chan, Kai-Chow Choi, Carmen WH Chan, Diana TF Lee, The Nethersole School of Nursing, The Chinese University of Hong Kong, New Territories, Hong Kong, China

Yan Lou, Nursing Department, School of Medicine, Hangzhou Normal University, Hangzhou 311121, Zhejiang Province, China

Kristina Shin, Institute of Textiles and Clothing, the Hong Kong Polytechnic University, Kowloon, Hong Kong, China

Ava Kwong, Breast Surgery Division, Department of Surgery, Queen Mary Hospital, Hong Kong, China

Author contributions: So WKW, Chan DNS, Choi KC and Chan CWH had contributed to the planning; the manuscript was conducted by So WKW and Chan DNS; the meta analysis was done by Choi KC; So WKW, Chan DNS, and Lou Y had drafted the manuscript; So WKW, Lou Y, Choi KC, Chan CWH, Shin K, Kwong A and Lee DTF had contributed to the critical review of the manuscript.

Conflict-of-interest statement: The authors declare there is no conflict of interest.

Data sharing statement: On behalf of the authors, technical appendix, statistical code, and dataset available from Prof Winnie So at winnieso@cuhk.edu.hk will provide a permanent, citable and open-access home for the dataset.

Correspondence to: Dr. Winnie KW So, Associate Professor, The Nethersole School of Nursing, The Chinese University of Hong Kong, New Territories, Rm 731, 7/F, Esther Lee Building, Hong Kong, China. winnieso@cuhk.edu.hk

Telephone: +852-39431072
Fax: +852-26036041

Received: October 27, 2014
Peer-review started: October 28, 2014
First decision: November 28, 2014
Revised: June 30, 2015
Accepted: July 11, 2015
Article in press: July 13, 2015
Published online: August 26, 2015

Abstract

AIM: To evaluate existing evidence for the association between different type of brassiere exposures and the risk of breast cancer.

METHODS: Ovid Medline, CINAHL, Cochrane Data Base of Systematic Reviews, Pubmed, Scopus, Proquest, Sciencedirect, Wiley Online Library, WanFang Data, Hong Kong Index to Chinese Periodicals, China Journal Net, Chinese Medical Current Contents, Chinese Biomedical Literature Database, China Academic Journals Full-Text database, Taiwan Electronic Periodical Services and HyRead; reference lists of published studies; original research studies published in English or Chinese examining the association between type and duration of brassiere-wearing and breast cancer risk. Data were abstracted by a first reviewer and verified by a second. Study quality was rated according to predefined criteria. "Fair" or "good" quality studies were included. Results were summarised by meta-analysis whenever adequate material was available.

RESULTS: Twelve case-control studies were included in the review. Meta-analysis showed brassiere wearing during sleep was associated with a 1.3 times of increased risk. The odd ratio for more than 12 h of daily brassiere use in relation to breast cancer risk was 1.08

(95%CI: 1.01-1.14).

CONCLUSION: The present review demonstrates insufficient evidence to establish a positive association between the duration and type of brassiere wearing and breast cancer. Further research is essential; specifically, a large-scale epidemiological study of a better design is needed to examine the association between various forms of brassiere exposure in detail and breast cancer risk, with adequate control of confounding variables.

Key words: Breast cancer; Brassiere; Risk factors; Systematic review; Meta-analysis

Core tip: This systematic review and meta-analysis aimed to evaluate the association between 8 areas of brassiere-wearing practices and the risk of breast cancer. Twelve case-control studies met inclusion criteria for review. Although the meta-analysis shows statistically significant findings to support the association between brassiere wearing during sleep and breast cancer risk, evidence was insufficient to establish a positive association between brassiere wearing (duration and type) and breast cancer risk. A large-scale epidemiological study is needed to examine the relationship between various forms of brassiere exposure and breast cancer risk.

So WKW, Chan DNS, Lou Y, Choi KC, Chan CWH, Shin K, Kwong A, Lee DTF. Brassiere wearing and breast cancer risk: A systematic review and meta-analysis. *World J Meta-Anal* 2015; 3(4): 193-205 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i4/193.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i4.193

INTRODUCTION

Breast cancer is the most prevalent invasive cancer in women, comprising 22.9% of all such cancers in women worldwide and causing 458503 deaths in 2008. The incidence rate of the disease varies across countries, ranging from 18 to 90 per 100000, with the lowest rate found in developing countries and the highest in the developed world[1]. A number of factors have been examined for their association with the risk of breast cancer: demographic factors (gender and age); heredity and disease history (history of benign breast conditions); reproductive and hormonal factors (menstrual periods and diethylstilbestrol exposure); lifestyle and environmental factors (handedness, smoking and alcohol consumption, physical activity, previous chest radiation, body weight and night work)[2-5].

Among these, lifestyle-related factors are modifiable and thus are the most common areas to be targeted in breast cancer prevention. Apart from those listed above, brassiere wearing seems to be a subtle lifestyle-related factor that constantly arouses discussion among researchers. The brassiere is designed to support and uplift the breasts by utilising the tension of elastic materials, and this has now become a widespread habit where a brassiere is considered as a kind of fashion item with a protective function[6,7]. It is commonly thought that brassieres help to give women a better body shape, and that the underwired type provides better support for the breasts and prevents them from sagging[6,7]. Women usually wear a brassiere during normal daily activity for the purpose of breast support, and it has become an indispensable part of everyday attire. Furthermore, brassiere wearing seems to conform to a social norm whereby women cover their breasts to avoid any embarrassment and to improve their self-confidence[6].

Given the increasing incidence of breast cancer, attention has been paid to investigating this increasingly common habit. Studies have shown that wearing an underwired brassiere, sleeping with a brassiere, wearing one for more than 12 h a day, and incorrect brassiere wearing are potential risk factors[6-14]. However, such studies have failed to control risk factors such as body weight and epidemiological data, thereby diminishing the validity of their results. Findings from a multicentre study, though inconsistent, showed a non-significant association between daily use of a brassiere and breast cancer in premenopausal women[3,8-16].

In view of emerging concerns about breast cancer in relation to brassiere wearing and about the inconsistent results available from previous research, the purpose of this systematic review and meta-analysis is to evaluate the existing evidence for the association between different brassiere-wearing practices and breast cancer risk, to provide a clearer picture of the evidence and in that way to inform breast cancer prevention planning.

MATERIALS AND METHODS

Data sources and searches

Two investigators made a comprehensive literature search in January 2015 of relevant articles published in English or Chinese, using Ovid Medline (since 1946), Cumulative Index to Nursing and Allied Health Literature (CINAHL; since 1937), Cochrane Data Base of Systematic Reviews (since 1995), PubMed, Scopus (since 1823), Proquest (since 1923), Sciencedirect, Wiley Online Library, WanFang Data (since 1993), Hong Kong Index to Chinese Periodicals, China Journal Net (since 1915), Chinese Medical Current Contents (CMCC; since 1994), Chinese Biomedical Literature Database (since 1980), China Academic Journals Full-Text database (since 1994), Taiwan Electronic Periodical Services and HyRead (since 1974). Relevant keywords and search terms used were "breast cancer, breast neoplasm, breast carcinoma, bra, brassiere, constrictive clothing, underwear, undergarment, risk factor". A

Table 1 United States Preventive Services Task Force Quality Rating Criteria (Case-Control study)

Criteria
Accurate ascertainment of cases
Nonbiased selection of cases/controls with exclusion criteria applied equally to both
Response rate
Diagnostic testing procedures applied equally to each group
Measurement of exposure accurate and applied equally to each group
Appropriate attention to potential confounding variable
Definition of ratings based on above criteria
Good: Appropriate ascertainment of cases and non-biased selection of case and control participants; exclusion criteria applied equally to cases and controls; response rate equal to or greater than 80%; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to confounding variables
Fair: Recent, relevant, without major apparent selection or diagnostic work-up bias but with response rate less than 80% or attention to some but not all important confounding variables
Poor: Major selection or diagnostic work-up biases, response rates less than 50%, or inattention to confounding variables

secondary search for studies not identified through databases was conducted by manually reviewing reference lists of the twelve studies finally chosen.

Study selection

The authors developed inclusion and exclusion criteria for abstracts and articles based on the target population, risk factor and outcome measure. Only original research studies in either English or Chinese were included. The target population (case) consisted of women with breast cancer. Measurements of exposure to brassiere wearing are included in the data collection section. The outcome measure covered risk factors of breast cancer, and is reported in the results section, either in the text or in a table. We included original research studies with the full text available, and excluded abstracts, unpublished studies and articles written in languages other than English or Chinese. When a study met the inclusion criteria but there was insufficient data in the paper itself, the corresponding author was contacted for further information.

Data extraction and quality assessment

A reviewer abstracted data from the studies identified and transferred it into a structured form, which included the following information: year and country of study, study design, sample size, participant characteristics (including age, sources and residential status), diagnostic method, information on brassiere exposure, adjusted covariates, outcome results and study quality. A second reviewer confirmed the accuracy of the data.

Two reviewers assessed the quality of the studies according to predefined criteria developed by the US Preventive Services Task Force[17], which specifically grades the internal validity of case-control studies. Six aspects were evaluated: accurate ascertainment of cases, non-biased selection of cases/controls with exclusion criteria applied equally to both, response rate, diagnostic testing procedures applied equally to each group, accurate measurement of exposure applied equally to each group, and appropriate attention to potential confounding variable A "good" study meets all these criteria, while one of "fair" quality does not meet all criteria but is without a fatal flaw invalidating its results. Two reviewers independently rated the quality of each study as "good", "fair" or "poor" (Table 1) and resolved discrepancies by consensus in the presence of a third reviewer. Only studies of fair or good quality were included in the review.

Data synthesis and analysis

Results of the studies were synthesised in a narrative way in an evidence table corresponding to the topic of the review. We assessed the heterogeneity of the studies qualitatively by their study design, participant characteristics, data collection and analytic methods. Specific quantitative results were synthesised narratively. To define the suitability of the risk factors to be included in the meta-analysis, the homogeneity of study method, participant background and statistical methods were considered. Specifically, meta-analysis was conducted for the effect on the risk of breast cancer entailed by wearing a brassiere during sleep. This was the only factor assessed in seven studies, with the results shown in six of them. Another factor, tightness of the brassiere was assessed in three studies, but not considered further, as multivariable adjusted analysis produced insignificant results. The combined estimate of risk effects, together with its 95%CI, was calculated by using the estimates in individual studies, producing the best control over other potential risk factors and confounders. In particular, the odds ratio of each study, with as many other risk factors and/or confounders as possible adjusted, was chosen for meta-analysis. A total of six studies were judged appropriate for analysis, with five having adjusted odds ratios. Heterogeneity among the studies was assessed by Cochrane Q-test and I^2 statistics[18]. As neither the Q test ($P < 0.05$) nor $I^2 > 40\%$ were significant[18], the fixed-effects (weighted inverse variance) method was applied to the meta-analysis. Review Manager (RevMan5.3, Cochrane Collaboration, Oxford, England) was used for the meta-analysis. The statistical method of this study was reviewed by Choi KC who is a biostatistician and one of

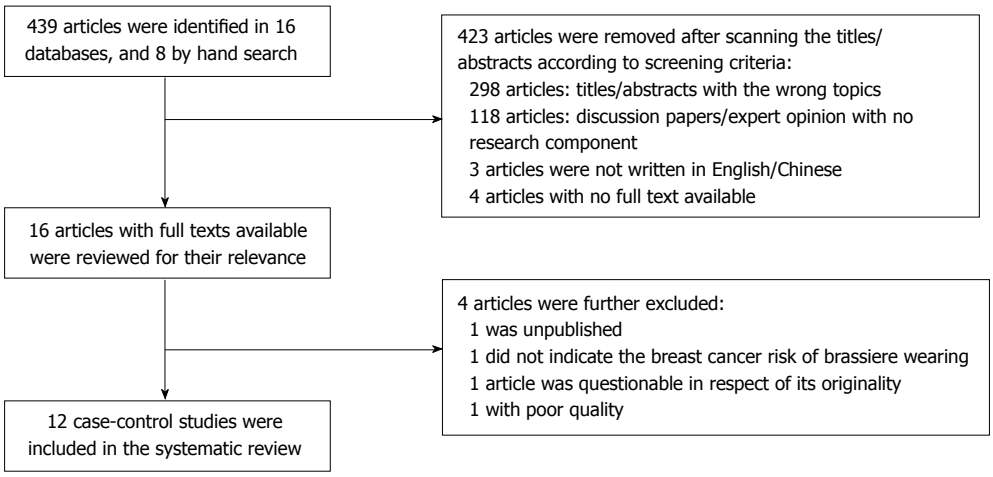


Figure 1 Flow chart of study selection.

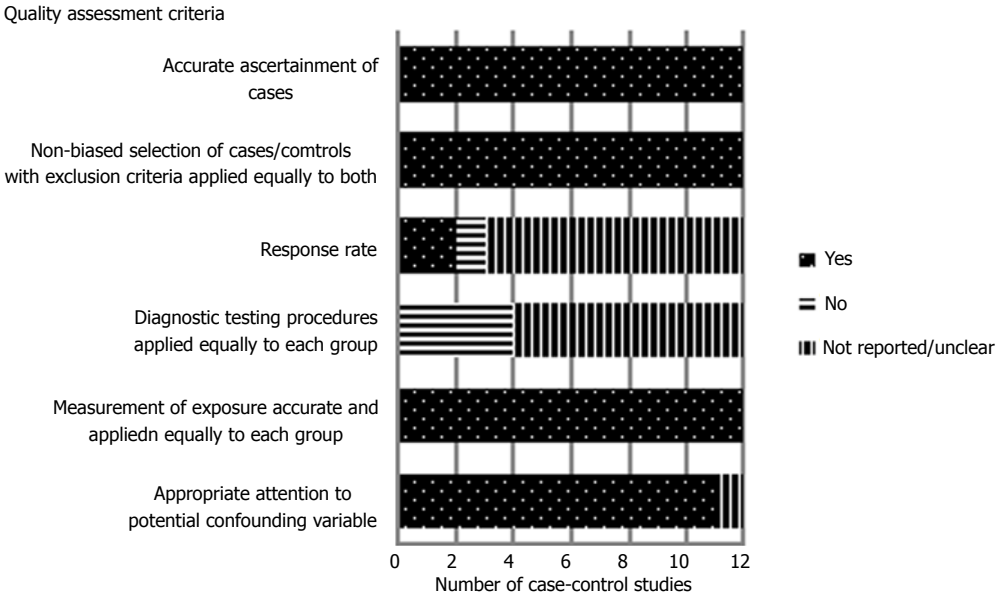


Figure 2 Quality assessment of 12 Case-control studies examining the association between brassiere exposure and breast cancer risk.

the authors.

RESULTS

Study inclusion

A total of 439 studies were identified through 16 databases. Of these, 423 were removed, after scanning of abstracts and titles, because they were concerned with an irrelevant topic, discussion paper or expert opinion, or were not written in English or Chinese, or the full text was not accessible (Figure 1). This left sixteen articles with full texts, which were further reviewed for relevance. After a detailed examination, four further articles were excluded, one of which, with the same statistical results as another article published in 2006[9], raised concerns about its originality. The twelve remaining studies that met all inclusion criteria were further assessed for quality. Two studies of good and ten of fair quality examining brassiere exposure and breast cancer risk were finally included in the systematic review and meta-analysis. Characteristics, results and quality assessment of the twelve studies are illustrated in Tables 2 and 3 and Figure 2, respectively.

All twelve studies were published between 1991 and 2014, and all were of a case-control design, where people with breast cancer (case participants) were compared with others free of the disease (control participants), with respect to the risk of breast cancer from wearing a brassiere (risk factor). One multicentre study was conducted across Europe, Asia and North and South America[3], one in the metropolitan areas in the United State[19] and ten in Asia[9-16,20,21]. The ratio of case to control was 1:1 (n = 8)[9,10,12,13,15,16,19,20], 1:2 (n = 2)[14,21] or 1:3 (n = 2)[3,11]. The number of participants varied across studies, from 190 to 9333. Seven reported the mean age of participants, varying from 45.6 to 50.4[9,10,12-16]. Two studies only stated that they included participants of 35 or more[3,11], one reported participants

Table 2 Study characteristics

Ref.	Design	Country	No. of participants		Source of participants		Age of participants		Years of residence of participants	Residential status of participants	Age-match between participants	Ratio of cases to controls	Diagnosis of breast cancer status	
			Cases	Controls	Cases	Controls	Cases	Controls					Cases contc	Controls
Hsieh *et al*[3]	Case-control	Greece, United States, Wales, Brazil, Yugoslavia, Taiwan, Japan	2325	7008	Hospital	Same hospital as cases	Over 35 years old	Over 35 years old	Not reported	Yes	± 2 yr	1:03	Diagnosed in hospital	Not reported
Feng *et al*[20]	Case-control	China	262	262	Hospital	Same hospital as cases, relative/ neighbor of index case	Not reported	Not reported	Not reported	Yes	± 5 yr	1:01	Diagnosed in hospital	Not reported
Lee *et al*[10]	Case-control	Taiwan	250	219	Out-patient clinic or cancer center	Same as cases	Mean: 47.2;	Mean: 46.4	Not reported	Yes	± 5 yr	1:01	Pathological examination	Mammogram
Zhu *et al*[9]	Case-control	China	246	246	Hospital	Same hospital as cases; relative/ neighbor of index case	Mean: 50.4	Mean: 50.0	Not reported	Not reported	± 5 yr	1:01	Diagnosed in hospital	Not reported
Zhang *et al*[11]	Case-control	China	284	669	Hospital	Health check program	Over 35 years old	Over 35 years old	Not reported	Yes	Not reported	1:03	Pathological examination	Not reported
Chen[21]	Case-control	China	167	334	Hospital	Neighbor	Not reported	Not reported	Over 10 years	Yes	± 2 yr	1:02	Diagnosed in hospital	Not reported
Hu and Lin[16]	Case-control	China	95	95	Hospital	Same hospital as cases	Mean: 48	Mean: 48	Over 10 years	Yes	Not reported	1:01	Pathological examination	Not reported
Liu *et al*[15]	Case-control	China	365	365	Not reported	Randomly obtained from population	Mean: 46.4	Mean: 45.6	Over 10 years	Yes	± 3 yr	1:01	Pathological examination	Not reported
Yao *et al*[13]	Case-control	China	200	200	Hospital	Same hospital	Mean: 47.30	Mean: 45.88	Over 10 years	Yes	± 2 yr	1:01	Pathological examination	Not reported
Chen *et al*[19]	Case-control	United States	Ducta: 454 Lobular: 590	469	Cancer Surveillance System	Randomly obtained from population	Between 55-74	Between 55-74	Not reported	Yes	± 5 yr	1:01:01	Pathological examination and tumor tissue specimens	Not reported
Liu *et al*[14]	Case-control	China	208	416	Hospital	Coworker/ neighbor	Mean: 50.1	Mean: 49.2	Over 5 years	Yes	± 5 yr	1:02	Pathological examination	Not reported
Shen *et al*[12]	Case-control	China	275	275	Hospital	Not reported	Mean: 45.6	Mean: 48.5	Not reported	Yes	Not reported	1:01	Pathological examination	Physical examination

aged between 55-74[19], while two did not report the age[20,21]. The control participants were age-matched with the case participants within 2-5 years. Two studies reported the response rate (83%-87%)[10,19]. Data on participants' demographic data and information on brassiere exposure were collected by means of either questionnaire or interview. Eleven studies indicated that their participants (both case and control) were resident in the study area[3,10-16,19-21], one stated over 5 years of residence[14], and four

Table 3 Evidence table of brassiere wearing and breast cancer risk

Ref.	Study years	Area of brassiere exposure								Effect size: OR (95%CI)	Adjusted covariates		Study quality
		Duration of Brassiere wearing per day	Brassiere users	Sleeping with brassiere	Tightness of brassiere wearing	Wear underwired brassiere	Age begun brassiere wearing	Brassiere cup size 1 yr before reference data	Appropriateness of brassiere wearing		Modifiable	Non-modifiable	
Hsieh *et al*[3]	Not reported		Yes							(Premenopausal) non-brassiere-users *vs* brassiere-users: 0.44 (0.17-1.15)	Age at first birth Parity Obesity	Age at interview Study center	Fair
Feng *et al*[20]	1 yr			Yes	Yes					Sleep without brassiere *vs* with brassiere: 0.26 (0.09-0.77) Tightness of brassiere wearing: Not significant	Occupational contact of chemicals Emotional adjustment Psychological distress History of abortion Diet (high fat, rich in beans, take breakfast, drink tea) Night work Personality	History of benign breast disease Family history of breast cancer Menarche history	Fair
Lee *et al*[10]	3 yr	Yes		Yes						Wear brassiere (≥ 13 h/d) *vs* (≤ 13 h/d) (> age 40): 3.0 (1.6-5.7) Wear brassiere (≥ 13 h/d) *vs* (≤13 h/d) (all ages): 2.1 (1.3-3.3) Sleep with brassiere: Not significant	Physical activity Education High fat food Supplement and allium use Total calories	Age	Good
Zhu *et al*[9]	1 yr			Yes	Yes					Sleep with brassiere *vs* without brassiere: 2.32 (1.32-4.10) Tightness of brassiere wearing: Not significant	Parity Diet (high fat, rich in beans, drink tea) Psychological distress Personality Occupational contact of chemicals	Menarche history Family history of breast cancer History of benign breast disease	Fair
Zhang *et al*[11]	10 mo			Yes						(Premenopausal) sleep without brassiere *vs* with brassiere: 0.401 (0.250-0.644) Post-menopausal: not significant	Oral contraceptive use Physical activity Education Emotional problem/adjustment	Family history of breast cancer History of benign breast disease or breast biopsy Breast pain during menstruation	Fair
Chen[21]	2 yr			Yes	Yes					Both are not significant	Body mass index Lactation Oral contraceptive use Education Psychological distress Smoking history Night work Diet Occupational contact of chemicals or radiation	Age Family history of breast cancer History of benign breast disease Age at menarche and menopause	Fair

Hu and Lin[16]	3 yr			Yes			Wear hard brassiere (Yes *vs* No): 6.729 (2.001-22.635)	Lactation Oral contraceptive use Sexual life Passive smoking Psychological distress Diet (fried and oily food) Sleeping hours History of abortion	History of benign breast disease, cervical disease, ovarian cancer, hepatitis -	Fair
Liu *et al*[15]	3 yr		Yes				Sleep with brassiere *vs* without brassiere: 2.313 (1.323-4.121)	Lactation Parity History of abortion Age at first birth Occupational contact of chemicals Passive smoking Diet (high fat, spicy/salty food, preserved food, seafood, drink water) Psychosocial problem (trauma, family problem, distress, anger)	Age at menarche Family history of breast cancer -	Fair
Yao *et al*[13]	15 mo		Yes	Yes			Sleep with brassiere *vs* without brassiere: 1.902 (1.177-3.072) Wear hard brassiere (Yes *vs* No): Not significant	Large scale renovation Non-environmental friendly decoration materials Interval between renovations Nature of occupation Lactation Number of birth labour Job-related life events Fruit intake High fat intake Salted food intake Workplaces condition Sleeping hours and quality -	Family history of other tumors Family history of breast cancer History of cancer in first degree relatives Mammary hyperplasia The death of a loved one	Fair
Chen *et al*[19]	4 yr 3 mo	Yes		Yes	Yes	Yes	All are not significant	Education Annual parity BMI Bra band size Age at first full-term pregnancy Mammogram screening	Race/ethnicity Family history of breast cancer Typs of menopause	Good

Liu *et al*[14]	2 yr 7 mo	Yes			Wear brassiere (≥ 13 h/d) *vs* (≤13 h/d): 1.064 (1.001-1.132)	Parity Occupation Oral contraceptive use Passive smoking Use of royal jelly Use of other supplement Work-related stress Personality Emotional adjustment Interpersonal relationship Contact of chemicals Psychological distress Adverse life events Sleeping hours	History of benign breast disease	Fair
Shen *et al*[12]	4 yr			Yes	Appropriately wearing *vs* inappropriately wearing: 2.313 (1.112-5.43)	History of abortion Lactation Work intensity Long working hours Poor sleep quality Family discord Satisfaction towards life Fatigue strength Smoking	Family history of other tumors Family history of breast cancer Mammary hyperplasia	Fair

further stated that participants had lived in the area for more than 10 years[13,15,16,21]. Nearly all case participants were recruited from hospital, outpatient clinic or cancer centre, with control participants recruited from the same hospital[3,9,13,16,20], outpatient clinic or cancer centre[10], health-check programme[11], among neighbours/relatives of the case participants[9,14,20,21], randomly obtained from the local population[15,19], or from the same residential area[12]. All case participants had their cancer diagnosis confirmed in hospital and eight studies specifically reported that the diagnosis was confirmed by pathological examination. Nevertheless, eleven studies did not clearly state what diagnostic tests had been carried out on participants in the control group[3,9,11-16,19-21], reporting only that they were found to be healthy and were not currently suffering from breast cancer. One study had conducted mammography screening[10] and one had underwent physical examination[12] to exclude the presence of breast cancer among control participants.

Brassiere exposure

Eight aspects of brassiere exposure have been examined previously and included in the present review: brassiere users and non-users[3], duration of brassiere wearing per day[10,14,19], sleeping with or without a brassiere[9-11,13,15,20,21], tightness of the brassiere[9,20,21], wearing an underwired brassiere[13,16,19], age began brassiere wearing[19], brassiere cup size 1 year before reference data[19], and appropriateness of brassiere wearing[12].

Analytic methods

Univariate and multivariable logistic regression analysis of the data was used to explore 8 areas of brassiere wearing practice and breast cancer risk. All studies used multivariable logistic regression analysis, with the results shown in the Results section, in the text and/or in table form. Of the twelve studies, nine presented the univariate results in tabular form[9-16,19]. Two reported that they had derived significant results from the univariate analysis. However, no detailed figures were given in either text or tabular form[20,21]. Ten studies reported that significant variables detected in the univariate analysis were further included in the multivariable logistic regression model[9-16,20,21].

Attention to covariates

To identify any significant relationship between brassiere wearing and breast cancer risk, all the studies paid appropriate attention to the covariates which also provided reliable and comparable data considered as confounding variables during statistical analysis, although not all known non-modifiable and modifiable factors were included.

Non-modifiable factors

Eight known non-modifiable risk factors were considered in the analysis of the twelve studies. They were age[3,10,21], history of benign breast disease[9,11,14,16,20,21], family history of breast cancer[9,11-13,15,19-21], family history of other tumors[12,13], age at menarche and menarche history[9,15,20,21], age at menopause[21], mammary hyperplasia[12,13], death of a loved one[13]. Factors that had an uncertain effect on breast cancer risk were also reported, such as breast pain during menstruation[11]. One study considered the study centre to be one of the covariates, since it was conducted in seven countries[3]. Chen *et al*[19] also evaluated race/ethnicity and types of menopause as covariates as potential confounders.

Modifiable factors

Several known modifiable risk factors were considered in the analysis. These factors are grouped under 5 broad categories, namely lifestyle-related, reproductive, nutritional, psychological and emotional, and others modifiable factors with uncertain effects. Lifestyle-related factors such as obesity[3], body mass index[19,21], physical activity[10,11], oral contraceptive use[11,14,16,21] and lactation[12,13,15,16,21], and reproductive factors such as parity[3,9,14,15] and age at first birth[3,15,19]. Additionally, dietary habits, being uncertain factors of breast cancer risk, were commonly referred to in the studies, including diets high in fat[9,10,13,15,20], beans[9,20], dried or oily food[16], spicy, salty, preserved seafood or other foods[13,15], breakfast[20], tea or water[9,15,20], supplement and allium use[10,14], and total calorie intake[10]. Psychological and emotional factors such as emotional adjustment[11,14,20], psychological distress[9,14-16,20,21], and psychological problems (trauma, family problems, anger, fatigue, and stress)[12,14,15] constituted another uncertain category commonly found in the studies. Some factors with uncertain effects on breast cancer risk were also considered, such as occupational contact with chemicals or radiation[9,15,20,21], night work[20,21], personality[9,14,20], smoking history or passive smoking exposure[12,14-16,21], sex life[16], sleeping hours[13,14,16], education[10,11,19,21], and any history of abortion[12,15,16,20].

Association between exposure of brassiere and breast cancer risk

Univariate analyses were conducted to identify candidate variables for further multivariable logistic regression analysis. Among eight areas of brassiere exposures that has been studied, brassiere wearing more than twelve hours[10,14], sleeping with a brassiere on[9-11,15], tightness of the brassiere[9], wearing an underwired type[16], and incorrect brassiere wearing[12] were found to have a significant association with breast cancer. Two studies reported in their Results sections that sleeping with a brassiere on and tightness of the brassiere were significantly associated with the occurrence of breast cancer, but no data or figures were provided either in the text or in tabular form[20,21].

After univariate analyses, multivariable analysis was carried out to investigate further brassiere exposure and the risk of breast cancer. One study found that pre-menopausal women who were not brassiere users were less than half as likely to contract the disease as those who did wear one (OR = 0.44; 95%CI: 0.17-1.15)[3]. Hu and Lin[16] studied the effect of wearing an underwired brassiere on the incidence of breast cancer, and claimed that women who wore that type were 6.7 times more likely to have breast cancer than those who did not (OR = 6.7; 95%CI: 2.0-22.6)[16]. However, insignificant association was found in other 2 studies[13,19].

Inconsistent results were noted in the case of brassiere wearing during sleep. Two studies found no significant difference in the breast cancer risk between sleeping with or without a brassiere on[10,21]. However, five others found that sleeping with a brassiere on was a significant risk factor[9,11,13,15,20], interpreting the results in different ways. Feng *et al*[20] reported that sleeping without a brassiere was much less risky than sleeping with one (OR= 0.26; 95%CI: 0.09-0.77)[20]. Zhang *et al*[11] investigated this by classifying the participants into pre- and post-menopausal groups. They found that the former, sleeping without a brassiere, were less than half as likely to contract breast cancer than those who did wear one (OR = 0.40; 95%CI: 0.25-0.64)[11]. Three studies showed that women who slept with a brassiere on were 1.9-2.3 times more likely to have breast cancer than those who did not[9,13,15]. In three studies, the tightness of the brassiere was found to have an association but was not a significant risk factor[9,20,21].

The results between duration of daily brassiere use and breast cancer risk are in disagreement. Chen *et al*[19] found insignificant association between duration of brassiere wearing per day and the risk the breast cancer, but Liu *et al*[14] and Lee *et al*[10] reported more than 12 h or daily brassiere wear increased the risk of breast cancer. Meta-analyses were conducted to evaluate the findings on brassiere wearing during sleep and duration of daily brassiere use. Although a total of seven studies reported findings related to brassiere wearing during sleep[9-11,13,15,20,21], only six provided numerical results[9-11,13,15,20], and only data from these six are therefore included in the meta-analysis (Figure 3). The total case-to-control ratio was 1607:1961. The meta-analysis of the five studies with adjusted odds ratios and the other with an unadjusted ratio showed significant heterogeneity [heterogeneity: χ^2 = 43.51, *df* = 5 (P < 0.001); I^2 = 89%]. The fixed-effects method was applied to the meta-analysis, showing a significant association between brassiere wearing during sleep

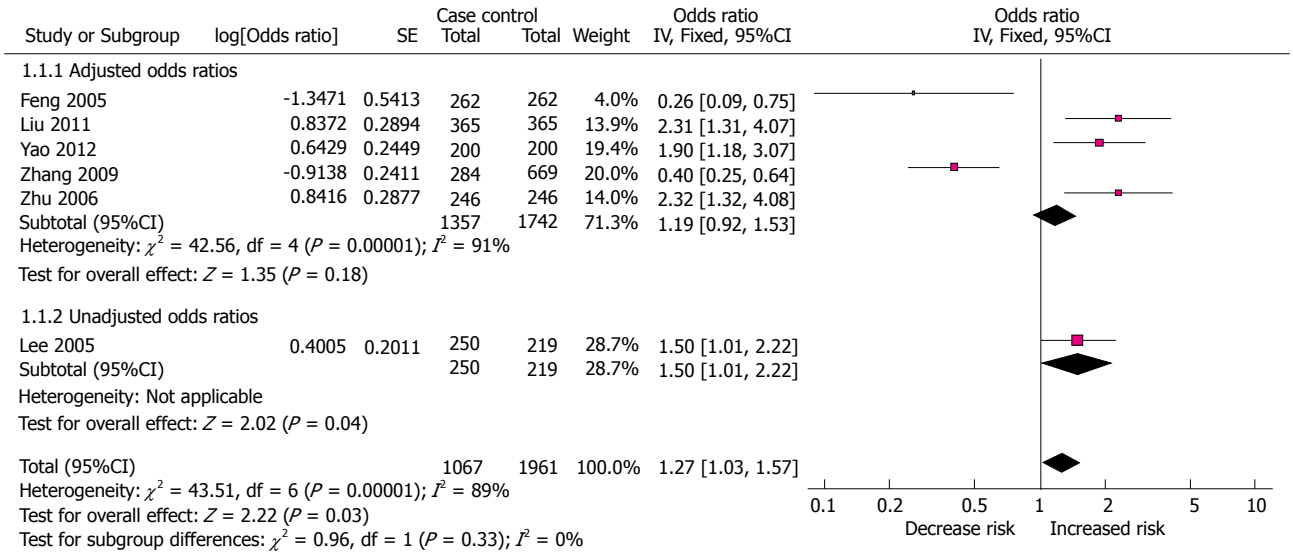


Figure 3 Meta-analysis of the findings regarding the brassiere wearing during sleep.

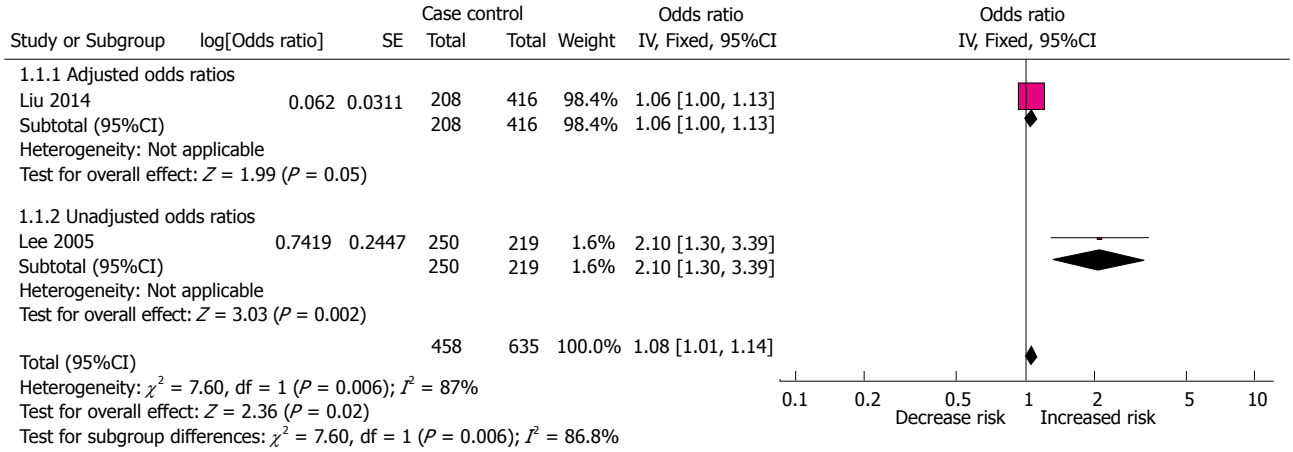


Figure 4 Meta-analysis of the findings regarding the brassiere wearing more than 12 h/d.

and breast cancer (summary OR = 1.27; 95%CI: 1.03-1.57).

Two studies that provided numerical results regarding to duration of daily brassiere use were entered into the meta-analysis (Figure 4). The total case-to-control ratio was 458:635. The odd ratio for more than 12 h of daily brassiere use in relation to breast cancer risk indicated a significant association used fixed-effects method (summary OR = 1.08; 95%CI: 1.01-1.14). Results for the heterogeneity analysis shown significant high heterogeneity for risk of breast cancer [heterogeneity: χ^2 = 7.6, *df* = 1 (*P* = 0.006); I^2 = 87%].

DISCUSSION

According to the American Cancer Society, a number of modifiable and non-modifiable factors that may contribute to the incidence of breast cancer have been identified[22]. However, there are still many uncertain factors awaiting clarification and confirmation in respect of breast cancer risk. The growing popularity of brassiere wearing and public awareness of breast cancer prevention have led researchers to investigate the possible association between brassiere wearing and breast cancer risk. An initial examination of 439 published studies investigating that association was conducted, and as a result a total of twelve "fair" to "good" quality case-control studies involving over 10000 participants were included in the present systematic review.

This review shows that 8 areas of brassiere exposure[12,13,19,20] have been studies in relation to breast cancer risk. There are inconsistent results in respect of the association between brassiere wearing during sleep, duration of daily brassiere use and breast cancer. Meta-analyses, when used to evaluate these associations further, show a positive relation between brassiere wearing during sleep and breast cancer with an OR of 1.27 (95%CI: 1.03-1.57), and between duration of daily brassiere use and breast cancer with an OR of 1.08 (95%CI: 1.01-1.14). These may be

simply the result of prolonged wearing of a brassiere - a finding consistent with what Singer and Grismaijer concluded in 1995[8]. The assumption of both Singer and Grismaijer[8] and Kumar[23] was that brassiere wearing might inhibit the temperature regulatory system of the breast and engender breast neoplasm. However, the idea of changes in surface temperature over the breast area is challenged by King[24], who suggests that the temperature may correlate with the amount of clothing worn, but it is hard to determine how much breast glandular tissue has been exposed to abnormal temperature as the result of clothing. To be more accurate, the temperature should be measured at the breast tissue itself rather than the surface[24].

In addition to the duration of brassiere wearing, one study included in this review also showed that wearing an underwired brassiere had a 6.7 times increased risk of breast cancer[16] although 2 studies reported no association between the two[13,19]. This raises further consideration of the influence of different brassiere designs. Underwired types have thin rigid material sewn into the underside of the cup of the bra. The material may be metal, plastic or resin. Compared to other types, underwired brassieres provide better lifting and shaping effects and are thus preferred by many women. However, the underwired type is more uncomfortable in comparison with its wire-free counterpart, as it puts added pressure on the breasts, potentially clogging the milk duct and lymphatic system. Obstructing the latter over the breast area may cause fluid, toxin and carcinogen to accumulate, and cause breast cancer as a result[8,23]. Nevertheless, various studies have opposed the suggestion of a blockage in the lymphatic drainage system, on the grounds that lymphatic fluid drains upward to the armpit but not down to the bottom of brassiere, which is considered to be the constricted region, and in fact it has been shown that lymphatic drainage is not affected even if a tight-fitting or underwired brassiere is worn[25-28]. This may be confirmed by the results of three other studies included in this review, which found a non-significant association between the tightness of the brassiere and breast cancer risk[9,20,21]. Additionally, the one study that gives a certain risk factor connected with underwired brassieres does not specifically report any history of using such types, which generates further concern about its results. Interestingly, an investigation into awareness and knowledge of breast cancer risk among Malaysian women revealed that more than a third of participants linked the wearing of underwire brassieres with breast cancer[29], indicating that women paid attention to brassiere wearing in the prevention of breast cancer even without robust evidence of its risks.

Apart from the association between the kind of brassiere, the duration of wearing one and breast cancer, Hsieh and Trichopoulos's study reported an association between brassiere use and breast cancer among pre-menopausal women (OR = 0.44; 95%CI: 0.17-1.15; *P* about 0.09)[3]. However, the authors went on to speculate that this might have been related to obesity, as the brassiere users in the study were substantially heavier than non-users.

Though not conclusive, findings in this review raise concerns about brassiere wearing in general - not only about the duration, but also the nature of the garment itself (*e.g.*, construction method and materials) and how and when it is worn. Brassiere-wearing practice is undoubtedly a serious issue whose relation to breast cancer risk needs to be explored. However, the evidence accumulated in previous studies, as discussed in the current review, is not strong enough to draw reliable conclusions. To explore in greater depth the effects of brassiere wearing on the risk of breast cancer, further research is recommended to examine the physiological responses of the breast to brassiere exposure in association with cancer risks, with proper adjustment for confounding variables during the analysis.

Limitations

There are several limitations to the present review that may affect the validity and generalisability of its findings. The eight studies included were all of the case-control design. Although such design is commonly used to identify risk factors contributing to a medical condition, its potential bias may influence the reliability of results. This includes selection, information and confounding bias.

First, there was selection bias in respect of the source of participants. Of the current twelve studies, two are the population-based case-control type[15, 19], another seven recruited case and control participants from the same hospitals[3,9-11,13,16,20], two matched control participants from the neighbourhood[14,21], and the remaining one with unclear source of sample[12]. This could lead to admission rate bias. The control participants from hospital are not as representative as those from the community. For example, people in areas remote from hospital and people unable to afford medical care may not be included in the study. Furthermore, only two studies in this review recruited incident cases[3,9], while the others might have prevalent-incidence bias, whereby breast cancer patients reduce the use of brassiere after diagnosis.

Second, information bias might dilute the results of the review. Case-control studies usually suffer from measurement bias due to participants' failure to recall past events[30]. Only two studies in this review reported the response rate[10,19], and this limited our ability to evaluate the bias caused by response rates in different studies. Additionally, of the twelve studies, only two indicated the use of mammography or physical examination and pathological examination to confirm breast cancer status in both case and control groups[10,12]. It would certainly be difficult and might be unethical to ask control participants to undergo pathological examination or mammography, but mis-classification bias might otherwise be introduced, as it is possible that some control participants are potential

breast cancer patients. Furthermore, the search for studies in this review was limited to those written in English or Chinese, and those in other languages were not considered, potentially limiting the number included and thus affecting the generalisability of the findings to different populations. It was also possible that certain related studies could not be identified, despite the use of multiple databases.

Third, confounding bias might influence the association detected between brassiere-wearing practices and breast cancer risk. Studies varied in the degree to which confounding variables were adjusted during analysis, and none was able to control all known non-modifiable and modifiable factors, leading to potential bias in the estimates of odds ratios. Obesity was identified as a confounding factor that related to the use of a brassiere and the occurrence of breast cancer. However, only three studies controlled the factor of obesity or BMI during the analysis stage[3,19,21]. Also, the use of a brassiere and a preference for its particular design are related to socio-economic status, which itself is associated with breast cancer risk. Three studies in the review adjusted for the education factor[11,19,21], and eight considered the effect of diet[9,10,13-16,20,21]. Further studies should take annual income into account to eliminate the confounding influence of socio-economic status.

In conclusion, the present review demonstrates insufficient evidence to establish a positive association between brassiere wearing (duration and type) and breast cancer risk, although the meta-analysis shows statistically significant findings which support the association between brassiere wearing during sleep and breast cancer risk. Twelve studies reviewed suffered from selection bias, information bias or improper adjustment of confounding variables, which all affect the validity and generalisability of the findings to different populations. Further research is essential - specifically, a large-scale epidemiological study of a better design is needed to examine the association between various forms of brassiere exposure in detail and breast cancer risk, with adequate control of confounding variables. In this way, women may be granted the informed view of brassiere usage that they need and deserve. Since wearing brassieres has already become common among women in developed countries, education on the proper use of such garments can help maintain better breast health.

COMMENTS

Background

Given an increasing incidence of breast cancer, attention has been paid to investigating common modifiable lifestyle-related factors including brassiere wearing practice for breast cancer prevention. However, inconsistent results have been reported in the association between use of a brassiere and risk of breast cancer.

Research frontiers

The authors performed the first systematic review and meta-analysis to investigate the association between different type of brassiere exposures and the risk of breast cancer.

Innovations and breakthroughs

Eight aspects of brassiere exposure have been examined by 12 Case-control studies were included in the present review: brassiere users and non-users, duration of brassiere wearing per day, sleeping with or without a brassiere, tightness of the brassiere, wearing an underwired brassiere, age began brassiere wearing, brassiere cup size one year before reference data, and appropriateness of brassiere wearing. Meta-analyses were conducted to evaluate the findings on brassiere wearing during sleep and duration of daily brassiere use. The results showed that brassiere wearing during sleep was associated with a 1.3 times of increased risk. The odd ratio for more than 12 h of daily brassiere use in relation to breast cancer risk was 1.08 (95%CI: 1.01-1.14).

Applications

The meta-analysis shows statistically significant findings to support the association between brassiere wearing during sleep and breast cancer risk, however, all the twelve studies reviewed suffered from selection bias, information bias or improper adjustment of confounding variables, which all affect the validity and generalisability of the findings to different populations. A large-scale epidemiological study is needed to examine the relationship between various forms of brassiere exposure and breast cancer risk, with adequate control of confounding variables.

Terminology

Brassiere is designed to support and uplift the breasts by utilising the tension of elastic materials, and this has now become a widespread habit where a brassiere is considered as a kind of fashion item with a protective function.

Peer-review

This paper addresses the breast cancer risk associated with wearing brassiere. The paper is very well written and well structured. The article also has the potential to add to what the we know about risk factors of breast cancer.

REFERENCES

1 **International Agency for Research on Center**. World Cancer Report. World Health Organisation, 2008. Available from: URL: http://www.iarc.fr/en/publications/pdfs-online/wcr/2008/wcr_2008.pdf

2 **Chen X**, Kong Z. Chinese Cancer Registry Annual Report 2009. Beijing: Military Medical Science Press, 2010

3 **Hsieh CC**, Trichopoulos D. Breast size, handedness and breast cancer risk. *Eur J Cancer* 1991; **27**: 131-135 [PMID: 1827274 DOI: 10.1016/0277-5379(91)90469-T]

4 **American Cancer Society**. Breast cancer overview, 2012. Available from: URL: http://www.cancer.org/cancer/breastcancer/overviewguide/index

5 **Megdal SP**, Kroenke CH, Laden F, Pukkala E, Schernhammer ES. Night work and breast cancer risk: a systematic review and meta-analysis. *Eur J Cancer* 2005; **41**: 2023-2032 [PMID: 16084719 DOI: 10.1016/j.ejca.2005.05.010]

6 **Risius D**, Thelwell R, Wagstaff C, Scurr J. Influential factors of bra purchasing in older women. *J Fash Market Manag* 2012; **16**: 366-380 [DOI: 10.1108/13612021211246099]

7 **Liao CS**, Lee CW. The application of codesign in new bra product innovations. *Int J Cloth Sci Technol* 2010; **22**: 211-227 [DOI: 10.1108/09556221011018676]

8 **Singer S**, Grismaijer S. Dressed to kill: The link between breast cancer and bras. New York: Avery Publishing Group, 1995

9 **Zhu LH**, Li SF. A case-control study on risk factors of female breast cancer in Zhengzhou City. *J Chin Prim Med Pharm* 2006; **13**: 679-680

10 **Lee MM**, Chang IY, Horng CF, Chang JS, Cheng SH, Huang A. Breast cancer and dietary factors in Taiwanese women. *Cancer Causes Control* 2005; **16**: 929-937 [PMID: 16132802 DOI: 10.1007/s10552-005-4932-9]

11 **Zhang AQ**, Xia JH, Wang Q, Li WP, Xu J, Chen ZY, Yang

JM. Risk factors of breast cancer in women in Guangdong and the countermeasures. *Nanfang Yike Daxue Xuebao* 2009; **29**: 1451-1453 [PMID: 19620080]
12 **Shen M**, Gu JF, Dong JY. Risk factors of breast cancer incidence and nursing strategies. *Chin J Gen Pract* 2014; **12**: 782-785
13 **Yao XY**, Ni SS, Zhou J, Hu HY, Li LL, Wan F, Wang YK, Chen YD. [A case-control study on risk factors of female breast cancer in Zhejiang province]. *Zhejiang Daxue Xuebao Yixueban* 2012; **41**: 512-518 [PMID: 23086643 DOI: 10.3785/j.issn.1008-9292.2012.05.008]
14 **Liu L**, Ding H, Jia ZX, Qiu LP, Tao MF, Yan M. A case-control on risk factors of female breast cancer in Beijing. *Matern Child Health Care China* 2014; **21**: 3407-3408 (in Chinese) [DOI: 10.7620/zgfybj.j.issn.1001-4411.2014.21.16]
15 **Liu SP**, Qiu YX, Li YL. A study on risk factors of breast cancer in the Zhoushan island. *J Hebei Med Univ* 2011; **32**: 1336-1338
16 **Hu PX**, Lin FC. A case-control study on risk factors of female breast cancer in the city of Shenzhen: A report of 95 cases. *New Med* 2011; **42**: 291-294
17 **Harris RP**, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, Atkins D. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med* 2001; **20**: 21-35 [PMID: 11306229 DOI: 10.1016/S0749-3797(01)00261-6]
18 **Higgins J**, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0, 2011. Available from: URL: http://handbook.cochrane.org/
19 **Chen L**, Malone KE, Li CI. Bra wearing not associated with breast cancer risk: a population-based case-control study. *Cancer Epidemiol Biomarkers Prev* 2014; **23**: 2181-2185 [PMID: 25192706]
20 **Feng Y**, Wu JG, Shi LY. A case-control study with multivariate analysis for 262 cases of female breast cancer. *Chin J Epidemiol* 2005; **26**: 925
21 **Chen HM**. A case-control study on risk factors of breast cancer for women in Shantou in recent years. *Med Inf* 2011; **24**: 1607-1608
22 **American Cancer Society**. Breast cancer, 2012. Available from: URL: www.cancer.org/acs/groups/cid/documents/webcontent/003090-pdf.pdf
23 **Kumar A**. Burn the bra! (and men's tight underpants too): compromised 'chaotic' cooling by constrictive clothing in the causation of testicular and breast cancers. *Med Hypotheses* 2009; **73**: 1079-1080 [PMID: 19833445 DOI: 10.1016/j.mehy.2008.12.007]
24 **King CR**. Bras, breast carcinoma, and cryptorchid testis. *Lancet* 1979; **1**: 45 [PMID: 83491 DOI: 10.1016/S0140-6736(79)90486-0]
25 **Monson N**. 7 cancer facts you need to know now. *Redbook* 2001; **197**: 36-38
26 **Levine H**. Breast cancer update: What really reduces your risk. *Redbook* 2002; **199**: 36-43
27 **Swenson KK**, Nissen MJ, Leach JW, Post-White J. Case-control study to evaluate predictors of lymphedema after breast cancer surgery. *Oncol Nurs Forum* 2009; **36**: 185-193 [PMID: 19273407 DOI: 10.1188/09.onf.185-193]
28 **Stenson J**. What doesn't cause cancer? *Shape* 2005; **24**: 106-108
29 **Al-Dubai SA**, Qureshi AM, Saif-Ali R, Ganasegeran K, Alwan MR, Hadi JI. Awareness and knowledge of breast cancer and mammography among a group of Malaysian women in Shah Alam. *Asian Pac J Cancer Prev* 2011; **12**: 2531-2538 [PMID: 22320951]
30 **Pagano M**, Gauvreau K. Principles of Biostatistics. 2nd ed. Pacific Grove, California: Duxbury, 2000

P- Reviewer: Khajehei M, Yao HR **S- Editor**: Ji FF
L- Editor: A **E- Editor**: Wu HL

Published by **Baishideng Publishing Group Inc**
8226 Regency Drive, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

ISSN 2308-3840 (online)

World Journal of *Meta-Analysis*

World J Meta-Anal 2015 October 26; 3(5): 206-224

Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

A peer-reviewed, online, open-access journal of meta-analysis

Editorial Board

2013-2018

The *World Journal of Meta-Analysis* Editorial Board consists of 400 members, representing a team of worldwide experts in clinical meta-analysis research. They are from 41 countries, including Argentina (2), Australia (3), Austria (1), Belgium (5), Brazil (11), Canada (16), Chile (3), China (121), Croatia (1), Egypt (1), Finland (4), France (2), Germany (9), Greece (11), Hungary (2), India (12), Iran (2), Ireland (1), Israel (2), Italy (41), Japan (5), Lithuania (1), Malaysia (1), Netherlands (8), New Zealand (1), Norway (1), Peru (1), Poland (4), Portugal (6), Romania (1), Saudi Arabia (4), Singapore (3), South Africa (1), South Korea (7), Spain (8), Sri Lanka (2), Sudan (1), Switzerland (2), Thailand (4), Turkey (3), United Kingdom (22), and United States (64).

Alessandro Cucchetti, *Bologna*
Gianfranco Damiani, *Rome*
Fabrizio D'Ascenzo, *Turin*
Massimo Del Fabbro, *Milano*
Valeria Fadda, *Florence*
Alessandro Fancellu, *Sassari*
Giuseppe Ferrante, *Milan*
Virginia Festa, *Rome*
Francesco Fiorica, *Ferrara*
Guglielmo Giraldi, *Rome*
Jenny Guidi, *Bologna*
Lorenzo Loffredo, *Rome*
Andrea Messori, *Firenze*
Eliano Pio Navarese, *Bydgoszcz*
Stefano Omboni, *Solbiate Arno*
Alvisa Palese, *Udine*
Stefano Palomba, *Reggio Emilia Stefano*
Carlo Perricone, *Rome*
Mario Petretta, *Naples*
Alessandro Pezzini, *Brescia*
Gianluca Pontone, *Milan*
Paolo Emilio Puddu, *Rome*
Andrea Rognoni, *Novara*
Giuseppe Scalabrino, *Milan*
Città della Salute e della Scienza, *Turin*
Fabrizio Sgolastra, *L'Aquila*
Maria Lucia Specchia, *Rome*
Stefano Trastulli, *Terni*
Fabio Tine, *Palermo*
Nereo Vettoretto, *Chiari*
Alberto Vianello, *Perugia*
Luigi Zorcolo, *Cagliari*

Japan

Nguyen Tien Huy, *Nagasaki*
Hiroharu Kamioka, *Tokyo*
Koji Kawakami, *Kyoto*
Keitaro Matsuo, *Nagoya*
Kazushi Okamoto, *Nagoya*

Lithuania

Edmundas Kadusevicius, *Kaunas*

Malaysia

SP Pani, *Ipoh*

Netherlands

Michel van den Bekerom, *Amsterdam*
Dimitra Dodou, *Delft*
D Haverkamp, *Amsterdam*
Vassilios Koussoulas, *Drachten*
BJ Polder, *Nijmegen*
Theo Stijnen, *Leiden*
RNM Weijers, *Amsterdam*
Joost de Winter, *Delft*

New Zealand

Shaofeng Li, *Auckland*

Norway

Eivind Berge, *Oslo*

Peru

Rafael Bolaños Díaz, *Lima*

Poland

Maciej Banach, *Lodz*
Krzysztof Jonderko, *Sosnowiec*
Jolanta Lissowska, *Warsaw*
Maciej Plaszewski, *Warsaw*

Portugal

Daniel Caldeira, *Lisboa*
J Costa, *Lisbon*
Ana Miguel, *Coimbra*
Manuel Morgado, *Covilh*
Bárbara Peleteiro, *Porto*
Rui Torres, *Paredes*

Romania

Fratila Ovidiu, *Oradea*

Saudi Arabia

Hazem M Al-Mandeel, *Riyadh*
Ezzeldin M Ibrahim, *Jeddah*
Mutahir A Tunio, *Riyadh*
Hayfaa A Wahabi, *Riyadh*

Singapore

Nikos LD Chatzisarantis, *Singapore*
Edwin Choon Wyn Lim, *Singapore*
Roger Ho, *Singapore*

South Africa

Alaine Umubyeyi Nyaruhirira, *Pretoria*

South Korea

Jung-Hee Kim, *Cheonan*
Hyangsook Lee, *Seoul*
Myeong Soo Lee, *Daejeon*
Chi-Un Pae, *Bucheon*
Yong Hyun Park, *Yangsan*
Jae Hong Seo, *Seoul*
Yong Sang Song, *Seoul*

Spain

Pablo Avanzas, *Oviedo*
Joan Cid, *Barcelona*
Joaquin de Haro, *Madrid*
Joan Guardia-Olmos, *Barcelona*
Nabil Halaihlel, *Zaragoza*
JA Monge-Argilés, *Alicante*
Raul Moreno, *Madrid*
Inés Velasco, *Aracena*

Sri Lanka

Ranil Jayawardena, *Colombo*
Priyanga Ranasinghe, *Colombo*

Switzerland

Jay P Singh, *Zurich*
Giorgio Treglia, *Bellinzona*

Sudan

Samir MH Shaheen, *Khartoum*

Thailand

Chuenjid Kongkaew, *Phitsanulok*
Manop Pithukpakorn, *Bangkok*
Piyamitr Sritara, *Bangkok*
Surasak Saokaew, *Phayao*

Turkey

Nese Demirturk, *Afyonkarahisar*
Nilüfer Ozabaci, *Eskisehir*
Ilke Sipahi, *Istanbul*

United Kingdom

Omar M Aboumarzouk, *Wales*
Abeer Al-Namankany, *London*
Ernest A Azzopardi, *Cardif*
Umberto Benedetto, *Cambridge*
Joanne Brooke, *London*
Noriko Cable, *London*
David Chan, *Cardiff*
YC Cheong, *Southampton*
Andrew Currie, *Harrow*
G Nabi, *Dundee*

Lesley A Anderson, *Belfast*
Valentina Gallo, *London*
Gianpiero Gravante, *Leicester*
Peter N Lee, *Surrey*
Igho Onakpoya, *Oxford*
Ashish Pradhan, *Huntingdon*
Evridiki Patelarou, *London*
Jian-Qing Shi, *Newcastle*
Surendra P Singh, *Wolverhampton*
Natalie Taylor, *Leeds*
Yousef Shahin, *Hull*
Zheng Ye, *Cambridge*

United States

L Joseph Su, *Rockville*
Olusola Adesope, *Pullman*
Mike Allen, *Milwaukee*
Bhupinder Anand, *Bellaire*
Stephen Aronoff, *Philadelphia*
KoKo Aung, *San Antonio*
William L Baker, *Farmington*
Moritz C Wyler von Ballmoos, *Milwaukee*
Matthew L Bechtold, *Columbia*
Atul Bhardwaj, *Hershey*
Somjot S Brar, *Los Angeles*
Hui Cai, *Nashville*
Subhash Chandra, *Towson*
Wen-Pin Chang, *Omaha*
Yong Chen, *North Wales*
Myunghan Choi, *Phoenix*
John Coverdale, *Houston*
Prakash C Deedwania, *Fresno*
Eugene Demidenko, *Hanover*
Hong-Wen Deng, *New Orleans*
Eric M Deshaies, *Syracuse*
Tao Fan, *Whitehouse Station*
Shinga Feresu, *Bloomington*
Janvier Gasana, *Miami*
Kaveh Hajifathalian, *Boston*
Mohammad Obaidul Hoque, *Baltimore*
Larissa R Brunner Huber, *Charlotte*
Imran H Iftikhar, *Columbia*
Vijayvel Jayaprakash, *Buffalo*
Xuezhi Jiang, *Weat Reading*
Shuo Jiao, *Seattle*
Evelyn Johnson, *Boise*
Le Kang, *Silver Spring*
SR Kapadia, *Cleveland*
Lior Katz, *Houston*
Daniel M Laskin, *Richmond*
Yu Liang, *Foster City*
Paul Ellis Marik, *Norfolk*
Lynne V McFarland, *Seattle*
Marcovalerio Melis, *New York*
Brian J Miller, *Augusta*
Pavlos Msaouel, *New York*
Joshua E Muscat, *Hershey*
Chee Yuan Ng, *Loma Linda*
Nghi C Nguyen, *Saint Louis*
Brandi S Niemeier, *Whitewater*
Thomas D Parsons, *Denton*
Nidal Abi Rafeh, *New Orleans*
Praveen Roy, *Marshfield*
Ali Salavati, *Philadelphia*
Ankur Sethi, *Chicago*
Tatyana A Shamliyan, *Minneapoli*
Qian Shi, *Rochester*
Zhongjie Shi, *Philadelphia*
Param Puneet Singh, *Chicago*
KV Slavin, *Chicago*
Ali El Solh, *Buffalo*
Jielin Sun, *Winston-Salem*
Richard G Trohman, *Chicago*
Laurah Turner, *Cincinnati*
Sheila Wilhelm, *Detroit*
Alex K Wong, *Los Angeles*
Xiaohui Xu, *Gainesville*
Lu Yin, *Nashville*

WJMA
World Journal of Meta-Analysis

Contents

Bimonthly Volume 3 Number 5 October 26, 2015

SYSTEMATIC REVIEW

META-ANALYSIS

NAME OF JOURNAL
World Journal of Meta-Analysis

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Bimonthly

EDITOR-IN-CHIEF
Giuseppe Biondi-Zoccai, MD, Assistant Professor, Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
World Journal of Meta-Analysis
Room 903, Building D, Ocean International Center,
No. 62 Dongsihuan Zhonglu, Chaoyang District,
Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
8226 Regency Drive,
Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
October 26, 2015

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i5.206

World J Meta-Anal 2015 October 26; 3(5): 206-214
ISSN 2308-3840 (online)

Traditional Chinese manual acupuncture for management of obesity: A systematic review

Kang Xiao Li, Angela Weihong Yang, Charlie CL Xue, George Binh Lenon

Kang Xiao Li, Angela Weihong Yang, Charlie CL Xue, George Binh Lenon, Traditional and Complementary Medicine Research Program, Health Innovations Research Institute, School of Health Sciences, RMIT University, Bundoora, VIC 3083, Australia

Author contributions: Li KX and Yang AW made equal contribution; Li KX and Yang AW searched, entered and analysed the data and draft the manuscript; Xue CC and Lenon GB checked data and revised the manuscript; all authors approved the final manuscript.

Conflict-of-interest statement: The authors declare that they have no conflict-of- interests.

Data sharing statement: No additional data are available.

Correspondence to: George Binh Lenon, PhD, Traditional and Complementary Medicine Research Program, Health Innovations Research Institute, School of Health Sciences, RMIT University, Crn of Plenty Rd and Clement Drive, Bundoora, VIC 3083, Australia. george.lenon@rmit.edu.au
Telephone: +61-3-99256587
Fax: +61-3-99257178

Received: January 20, 2015
Peer-review started: January 30, 2015
First decision: March 6, 2015
Revised: June 5, 2015
Accepted: September 16, 2015
Article in press: September 18, 2015
Published online: October 26, 2015

Abstract

AIM: To evaluate the effectiveness and safety of acupuncture for the treatment of obesity by reviewing currently available randomised controlled trials.

METHODS: This review followed the Cochrane Handbook for Systematic Reviews of Interventions. Fifteen English and three Chinese databases were searched from their respective inceptions until July 2014. Key words used in the search consisted of acupuncture, needles, obesity, overweight, randomised trial and their synonyms. The risk of bias of included studies was assessed. The differences in effect size between acupuncture and control (including sham, no treatment, western medicine and dietary therapy/exercise) groups were compared using Cochrane Collaboration's RevMan 5.3 software.

RESULTS: Two thousand six hundred and twenty-one records were identified; after full-text articles assessed for eligibility, 9 of them met inclusion criteria. Majority of included studies had unclear or high risk of bias across all domains. All included studies had high or unclear risk of bias in randomisation, blinding and outcome data. Meta-analysis showed that acupuncture was more effective for reducing body weight and body mass index than no treatment group. Manual acupuncture was also superior to dietary therapy alone for decreasing body weight. With dietary therapy as co-intervention, combined acupuncture group achieved lower body mass index than combined sham acupuncture group or dietary therapy alone group at the end of treatment period. No severe adverse events from acupuncture group were reported from all included studies.

CONCLUSION: Due to the poor quality of included studies the effectiveness of acupuncture cannot be concluded. Better-designed, large-scale, randomised, sham-controlled clinical trials with long-term follow-up are needed.

Key words: Acupuncture; Obesity; Weight loss; Overweight; Body weight; Body mass index; Randomised clinical trial

Core tip: This systematic review identified the benefit of traditional Chinese manual acupuncture in the management of obesity. However, the effectiveness cannot be confirmed due to poorly-design randomised clinical trials.

Li KX, Yang AW, Xue CC, Lenon GB. Traditional Chinese manual acupuncture for management of obesity: A systematic review. *World J Meta-Anal* 2015; 3(5): 206-214 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i5/206.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i5.206

INTRODUCTION

Obesity is a medical condition involving more than one kind of biological basis or metabolic disease[1]. It is often accompanied with fatigue, lassitude, hydrosis, various neurosis, headache, palpitation, and abdominal distension[2]. The World Health Organisation (WHO) estimates more than one billion overweight adults worldwide out of which at least 500 million are obese[3]. The prevalence of obesity is a major public health concern since obesity is often associated with cerebrovascular and cardiovascular diseases, hypertension, arteriosclerosis, diabetes and speeding up the aging process[4]. Obesity places considerable economic burden on already strained healthcare systems as it reduces quality of life and leads to premature mortality[5].

However, the treatment for obesity without obvious causes is challenging. Currently, the main management of obesity includes dietary therapy, physical exercise, behaviour treatment, pharmacotherapy (such as Sibutramine and Orlistat) and surgery[6]. However, it is difficult to change behaviours to maintain the weight loss by the same dietary changes and physical exercises in a long term[7]. Pharmacotherapy often associates with undesired effects[8]. Surgery is usually considered only when there are serious medical conditions, for example, a high risk of obesity-related illness and death[9]. Recently, there is a global trend for obesity sufferers to seek treatment from complementary and alternative medicine including acupuncture.

Acupuncture for weight management has been practicing for many centuries. It involves inserting thin and solid needles to certain points on the body to achieve therapeutic effects. A number of reviews involving different types of needling and needle stimulation techniques (including traditional Chinese acupuncture, electro-acupuncture, dry needling, ear-acupuncture and transcutaneous electrical nerve stimulation) for weight loss have been published[10-12]. However, the effects of traditional Chinese manual acupuncture are unclear as its effects have not been reviewed separately. Traditionally, acupuncture applies manual stimulation based on Chinese medicine theory such as point selections and *De Qi* sensation without electrical stimulation. Therefore, this study aimed to evaluate the therapeutic benefits and safety of traditional manual stimulation acupuncture for the management of obesity by systematically reviewing the currently available randomised clinical trials (RCTs).

MATERIALS AND METHODS

This review was conducted following the instructions specified in the Cochrane Handbook for Systematic Reviews of Interventions 5.2[13].

Literature search

We searched 15 electronic English and three Chinese databases, including Cochrane Central Register of Controlled Trials, PubMed, EMBASE, CINAHL, Science Direct, LILACS (Latin American and Caribbean Health Sciences), ProQuest, Web of Science, Informit, Psycinfo, Blackwell Synergy, KoreaMed, INDMED, Ingenta, mRCT, VIP Information (www.cqvip.com), China National Knowledge Infrastructure (www.cnki.net), and Wanfang Data (www.wanfangdata.com), from their respective inceptions to July 2014. Key words used consisted of acupuncture, needles, obesity, overweight, weight loss, randomised trials and their synonyms. We also hand searched the reference lists of clinical trials and reviews as well as the conference proceedings of the World Congress of Chinese Medicine from 2003 onwards to identify the relevant studies.

Study selection

RCTs, with or without blinding, were considered if they investigated the effects of manual acupuncture for adult patients of any gender, diagnosed as obesity or overweight according to WHO criteria [overweight: body mass index (BMI) $\geqslant$ 25; obesity: BMI $\geqslant$ 30]. A study was included when it compared traditional manual acupuncture with sham, no treatment, Western medicine, dietary therapy and/or exercise. Co-intervention was allowed as long as the same co-intervention was used in both arms.

Studies of non-traditional acupuncture such as electro-acupuncture, dry needling, ear-acupuncture or transcutaneous electrical nerve stimulation were excluded.

Two reviewers (Li KX and Yang AW) independently screened the titles and abstracts. When lack of adequate information for judgement identified, full texts were obtained for further assessment. Any discrepancies between two reviewers were resolved by discussion or consultation with the third party (Lenon GB).

Data extraction

A predefined data extraction form was used to extract the data of included studies, including characteristics of participants, interventions and outcomes. Two reviewers

(Li KX and Yang AW) independently extracted the data and discussed the differences. Unresolved issues were consulted with the third party (Lenon GB).

Risk of bias assessment

We assessed the risk of bias of each study including random sequence generation, allocation concealment, blinding of participants and outcome assessors, incomplete outcomes, selective reporting and other bias (*e.g.*, baseline imbalance). Each bias was classified as low, unclear or high risk.

Statistical analysis

The primary outcomes included effectiveness rate, body weight and BMI and the secondary outcome was adverse events. The meta-analysis on effectiveness rate, body weight and BMI were performed using Review Manager 5.3 developed by the Cochrane Collaboration[14]. The effectiveness rate was calculated as (scores before treatment - scores after treatment)/scores before treatment × 100%. The effects assessed immediately after the treatment period were considered as short-term effects whilst the effects followed up for more than three months after the last treatment were regarded as long-term. The effectiveness rate was analysed by risk ratio (RR). Body weight and BMI were calculated by difference in means (MD). All were with 95%CI.

RESULTS

A total of 2621 potential studies were identified following search strategy and 2427 of them were excluded after screening titles and abstracts. Full texts were obtained for the remaining 194 studies and nine of them met inclusion criteria[15-23]. The main reasons for exclusion were non-obesity, non-acupuncture and non-RCT. Figure 1 illustrates the study selection process and provides the exclusion reasons.

Characteristics of included studies

Eight included studies were conducted in mainland China[15,16,18-23] and one in Turkey[17]. Except two studies published in English[17,23], the rest seven papers were published in Chinese. The total sample size is 556, ranging from 33 to 118 with an average of 61 per trial. A total of 110 males and 446 females were involved in the included studies. All the participants in the RCTs were clearly diagnosed as obesity and/or overweight. The duration of treatment lasted from 30 d to 65 d and the treatment sessions varied from 10 to 40.

Two studies used three-arm design[18,23] and the rest RCTs were two-arm trials. Although Xu *et al*[23]'s trial claimed it was a three-arm trial, the same treatment protocol was used in the two treatment groups by two different acupuncturists. In terms of interventions, all of them involved manual acupuncture in the treatment groups. Seven RCTs adopted standard acupoints to treat all participants in the acupuncture group. Only two studies[16,22] applied different groups of acupoints to treat participants with different syndromes (a complex of symptoms and signs according to Chinese medicine theory). Acupuncture treatment in seven trials obtained *De Qi* sensation[15-21]. Across the nine trials, different acupuncture point formulae were prescribed. The top five frequently used acupoints were Zusanli (ST36, 6 trials), Sanyinjiao (SP6, 6 trials), Zhongwan (CV12, 5 trials), Xuehai (SP10, 4 trials) and Tianshu (ST25, 4 trials).

In the control groups, comparators were various, including sham acupuncture, no treatment, Western medicine, dietary therapy and exercise. Five RCTs[16,17,19,22,23] and one arm of Luo *et al*[18]'s study did not involve any co-intervention. One compared manual acupuncture with sham acupuncture[17]. Two studies[22,23] and one arm of Luo *et al*[18]'s study compared manual acupuncture with no treatment and another two studies[16,19] compared acupuncture with dietary therapy. The remaining three RCTs involved co-interventions in both groups. Acupuncture was assessed as an adjunct treatment to dietary therapy and/or exercise[15,20,21] when compared to dietary therapy and/or exercise alone.

With regard to outcome measures, five RCTs used the effectiveness rate to present the effects of acupuncture on reduction of body weight of people with obesity[15,16,18,19,22]. Five RCTs evaluated BMI[17,18,20-22] and four studies assessed body weight[15,17,20,22] at the end of treatment period. Only one included study monitored adverse events[15].

The detailed characteristics of participants, interventions and outcome measures are summarised in Table 1.

Summary of risk of bias assessment

Randomisation was claimed in all included RCTs. The risk of bias assessment in illustrated in the Figures 2 and 3. However, only four trials clearly provided appropriate methods for random sequence generation, one using three colour cards[18], another using computer-generated randomisation numbers[17] and the other two applying random number table[20,21]. The rest of five RCTs did not provide the methods for randomisation. All but one did not state the methods for allocation concealment. Only Güçel *et al*[17]'s trial indicated that they used urn for allocating the randomisation numbers. The performance biases of seven RCTs were classified as high risk because they applied different forms of treatments to two groups which made blinding of participants impossible[15,16,18-20,22,23]. In addition, two of them only involved one author and this implied that it was unlikely to blind the outcome assessors in these trials[15,20]. One study clearly indicated that all participants completed the trial[17]. Another study[15] indicated two dropouts from the control group due to side effects. The rest of seven trials did not provide sufficient information for assessment. Dong[15]'s trial mentioned it followed up the participants for three months however, no results at the end of follow-up period were reported. In Wang *et*

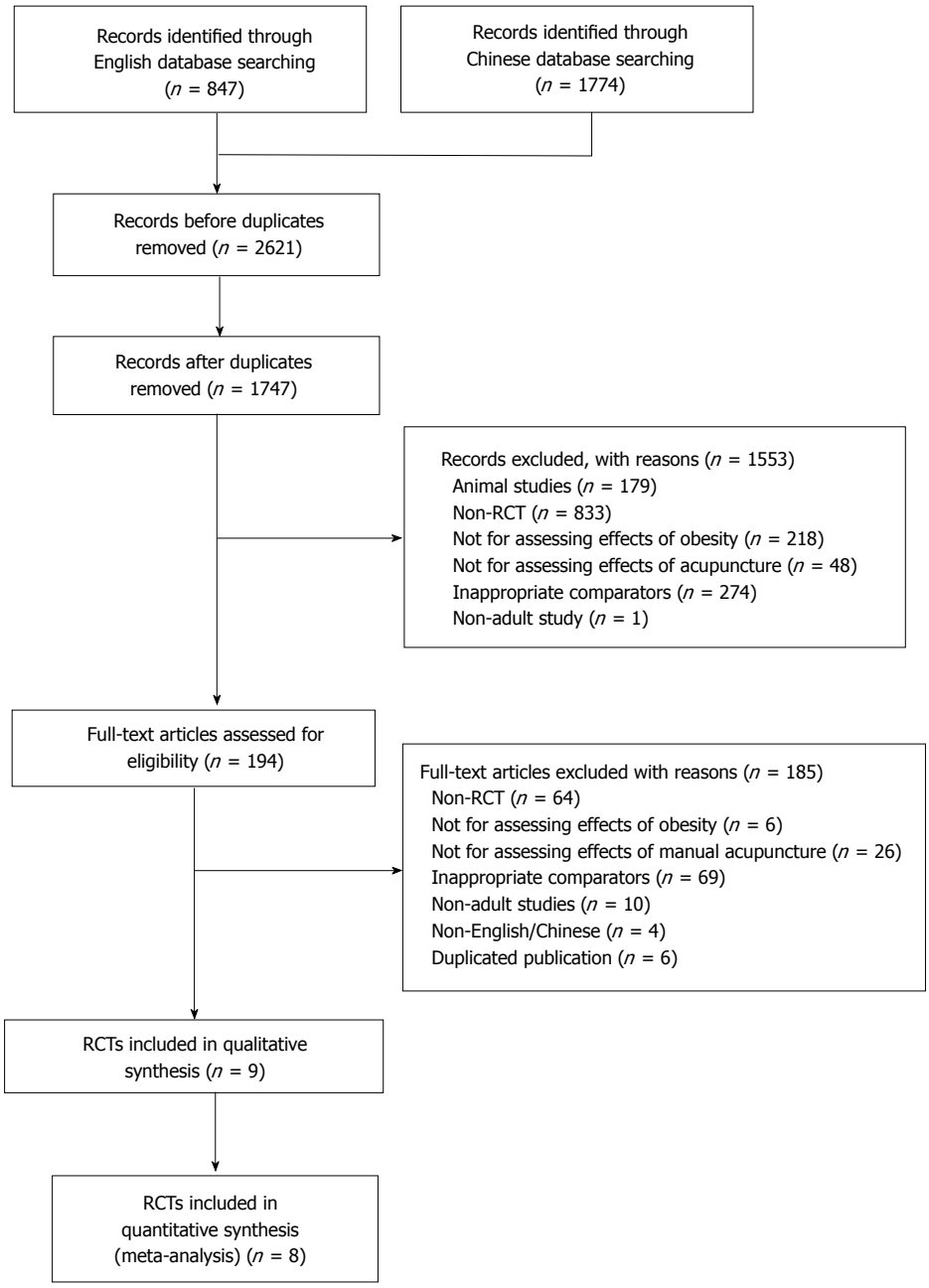


Figure 1 Flow diagram of study selection process. RCT: Randomised clinical trial.

al[22]'s study, the results of body weight and BMI for the control group were not reported. The remained seven included studies reported all the outcome measures listed in the methods section of their published papers. Four trials did not report if the baseline data were comparable[15,19,22,23].

Clinical effectiveness

Effectiveness rate: Five RCTs assessed the short-term effects of acupuncture[15,16,18,19,22]. Long-term effects of acupuncture were not assessed due to lack of data in all the included studies. As one study used the different criteria for effectiveness rate, its data were not included in the meta-analysis[19]. The pooled data from two studies[18,22] showed that acupuncture was significantly more effective than no treatment at the end of treatment period (RR = 8.28; 95%CI: 1.04-65.97). When compared to dietary therapy, Fan *et al*[16]'s study indicated that acupuncture had better effects at the end of 20 treatments (RR = 1.93; 95%CI: 1.54-2.40). With dietary therapy and exercise as co-intervention[15], acupuncture group demonstrated similar effects to western medication (sibutramine hydrochloride) group at the end of 40-d treatment (RR = 1.09; 95%CI: 0.42-2.85).

Body weight: Only three studies reported body weight of two groups at the end of treatment period[15,17,20]. No significant difference was found in the three comparisons. There was no significant difference between real and sham acupuncture after 10 treatments[17] (MD: -4.40; 95%CI: -11.21-2.41). Similarly, combination of

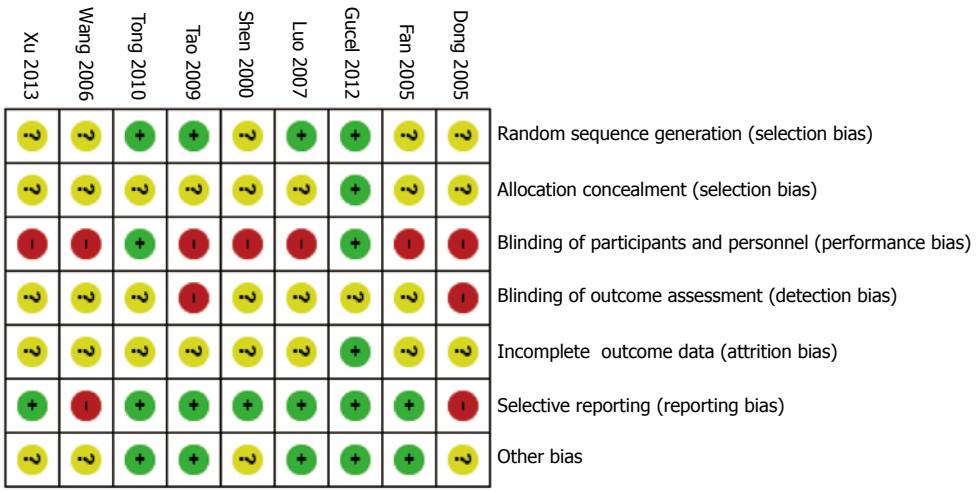


Figure 2 Risk of bias summary of included studies.

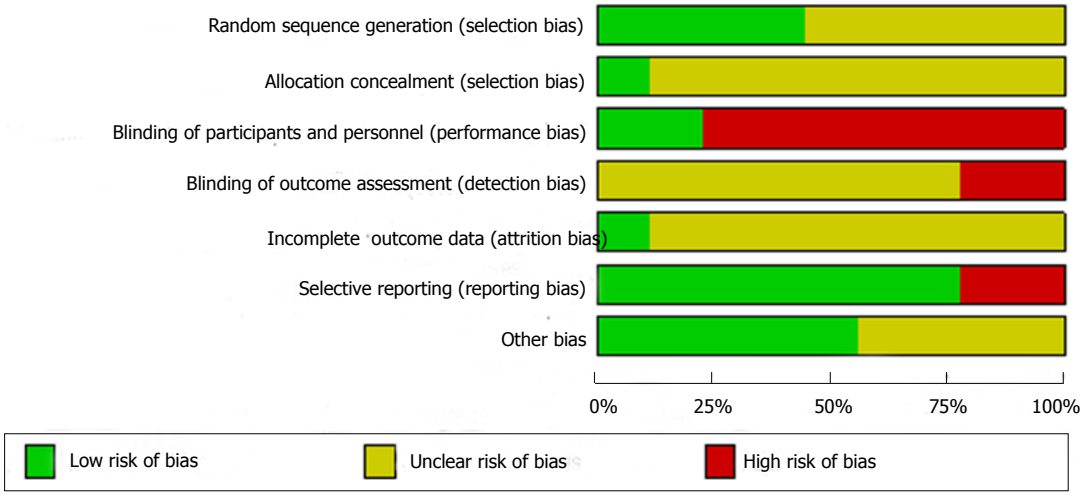


Figure 3 Risk of bias graph of included studies.

acupuncture and dietary therapy showed no difference in body weight from dietary alone group at the end of 20 sessions (MD: -0.93; 95%CI: -3.75-1.89)[20]. With dietary therapy and exercise as co-intervention, the findings from Dong[15]'s trial indicated that the body weight of patients treated by acupuncture combined group was not superior to that of patients treated by western medication combined group after 40-d treatment (MD: -2.44; 95%CI: -10.57-5.69).

BMI: BMI data from four trials were included in the meta-analysis[17,18,20,21]. There was no significant difference in BMI between real and sham acupuncture at the end of 10 treatments (MD: -2.56; 95%CI: -5.37-0.25)[17]. The pooled data from two studies[18,23] showed that acupuncture was more effective for BMI than no treatment (MD: -1.88; 95%CI: -2.67 to -1.08). With dietary therapy as co-intervention, real acupuncture group was superior to sham acupuncture group for BMI at the end of five-week treatment (MD: -1.97; 95%CI: -3.15 to -0.79)[21]. Combining acupuncture with dietary therapy was more effective on BMI after 20-session treatment than dietary therapy alone group (MD: -0.55; 95%CI: -1.03 to -0.07)[20].

Adverse events reported in the included studies

Only one out of nine included studies reported adverse events. Dong[15]'s study mentioned that 16.67% of patients in treatment group and 42.86% in control group occurred adverse events, including constipation, dizziness, insomnia, palpitation and high blood pressure. Two patients dropped out from the control group due to side effects.

DISCUSSION

This project reviewed the effects of traditional Chinese manual acupuncture for the treatment of obesity. The focus on the manual acupuncture avoids other confounding variables involved in the intervention. Although the meta-analysis did not show any significant difference between two groups for body weight at

Table 1 Characteristics of participants and interventions in the included studies

	Ref.	Setting	Randomised sample size		Intervention		Outcome measure	Adverse events
			T (*n*, age)	C (*n*, age)	T	C		
Acu *vs* sham acu	Güçe *et al*[17]	Clinic, outpatients	20, NS M/F: 0/20	20, NS M/F: 0/20	Acupoints: Hegu (LI4), Shenmen (HT7), Zusanli (ST36), Neiting (ST44), Sanyinjiao (SP6) Insertion depth: 5-10 mm; Duration: 5 wk; Sessions: 10; Frequency: 2/wk; Retention time: 20 min; Follow-up: NS	Sham acu: Same procedures as T group	BW BMI Blood test	NS
Acu *vs* no treatment	Luo *et al*[18]	NS	T1: 20, 21-48 yr; M/F: 8/12 T2: 20, 21-48 yr; M/F: 10/10	20, 21-48 yr M/F: 6/14	T1: Electro acu; T2: Manual acu Acupoints (2 groups used alternatively): (1) Liangqiu (ST34), Xuehai (SP10), Zhigou (TE6), Waiguan (TE5), Sanyinjiao (SP6), Zusanli (ST36), Fujue (SP14), Daimai (GB26), Tianshu (ST25), Wailing (ST26), Guanyuan (CV4), Taixi (KI3), Shangjuxu (ST37) (2) Gongsun (SP4), Neiting (ST44), Fenglong (ST40), Neiguan (PC6), Yinlingquan (SP9), Zusanli (ST36), Shuifen (CV9), Zhongwan (CV12), Daju (ST27), Shuidao (ST28), Guanyuan (CV4), Taichong (LR3), Quchi (LI11), Xiajuxu (ST39) Insertion depth: 10-35 mm; Duration: 65 d Sessions: 27; Frequency: once every other day; Retention time: 30 min; Follow-up: NS	No treatment	ER BMI WHR Blood test	NS
	Wang *et al*[22]	NS	18, 20-55 yr M/F: NS	18, 20-55 yr M/F: NS	Acupoints: (1) Stomach heat: 10 abdominal points + Quchi (LI11), Neiting (ST44), Shangjuxu (ST37), Pishu (BL20), Xuehai (SP10), Fenglong (ST40) (2) Qi deficiency: 10 abdominal points + Shangjuxu (ST37), Sanyinjiao (SP6), Yinlingquan (SP9), Pishu (BL20), Xinshu (BL15), Taixi (KI3), Guanyuan (CV4) Insertion depth: NS; Duration: 30 d; Sessions: 30; Frequency: 1/d; Retention time: 50 min; Follow-up: NS	No treatment	BW BMI F% Electrogas-trography	NS
	Xu *et al*[23]	Hospital, outpatients	T1: 15, 29-63 yr; M/F: 0/15 T2: 15, 29-63 yr; M/F: 0/15	15, 29-63 yr M/F: 0/15	T1 and T2 on same acupoints: Zhongwan (CV12), Shuifen (CV9), Qihaishu (BL24), Guanyuanshu (BL26), Shuidao (ST28), Tianshu (ST25), Zusanli (ST36), Sanyinjiao (SP6) Insertion depth: 1.5 inch; Duration: 40 d; Sessions: 20; Frequency: once every other day Retention time: 30 min; Follow-up: NS	No treatment	BMI Faecal microbial flora	NS

Acu *vs* diet	Fan *et al*[16]	Hospital, outpatients	50, NS M/F: 11/39	50, NS M/F: 10/40	Primary acupoints: Xuehai (SP10), Sanyinjiao (SP6), Tianshu (ST25), Zusanli (ST36), Hegu (LI4) Insertion depth: NS; Duration: 50 d; Sessions: 30; Frequency: 1/d for the first 10 sessions; then once every other day for the last 20 sessions; Retention time: 30 min; Follow-up: NS	Dietary therapy for 50 d	ER Blood test	NS
	Shen *et al*[19]	NS	30, 18-52 yr; M/F: 1/29	30, 18-52 yr; M/F: 1/29	Acupoints: Huatuojiaji T3 to L5 Insertion depth: 1-1.5 cun; Duration: 30 d; Sessions: 30; Frequency: 1/d; Retention time: 30 min; Follow-up: NS	Dietary therapy for 30 d	ER	NS
Acu + diet *vs* sham acu + same diet	Tong *et al*[21]	Hospital, outpatients	76, 18-60 yr M/F: 18/58	42, 18-60 yr M/F: 12/30	Manual acu + dietary therapy Acupoints: Zhongwan (CV12), Zhongji (CV3), Daheng (SP12), Xiawan (CV10), Shimen (CV5), Tianshu (ST25), Liangqiu (ST34), Zusanli (ST36), Yinlingquan (SP9) Insertion depth: subject to De Qi; Duration: 5 wk; Sessions: 12; Frequency: once every other day; Retention time: 30 min; Follow-up: NS	Sham acu + same dietary therapy	BMI Adipose layer score	NS
Acu + diet *vs* Same diet	Tao[20]	NS, outpatients	20, 20-27 yr M/F: NS	20, 20-27 yr M/F: NS	Manual acu + dietary therapy Acupoints: Zhongwan (CV12), Qihai (CV6), Tianshu (ST25), Zhigou (TE6), Fenglong (ST40), Liangqiu (ST34), Yinlingquan (SP9), Zusanli (ST36), Sanyinjiao (SP6), Neiting (ST44) Insertion depth: NS; Duration: 37 d; Sessions: 20; Frequency: 1/d; Retention time: 30 min; Follow-up: NS	Dietary therapy: same as T group	BW BMI Blood test	NS
Acu + diet/ exercise *vs* WM + same diet/exercise	Dong[15]	Hospital, outpatients	30, 18-65 yr M/F: NS	30, 18-65 yr M/F: NS	Manual acu + dietary therapy and exercise Acupoints: Xuehai (SP10), Pishu (BL20), Weishu (BL21), Zhongwan (CV12), Guanmen (ST22) Insertion depth: NS; Duration: 40 d; Sessions: 40; Frequency: 1/d; Retention time: 30 min; Follow-up: 3 mo	WM + same dietary therapy and exercise WM: Sibutramine hydrochloride 10 mg/d for 40 d	ER BW F%	T: 16.67% C: 42.86%; dry mouth, constipation, dizziness, headache, insomnia, palpitation, poor appetite

Acu: Acupuncture; BMI: Body mass index; BW: Body weight; C: Control group; ER: Effectiveness rate; F: Female; F%: Body fat composition; M: Male; NS: Not stated; T: Treatment group; WHR: Waist-hip circumference ratio; WM: Western medicine.

the end of treatment, it demonstrated that manual acupuncture was more effective than no treatment for the reduction of body weight and BMI at the end of treatment period. Acupuncture could also significantly benefit for weight loss in a short term when compared to diet control (no co-intervention). With dietary therapy as co-intervention, combined acupuncture group was more effective for BMI than diet alone group or combined sham acupuncture group.

Future research may consider using the reduction of body weight instead of body weight itself as the primary outcome measure. Effects of acupuncture are probably resulted from improving the sense of wellbeing which may suppress the excessive desire of food[24], regulating metabolism and thereby enable the body to utilise food efficiently instead of storing it as fat[25,26]. It may also stimulate the nervous system to release chemicals which may in turn release other neurotransmitters or

hormones producing the desired effects[27].

Although different acupoint prescriptions were used, the commonly used acupoints were identified, including Zusanli (ST36), Sanyinjiao (SP6), Zhongwan (CV12), Yinlingquan (SP9) and Tianshu (ST25). These acupoints have been widely used in Chinese medicine clinical practice due to their actions to remove dampness-heat and regulate digestive system. In Chinese medicine, obesity is considered resulted from excessively consuming high energy nutrition and causing accumulation of dampness and heat in the body[28]. Thus, the above-mentioned acupoints are appropriate for reducing weight.

Three of nine included RCTs[15,20,21] involved co-intervention (dietary therapy or dietary therapy with exercise) which reflects the fact that acupuncture is commonly used as an adjunct therapy to other modalities for weight loss in clinical practice. Our findings were consistent with others' statement that patients with obesity tend to use combination of different therapies for weight loss[29] (Akilen, 2014 #44). This review could not identify an ideal treatment regimen for treating obesity because there were numerous confounders existed in the included studies, such as sample size (33 to 118 subjects), number of acupoints (5 to 14 points), number of needles (9 to 25 needles), and treatment duration (30 to 65 d).

In addition, the majority of the included studies had methodological issues, that is, high and/or unclear risk of bias in selection, performance, detection and attrition. The quality of evidence was classified as low or very low.

In conclusion, acupuncture showed certain beneficial effects for reducing body weight and BMI in patients with obesity. However, due to the high or unclear risk of bias, the results need to be interpreted with caution. More well-designed, large-scale, randomised, sham-controlled clinical trials with long-term follow-up are needed to assess the efficacy and safety of acupuncture for weight loss.

ACKNOWLEDGMENTS

The authors would like to thank RMIT for Kang Xiao Li's PhD scholarship.

COMMENTS

Background

Chinese acupuncture is becoming popular for weight management in the western countries. Many randomised controlled trials reported positive effects on electro-acupuncture and acupuncture microsystem. Traditional manual acupuncture trials are mainly conducted in China where obesity is growing rapidly; however, its effects have not been systematically evaluated.

Research frontiers

To the best of our knowledge, no systematic review of traditional Chinese manual acupuncture has previously been reported. The aim of this study was to systematically review effectiveness of Chinese manual acupuncture in the management of overweight and obesity.

Innovations and breakthroughs

Prevalence of global obesity is on the rise which leads to many serious health complications. This creates an economic burden and quality of life. Traditional manual acupuncture is promising and safe for individuals with overweight and obesity.

Applications

Acupuncture can be applied to reduce weight in both conditions: Overweight and obesity. It is in everyday practice of traditional Chinese medicine clinics and it is becoming more popular in many countries.

Terminology

Traditional Chinese manual acupuncture refers to body acupuncture without any electrical stimulation. This type of acupuncture has been practicing for thousands of years in China and Asian countries.

Peer-review

The paper is well written and highlights an extraordinary, but interesting theme.

REFERENCES

1 **Bray GA**, Champagne CM. Obesity and the Metabolic Syndrome: implications for dietetics practitioners. *J Am Diet Assoc* 2004; **104**: 86-89 [PMID: 14702589 DOI: 10.1016/j.jada.2003.10.041]

2 **Patterson RE**, Frank LL, Kristal AR, White E. A comprehensive examination of health conditions associated with obesity in older adults. *Am J Prev Med* 2004; **27**: 385-390 [PMID: 15556738 DOI: 10.1016/j.amepre.2004.08.001]

3 **Anderson AS**, Caswell S. Obesity management--an opportunity for cancer prevention. *Surgeon* 2009; **7**: 282-285 [PMID: 19848061 DOI: 10.1016/s1479-666x(09)80005-x]

4 **Sharma AM**, Iacobellis G. Treatment of obesity: a challenging task. *Contrib Nephrol* 2006; **151**: 212-220 [PMID: 16929144 DOI: 10.1159/000095331]

5 **Brown A**, Siahpush M. Risk factors for overweight and obesity: Results from the 2001 National Health Survey. *Public Health* 2007; **121**: 603-613. Available from: URL: http://www.sciencedirect.com/science/article/B73H6-4NYJS2V-1/2/c716d2a7e74abca3d87b2fa65c6048bd

6 **Eby JG**, Colditz GA. Obesity/overweight: Prevention and weight management. In: Kris H, editor. International Encyclopedia of Public Health. Oxford: Academic Press, 2008

7 **Rippe JM**, Crossley S, Ringer R. Obesity as a chronic disease: modern medical and lifestyle management. *J Am Diet Assoc* 1998; **98**: S9-15 [PMID: 9787730]

8 **Ballinger A**, Peikin SR. Orlistat: its current status as an anti-obesity drug. *Eur J Pharmacol* 2002; **440**: 109-117 [PMID: 12007529 DOI: 10.1016/s0014-2999(02)01422-x]

9 **Shippey SH**, Macedonia CR. Surgical treatment of extreme obesity. *Primary Care Update for OB/GYNS* 2003; **10**: 278-283

10 **Lin XM**, Li B, Du YH, Xiong J, Sun P. [Systematic evaluation of therapeutic effect of acupuncture for treatment of simple obesity]. *Zhongguo Zhen Jiu* 2009; **29**: 856-860 [PMID: 19873927]

11 **Cho SH**, Lee JS, Thabane L, Lee J. Acupuncture for obesity: a systematic review and meta-analysis. *Int J Obes* (Lond) 2009; **33**: 183-196 [PMID: 19139756 DOI: 10.1038/ijo.2008.269]

12 **Sui Y**, Zhao HL, Wong VC, Brown N, Li XL, Kwan AK, Hui HL, Ziea ET, Chan JC. A systematic review on use of Chinese medicine and acupuncture for treatment of obesity. *Obes Rev* 2012; **13**: 409-430 [PMID: 22292480 DOI: 10.1111/j.1467-789X.2011.00979.x]

13 **Higgins JPT**, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011

14 **Review Manager (RevMan) [Computer program]**. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014

15 **Dong XY**. Clinical study of acupuncture for the treatment of simple

obesity. *Liaoning Zhongyi Zazhi* 2005; **32**: 1070-1071
16 **Fan YX**, Bai YH, Gui H. Effects of acupuncture for the treatment of simple obesity complicated with hyperlipedia. *Ningxia Yike Daxue Xuebao* 2005; **27**: 144-145
17 **Güçel F**, Bahar B, Demirtas C, Mit S, Cevik C. Influence of acupuncture on leptin, ghrelin, insulin and cholecystokinin in obese women: a randomised, sham-controlled preliminary trial. *Acupunct Med* 2012; **30**: 203-207 [PMID: 22729015 DOI: 10.1136/acupmed-2012-010127]
18 **Luo HL**, Li RH. Dianzhen dui danchunxing feipang shoushu ji zhiliansu de yingxiang. *Acupunct Res* 2007; **32**: 264-267
19 **Shen J**, Hou Q. Effects of needling huatuojiaji for the treatment of simple obesity. *Shanghai Zhenjiu Zazhi* 2000; **19**: 29
20 **Tao HX**. Clinical observation on simple obesity in Stomach-heat and dampness-stagnation type mainly by acupuncture. *Liaoning Zhongyiyao Daxue Xuebao* 2009; **11**: 169-170
21 **Tong J**, Chen JX, Zhang ZQ, Pan Y, Zheng J, Yao H. Therapeutic effect of acupuncture therapy for simple obesity. *Guangzhou Zhongyiyao Daxue Xuebao* 2010; **27**: 579-582
22 **Wang XY**, Li JZ. Study on effects of electrogastrography of simple obesity with acupuncture. *Zhongguo Meirong Yixue* 2006; **15**: 570-572
23 **Xu Z**, Li R, Zhu C, Li M. Effect of acupuncture treatment for weight loss on gut flora in patients with simple obesity. *Acupunct Med* 2013; **31**: 116-117 [PMID: 22961606 DOI: 10.1136/acupmed-2012-010209]
24 **Eich H**, Hannig M, Zimmermann E, Klieser E. Acupuncture in the treatment of psychoactive-drug-induced obesity - An experimental study. *Deutsche Zeitschrift fur Akupunktur* 2005; **48**: 6-11
25 **Hsu CH**, Hwang KC, Chao CL, Lin JG, Kao ST, Chou P. Effects of electroacupuncture in reducing weight and waist circumference in obese women: a randomized crossover trial. *Int J Obes* (Lond) 2005; **29**: 1379-1384 [PMID: 15953937 DOI: 10.1038/sj.ijo.0802997]
26 **Wei QL**, Liu ZC. Treatment of simple obesity with auricular acupuncture, body acupuncture and combination of auricular and body acupuncture. *Zhongguo Linchuang Kangfu* 2004; **8**: 4357-4359
27 **Lacey JM**, Tershakovec AM, Foster GD. Acupuncture for the treatment of obesity: a review of the evidence. *Int J Obes Relat Metab Disord* 2003; **27**: 419-427 [PMID: 12664074]
28 **Deadman P**. A manual of acupuncture. England: Journal of Chinese Medicine Publications, 2009
29 **Akilen R**, Pimlott Z, Tsiami A, Robinson N. The use of complementary and alternative medicine by individuals with features of metabolic syndrome. *J Integr Med* 2014; **12**: 171-174 [PMID: 24735789 DOI: 10.1016/s2095-4964(14)60012-1]

P- Reviewer: Shin BC, Xiong XJ **S- Editor**: Gong XM
L- Editor: A **E- Editor**: Wu HL

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i5.215

World J Meta-Anal 2015 October 26; 3(5): 215-224
ISSN 2308-3840 (online)

META-ANALYSIS

How to impute study-specific standard deviations in meta-analyses of skewed continuous endpoints?

Teresa Greco, Giuseppe Biondi-Zoccai, Marco Gemma, Claude Guérin, Alberto Zangrillo, Giovanni Landoni

Teresa Greco, Laboratorio di Statistica Medica, Biometria ed Epidemiologia "G. A. Maccacaro", Dipartimento di Scienze Cliniche e di Comunità, University of Milan, 20133 Milan, Italy

Teresa Greco, Marco Gemma, Alberto Zangrillo, Giovanni Landoni, Anaesthesia and Intensive Care Department, IRCCS San Raffaele Scientific Institute, 20132 Milan, Italy

Giuseppe Biondi-Zoccai, Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, 04100 Latina, Italy

Giuseppe Biondi-Zoccai, Eleonora Lorillard Spencer Cenci Foundation, 00185 Rome, Italy

Giuseppe Biondi-Zoccai, Meta-analysis and Evidence Based Medicine Training in Cardiology (METCARDIO), 18014 Ospedaletti, Italy

Claude Guérin, Medical Intensive Care, Hospital de La Croix Rousse, 69317 Lyon, France

Author contributions: All authors contributed equally to this work; Greco T conceived the study, participated in its design and coordination, did the analyses and revised the manuscript critically; Biondi-Zoccai G and Gemma M conceived the study, helped in interpretation of data and to draft the manuscript; Guérin C, Zangrillo A and Landoni G conceived the study, participated in its design and coordination, and drafted the manuscript; all authors read and approved the manuscript.

Conflict-of-interest statement: The authors declare that there are no conflicts of interest. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. This work is part of the PhD program in Biomedical Statistics, University of Milan, Italy.

Data sharing statement: Technical appendix, statistical code, and dataset are available in the Supplemental Material and from the corresponding author (at greco.teresa@hotmail.it), who will provide a permanent, citable, and open-access home for the dataset.

Correspondence to: Teresa Greco, MSc, Laboratorio di Statistica Medica, Biometria ed Epidemiologia "G. A. Maccacaro", Dipartimento di Scienze Cliniche e di Comunità, University of Milan, Via Festa del Perdono 7, 20133 Milan, Italy. greco.teresa@hotmail.it
Telephone: +39-02-26436153

Received: January 8, 2015
Peer-review started: January 10, 2015
First decision: June 3, 2015
Revised: June 26, 2015
Accepted: July 24, 2015
Article in press: July 27, 2015
Published online: October 26, 2015

Abstract

AIM: To compare four methods to approximate mean and standard deviation (SD) when only medians and interquartile ranges are provided.

METHODS: We performed simulated meta-analyses on six datasets of 15, 30, 50, 100, 500, and 1000 trials, respectively. Subjects were iteratively generated from one of the following seven scenarios: five theoretical continuous distributions [Normal, Normal (0, 1), Gamma, Exponential, and Bimodal] and two real-life distributions of intensive care unit stay and hospital stay. For each simulation, we calculated the pooled estimates assembling the study-specific medians and SD approximations: Conservative SD, less conservative SD, mean SD, or interquartile range. We provided a graphical evaluation of the standardized differences.

To show which imputation method produced the best estimate, we ranked those differences and calculated the rate at which each estimate appeared as the best, second-best, third-best, or fourth-best.

RESULTS: Our results demonstrated that the best pooled estimate for the overall mean and SD was provided by the median and interquartile range (mean standardized estimates: 4.5 ± 2.2, P = 0.14) or by the median and the SD conservative estimate (mean standardized estimates: 4.5 ± 3.5, P = 0.13). The less conservative approximation of SD appeared to be the worst method, exhibiting a significant difference from the reference method at the 90% confidence level. The method that ranked first most frequently is the interquartile range method (23/42 = 55%), particularly when data were generated according to the Standard Normal, Gamma, and Exponential distributions. The second-best is the conservative SD method (15/42 = 36%), particularly for data from a bimodal distribution and for the intensive care unit stay variable.

CONCLUSION: Meta-analytic estimates are not significantly affected by approximating the missing values of mean and SD with the correspondent values for median and interquartile range.

Key words: Imputation; Interquartile range; Meta-analysis; Randomized controlled trial; Standard deviation

Core tip: Meta-analyses of continuous endpoints are generally supposed to deal with normally distributed data and the pooled estimate of the treatment effect relies on means and standard deviations. However, if the outcome distribution is skewed, some authors correctly report the median together with the corresponding quartiles. In the present work, we compared methods for the approximation of means and standard deviations when only medians with quartiles are provided. Our results demonstrate that meta-analytic estimates are not significantly affected by approximating the missing values of mean and standard deviation with the correspondent values for median and interquartile range.

Greco T, Biondi-Zoccai G, Gemma M, Guérin C, Zangrillo A, Landoni G. How to impute study-specific standard deviations in meta-analyses of skewed continuous endpoints? *World J Meta-Anal* 2015; 3(5): 215-224 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i5/215.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i5.215

INTRODUCTION

Meta-analysis (MA) is a powerful statistical method that merges the results of different studies considering the same outcome variables. The included studies are mainly randomized controlled trials with experimental and control arms. MA aims at assessing the treatment effect size under scrutiny, at identifying sources of heterogeneity among the included studies, at unrevealing patterns behind the available data, and sometimes at identifying new subgroup associations. Some authors believe that MA represents the highest level of evidence to provide recommendations on clinical issues.

Just as many other innovative statistical techniques, MA is still a matter of intense debate, since many of its assumptions are critical and even small violations of them can lead to misleading conclusions[1]. The appeal of MA resides also in the quick and cost-effective way it yields useful pieces of information for clinical decision making. Hence, MAs are published and quoted with an impressive increasing frequency and it seems evident that, despite their limitations, they will continue to play a crucial role in medical decision-making in the foreseeable future.

Meta-analyses of continuous outcomes exploit data with a Gaussian distribution, so that the pooled estimate computation requires the study-specific mean, standard deviation (SD), and sample size of the variable at stake. The easiest way to compare the outcomes of two treatment groups is to evaluate the difference between their means[2]. If measurements are expressed in the same unit, the mean difference between the treatment and control groups can be used. Results from trials in which the same outcome is measured in different units can be compared by using SD units rather than absolute differences[2,3]. However, if data are reported in a limited or incomplete way, it can be difficult or impossible to obtain sufficient information to perform a correct summary of the results. Missing SD and non-compliance in reporting collected data are common limitations in MAs of continuous outcomes.

If the outcome has a skewed distribution, authors often report in the original paper the median together with the 1st and 3rd quartiles, rather than the mean with its SD. MA authors arbitrarily combine these study-specific estimates to approximate the missing SD. In addition, some authors combine together means (SD) and medians (1st and 3rd quartiles) from different studies.

In this setting, the most immediate question is whether it is legitimate to approximate study-specific means and SDs from study-specific medians and quartiles and how to do it in the most appropriate way.

In the present work we simulated MA of continuous outcomes generated from seven different distributions and we compared four methods to approximate SDs when only study-specific medians and quartiles are available. For each simulation we calculated a pooled estimate by assembling the individual medians and, in turn, all four SD approximations. Finally, we compared these results with those obtained by pooling the individual means and SDs.

Table 1 Distributions of the continuous endpoint for treatment and control groups

Scenario	Endpoint distribution	
	Treatment group	Control group
Normal	Mean = 5 and SD = 2	Mean = 7 and SD = 2
Standard normal	Mean = 0 and SD = 1	Mean = 0 and SD = 1
Gamma	Alpha = 2 and beta = 5	Alpha = 2 and beta = 7
Exponential	Mean = 5 and lambda = 0.2	Mean = 7 and lambda = 0.14
Bimodal	50% Normal distribution with mean = 5 and SD = 2 and 50% standard normal distribution	50% Normal distribution with mean = 7 and SD = 2 and 50% standard normal distribution
ICU stay	Real-life data	Real-life data
Hospital stay	Real-life data	Real-life data

ICU: Intensive care unit.

To our knowledge, this is the first study on how to impute the study-specific mean and SD in meta-analyses of skewed outcomes. After a failed tentative of comparing results coming from published studies, the present paper compares the available methods making use of both simulations and real-life set-ups to identify the best and the worst SD approximation.

MATERIALS AND METHODS

Simulation algorithm

We generated six datasets of 15, 30, 50, 100, 500, and 1000 trials, respectively. Each trial is based on an equal number of treated and control subjects which is fixed at the number of trials included in the MA under examination. The distributions of the continuous endpoints for the treatment and control groups are generated according to Table 1. The first five scenarios provided the basis for our analysis on simulated data. The last two scenarios of Table 1 represent our real-life data and are randomly extracted from an Italian observational study with more than 7000 patients with cardiovascular disease.

In summary, we generated 1695 (15 + 30 + 50 + 100 + 500 + 1000) trials for each of the seven distribution scenarios for a total of 11865 trials. For each trial we calculated the principal measures of position, mean and median, and variability, SD and interquartile range (IQR).

All simulations and analyses were performed using SAS (release 9.2, 2002-2008 by SAS Institute Inc., Cary, NC, United States)[4]. An example of SAS code is reported in Table 2.

Simulated MA

For each of the six datasets we carried out a series of MAs differing with respect to the method of imputation of the study-specific SD[2], as described in Table 3.

For each distribution scenario, 30 MAs were therefore performed (6 datasets × 5 methods).

Each MA was carried out using a fixed effect model by the Inverse-Variance method[3].

The global estimate is the mean difference between treatment and control groups, obtained by pooling individual means and SDs.

Comparison of estimates

For each distribution scenario we computed four standardized estimates, $\theta stand_{ijk}$, calculated as:

$$\theta stand_{ijk} = (\theta ijk - \theta reference_{ijk})/se(\theta ijk)$$

for

i = 1, 2, 3, 4; j = 15, 30,..., 1000 and k = 1, 2,..., 7.

where θ_{ijk} is the pooled mean difference resulting from the performed MA, $se(\theta_{ijk})$ is the corresponding standard error and $\theta reference_{ijk}$ is the global estimate obtained by pooling individual means and SDs for each dataset and distribution scenario.

We obtained a total of 168 standardized estimates (7 distributions × 6 datasets × 4 methods).

Statistical model

After blocking for dataset and distribution, we evaluated if the standardized estimates, calculated by each of the four methods, were different from zero in the framework of a repeated measures model using the mixed procedure implemented in the SAS software. Statistical significance was set at the two-tailed 0.05 level for hypothesis testing. Adjustment for multiple comparisons was performed with Tukey-Kramer, Bonferroni and Scheffé corrections[5,6].

Ranking

To identify which imputation method produced the pooled estimate θ_{ijk} with the minimum difference from the reference one, $\theta reference_{ijk}$, we ranked each standardized estimate $\theta stand_{ijk}$ for each dataset and for each distribution scenario. According to this ranking, we calculated the rate at which each estimate appeared as the best, second-best, third-best, or fourth-best[7] and the area under the cumulative ranking curve[8].

Literature screening

We screened all the best published studies reported in the medical literature on critically ill patients included in the 1st web based International Consensus Conference[9] to assess which information authors of original papers provided in the case of intensive care unit (ICU) and hospital stay outcome. We contacted the corresponding authors by e-mail when one or more of the following information was missing: Mean, SD, median, 1st-3rd quartiles. No manuscript provided all of the information requested and only 3 authors replied to our e-mail and provided the extra information needed.

Statistical analysis

The statistical methods of this study were reviewed by

Table 2 Example of SAS code to simulate a meta-analysis on 15 datasets with 15 records generated from a Gamma distribution (alpha = 2 and beta = 5 *vs* alpha = 2 and beta = 7 for the treatment and control groups, respectively)

```
* q is the assigned library;
************************************************;
* SIMULATIONS;
************************************************;
%let s = gamma;
%let ndset = 15;
* Simulation of n = 15 dataset using the Gamma distributions;
%macro simul;
%do q = 1 %to &ndset;
%let seed = %sysevalf(1234567 + &q);
%let num_i = %sysevalf(&ndset);
%let v = %sysevalf(0 + &q);
data s&q;
k = &q;
  %do i = 1%to &num_i;
   var1 = 5*rangam(&seed,2);
   var2 = 7*rangam(&seed,2);
   output;
%end;
run;
%end;
* Dataset combining;
data simul_&s;
set
   %do w = 1%to &ndset;
   s&w
   %end;
   ;
 run;
%mend;
%simul;
* Descriptive statistics for each dataset;
ods trace on;
ods output summary = summary_&s;
proc means data = simul_&s mean std median q1 q3;
   class k;
   var var1 var2;
run;
ods trace off;
data summary_&s;
set summary_&s;
l1 = (var1_Median-var1_Q1)/0.6745;
l2 = (var2_Median-var2_Q1)/0.6745;
u1 = (var1_Q3-var1_Median)/0.6745;
u2 = (var2_Q3-var2_Median)/0.6745;
if l1 > u1 then MeSD_v1_cons=l1; else MeSD_v1_cons=u1;
if l2 > u2 then MeSD_v2_cons=l2; else MeSD_v2_cons=u2;
if l1 > u1 then MeSD_v1_prec=u1; else MeSD_v1_prec=l1;
if l2 > u2 then MeSD_v2_prec=u2; else MeSD_v2_prec=l2;
MeSD_v1_mean=(var1_Q3-var1_Q1)/1.349;
MeSD_v2_mean=(var2_Q3-var2_Q1)/1.349;
* Median difference;
MeD = var1_Median-var2_Median;
*1 conservative estimate of standard deviation;
a1sd = ((MeSD_v1_cons)**2)/NObs;
b1sd = ((MeSD_v2_cons)**2)/NObs;
MeSD_cons=sqrt(a1sd + b1sd);
*2 less conservative estimate of standard deviation;
a2sd = ((MeSD_v1_prec)**2)/NObs;
b2sd = ((MeSD_v2_prec)**2)/NObs;
MeSD_prec = sqrt(a2sd + b2sd);
*3 mean estimate of standard deviation;
a3sd = ((MeSD_v1_mean)**2)/NObs;
b3sd = ((MeSD_v2_mean)**2)/NObs;
MeSD_mean = sqrt(a3sd + b3sd);
*4 Interquartile range;
a4sd = ((var1_Q3-var1_Q1)**2)/NObs;
b4sd = ((var2_Q3-var2_Q1)**2)/NObs;
MeSD_iqr = sqrt(a4sd + b4sd);
* Mean difference and pooled standard deviation;
MD = var1_Mean-var2_Mean;
asd = ((var1_StdDev)**2)/NObs;
bsd = ((var2_StdDev)**2)/NObs;
SD = sqrt(asd + bsd);
drop l1 l2 u1 u2 asd bsd a1sd b1sd a2sd b2sd a3sd b3sd a4sd b4sd;
run;
************************;
* Meta-analyses;
data sum_&s;
set summary_&s;
keep k NObs MeD MeSD_cons MeSD_prec MeSD_mean MeSD_iqr MD SD qq;
run;
*1 Median and conservative estimate of standard deviation;
data meta_&s.1;
set sum_&s;
model = "Conservative SD";
MDz = MeD;
SDz = MeSD_cons;
w = 1/(SDz**2);
MDw = MDz*w;
keep model k NObs MDz SDz w MDw;
run;
*2 Median and less conservative estimate of standard deviation;
data meta_&s.2;
set sum_&s;
model = "Less Conservative SD";
MDz = MeD;
SDz = MeSD_prec;
w = 1/(SDz**2);
MDw = MDz*w;
keep model k NObs MDz SDz w MDw;
run;
*3 Median and mean estimate of standard deviation;
data meta_&s.3;
set sum_&s;
model = "Mean SD";
MDz = MeD;
SDz = MeSD_mean;
w = 1/(SDz**2);
MDw=MDz*w;
keep model k NObs MDz SDz w MDw;
run;
*4 Median and interquartile range;
data meta_&s.4;
set sum_&s;
model = "IQR";
MDz = MeD;
SDz = MeSD_iqr;
w = 1/(SDz**2);
MDw = MDz*w;
keep model k NObs MDz SDz w MDw;
run;
*Mean and standard deviation (reference);
data meta_&s.5;
set sum_&s;
model = "Reference";
MDz = MD;
SDz = SD;
w = 1/(SDz**2);
MDw = MDz*w;
keep model k NObs MDz SDz w MDw;
run;
proc format;
value model
1 = "conservative SD"
2 = "Less Conservative SD "
3 = "Mean SD "
4 = "IQR"
```

```
5 = "Reference"
;
run;
*** Fixed effect model meta-analysis - Inverse of Variance method;
%macro meta_iv;
%do i = 1%to 5;
ods output  Summary = somme&i;
proc means data = meta_&s&i sum;
var MDw w;
run;
data somme&i;
set somme&i;
model = &i;
format model model.;
theta = MDw_Sum/w_Sum;
se_theta = 1/(sqrt(w_sum));
lower = theta - (se_theta*1.96);
upper = theta + (se_theta*1.96);
mtheta = sqrt(theta**2);
CV = se_theta/mtheta;
keep model theta se_theta lower upper cv;
run;
%end;
data aaMeta_&s;
set
  %do w = 1% to 5;
  somme&w
  %end;
  ;
run;
title "distr = &s - k = &ndset";
proc print; run;
%mend;
%meta_iv;
```

Rosalba Lembo from San Raffaele Scientific Institute, Via Olgettina 60, 20132 Milan, Italy.

RESULTS

Table 4 reports the $\theta_{stand_{ijk}}$ means and the corresponding unadjusted and adjusted *P* values to assess the presence of a significant difference from zero. Conservative SD and IQR methods showed the lowest mean difference (with mean standardized estimates equal to 4.5 ± 3.5 for conservative SD and 4.5 ± 2.2 for IQR methods, respectively). The less conservative SD method appeared to be the worst, exhibiting the highest difference from the reference.

Table 5 shows the $\theta_{stand_{ijk}}$ estimates for each distribution, dataset, and imputation method and some descriptive statistics. For each distribution, the number of times each standardized estimate ranked first is indicated. The method that ranked first most frequently was IQR (23/42 = 55%), particularly when the data were generated according to the Standard Normal, Gamma, and Exponential distributions. The second best is the Conservative SD method (15/42 = 36%), which was particularly suitable for data with a bimodal distribution and for the ICU stay variable. The quartiles at the bottom of Table 5 were similar for these two methods: 1.62 (0.27-4.94) and 1.40 (0.34-4.86), respectively. Figures 1 and 2 show plots of the pooled estimates for all distribution scenarios. The difference between $\theta_{stand_{ijk}}$ and reference values increased together with the increased number of trials in each MA.

Table 6 reports the rates of occurrence as best, second-best, third-best and fourth-best, for the four imputation methods. The conservative SD and IQR methods most often appeared as best or second-best (cumulative frequencies: 35/42 and 41/42 respectively), and the less conservative SD and mean SD methods as third- or fourth-best. The less conservative SD method was identified as the worst method as it had by far the highest number of fourth positions. Even when it ranked first, the $\theta_{stand_{ijk}}$ was very similar to the IQR method, which is consistently suitable for any distribution scenario.

Figure 3 shows the areas under the cumulative ranking curve for each imputation method. IQR method yields the largest area.

DISCUSSION

The aim of our work was to clarify how to best impute study-specific mean and SD when only the median and 1st and 3rd quartiles are provided.

It is a notoriously good practice to report information on medians and quartiles for skewed distributions but standard meta-analytic approaches require study-specific means and SDs, so that careful evaluation of the best trade-off between these two approaches is needed.

This issue is of prominent interest when dealing with MA, but few scientific works addressed this topic. Hozo *et al*[10] described two formulas for estimating the mean from the median, range, and sample size values. Pigott[11] discussed and examined how to deal with missing data (studies, effect size, and methodological information) during MA. Furukawa *et al*[12,13] reported on the imputation of the missing response rate from the mean ± SD[12] and suggested that borrowing the missing SD from other studies included in the MA may be a valid solution[13]. Thiessen Philbrook *et al*[14] compared results from MAs that either were restricted to available data or imputed the missing variance with one of four methods (*P* values, nonparametric summaries, multiple imputation, or correlation coefficients). Robertson *et al*[15] highlighted and evaluated different ways to include in MAs studies in which the treatment effect was not provided. Weibe *et al*[16] conducted a systematic review of methods for handling missing variances in MAs of continuous outcomes and classified the relevant approaches into eight groups: algebraic recalculation, approximate algebraic recalculation, study-level imputation, study-level imputation from nonparametric summaries, study-level imputation of correlation (for change-from-baseline or crossover SD and to calculate the design effect for cluster studies), MA-level imputation of overall effect, MA-level tests, and no-imputation methods. Finally, Stevens[17] gave an overview of the Bayesian approach to deal with missing data in MA. However, authors who carry out MAs rarely

Table 3 Method for imputing the study-specific standard deviation

Method number	Method name	Mean imputation	Standard Deviation imputation[1]
0	Reference	Mean	SD
1	Conservative SD	Median	max[(3rd quartile - median)/0.6745; (median - 1st quartile)/0.6745]
2	Less Conservative SD	Median	min[(3rd quartile - median)/0.6745; (median - 1st quartile)/0.6745]
3	Mean SD	Median	(3rd quartile - 1st quartile)/(2 × 0.6745)
4	IQR	Median	(3rd quartile - 1st quartile)

[1]The 0.675 is the Z value corresponding to an area of 75% of the standard normal distribution function. IQR: Interquartile range.

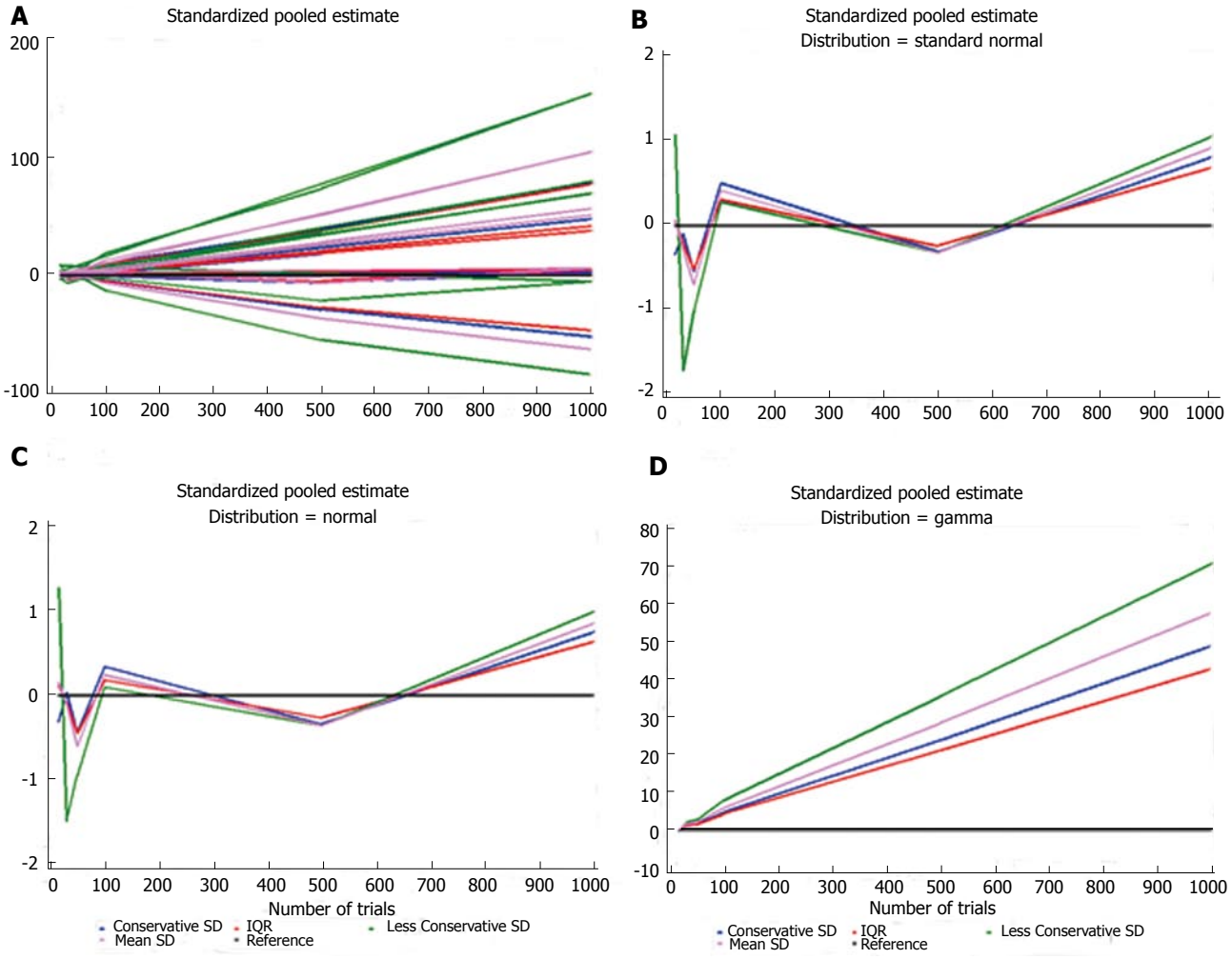


Figure 1 Value of each standardized estimate, $\theta_{stand_{ijk}}$, for each number of dataset included in the corresponding meta-analysis simulated. A: All distributions together; B: Normal (mean = 5 and SD = 2 for the treatment group; mean = 7 and SD = 2 for the control group); C: Standardized Normal; D: Gamma (alpha = 2 and beta = 5 for the treatment group; alpha = 2 and beta = 7 for the control group). The X-axis represents the number of studies (datasets) included in the meta-analysis. For each approximation method (blue line for the conservative estimate of SD, green line for the less conservative estimate of SD, pink line for the mean estimate of SD, and red line for the interquartile range), the Y-axis reports the difference between the standardized estimate and the reference (black line). IQR: Interquartile range.

adopt similar methods in their current clinical practice.

In the present study, we showed that MA pooled estimates are not significantly affected by approximating the missing study-specific mean and SD with the corresponding median and IQR, in both simulated and real-life set-ups. In comparison with the other methods, we found that the Median-IQR method has extra advantages, since it is the simplest one and it makes no assumption on the underlying distribution of the data. Furthermore, we showed how the use of a less conservative approximation of SD can bias the meta-analytic pooled estimate when authors work with skewed data. Nevertheless, it is well known that median and mean values are very different as the data distribution is skewed[2].

Our study has several limitations. First, we did not perform a sensitivity analysis with the random effects model. However, since data were generated from the

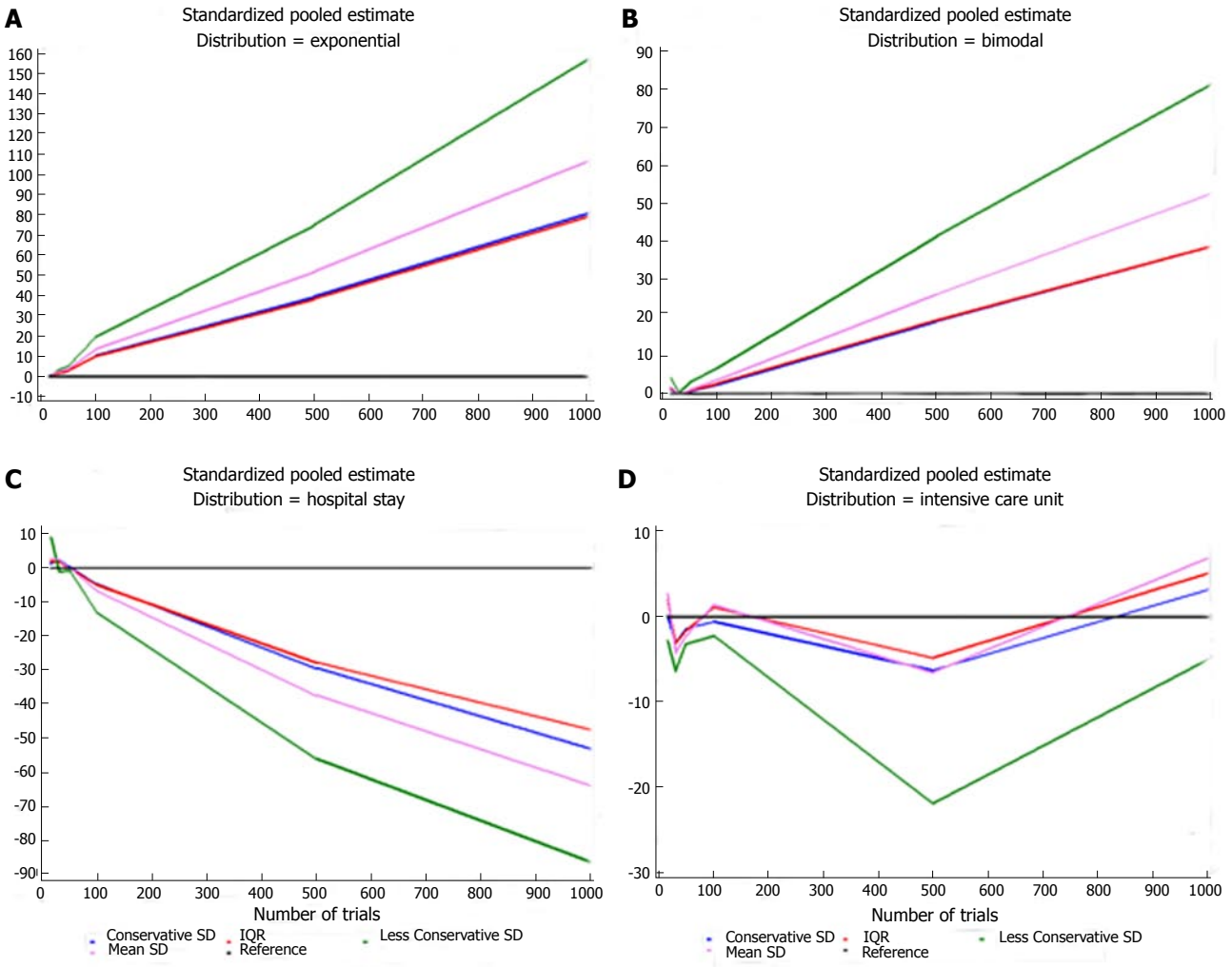


Figure 2 Value of each standardized estimate, $\theta stand_{ijk}$, for each number of dataset included in the corresponding meta-analysis simulated. A: Exponential (mean = 5 and lambda = 0.2 for the treatment group; mean = 7 and lambda = 0.14 for the control group); B: Bimodal (50% Normal distribution (with mean = 5 and SD = 2) and 50% Standardized Normal for the treatment group; 50% Normal distribution (with mean = 7 and SD = 2) and 50% standardized normal for the control group); C and D: Real distribution of intensive care unit stay and hospital stay of an Italian cardiology dataset with 7471 patients. The X-axis represents the number of studies (datasets) included in the meta-analysis. For each approximation method (blue line for the conservative estimate of SD, green line for the less conservative estimate of SD, pink line for the mean estimate of SD, and red line for the interquartile range), the Y-axis reports the difference between the standardized estimate and the reference (black line). IQR: Interquartile range.

Table 4 Comparison of results obtained from the four methods of approximation of study-specific means and standard deviations in a meta-analysis of a continuous outcome

	Standardized pooled estimate, (standard error)	*P* value[1] (unadjusted)	*P* value[1] (Tukey-Kramer adjustment)	*P* value[1] (Bonferroni adjustment)	*P* value[1] (Scheffé adjustment)
Conservative SD	4.5 (3.5)	0.14	0.7	0.9	0.8
Less conservative SD	7.8 (2.9)	0.01	0.056	0.07	0.12
Mean SD	6.1 (3.3)	0.04	0.3	0.6	0.5
IQR	4.5 (2.2)	0.13	0.2	0.4	0.4

[1]*P* values were derived from a test on the "method" effect in a repeated measures model including also adjustment by "dataset" and "distribution" effects. IQR: Interquartile range.

same distribution for each trial, we decided to apply fixed effect models. Second, we did not analyze mixed set-ups in which study-specific means and SDs are available for some studies and others are imputed. Third, we worked only on trials where the number of treated/control subjects was equal to the number of studies included in the MA. It follows that we did not consider set-ups with many studies and few subjects or vice versa. Although we present several distribution scenarios, the choice of their parameters was arbitrary. However, the performance is not expected to change with different parameters.

Table 5 Absolute differences between standardized estimates, $\theta_{stand_{ijk}}$, calculated by means of one of the four methods (conservative SD, less conservative SD, mean SD and interquartile range), and the reference

Distribution scenario	Dataset	Conservative SD	Less Conservative SD	Mean SD	IQR
Normal	15	-0.310	1.274	0.149	0.110
Normal	30	0.029	-1.483	-0.109	-0.081
Normal	50	-0.434	-0.946	-0.599	-0.444
Normal	100	0.340	0.095	0.243	0.180
Normal	500	-0.336	-0.357	-0.353	-0.261
Normal	1000	0.754	0.989	0.860	0.638
No. of times of beginning first in the ranking[1]		2	1	0	3
Standard normal	15	-0.335	1.072	0.062	0.046
Standard normal	30	-0.101	-1.710	-0.290	-0.215
Standard normal	50	-0.535	-1.013	-0.690	-0.511
Standard normal	100	0.502	0.281	0.416	0.308
Standard normal	500	-0.306	-0.314	-0.317	-0.235
Standard normal	1000	0.814	1.054	0.923	0.684
Normal					
No. of times of beginning first in the ranking[1]		1	1	0	4
Gamma	15	-0.283	0.054	-0.119	-0.088
Gamma	30	1.441	2.229	1.846	1.368
Gamma	50	1.795	2.929	2.218	1.644
Gamma	100	4.915	8.193	6.070	4.500
Gamma	500	24.081	35.799	28.753	21.314
Gamma	1000	49.089	71.072	58.012	43.002
No. of times of beginning first in the ranking[1]		0	1	0	5
Exponential	15	-0.150	-0.163	-0.157	-0.116
Exponential	30	1.880	2.958	2.301	1.706
Exponential	50	2.948	4.975	3.707	2.748
Exponential	100	10.213	19.490	13.493	10.002
Exponential	500	39.546	74.955	51.913	38.481
Exponential	1000	80.605	157.083	106.593	79.016
No. of times of beginning first in the ranking[1]		0	0	0	6
Bimodal	15	1.142	4.114	1.751	1.298
Bimodal	30	0.079	0.356	0.096	0.071
Bimodal	50	0.545	3.051	1.110	0.823
Bimodal	100	2.405	6.849	3.650	2.706
Bimodal	500	19.156	41.495	26.212	19.431
Bimodal	1000	38.825	81.301	52.527	38.938
No. of times of beginning first in the ranking[1]		5	0	0	1
ICU stay	15	0.076	-2.816	2.667	1.977
ICU stay	30	-3.011	-6.341	-4.201	-3.114
ICU stay	50	-1.361	-3.162	-2.163	-1.603
ICU stay	100	-0.58	-2.205	1.393	1.032
ICU stay	500	-6.218	-21.788	-6.462	-4.790
ICU stay	1000	3.162	-5.020	6.801	5.042
No. of times of beginning first in the ranking[1]		5	0	0	1
Hospital stay	15	1.437	8.948	2.777	2.058
Hospital stay	30	2.603	-1.088	2.595	1.924
Hospital stay	50	0.297	-0.839	-0.055	-0.041
Hospital stay	100	-4.674	-13.063	-6.734	-4.992
Hospital stay	500	-29.239	-55.703	-37.170	-27.554
Hospital stay	1000	-52.720	-85.673	-63.453	-47.038
No. of times of beginning first in the ranking[1]		2	1	0	3
1st quartile		0.337	1.023	0.368	0.273
Median		1.399	2.944	2.190	1.624
3rd quartile		4.855	12.034	6.666	4.942
Total number of times of beginning first in the ranking[1]		15	4	0	23

[1]The ranking procedure was based on differences in absolute value. Number of times that each standardized estimate ranked as for first, for each simulated distribution. Median (interquartile range) of overall standardized estimate distributions. IQR: Interquartile range.

Table 6 Absolute and relative frequencies of occurrence of the four methods to approximate study-specific means and standard deviations in a meta-analysis of a continuous outcome *n* (%)

	No. of first ranking	No. of second ranking	No. of third ranking	No. of fourth ranking
Conservative SD	15 (35.7)	20 (47.6)	3 (7.1)	4 (9.5)
Less Conservative SD	4 (9.5)	1 (2.4)	1 (2.4)	36 (85.7)
Mean SD	0	3 (7.1)	37 (88.1)	2 (4.8)
IQR	23 (54.8)	18 (42.9)	1 (2.4)	0

IQR: Interquartile range.

It is not surprising that many papers report median and IQR, rather than mean and SD. Actually this is considered good practice when dealing with non-normally distributed data. As an example, since the distributions of ICU or hospital stay are skewed, it might be worth using median and IQR when setting up any MA with these end-points.

Conclusion

Our work supports the procedure of using study-specific medians and quartiles to impute means and standard deviations. This avoids the dangerous practice of not including in the MA studies with missing information. Nevertheless, we recognize that, in order to improve the quality of future MAs, authors of research papers should report as much information as possible at least for concerning their primary outcomes. We suggest that authors who use median and interquartile range in MAs with continuous endpoints, perform a sensitivity analysis

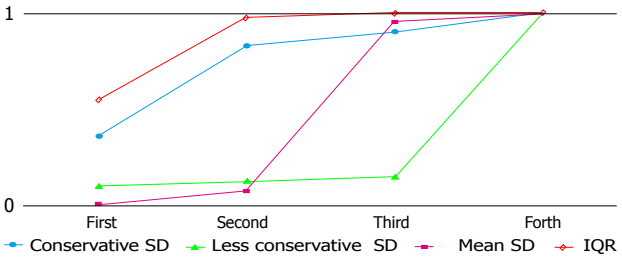


Figure 3 Area under the cumulative ranking curve for the four methods to approximate mean and standard deviation in a meta-analysis of continuous outcome. IQR: Interquartile range.

in which trials not providing study-specific means and SDs are excluded.

COMMENTS

Background

Meta-analyses (MAs) of continuous outcomes exploit data with a Gaussian distribution, so that the pooled estimate computation requires the study-specific means, standard deviations (SDs), and sample sizes. However, should data be reported in a limited or incomplete way, it can be difficult or impossible to obtain sufficient information to perform a correct summary of the results. Missing standard deviations and non-compliance in reporting collected data are common limitations in MAs of continuous outcomes. At the simpler level, an interest emerged as to the opportunity and the best way to approximate on study-specific means and SDs from study-specific medians and quartiles.

Research frontiers

The aim of the work was to clarify how to best impute study-specific mean and SD when only the median and 1st and 3rd quartiles are provided. The authors compared the available methods in simulated and real-life set-ups to identify the best and the worst one. In the present study, authors showed that MA pooled estimates are not significantly affected by approximating the missing study-specific mean and SD with the corresponding median and interquartile range. Among the four methods proposed, the Median-IQR method has the extra advantages since it is the simplest one and that makes no assumption on the underlying distribution of the data.

Innovations and breakthroughs

For the first time, to our knowledge, this manuscript provides a list of four available approximation methods in MAs of skewed outcomes. The authors became aware of these methods in the clinical practice and we strongly believe in the good practice to report information on medians and quartiles when a distribution is skewed. However, standard meta-analytic approaches require study specific means and SDs and it was natural to us to find out a compromise between these two needs. This issue is important when carrying out a MA, but few scientific works addressed this topic. Authors described formulas and estimating procedures to work with missing data in an MA, in either frequentist or Bayesian approach. However, authors who carry out MAs rarely adopt similar methods in their current clinical practice.

Applications

The work gives support to the procedure of using study-specific medians and quartiles to impute means and SDs in an MA of skewed outcome. This avoids the dangerous practice of not including in the MA studies with missing information.

Terminology

Approximation: An estimate of the value of a quantity to a desired degree of accuracy; Conservative estimate: Estimate that avoids excess in approximating the quantity or worth of the list of (potentially infinite) values identified; Distribution: List of the values in a population, or sample, with the corresponding frequency or probability of occurrence; Estimate: A value (a point estimate) or range of values (an interval estimate) to a parameter of a population on the basis of sampling statistics; Interquartile range: Measure of dispersion around the median given by the difference between the 3rd quartile and 1st quartile of the distribution; these quartiles can be clearly seen on a box-plot of the data; Study-specific: Related to a specific study included in a more comprehensive MA (*i.e.*, group of study).

Peer-review

The study addresses a very interesting topic and deals with a major concern for investigators who have to perform a meta-analysis from published studies.

REFERENCES

1 **Greco T**, Zangrillo A, Biondi-Zoccai G, Landoni G. Meta-analysis: pitfalls and hints. *Heart Lung Vessel* 2013; **5**: 219-225 [PMID: 24364016]

2 **Higgins JPT**, Green S. Cochrane handbook for systematic reviews of interventions. Version 5.1.0. The Cochrane collaboration 2011. [Accessed 2015 Jun]. Available from: URL: http://www.cochrane-handbook.org/

3 **Rothstein HR**, Sutton AJ, Borenstein M. Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. Chapter 8. The Trim and Fill Method. New York: John Wiley & Sons Ltd, 2006 [DOI: 10.1002/0470870168.ch8]

4 **Moser EB**. Repeated measures modeling with PROC MIXED. In: Proceedings of the 29th SAS Users Group International Conference. Montreal, Canada 2004. [Accessed 2015 Jun]. Available from: URL: http://www2.sas.com/proceedings/sugi29/188-29.pdf

5 **Edwards D**, Berry JJ. The efficiency of simulation-based multiple comparisons. *Biometrics* 1987; **43**: 913-928 [PMID: 3427176]

6 **Rafter JA**, Abell ML, Braselton JP. Multiple Comparison Methods for Means. *SIAM Review* 2002; **44**: 259-278

7 **Dias S**, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials 2011. [Accessed 2015 Jun]. Available from: URL: http://www.nicedsu.org.uk

8 **Flach P**, Matsubara ET. On classification, ranking, and probability estimation. Dagstuhl Seminar Proceedings 07161. Probabilistic, Logical and Relational Learning - A Further Synthesis 2008. [Accessed 2015 Jun]. Available from: URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.183.6043&rep=rep1&type=pdf

9 **Landoni G**, Augoustides JG, Guarracino F, Santini F, Ponschab M, Pasero D, Rodseth RN, Biondi-Zoccai G, Silvay G, Salvi L, Camporesi E, Comis M, Conte M, Bevilacqua S, Cabrini L, Cariello C, Caramelli F, De Santis V, Del Sarto P, Dini D, Forti A, Galdieri N, Giordano G, Gottin L, Greco M, Maglioni E, Mantovani L, Manzato A, Meli M, Paternoster G, Pittarello D, Rana KN, Ruggeri L, Salandin V, Sangalli F, Zambon M, Zucchetti M, Bignami E, Alfieri O, Zangrillo A. Mortality reduction in cardiac anesthesia and intensive care: results of the first International Consensus Conference. *Acta Anaesthesiol Scand* 2011; **55**: 259-266 [PMID: 21288207 DOI: 10.1111/j.1399-6576.2010.02381.x]

10 **Hozo SP**, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol* 2005; **5**: 13 [PMID: 15840177]

11 **Pigott TD**. Missing predictors in models of effect size. *Eval Health Prof* 2001; **24**: 277-307

12 **Furukawa TA**, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analyses. *Int Clin Psychopharmacol* 2005; **20**: 49-52 [PMID: 15602117]

13 **Furukawa TA**, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 2006; **59**: 7-10 [PMID: 16360555]

14 **Thiessen Philbrook H**, Barrowman N, Garg AX. Imputing variance estimates do not alter the conclusions of a meta-analysis with continuous outcomes: a case study of changes in renal function after living kidney donation. *J Clin Epidemiol* 2007; **60**: 228-240 [PMID: 17292016]

15 **Robertson C**, Idris NR, Boyle P. Beyond classical meta-analysis: can inadequately reported studies be included? *Drug Discov Today* 2004; **9**: 924-931 [PMID: 15501727]
16 **Wiebe N**, Vandermeer B, Platt RW, Klassen TP, Moher D, Barrowman NJ. A systematic review identifies a lack of standardization in methods for handling missing variance data. *J Clin Epidemiol* 2006; **59**: 342-353 [PMID: 16549255]
17 **Stevens JW**. A note on dealing with missing standard errors in meta-analyses of continuous outcome measures in WinBUGS. *Pharm Stat* 2011; **10**: 374-378 [PMID: 21394888 DOI: 10.1002/pst.491]

P- Reviewer: Naugler C, Omboni S, Trkulja V **S- Editor**: Ji FF
L- Editor: A **E- Editor**: Wu HL

Published by **Baishideng Publishing Group Inc**
8226 Regency Drive, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

ISSN 2308-3840 (online)

World Journal of *Meta-Analysis*

World J Meta-Anal 2015 December 26; 3(6): 225-303

Published by Baishideng Publishing Group Inc

WJMA World Journal of Meta-Analysis

A peer-reviewed, online, open-access journal of meta-analysis

Editorial Board

2013-2018

The *World Journal of Meta-Analysis* Editorial Board consists of 400 members, representing a team of worldwide experts in clinical meta-analysis research. They are from 41 countries, including Argentina (2), Australia (3), Austria (1), Belgium (5), Brazil (11), Canada (16), Chile (3), China (121), Croatia (1), Egypt (1), Finland (4), France (2), Germany (9), Greece (11), Hungary (2), India (12), Iran (2), Ireland (1), Israel (2), Italy (41), Japan (5), Lithuania (1), Malaysia (1), Netherlands (8), New Zealand (1), Norway (1), Peru (1), Poland (4), Portugal (6), Romania (1), Saudi Arabia (4), Singapore (3), South Africa (1), South Korea (7), Spain (8), Sri Lanka (2), Sudan (1), Switzerland (2), Thailand (4), Turkey (3), United Kingdom (22), and United States (64).

Alessandro Cucchetti, *Bologna*
Gianfranco Damiani, *Rome*
Fabrizio D'Ascenzo, *Turin*
Massimo Del Fabbro, *Milano*
Valeria Fadda, *Florence*
Alessandro Fancellu, *Sassari*
Giuseppe Ferrante, *Milan*
Virginia Festa, *Rome*
Francesco Fiorica, *Ferrara*
Guglielmo Giraldi, *Rome*
Jenny Guidi, *Bologna*
Lorenzo Loffredo, *Rome*
Andrea Messori, *Firenze*
Eliano Pio Navarese, *Bydgoszcz*
Stefano Omboni, *Solbiate Arno*
Alvisa Palese, *Udine*
Stefano Palomba, *Reggio Emilia Stefano*
Carlo Perricone, *Rome*
Mario Petretta, *Naples*
Alessandro Pezzini, *Brescia*
Gianluca Pontone, *Milan*
Paolo Emilio Puddu, *Rome*
Andrea Rognoni, *Novara*
Giuseppe Scalabrino, *Milan*
Città della Salute e della Scienza, *Turin*
Fabrizio Sgolastra, *L'Aquila*
Maria Lucia Specchia, *Rome*
Stefano Trastulli, *Terni*
Fabio Tine, *Palermo*
Nereo Vettoretto, *Chiari*
Alberto Vianello, *Perugia*
Luigi Zorcolo, *Cagliari*

Japan

Nguyen Tien Huy, *Nagasaki*
Hiroharu Kamioka, *Tokyo*
Koji Kawakami, *Kyoto*
Keitaro Matsuo, *Nagoya*
Kazushi Okamoto, *Nagoya*

Lithuania

Edmundas Kadusevicius, *Kaunas*

Malaysia

SP Pani, *Ipoh*

Netherlands

Michel van den Bekerom, *Amsterdam*
Dimitra Dodou, *Delft*
D Haverkamp, *Amsterdam*
Vassilios Koussoulas, *Drachten*
BJ Polder, *Nijmegen*
Theo Stijnen, *Leiden*
RNM Weijers, *Amsterdam*
Joost de Winter, *Delft*

New Zealand

Shaofeng Li, *Auckland*

Norway

Eivind Berge, *Oslo*

Peru

Rafael Bolaños Díaz, *Lima*

Poland

Maciej Banach, *Lodz*
Krzysztof Jonderko, *Sosnowiec*
Jolanta Lissowska, *Warsaw*
Maciej Plaszewski, *Warsaw*

Portugal

Daniel Caldeira, *Lisboa*
J Costa, *Lisbon*
Ana Miguel, *Coimbra*
Manuel Morgado, *Covilh*
Bárbara Peleteiro, *Porto*
Rui Torres, *Paredes*

Romania

Fratila Ovidiu, *Oradea*

Saudi Arabia

Hazem M Al-Mandeel, *Riyadh*
Ezzeldin M Ibrahim, *Jeddah*
Mutahir A Tunio, *Riyadh*
Hayfaa A Wahabi, *Riyadh*

Singapore

Nikos LD Chatzisarantis, *Singapore*
Edwin Choon Wyn Lim, *Singapore*
Roger Ho, *Singapore*

South Africa

Alaine Umubyeyi Nyaruhirira, *Pretoria*

South Korea

Jung-Hee Kim, *Cheonan*
Hyangsook Lee, *Seoul*
Myeong Soo Lee, *Daejeon*
Chi-Un Pae, *Bucheon*
Yong Hyun Park, *Yangsan*
Jae Hong Seo, *Seoul*
Yong Sang Song, *Seoul*

Spain

Pablo Avanzas, *Oviedo*
Joan Cid, *Barcelona*
Joaquin de Haro, *Madrid*
Joan Guardia-Olmos, *Barcelona*
Nabil Halaihlel, *Zaragoza*
JA Monge-Argilés, *Alicante*
Raul Moreno, *Madrid*
Inés Velasco, *Aracena*

Sri Lanka

Ranil Jayawardena, *Colombo*
Priyanga Ranasinghe, *Colombo*

Switzerland

Jay P Singh, *Zurich*
Giorgio Treglia, *Bellinzona*

Sudan

Samir MH Shaheen, *Khartoum*

Thailand

Chuenjid Kongkaew, *Phitsanulok*
Manop Pithukpakorn, *Bangkok*
Piyamitr Sritara, *Bangkok*
Surasak Saokaew, *Phayao*

Turkey

Nese Demirturk, *Afyonkarahisar*
Nilüfer Ozabaci, *Eskisehir*
Ilke Sipahi, *Istanbul*

United Kingdom

Omar M Aboumarzouk, *Wales*
Abeer Al-Namankany, *London*
Ernest A Azzopardi, *Cardif*
Umberto Benedetto, *Cambridge*
Joanne Brooke, *London*
Noriko Cable, *London*
David Chan, *Cardiff*
YC Cheong, *Southampton*
Andrew Currie, *Harrow*
G Nabi, *Dundee*

Lesley A Anderson, *Belfast*
Valentina Gallo, *London*
Gianpiero Gravante, *Leicester*
Peter N Lee, *Surrey*
Igho Onakpoya, *Oxford*
Ashish Pradhan, *Huntingdon*
Evridiki Patelarou, *London*
Jian-Qing Shi, *Newcastle*
Surendra P Singh, *Wolverhampton*
Natalie Taylor, *Leeds*
Yousef Shahin, *Hull*
Zheng Ye, *Cambridge*

United States

L Joseph Su, *Rockville*
Olusola Adesope, *Pullman*
Mike Allen, *Milwaukee*
Bhupinder Anand, *Bellaire*
Stephen Aronoff, *Philadelphia*
KoKo Aung, *San Antonio*
William L Baker, *Farmington*
Moritz C Wyler von Ballmoos, *Milwaukee*
Matthew L Bechtold, *Columbia*
Atul Bhardwaj, *Hershey*
Somjot S Brar, *Los Angeles*
Hui Cai, *Nashville*
Subhash Chandra, *Towson*
Wen-Pin Chang, *Omaha*
Yong Chen, *North Wales*
Myunghan Choi, *Phoenix*
John Coverdale, *Houston*
Prakash C Deedwania, *Fresno*
Eugene Demidenko, *Hanover*
Hong-Wen Deng, *New Orleans*
Eric M Deshaies, *Syracuse*
Tao Fan, *Whitehouse Station*
Shinga Feresu, *Bloomington*
Janvier Gasana, *Miami*
Kaveh Hajifathalian, *Boston*
Mohammad Obaidul Hoque, *Baltimore*
Larissa R Brunner Huber, *Charlotte*
Imran H Iftikhar, *Columbia*
Vijayvel Jayaprakash, *Buffalo*
Xuezhi Jiang, *Weat Reading*
Shuo Jiao, *Seattle*
Evelyn Johnson, *Boise*
Le Kang, *Silver Spring*
SR Kapadia, *Cleveland*
Lior Katz, *Houston*
Daniel M Laskin, *Richmond*
Yu Liang, *Foster City*
Paul Ellis Marik, *Norfolk*
Lynne V McFarland, *Seattle*
Marcovalerio Melis, *New York*
Brian J Miller, *Augusta*
Pavlos Msaouel, *New York*
Joshua E Muscat, *Hershey*
Chee Yuan Ng, *Loma Linda*
Nghi C Nguyen, *Saint Louis*
Brandi S Niemeier, *Whitewater*
Thomas D Parsons, *Denton*
Nidal Abi Rafeh, *New Orleans*
Praveen Roy, *Marshfield*
Ali Salavati, *Philadelphia*
Ankur Sethi, *Chicago*
Tatyana A Shamliyan, *Minneapoli*
Qian Shi, *Rochester*
Zhongjie Shi, *Philadelphia*
Param Puneet Singh, *Chicago*
KV Slavin, *Chicago*
Ali El Solh, *Buffalo*
Jielin Sun, *Winston-Salem*
Richard G Trohman, *Chicago*
Laurah Turner, *Cincinnati*
Sheila Wilhelm, *Detroit*
Alex K Wong, *Los Angeles*
Xiaohui Xu, *Gainesville*
Lu Yin, *Nashville*

WJMA World Journal of Meta-Analysis

Contents Bimonthly Volume 3 Number 6 December 26, 2015

EVIDENCE-BASED MEDICINE

META-ANALYSIS

World Journal of Meta-Analysis
Volume 3 Number 6 December 26, 2015

Contents

NAME OF JOURNAL
World Journal of Meta-Analysis

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Bimonthly

EDITOR-IN-CHIEF
Giuseppe Biondi-Zoccai, MD, Assistant Professor, Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
World Journal of Meta-Analysis
Room 903, Building D, Ocean International Center, No. 62 Dongsihuan Zhonglu, Chaoyang District, Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Inc
8226 Regency Drive,
Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
December 26, 2015

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i6.225

World J Meta-Anal 2015 December 26; 3(6): 225-231
ISSN 2308-3840 (online)

Towards better meta-analyses in assisted reproductive technology: Fixed, random or multivariate models?

Philippe Lehert

Philippe Lehert, Faculty of Medicine, the University of Melbourne, Southbank 3006, Victoria, Australia

Philippe Lehert, Faculty of Economics, UCL Louvain University, B-7000 Mons, Belgium

Author contributions: This author is the exclusive author of this whole research.

Conflict-of-interest statement: The author declares no competing interests.

Data sharing statement: None.

Correspondence to: Dr. Philippe Lehert, PhD, Professor of Statistics, Faculty of Medicine, the University of Melbourne, 801/250 St Kilda Rd, Southbank 3006, Victoria, Australia. philippe.lehert@gmail.com
Telephone: +61-3-96999411
Fax: +61-3-96999411

Received: May 15, 2015
Peer-review started: May 20, 2015
First decision: July 26, 2015
Revised: September 27, 2015
Accepted: October 16, 2015
Article in press: October 19, 2015
Published online: December 26, 2015

Abstract

AIM: To study the validity of the fixed, random, and multivariate meta-analytical models applied in meta-analyses in artificial reproduction technique.

METHODS: Based on common characteristics of *in vitro* fertilization (IVF) meta-analyses, we simulated a large number of data to compare results issued from the fixed model (FM) with the random model (RM). For multiple endpoints meta-analysis (MA), we compared the univariate RM with the multivariate model (MM). Finally, we illustrate our findings in re-analyzing a recent MA.

RESULTS: In our review, although a homogeneous effect was excluded in 89% of the MAs (11%), FM was utilized in 41 studies (82%). From simulations, a concordance of 59% ± 6% was found between the two tests, with up to 65% of falsely significant results with FM. The *Q*-test on studies characterized by substantial heterogeneity falsely accepted homogeneity in 46% of studies. Comparing separate univariate RM and MM on multiple endpoints studies, MM reduces the between endpoint discrepancy (BED) of 68%, and increases the power of 57% ± 8%. In the example dealing with the controversial effect of luteneizing hormone supplementation to follicle stimulating hormone during ovarian stimulation in IVF cycles, MM reduced BED by 66%, and consistent effects were found for all the endpoints, irrespective of partial reporting.

CONCLUSION: The FM generally may produce falsely significant differences. The RM should always be used. For multiple endpoints, the MM constitutes the best option.

Key words: Meta-analysis; Random model; Fixed model; Assisted reproductive techniques; *In vitro* fertilization

Core tip: The numerous meta-analyses (MA) published in assisted reproduction technology (ART) are often

characterized by conflicting results. This paper provides evidence that the choice of the meta-analytical model constitutes a major concern. We first identified a general profile of characteristics of the ART studies, compare different models by simulation and resolve a practical case. MA based on the fixed model produce severe biases and falsely significant differences. Better results derive from the random model. For partially reported multiple endpoints, the multivariate model takes advantage of the between-endpoint inter-correlation and provides consistent estimates, better precision, and higher power.

Lehert P. Towards better meta-analyses in assisted reproductive technology: Fixed, random or multivariate models? *World J Meta-Anal* 2015; 3(6): 225-231 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i6/225.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i6.225

INTRODUCTION

Highly controversial in the 1990s, meta-analysis (MA) has become a widely recognized technique to synthesize evidence from clinical trials. Although considered by many clinicians as mixing apples and oranges[1], evidence based medicine groups have contributed greatly to the acceptability of the approach. As a result, MA considerably impacts drug prescription and clinical practice maybe even more than isolated trials. Thus, like clinical trials, meta-analyses should be conducted with the highest quality and methodological standards.

Trials may provide conflicting results, due to various reasons such as patient selection, trial conduct, duration, and sample size. Controversial results between MAs conducted on the same subject are more worrisome, as a MA constitutes, in essence, a synthesis from existing evidence. And yet, such differences are often observed among MAs, generating doubts on the validity of the results and the way they were identified.

The most known reasons of controversy are differences in study selection or elimination depending on whether or not they were published, blind, randomized, with a sufficient methodological quality risk of bias between studies.

Much less discussed, the choice of the meta-analytical method may involve potentially strong differences on results. In restraining to assisted reproduction technology (ART) context, at least two important concerns may be mentioned:

(1) A majority of MAs used the traditional and simplest fixed model (FM), in which the studied treatment effect is assumed constant across any study. An essential specificity of ART is the considerable difference of practice, procedures, medication use and knowhow among studies, countries or centres causing very heterogeneous performances[2,3]. To which extent this assumption of constant effect remains bearable, although in most MAs this assumption was not tested? Should other models such the random model (RM) admitting an heterogeneous treatment effect be more adapted, while being more conservative?

And (2) In most of the studies, several endpoints are evaluated, such as the number of retrieved oocytes, embryos, implantation rate, and pregnancy ratios. Their separate analyses involve difficulties in the discussion and become non-comparable for endpoints reported by different the number of selected trials (NST). This is called partial reporting and is very common in ART MAs. Separate MAs for each endpoint where high correlation and partial reporting co-exist produce a frequent paradox characterized by conflicting results on correlated endpoints, simply due to non-comparable power. The trickiest case is live birth, the ultimate endpoint in ART, much less reported than other endpoints, compared with earlier markers like clinical pregnancy necessitating much less follow up. Unlike univariate MA, the multivariate MA is a recent proposal briefly introduced here below, taking advantage of between endpoints correlations. To which extent using simple univariate models remains acceptable in such conditions, or is it worthwhile to turn to multivariate approach?

This research is based on the hypothesis that ART studies are characterized by a homogeneous profile of characteristics enabling the adapted choice of a meta-analytical model. In a first stage, we determined this profile based on a sample of MAs selected from a literature review, at a second stage we conducted simulation studies based on this profile to compare the models, and we apply these principles on a study case.

MATERIALS AND METHODS

We first attempted to identify a general specific profile of MAs in ART, including the NST, level of heterogeneity, number and kind of endpoints, the used model and its options. As lots of recalculations were necessary on each study, a random sample of 50 MAs was extracted from a list found from literature research (MEDLINE, EmBase, Google) irrespective of publication (paper/abstract), date or language, by using the key words list [(MA or systematic review) and (IUI or IVF or ICSI)].

We assessed whether a specific profile of MA studies has incidence on the adequacy of MA models, by using the specific profile collected from our MA sample, and conducting simulations in replicating these specific conditions on a multitude of generated samples. For all these samples we compared the FM with the RM and for multiple endpoints, the simple univariate calculation with the multivariate MA, on various criteria: (1) the magnitude of the difference between the estimates, and the direction of bias; (2) the difference in precision of estimates, and consequences on statistical power; and (3) for multiple endpoints, consistency between endpoints with respect to correlations between endpoints.

Finally, we illustrate our findings on a real study

case in re-analyzing a recently published MA in which we compare the results found according to the studied models.

Statistical analysis

Our simulation program was carried out with the statistical package R (release 3.01)[4], univariate and multivariate models (MM) calculated with metafor[5] and mvmeta[6] packages, respectively.

We compared models on the following characteristics: the relative deviation between the estimated effects (EE) derived from two models A and B was calculated as $RD_{AB} = 100.|EE_A - EE_B|/EE_A$. The concordance between two tests was defined as the mean proportion of concordant decisions for the same data ($P < 0.05$ cutoff) over all the simulated tests, the relative precision of an EE as the ratio RP = EE/CIL (CIL = EE 95% half confidence interval length) over all the simulated data. For multiple endpoints study we defined two indexes: (1) provided that the highest the correlation between any two endpoints X and Y, the smallest the difference $EE_x - EE_y$ should be, we define the between endpoint discrepancy (BED) index BDI as the mean of deviations $|EE_x-EE_y|$ over all pairs of endpoints (X, Y) and all the studies weighted by the coefficient of determination R^2 (X, Y); and (2) As partial report of endpoints affects the comparability between endpoints, we determined the sensitivity of power to NST by the correlation coefficient R (RP, NST).

RESULTS

Description of the MA sample

Our sample consisted of MAs published between 1997 until 2014, 12 were congress abstracts out of which 9 were available in poster proceedings. Thirty-seven limited selection to randomized controlled trials. Eight treatments were compared. Forty-eight are based on literature findings, and 2 on individual patient data. The median NST was 13 (IQ = 7-17). Forty-three studies analyzed multiple endpoints. These endpoints were either continuous (drug dosage and duration, estradiol, hormonal values), counts (number of oocytes, metaphase Ⅱ oocytes, embryos, transferred embryos), ratios (embryo quality, implantation rate, *etc.*) or binary endpoints: Biochemical pregnancy [positive pregnancy test (β-hCG) 15-20 d post-hCG administration], clinical pregnancy (ultrasound scan with at least one sac with heartbeat 35-42 d post-hCG administration), ongoing pregnancy (viable pregnancy 10-12 wk after embryo transfer), live birth, multiple birth, occurrence of ovarian hyperstimulation syndrome, ectopic pregnancy or miscarriage. The median effect size converted to Risk Ratio was RR = 1.35 (IQ = 1.11, 1.53). The most referenced endpoints were the number of oocytes, biochemical or clinical pregnancy (76%), the least referenced was live birth (21%). The median number of referenced endpoints was 7 (IQ = 4, 12).

Comparing RM and FM

In our study sample, FM was used as the main model in 41 studies (82%). This choice was not justified for 22 MAs (54%). *Q*-test was mentioned in 34 studies but discussed only in 8 studies, and FM was used in spite of a detected significant heterogeneity in 13 studies. The forest plot was available for almost all ($n = 48$) MAs and allowed recalculation of the non-reported *Q*-test and I^2 statistics. The homogeneity assumption was rejected by the *Q*-test in 35 studies (70%), although in the 15 other MAs, NST was less than 7. The overall mean I^2 statistic was 58% (SD = 12) and 52% (SD = 12) for the 15 studies not rejected by the *Q*-test. Out of the 41 studies based on the FM, 29 (71%) were characterized by $I^2 >$ 40. Based on 10000 simulated samples based on the identified profile, our results are summarized as follows:

Difference of EEs: The mean relative deviation between FM and RM was RD = 4.3% + 2.1%, 46% exceeding a deviation of 5%, higher differences observed for larger heterogeneity (RD = 8.3% + 3.2% when $I^2 > 60\%$).

Power and precision: The mean concordance between RM and FM was 59% ± 6%, the ratio of the relative precision to $RP_{RM/FM} = 1.39 \pm 0.12$, and the ratio of power to $P_{RM/FM} = 1.33 + 0.09$, thus expected 33% more significant results found with FM. For NST ⩽ 7, this difference was larger ($P_{RM/FM} = 1.65 \pm 0.13$).

***Q*-test fallacy:** For MAs characterized by $I^2 = 30\%$, 50% and 75% considered as a moderate, substantial and strong heterogeneity, the mean power (at 0.05 level) to reject homogeneity was 0.32, 0.54 and 0.89 when NST = 7, and 0.53, 0.66 and 0.97, when NST = 17. Thus using *Q*-test on studies characterized by at least substantial heterogeneity falsely accepted homogeneity in 46% and 34% of studies for NST = 7 and 17 respectively, corresponding to the quartiles of the NST distribution of our MA sample.

Partial reporting of multiple endpoints

For these studies (86% of our sample), separate univariate MAs were conducted for each endpoint. The median ratio between the NST available for the most and less reported endpoint in each study was 4.2 (IQ = 2.1, 8.4). The within-study correlations between endpoints were not available from our MA sample. We approximated these values in estimating correlations between endpoints in available retrospective studies.

EE difference: The mean relative differences between RM and MM were RD = 12.3% (IQ = 1.1, 35.3), RD strongly increasing with the magnitude of the correlations and the partiality of reporting.

Between endpoint discrepancy: By using RR to quantify EE, the mean BDI was 0.08 ± 0.04 and 0.25

Table 1 Luteneizing hormone supplementation effect: Comparison between fixed, random and multivariate models

		BPR				CPR				OPR				LBR			
Overall	FM	1.17	1.01	1.28	0.03	1.08	0.99	1.17	0.03	1.09	0.93	1.19	0.25	1.28	0.98	1.56	0.07
	RM	1.14	0.95	1.38	0.16	1.07	1	1.19	0.05	1.05	0.93	1.19	0.45	1.23	0.88	1.72	0.23
	MM	1.23	0.94	1.61	0.12	1.09	1	1.2	0.06	1.13	1.03	1.24	0.01	1.13	1.01	1.28	0.04
POR	FM	1.22	0.97	1.53	0.1	1.30	1.03	1.64	0.03	1.29	0.82	2.02	0.28	1.81	0.99	3.31	0.06
	RM	1.19	0.87	1.64	0.27	1.30	0.99	1.64	0.04	1.29	0.82	2.02	0.28	1.70	0.78	3.69	0.27
	MM	1.37	1.00	2.21	0.04	1.27	1.01	1.61	0.04	1.38	1.06	1.81	0.02	1.53	1.09	2.15	0.01

Luteneizing hormone supplementation effect for overall population and POR models. Comparison between fixed (FM), random (RM), and multivariate (MM) models. Values are risk ratio 95%CI, and *P*-value. BPR: Biochemical pregnancy rate; CPR: Clinical pregnancy rate; OPR: On going pregnancy rate; LBR: Live birth rate; POR: Poor ovarian responder.

Table 2 Between endpoint correlation

	CP	OP	LB	NST	I^2
BP	0.95 (0.91, 0.97)	0.91 (0.84, 0.95)	0.87 (0.79, 0.93)	22	41.2
CP	-	0.96 (0.92, 0.97)	0.92 (0.85, 0.95)	39	31.8
OP	-	-	0.96 (0.92, 0.98)	13	46.36
LB	-	-	-	8	40.1

Within-study correlations (95%CI). NST: Number of studies reporting each endpoint; I^2: Heterogeneity index; BP: Biochemical pregnancy; CP: Clinical pregnancy; OP: On going pregnancy; LB: Live birth.

± 0.06 for MM and RM, MM reducing the discrepancy index of 68% compared with RM.

Sensitivity to NST: The mean correlation R (RP, NST) across any two pairs of endpoints and for every study was 0.75 ± 0.07 and 0.27 ± 0.5 for RM and MM, respectively, thus MM having a beneficial effect on reduction of the sensitivity to NST of 64%.

Power and precision of estimates: The mean concordance of decision between RM and MM was 57% ± 8%. The ratio of the 95%CI length (CIL) of FM on RM was $CIL_{RM/MM}$ = 0.65 ± 0.13 and the mean power ratio $P_{MM/RM}$ = 1.57 ± 0.38, thus in average 57% more decisions of significant differences in using MM.

Study case

To illustrate these general principles, we re-analyzed the data of a recent MA[7]. Adding recombinant human luteinizing hormone (r-hLH) to recombinant human follicle-stimulating hormone (r-hFSH) during ovarian stimulation has motivated numerous studies and conflicting MAs, suggesting a particular benefit on poor responders (POR) compared with normal responders (NOR). The selection was comprised of RCTs on women (18-45 years) undergoing IVF/ICSI treated with r-hFSH plus r-hLH (FL group) or r-hFSH alone (F group). From 40 RCTs (6443 patients), we identified 45 separate reports on 31 NOR and 15 POR studies. Clinical pregnancy rate (CP) was significantly higher in FL group both for the overall sample and POR subgroup. However, the analysis failed to find significant results for the other pregnancy markers, due to partial reporting, biochemical BP, clinical CP, ongoing OP pregnancies and live birth LB, reported in the selected studies with 22, 39, 13 and 8 studies, respectively. We applied three methods on these data (Table 1): The univariate analyses based on FM and RM, and a multivariate meta-analysis (MM).

Through a bootstrapped replication (n = 1000), we assessed the superiority of the r-hLH supplementation effect in testing three models: (1) effect on overall population (intercept only); (2) restricted effect on POR only; and (3) effect on overall population with an additional effect on POR. By selecting our model based on a significant decrease of the Akaike Information Criteria value, model (2) was the more likely, but not significantly better than model (1), whereas model (3) was rejected, thus we restrained our analysis in testing the two first models (1) and (2).

Within-study correlations were unknown and handled following various techniques: Among possible handling techniques[8], we narrowed the range of possible values based on existing data on 10 centers at our disposal in conducting sensitivity analyses by imputing values within the 95%CI. Strong correlations were found between the endpoints (Table 2).

The proportion of concordant decisions (P < 0.05) was 5/8, 1/8, and 1/8 between FM-RM, RM-MM and FM-MM, respectively. The relative precision of estimates were RP = 0.26, 0.34 and 0.19 for FM, RM and MM, respectively. Among the three significant differences detected by RM, only one was confirmed by RM, highlighting FM anticonservatism, in this heterogeneous example (I^2 > 40%). RM detected only one significant difference for CPR, but this is also by far the most reported endpoint. MM detected consistent estimates with respect with high correlations and similar significant values unrelated to NST. Overall, the BED index of each model was BDI = 0.61, 0.43 and 0.16 for FM, and the sensitivity to NST was 0.85, 0.54 and 0.18. Thus MM reduced BDI and sensitivity to NST to 63% and 66% respectively. MM reducing sensitivity of 66% compared with RM. For LB in particular, characterized by the smallest NST, univariate models found a strong although non-significant values of 1.81 and 1.7, inconsistent with other endpoints, in spite of high correlations. Finally, we repeated this analysis in adding the number of oocytes, the number of embryos, and implantation rate (not reported). Very few differences were found

on these endpoints, RR on LB values were 1.51 (1.1, 2.03), P = 0.006, and 1.11 (0.97, 1.23), P = 0.06. For the POR and overall model, respectively. Our analysis provides evidence of a clinically relevant effect of live birth for POR supplemented with r-hLH, compared with r-hFSH alone, this effect remaining consistent for all the pregnancy endpoints.

DISCUSSION

Random or FM

In our MA sample, the FM was overwhelmingly preferred (82%), often unjustified, in spite of evidence of heterogeneity (only 7.2% of studies such that I^2 < 40%). The heterogeneity of effect among studies may be explained by the strong multifactorial variability observed in pregnancy predicting models[2,3] in which center variability was found as the major predictor, followed by the patient mix (age, ovarian reserve, *etc.*). The reasons behind center heterogeneity are many: Differences between used medication and dosage, staff expertise, differences in protocols and standard therapy, and patient mix differences. Although a fixed effect remains possible, the accumulation of causes potentially generating a variable treatment effect designates RM as the most likely model.

To which extent their results derived from FM can be considered as a reasonable approximation? Our simulation suggests a deviation higher than 5% for 46% of the studies and higher deviations expected for high heterogeneity (I^2 > 60%). The difference is more worrying concerning precision and power: The between-study τ^2 variance assumed in the RM increases the standard error and the confidence interval length, thus produces generally more conservative tests in particular for small NST. Previous papers warn that although the power of FM is generally better, the gain of power becomes uncertain for RM in particular when NST is small and I^2 increases[9]. Our simulation provides more accurate conclusions for ART: FM and RM provide inconsistent results with 33% and 65% more significant results in using FM for all the studies and small studies (NST < 7) respectively. The problem of discordance between the two tests is more concerning than the difference between estimates: FM is expected to falsely find significant differences in situations where heterogeneity is present as it is generally the case in ART studies.

Another crucial conclusion is the Q-test fallacy: In spite of its popularity, previous researches[9,10] warned on the high sensitivity of the power of this test with NST. Our simulation clearly highlights that, although it is the commonest test to select between RM and FM, the Q-test is not reliable in falsely accepted homogeneity in 46% and 34% of studies for NST = 7 and 17 respectively.

Univariate compared with multivariate approach

We limited the comparison between RM and MM. The median ratio of 4.2 between the available NST between the most and less reported endpoints implies expected consequences on precision, power and estimates consistency:

(1) Between endpoint discrepancies: MM model has been demonstrated to provide a unique solution offering optimal consistency between estimates[11]. In our results, considerable relative deviations between RM and MM were observed (RD = 18.3% + 7.3%). Moreover, taking into account the correlation between endpoints, the mean discrepancy index BDI was 0.08 ± 0.04 and 0.25 ± 0.06 for MM and RM, respectively. Thus MM increases the consistency between endpoint of 68% compared with RM.

(2) MM provides optimal estimates in reducing bias and sensitivity to partial reporting of multiple endpoints[11]. We confirm this result from our simulations: The mean correlation R (RP, NST) were 0.75 ± 0.07 and 0.27 ± 0.5 for RM and MM, respectively, thus MM reducing the undesired sensitivity to NST = 64%.

(3) The standard error of the estimates are always better in MM compared with univariate models MAs[11]. In our simulation, the power of MM is 57% more than RM. As both tests are applicable, univariate RM is more conservative than MM.

And (4) Feasibility and assumptions: MM requires more assumptions than univariate approach in particular the multivariate normal assumption, another important constraint is the knowledge of within-study correlations, which is generally difficult to obtain, as ideally these values necessitate individual patient data. Alternative techniques are possible to substitute approximations[11], and sensitivity analyses around these estimates are needed to assess the stability of the results to these approximations.

Although MM is characterized by multiple advantages, it was rarely used in practice in ART. The reasons include tradition, defiance against apparently more complex model, underlying hypotheses sometimes difficult to assess, more data necessary often non available (such as within study correlations for which simplifying their values needs further research), and the lack of easy software to implement this technique.

In conclusion, the concept of MA is now widely accepted, but many methodological aspects remain controversial and the choice of an unjustified model may result into strongly biased results. This paper highlights this particular aspect in ART where the frequent use of the fixed traditional FM is source of important biases, due to the strong observed heterogeneity and partial reporting of multiple endpoints. Based on our results, we suggest the following implications in practice:

(1) RM must be regarded as the appropriate model for MAs in ART research. FM should be considered only upon robust a priori justification concerning the homogeneity of the studied question and confirmed by the observed I^2. The Q-test should definitely be disregarded. More conservative than FM, RM has the same power in case of homogeneity of effect, in which case the two results coincide. Thus selecting RM does

not involve loss of power.

(2) The level of heterogeneity needs to be reported and discussed. The I^2 statistics and Tau value (standard deviation of the between study size) constitute good descriptive measurements of heterogeneity. Failing to provide heterogeneity level may induce mis-interpretations. An effect size of RR = 1.50 seems conclusive, however, a large value of I^2 means an important dispersion of this value. In that sense, the Tau value allows calculation of the proportion of studies for which RR becomes non-clinically meaningful.

(3) MAs may be characterized by very large between-study heterogeneity (I^2 > 75%); this happens particularly when important differences of selection criteria are observed between studies. These studies, not uncommon in ART (19% of our sample), are subject to controversy, some arguing that the summary of results is based on non comparable studies. The RM remains fully applicable in these cases, with an obvious loss of power, price to pay to demonstrate the generalizability of the efficacy of a treatment across heterogeneous situations.

And (4) MAs involving partially reported multiple endpoints suffer from chaotic difference between the effects on the studied endpoints, the precision of the endpoints badly affected by the available NST. MM takes advantage of the between-endpoint inter-correlation by borrowing strength from all the other endpoints, to provide a consistent, unique and comparable estimate for all the endpoints, thereby compensating for the effect due to unequal sample sizes. The MM is much more consistent, not affected by partial reporting, and with more accurate estimates on all the endpoints.

ACKNOWLEDGMENTS

The author would like to thank Dr. Dan Jackson (MRC Biostatistics Unit, Cambridge, United Kingdom) for fruitful discussion on multivariate MA questions.

COMMENTS

Background

Meta-analysis (MA) is a widely accepted technique employed in the synthesis and evaluation of evidence from past clinical trials. The growing impact of MA on clinical practice requires that they should be conducted according to high quality standards. Numerous MAs have been published in assisted reproductive technology (ART) and conflicting results are not uncommon. A rarely discussed reason is the choice of the meta-analytical model: Irrespective of expected heterogeneity of the studied effect and partially reported multiple endpoints, the fixed model (FM) is almost always used. The objective of the present study is to assess the extent to which this approach or other models are more appropriate.

Research frontiers

Although developed since more then 10 years, multivariate MA constitutes a promising approach for multiple endpoints MAs. In summary, multivariate model (MM) supposes a prior knowledge of each within-study correlation between endpoints, and assumes the existence of an additional unknown between-study random effect [as the univariate random model (RM)]. By fixing this random effect to zero, the MM generalizes the univariate fixed model (FM). The endpoints are assumed to be distributed according a multivariate normal distribution, the individual effect estimates are calculated as weighted mean over all the studies and in addition to univariate approach taking into account the correlation between endpoints. The multivariate MM model palliates the apparent deficiencies of separate univariate analysis for multiple endpoints MAs. However, these preliminary theoretical considerations show that the advantages of these models depend on study parameters, the number of studies (NST), considered endpoints, the level of partial reporting, the between study heterogeneity and the between endpoint correlations between effects.

Innovations and breakthroughs

The validity of models were discussed in statistical articles, rarely in applications, and this research provides evidence of the importance of this choice. The most known reasons of controversy in results are differences in study selection or elimination depending on whether or not they were published, blind, randomized, with a sufficient methodological quality risk of bias between studies. Much less discussed, the choice of the meta-analytical method may involve potentially strong differences on results. In restraining to ART context, at least two important concerns may be mentioned: (1) A majority of MAs used the traditional and simplest FM, in which the studied treatment effect is assumed constant across any study. An essential specificity of ART is the considerable difference of practice, procedures, medication use and know-how among studies, countries or centres causing very heterogeneous performances. To which extent this assumption of constant effect remains bearable, although in most MAs this assumption was not tested? Should other models such the RM admitting an heterogeneous treatment effect be more adapted, while being more conservative? (2) In most of the studies, several endpoints are evaluated, such as the number of retrieved oocytes, embryos, implantation rate, and pregnancy ratios. Their separate analyses involve difficulties in the discussion and become non-comparable for endpoints reported by different NST. This is called partial reporting and is very common in ART MAs. Separate MAs for each endpoint where high correlation and partial reporting co-exist produce a frequent paradox characterized by conflicting results on correlated endpoints, simply due to non-comparable power. The trickiest case is live birth, the ultimate endpoint in ART, much less reported than other endpoints, compared with earlier markers like clinical pregnancy necessitating much less follow up. Unlike univariate MA, the Multivariate MA is a recent proposal taking advantage of between endpoints correlations. To which extent using simple univariate models remains acceptable in such conditions, or is it worthwhile to turn to multivariate approach? This research is based on the hypothesis that ART studies are characterized by a homogeneous profile of characteristics enabling the adapted choice of a meta-analytical model. In a first stage, the author determined this profile based on a sample of MAs selected from a literature review, at a second stage the author conducted simulation studies based on this profile to compare the models, andthe authorsapply these principles on a study case.

Applications

These results may have a very important practical implication for ART/*in vitro* fertilization (IVF) researchers: The conclusions are very simple and strictly specific for MAs in this pathology: The FM has always been classically used and may provide important bias. The MM is the only model allowing a clear solution for multiple endpoints MA. For a better clinical understanding, I provide a practical example based on IVF data easily interpretable for clinicians.

Terminology

For reader less familiar with meta-analytical models,the authorssummarize the principles of the FM, RM and MM in appendix. In summary FM assumes a fixed effect across all the studies. The overall estimated effect of the studied treatment compared with control is estimated by the mean of all the study estimates weighted by the reciprocal of their variance. An alternative to FM is the RM generalizing FM by assuming that the effect varies across studies, according to a normal distribution with unknown mean and variance to be estimated. The Q-test can be used before and compares the adequacy of FM and RM in testing the significance of the dispersion of the EE_i, however this test is known oversensitive with the NST. Another heterogeneity measurement, I^2 statistic, evaluates the percentage of variation attributable to the between study heterogeneity. Various rules were based on I^2, in particular a value exceeding 40% evidencing a substantial heterogeneity should motivate the choice of RM. An introductory and seminal approach to MM can be found in. In summary,

MM supposes an prior knowledge of each within-study correlation between endpoints, and assumes the existence of an additional unknown between-study random effect (as the univariate RM). The multivariate MM model palliates the apparent deficiencies of separate univariate analysis for multiple endpoints MAs. However, these preliminary theoretical considerations show that the advantages of these models depend on study parameters, the NST, considered endpoints, the level of partial reporting, the between study heterogeneity and the between endpoint correlations between effects.

Peer-review

The author uses a mathematical approach to demonstrate existing concerns with meta-analyses conducted with studies on ART. In general, the author applies the findings from the analyses to support existing recommendations for "best practices" when performing a MA.

REFERENCES

1 **Eysenck HJ**. Meta-analysis and its problems. *BMJ* 1994; **309**: 789-792 [PMID: 7950571 DOI: 10.1136/bmj.309.6957.789]

2 **Arvis P**, Lehert P, Guivarc'h-Levêque A. Simple adaptations to the Templeton model for IVF outcome prediction make it current and clinically useful. *Hum Reprod* 2012; **27**: 2971-2978 [PMID: 22851717 DOI: 10.1093/humrep/des283]

3 **Paul SR**, Donner A. Small sample performance of tests of homogeneity of odds ratios in K 2 x 2 tables. *Stat Med* 1992; **11**: 159-165 [PMID: 1579755 DOI: 10.1002/sim.4780110203]

4 R, A Language and Environment for Statistical Computing, R Development Core Team, R Foundation for Statistical Computing. Vienna, Austria: 2010. Available from: URL: http://www.r-project.org

5 **Viechtbauer W**. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010; **36**: 1-48 [DOI: 10.18637/jss.v036.i03]

6 **Gasparrini A**, Armstrong B, Kenward MG. Multivariate meta-analysis for non-linear and other multi-parameter associations. *Stat Med* 2012; **31**: 3821-3839 [PMID: 22807043 DOI: 10.1002/sim.5471]

7 **Lehert P**, Kolibianakis EM, Venetis CA, Schertz J, Saunders H, Arriagada P, Copt S, Tarlatzis B. Recombinant human follicle-stimulating hormone (r-hFSH) plus recombinant luteinizing hormone versus r-hFSH alone for ovarian stimulation during assisted reproductive technology: systematic review and meta-analysis. *Reprod Biol Endocrinol* 2014; **12**: 17 [PMID: 24555766 DOI: 10.1186/1477-7827-12-17]

8 **Nam IS**, Mengersen K, Garthwaite P. Multivariate meta-analysis. *Stat Med* 2003; **22**: 2309-2333 [PMID: 12854095 DOI: 10.1002/sim.1410]

9 **Hardy RJ**, Thompson SG. Detecting and describing heterogeneity in meta-analysis. *Stat Med* 1998; **17**: 841-856 [PMID: 9595615 DOI: 10.1002/(SICI)1097-0258(19980430)17:8<841::AID-SIM781>3.0.CO;2-D]

10 **Borenstein M**, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods* 2010; **1**: 97-111 [PMID: 26061376 DOI: 10.1002/jrsm.12]

11 **Jackson D**, Riley R, White IR. Multivariate meta-analysis: potential and promise. *Stat Med* 2011; **30**: 2481-2498 [PMID: 21268052 DOI: 10.1002/sim.4172]

P- Reviewer: Wang R
S- Editor: Gong XM **L- Editor**: A **E- Editor**: Jiao XK

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i6.232

World J Meta-Anal 2015 December 26; 3(6): 232-253
ISSN 2308-3840 (online)

META-ANALYSIS

Role of self-expanding metal stents in patients with malignant colorectal obstruction: A systematic review and meta-analysis

Nirav Thosani, Subhas Banerjee, Vikesh Khanijow, Bhavana Rao, Priyanka Priyanka, Atilla Ertan, Sushovan Guha

Nirav Thosani, Subhas Banerjee, Division of Gastroenterology and Hepatology, Stanford University School of Medicine, Stanford, CA 94304, United States

Nirav Thosani, Vikesh Khanijow, Bhavana Rao, Priyanka Priyanka, Atilla Ertan, Sushovan Guha, Division of Gastroenterology, Hepatology and Nutrition, Department of Internal Medicine, the University of Texas Health Medical School at Houston, Houston, TX 77030, United States

Nirav Thosani, Priyanka Priyanka, Atilla Ertan, Sushovan Guha, Ertan Digestive Disease Center, Memorial Hermann Hospital-TMC, Houston, TX 77030, United States

Author contributions: Thosani N contributed to study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; statistical analysis; Banerjee S and Ertan A contributed to drafting of the manuscript; critical revision of the manuscript for important intellectual content; Banerjee S contributed to administrative, technical, or material support; study supervision; Khanijow V and Rao B contributed to acquisition of data; analysis and interpretation of data; drafting of the manuscript; Priyanka P contributed to statistical analysis; analysis and interpretation of data; drafting of the manuscript; Guha S contributed to study concept and design; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; administrative, technical, or material support; study supervision.

Conflict-of-interest statement: No potential conflict of interest exists between authors.

Data sharing statement: Data for this study is publicly available.

Correspondence to: Nirav Thosani, MD, MHA, Division of Gastroenterology, Hepatology and Nutrition, Department of Internal Medicine, the University of Texas Health Medical School at Houston, 4.234 Medical School Building, 6431 Fannin Street, Houston, TX 77030, United States. nirav.thosani@uth.tmc.edu
Telephone: +1-713-5006686
Fax: +1-713-5006699

Received: April 30, 2015
Peer-review started: April 30, 2015
First decision: June 24, 2015
Revised: October 7, 2015
Accepted: November 3, 2015
Article in press: November 4, 2015
Published online: December 26, 2015

Abstract

AIM: To assess the safety and efficacy of self-expandable metal stents (SEMSs) for malignant colorectal obstruction.

METHODS: Data regarding technical success, clinical success, and procedure related complications were collected from included studies. DerSimonian-Laird random effects model was used to generate the overall outcome. Thirty international studies with a total of 2058 patients with malignant colorectal obstruction were included.

RESULTS: The technical and clinical success rates for SEMS placement were 94% (95%CI: 92-96) and 91% (95%CI: 88-93), respectively. Overall complication rate for SEMS was 23% (95%CI: 18-29). Stent migration

8% (95%CI: 6-10) and stent obstruction 8% (95%CI: 6-11) were the most common complications, followed by perforation 5% (95%CI: 4%-7%). Surgical or endoscopic re-interventions were needed in 14% (95%CI: 10-18) of patients. Endoscopic repeat stent placement was required in 8% (95%CI: 6-10), while surgical intervention was needed in 6% (95%CI: 4-8).

CONCLUSION: SEMS are effective when used as palliation or bridge to surgery for malignant colorectal obstruction with high technical and clinical success. About 14% of patients require repeat endoscopic or surgical intervention for stent failure or to manage stent related complications.

Key words: Metal stent; Colorectal cancer; Colon cancer; Rectal cancer; Intestinal obstruction; Bowel obstruction; Malignant obstruction; Colonic obstruction

Core tip: The technical and clinical success rates for self-expandable metal stents (SEMSs) placement were 94% (95%CI: 92-96) and 91% (95%CI: 88-93), respectively. Overall complication rate for SEMS was 23% (95%CI: 18-29). Stent migration 8% (95%CI: 6-10) and stent obstruction 8% (95%CI: 6-11) were the most common complications, followed by perforation 5% (95%CI: 4%-7%). Surgical or endoscopic re-interventions were needed in 14% (95%CI: 10-18) of patients. Endoscopic repeat stent placement was required in 8% (95%CI: 6-10), while surgical intervention was needed in 6% (95%CI: 4-8). SEMS are effective when used as palliation or bridge to surgery for malignant colorectal obstruction with high technical and clinical success. About 14% of patients require repeat endoscopic or surgical intervention for stent failure or to manage stent related complications.

Thosani N, Banerjee S, Khanijow V, Rao B, Priyanka P, Ertan A, Guha S. Role of self-expanding metal stents in patients with malignant colorectal obstruction: A systematic review and meta-analysis. *World J Meta-Anal* 2015; 3(6): 232-253 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i6/232.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i6.232

INTRODUCTION

Colorectal cancer (CRC) is a devastating disease that impacts patients and healthcare systems significantly globally[1]. In United States, CRC is the second leading cause of cancer related mortality. In addition to early colon cancer screening, it is estimated that up to 50830 deaths will be attributed to CRC in the United States in 2013[1,2]. Approximately, 7% to 29% CRC patients presents with colon obstruction during clinical presentation[3-5]. Conventionally, these patients underwent emergency surgery (ES) with the creation of temporary or permanent colostomies to resolve symptoms. Despite an effective treatment alternative, surgery has been associated with high morbidity (32%-64%) and mortality rates (15%-34%)[4,5].

Spinelli *et al*[6] (1992) introduced self-expandable metal stents (SEMSs) to be used as palliative initial therapy for malignant rectal obstruction. On these lines, colonic stent placement may be used to palliate obstructive symptoms either for those patients where resection is not deemed an option, or to allow bowel preparation prior to elective surgery, popularly known as bridge to surgery (BTS). Early studies supported colonic stent placement as reduction in mortality, morbidity and required number of colostomies observed, along with its ability to prevent the need for ES in patients that have disseminated metastatic disease or critical surgical risks[7,8]. Pooled analysis by Sebastian *et al*[9] provided support for the placement of SEMS for neoplastic colonic obstruction, as technical and clinical success rates observed were 94% and 91% respectively. However, some studies had raised the concerns regarding safety of SEMS as high rate of long term adverse events including perforation were reported[5,10,11]. In 2007, a multicenter randomized trial that compared safety of SEMS over surgery for palliation of obstruction in stage Ⅳ left-sided CRC patient population was prematurely terminated, due to the unexpectedly high rate of perforation in the non-surgical patients[12]. In 2011, another multicenter randomized trial, assessing safety of SEMS over ES in acute left-sided malignant obstruction patient cohort was also terminated prematurely, as an interim analysis showed increased 30-d morbidity in the SEMS group with no significant increase in their mean global health status[5]. Given the significant heterogeneity with conflicting outcomes in published literature, the objective of the current study was to systematically review the safety, efficacy and overall clinical impact of SEMS in patient cohort with malignant colorectal obstruction with a comprehensive meta-analytic approach.

MATERIALS AND METHODS

Literature search strategy

Literature search for this systematic review was performed using the established guidelines (PRISMA) for systematic review[13]. Databases such as MEDLINE [Ovid MEDLINE(R) in-process and other non-indexed citations, Ovid MEDLINE(R) daily, Ovid MEDLINE(R) and Ovid OLD MEDLINE (R) 1946 to June 2013], SCOPUS (MEDLINE and EMBASE), Cochrane Database of Systematic Reviews, Google scholar, and CINAHL Plus were searched. A systematic literature search was performed by using following search terms: (1) "stent" and "CRC or colon cancer or rectal cancer or colonic obstruction or intestinal obstruction or bowel obstruction or malignant obstruction"; (2) "metal stent" and "CRC or colon cancer or rectal cancer or colonic obstruction or intestinal obstruction or bowel

obstruction or malignant obstruction"; (3) SEMS and "CRC or colon cancer or rectal cancer or colonic obstruction or intestinal obstruction or bowel obstruction or malignant obstruction"; and (4) SEMS and "CRC or colon cancer or rectal cancer or colonic obstruction or intestinal obstruction or bowel obstruction or malignant obstruction". Reference list of all the selected articles was screened to avoid exclusion of any potential article in the initial search. Literature search was limited to human subjects. Final screening to determine eligibility criteria for all the articles was performed independently by two investigators (NT and VK). All the differences were resolved by discussion with two senior investigators (MS and SG) on this study. After consensus, final report was retrieved and reviewed by the same two investigators (NT and VK).

Inclusion criteria

Study population: Patients diagnosed with primary or metastatic CRC having clinical and/or radiologic signs and symptoms of malignant bowel obstruction identified as cohort for this study.

Intervention: The intervention was endoscopic SEMS placement.

Study design: Both retrospective and prospective studies who identified patients with malignant bowel obstruction and underwent endoscopic placement of SEMS for palliation or BTS were included.

Outcome: The primary outcome of the study was to evaluate endoscopically placed SEMS technique for the following parameters: Technical and clinical success rate, rate of occurrence of adverse events (perforation, stent migration, and obstruction) and rate of need for re-interventions (surgical, endoscopic re-stenting, and other endoscopic interventions).

Exclusion criteria

Case reports, case series and studies with insufficient data were excluded. The studies where stents placement was non-endoscopic by intervention radiologists were also excluded from the analysis.

Data abstraction

Data was abstracted by two independent investigators (NT and VK) from the studies that met eligibility criteria. All the following extracted data was placed on standardized forms (Microsoft Excel, Microsoft Corporation, Redmond, Wash): (1) Study characteristics: study design, setting and criteria, country, year of publication, sample size, clinical context; (2) Demographics: age (mean), male and female patient proportion; and (3) Interventions: SEMS type and manufacture.

Outcomes: Technical and clinical success, complication, re-intervention rate.

Assessment of risk of bias

Assessment of study quality and risk of bias was performed as per guidelines established by Cochrane handbook[14]. Methods included were randomization schedule, conceal allocation, whether blinding was implemented, what proportion of patients completed follow-up, whether an intention-to-treat analysis was extractable, and whether there was evidence of selective reporting of outcomes. Two authors (NT and BR) evaluated independently and conflicting issues were resolved after discussion with senior investigators (SG).

Data synthesis and analysis

For the purpose of this meta-analysis, the technical success was defined as accurate endoscopic SEMS deployment, with adequate stricture coverage without any immediate procedure related adverse event. Decompression and relief of obstructive symptoms within 72 h of SEMS placement without any adverse events was identified as clinical success. Obstruction was defined as obstruction due to tumor ingrowth, tumor overgrowth and fecal impaction requiring intervention. Other adverse events assessed include perforation, migration, bleeding, tenesmus and any other related symptoms. Re-intervention was defined as surgical or endoscopic procedures done due to technical failure, clinical failure or adverse events.

DerSimonian-Laird random effects model was used to pool the data across the studies and to obtain overall estimates (in percentages) and the 95% confidence intervals (CIs). Robustness of the random-effects model is more than the fixed effects model as it incorporates weighing scheme both within and between-study variations[15]. Further subgroup analysis was performed for (1) indication of SEMS placement (palliation *vs* BTS); (2) prospective *vs* retrospective studies; (3) single center *vs* multicenter studies; and (4) geographical location (*i.e.*, North America, Europe, and Asia).

Cochran Q statistic was used to assess statistical heterogeneity with and between studies and further quantified using I^2 statistics[16]. We arbitrarily defined I^2 values of 25%, 50% and 75% for low, moderate and high heterogeneity, respectively and were generally used for descriptive purposes only[17]. To determine each study influence on pooled OR, removal of each study was done at a time in the meta-analysis during sensitivity analysis. Tools like Egger regression asymmetry test[18], Fail-safe N tests, and the trim-and-fill method[19] were used to assess the robustness of the meta-analysis for the publication bias. To further evaluate publication bias using the standard error and diagnostic odds ratio Funnel plot was constructed[20,21]. All statistical tests were performed with the Comprehensive Meta-analysis version 2.0 (Biostat, Englewood, NJ). A value of $P < 0.05$ was considered statistically significant for this meta-analysis. Statistical review of the study has been performed by an experienced biomedical statistician.

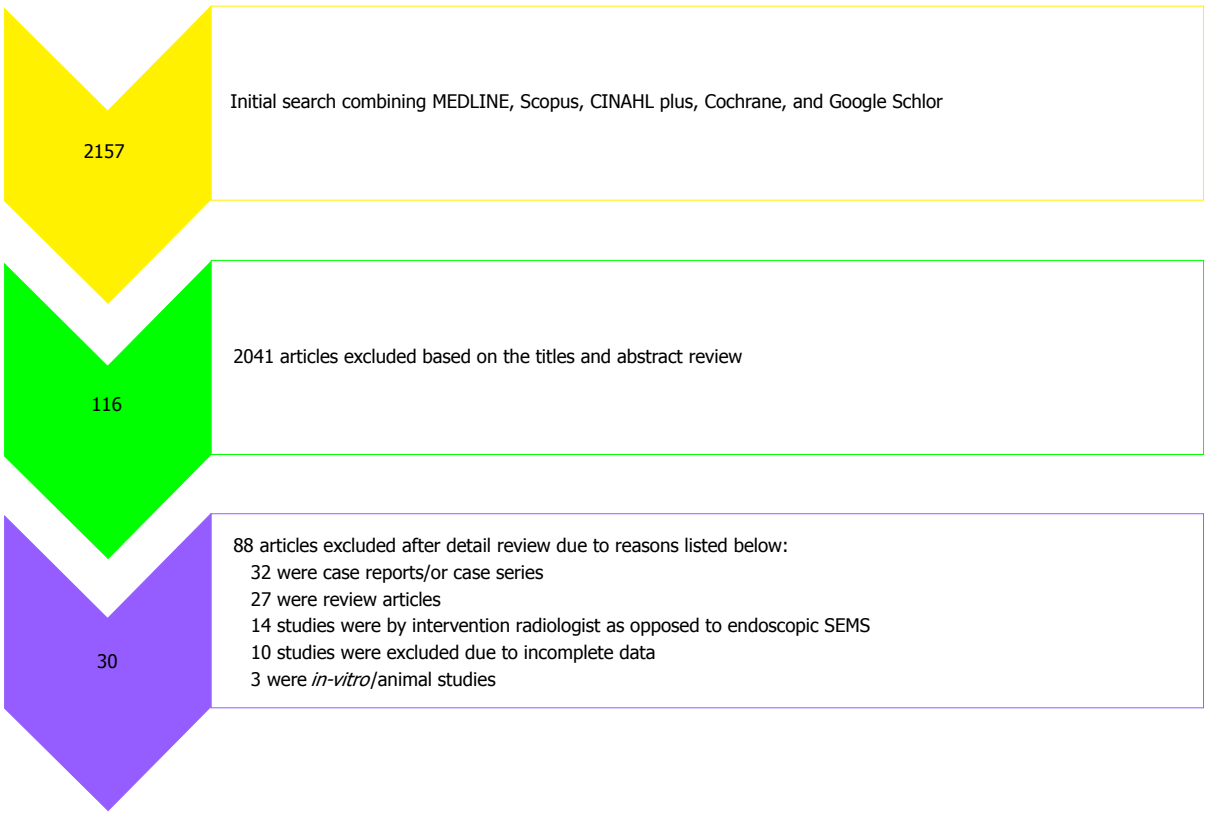


Figure 1 Study selection process. SEMS: Self-expandable metal stent.

RESULTS

Literature search

Figure 1 depicts the study selection process. Thirteen studies from Europe[5,10,22-32], 12 studies from Asia[7,33-43], 3 studies from North America[11,44,45] and 1 study each from Australia[8] and Africa[46] were included. There were 8 prospective studies[22-24,27,31,33,36-38] 17 retrospective studies[7,8,10,11,25,26,28-30,32,34,35,39,40,42,44,45], 4 randomized control trials (RCTs)[5,41,43,46] and one combined retrospective and prospective study[24]. Another RCT by Pirlet *et al*[47] was excluded from the meta analysis as approximately half of the SEMS in this study were placed by an interventional radiologist non endoscopically. In total, 2058 patients with malignant CRC obstruction were treated with endoscopic SEMS placement: 1313 for palliation and 745 as BTS. The study characteristics for all the included studies are shown in Table 1. The results regarding assessment for risk and bias and overall study quality are shown in Figure 2.

SEMS type

Multiple types of stents were used in these studies including Enteral Wallstent[5,7,8,10,11,23-26,29,40,41,44,45] (Boston Scientific, Natick, MA), uncovered esophageal Wallstent[7] (Boston Scientific, Natick, MA), Ultraflex Precision colonic stent[7,11,22,24,25,44] (Boston Scientific, Natick, MA), Polyflex esophageal stent[44] (Boston Scientific, Natick, MA), Wallflex colonic stent[5,10,11,27,31,35,38,39,44,45] (Boston Scientific, Natick, MA), Choo-stent[7,23,25] (M.I. Tech Co, Ltd, Seoul, South Korea), Memotherm-Stent[23,25] (Bard, Germany), Colonic Z-Stent[25] (Cook Medical, Inc., Bloomington, IN), EsophaCoili[25] (Medtronic/Instent, Eden Prairie, MN), Microtech[8,36] (Nanjing Microinvasive Co, China), Hanarostent[10,29,32-35] (M.I.Tech Co, Seoul, South Korea), Niti-S colonic covered[35,37] (Taewoong Medical Co, Seoul, South Korea), Niti-S Colonic uncovered stent[37,39] (Taewoong Medical Co, Seoul, South Korea) and Comvi stent[38,39] (Taewoong Medical Co, Gimpo, South Korea).

Meta-analysis

Technical and clinical success: Thirty studies reported technical success and were included in the analysis. Overall technical success rate was 94% (95%CI: 91.8-95.6) (Figure 3A). The I^2 value for heterogeneity analysis was 58. The technical success rate in the palliation group (10 studies) was 94.2% (95%CI: 91.3-96.1) and in the BTS group (8 studies) was 89.4% (95%CI: 79.5-94.8).

Twenty-nine studies reported clinical success and were included in the analysis. Overall clinical success rate was 90.6% (95%CI: 88.1-92.7) (Figure 3B). The I^2 value for heterogeneity analysis was 51. The clinical success rate in the palliation group (10 studies) was 91.7% (95%CI: 88.7-94) and in the BTS group (8 studies) was 87.9% (95%CI: 78.1-93.7). The results of the subgroup analysis for both technical and clinical success based on indication, center, design, and region are shown in detail in Table 2.

Table 1 Studies included in meta-analysis

	Ref.	Study design	SEMS		Total
			Palliation	BTS	
Palliation					
1	Ptok[23]-2006-Germany	Prospective, Single Center	48	0	48
2	Repici[22]-2007-Europe	Prospective, Multicenter	44	0	44
3	Im[33]-2008-South Korea	Prospective, Single Center	49	0	49
4	Suh[34]-2010-South Korea	Retrospective, Single Center	55	0	55
5	Jung[35]-2010-South Korea	Retrospective, Single Center	39	0	39
BTS					
6	Fregonese[24]-2008-United Kingdom	Prospective and Retrospective, Multicenter	0	36	36
7	Li[36]-2010-China	Prospective, Single Center	0	52	52
Palliation and BTS					
8	Meisner[25]-2004-Denmark	Retrospective, Single Center	51	38	89
9	Soto[26]-2006-Spain	Retrospective, Single Center	36	22	58
10	Lee[37]-2007-South Korea	Prospective, Single Center	37	43	80
11	Repici[27]-2008-Europe	Prospective, Multicenter	23	19	42
12	Stipa[28]-2008- Italy	Retrospective, Single Center	9	22	31
13	Branger[29]-2010-France	Retrospective, Single Center	66	27	93
14	Fernandez-Esparrach[10]-2010-Spain	Retrospective, Single Center	38	9	47
15	Lee[44]-2010-United States	Retrospective, Single Center	41	5	46
16	Park[38]-2010-South Korea	Prospective, Single Center	107	44	151
17	Small[11]-2010-United States	Retrospective, Single Center	168	65	233
18	West[30]-2010-United Kingdom	Retrospective, Multicenter	21	6	27
19	Meisner[31]-2011-Denmark	Prospective, Multicenter	257	182	439
Palliative SEMS *vs* palliative surgery					
20	Law[7]-2003-China	Retrospective, Single Center	30	0	61
21	Carne[8]-2004-New Zealand	Retrospective, Single Center	25	0	44
22	Suarez[32]-2009-Spain	Retrospective, Single Center	45	0	98
23	Vemulapalli[45]-2009-United States	Retrospective, Single Center	53	0	123
24	Lee[39]-2011-South Korea	Retrospective, Single Center	71	0	144
SEMS as BTS *vs* emergency surgery					
25	Ng[40]-2006-Hong Kong	Retrospective, Single Center	0	20	60
26	Cheung[41]-2009-China	Randomized Controlled Trial, Single Center	0	24	48
27	Guo[42]-2011-China	Retrospective, Single Center	0	34	92
28	vanHooft[5]-2011-The Netherlands	Randomized Controlled Trial, Multicenter	0	47	98
29	Ho[43]-2011-Singapore	Randomized Controlled Trial, Single Center	0	20	39
30	Abdel-Hamid[46]-2013-Egypt	Randomized Controlled Trial, Single Center	0	30	60
	Total		1313	745	2058

BTS: Bridge to surgery; SEMS: Self-expandable metal stent.

Complications

The overall complication rate included complications like perforation, migration, and stent obstruction and any other reported complication like stent related bleeding, tenesmus, *etc*. Twenty five studies reported all complications in detail and were included in the analysis of overall complication rate. The overall complication rate was 23.1% (95%CI: 18.5-28.6) (Figure 4A). The I^2 value for heterogeneity analysis was 82. The complication rate was higher for palliation group (27.3%, 95%CI: 18.7-38) compared to BTS group (13.8%, 95%CI: 8.3-22.2).

Stent occlusion due to tumor ingrowth, overgrowth or fecal impaction was the most common complication with overall stent occlusion rate of 8.3% (95%CI: 6.0-11.4) (Figure 4B). Stent occlusion was most likely function of tumor growth and disease progression over time, as for palliation group stent occlusion rate was 9.5% (95%CI: 5.4-16.4) compared to only 1.9% (95%CI: 0.5-7.1) for BTS group.

Stent migration was seen in 7.6% (95%CI: 5.7-10.0) of cases (Figure 4C). Once again, stent migration was more frequently seen in palliation group (10.2%, 95%CI: 7.1-14.5) compared to BTS group (4.1%, 95%CI: 2.0-8.1).

Perforation was seen in 4.9% (95%CI: 3.6-6.6) cases (Figure 4D). Perforations being early complication after stent placement, perforation rates were similar between palliation group 5.4% (95%CI: 2.9-9.8) and BTS group 4.0% (95%CI: 1.9-8.2). The results of the subgroup analysis for all complication rates based on indication, center, design, and region are shown in detail in Table 3.

Re-intervention

Twenty studies reported unplanned surgical or endoscopic re-interventions after SEMS placement and were included in the analysis. The overall re-intervention rate was 13.6% (95%CI: 10.1-18.0) (Figure 5A). The I^2 value for heterogeneity analysis was 69. Once again, re-interventions were required more frequently in palliation group (16.7%, 95%CI: 11.8-22.9) compared to BTS group (3.3%, 95%CI: 1.2-8.4).

Unplanned emergency surgeries were needed in

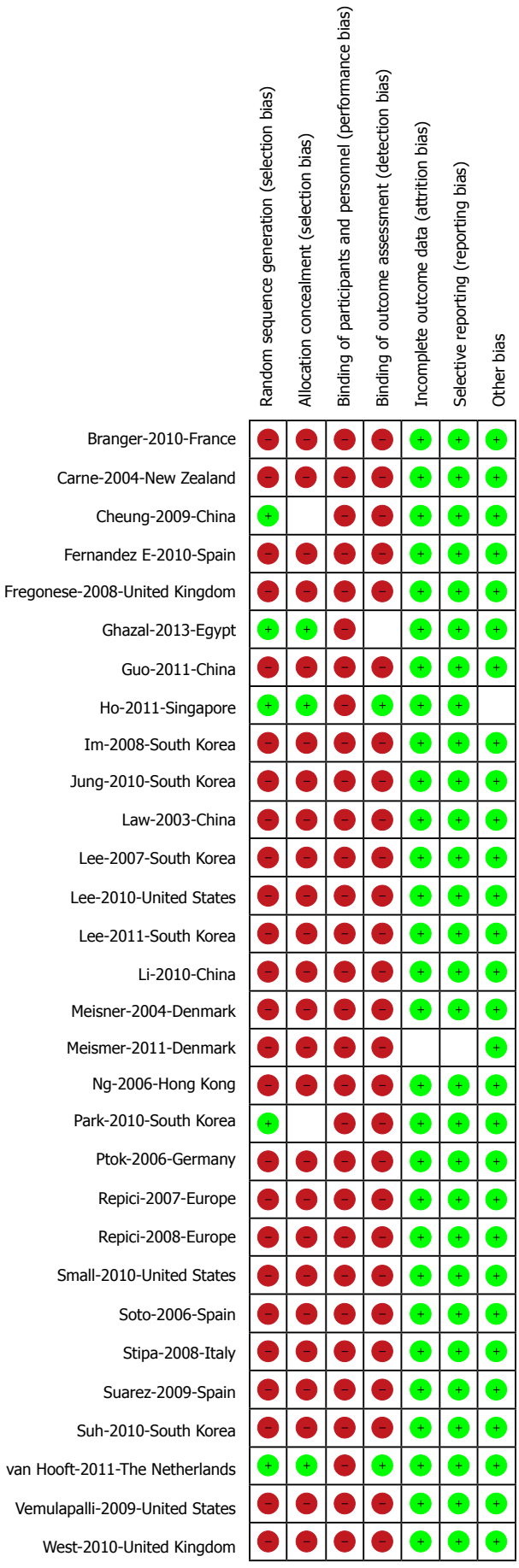


Figure 2 Summary of risk and bias along with overall quality assessments of included studies.

5.8% (95%CI: 3.9-8.5) of patients (Figure 5B). Rescue surgeries were more frequently needed in palliation group (7.6%, 95%CI: 4.1-13.8) compared to BTS group (3.2%, 1.1-8.7).

Repeat endoscopy with re-stent placement was needed in 8.0% (95%CI: 6.4-9.9) of cases (Figure 5C). Repeat endoscopy with interventions other than stent placement was observed in 3.8% (95%CI: 2.7-5.3) of cases (Figure 5D). These endoscopic intervention included endoscopy to control bleeding or to dis-impact fecal material blocking the stent. The results of the subgroup analysis for all re-interventions based on indication, center, design, and region are shown in detail in Table 4.

Heterogeneity and publication bias

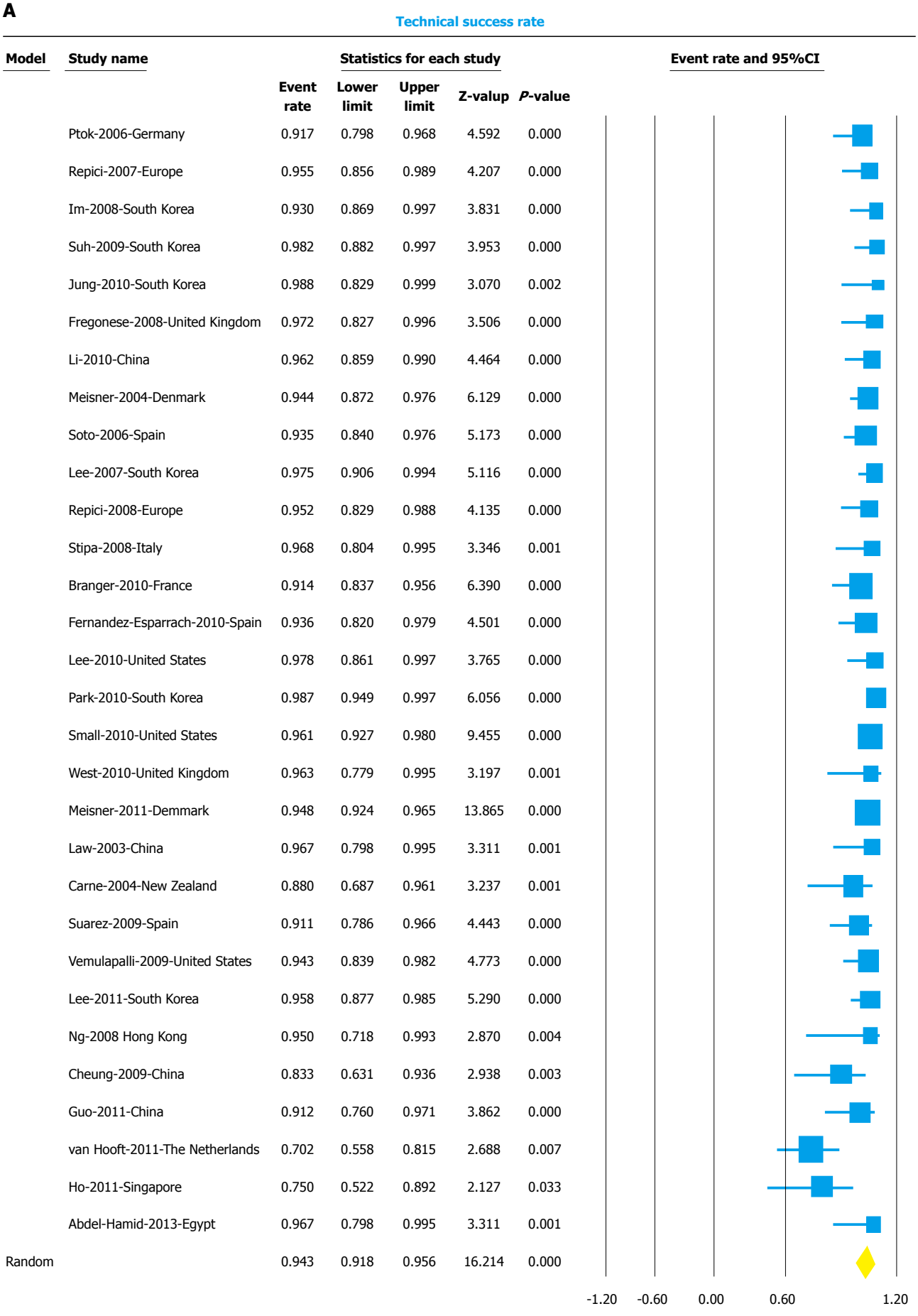
Heterogeneity was present in the analysis and was further explored by performing multiple subgroup analysis as shown in Tables 2, 3 and 4. For each analysis, sensitivity analysis by omitting one study at a time to evaluate the effect of single study on overall analysis was used to further explore heterogeneity. Sensitivity analysis did not reveal any particular study responsible for heterogeneity.

Funnel plots for the publication bias with respect to both technical and clinical success are shown in Figure 6, respectively. Rejection of N test indicated that for the combined two-tailed *P* value were non-significant ($P > 0.05$); and with no significant findings, it would take an additional 4517 studies. Using the random effects model the overall technical success rate was 94% (95%CI: 0.92-0.96). After publication bias adjustment using "Trim and Fill", the computed overall technical success rate was 93% (95%CI: 0.90-0.95).

DISCUSSION

Our meta-analysis of 30 included studies showed overall clinical and technical success rates of 90.6% and 94%, respectively. Further subgroup analysis showed different technical and clinical success rates respectively for studies that originated from Europe (92.8% and 89.1%), Asia (95.9% and 92.1%), and North America (96% and 92.5%). This could be related to the differences in patient population, extent and severity of colorectal obstruction, procedural volume, and training and experience of endoscopist. However, due to inconsistent reporting and lack of availability of these data, meta-regression analysis was not possible. The common reasons for technical failure were the following: Inability to pass guide wire through region of obstruction, occurrence of iatrogenic perforation during insertion, presence of metachronous bowel obstruction, miscalculation of length of stricture and presence of tortuousities or kinks in colonic lumen making insertions difficult. Based on a retrospective review of 412 patients where SEMSs were used for malignant colorectal obstruction, Yoon *et al*[48] reported that presence of carcinomatosis, extracolonic origin of tumor,and proximal site of obstruction were the factors related to

A


Technical success rate
Model
Study name
Statistics for each study
Event rate and 95%CI
Event rate
Lower limit
Upper limit
Z-valup
P-value
Ptok-2006-Germany 0.917 0.798 0.968 4.592 0.000
Repici-2007-Europe 0.955 0.856 0.989 4.207 0.000
Im-2008-South Korea 0.930 0.869 0.997 3.831 0.000
Suh-2009-South Korea 0.982 0.882 0.997 3.953 0.000
Jung-2010-South Korea 0.988 0.829 0.999 3.070 0.002
Fregonese-2008-United Kingdom 0.972 0.827 0.996 3.506 0.000
Li-2010-China 0.962 0.859 0.990 4.464 0.000
Meisner-2004-Denmark 0.944 0.872 0.976 6.129 0.000
Soto-2006-Spain 0.935 0.840 0.976 5.173 0.000
Lee-2007-South Korea 0.975 0.906 0.994 5.116 0.000
Repici-2008-Europe 0.952 0.829 0.988 4.135 0.000
Stipa-2008-Italy 0.968 0.804 0.995 3.346 0.001
Branger-2010-France 0.914 0.837 0.956 6.390 0.000
Fernandez-Esparrach-2010-Spain 0.936 0.820 0.979 4.501 0.000
Lee-2010-United States 0.978 0.861 0.997 3.765 0.000
Park-2010-South Korea 0.987 0.949 0.997 6.056 0.000
Small-2010-United States 0.961 0.927 0.980 9.455 0.000
West-2010-United Kingdom 0.963 0.779 0.995 3.197 0.001
Meisner-2011-Demmark 0.948 0.924 0.965 13.865 0.000
Law-2003-China 0.967 0.798 0.995 3.311 0.001
Carne-2004-New Zealand 0.880 0.687 0.961 3.237 0.001
Suarez-2009-Spain 0.911 0.786 0.966 4.443 0.000
Vemulapalli-2009-United States 0.943 0.839 0.982 4.773 0.000
Lee-2011-South Korea 0.958 0.877 0.985 5.290 0.000
Ng-2008 Hong Kong 0.950 0.718 0.993 2.870 0.004
Cheung-2009-China 0.833 0.631 0.936 2.938 0.003
Guo-2011-China 0.912 0.760 0.971 3.862 0.000
van Hooft-2011-The Netherlands 0.702 0.558 0.815 2.688 0.007
Ho-2011-Singapore 0.750 0.522 0.892 2.127 0.033
Abdel-Hamid-2013-Egypt 0.967 0.798 0.995 3.311 0.001
Random 0.943 0.918 0.956 16.214 0.000
-1.20
-0.60
0.00
0.60
1.20

B

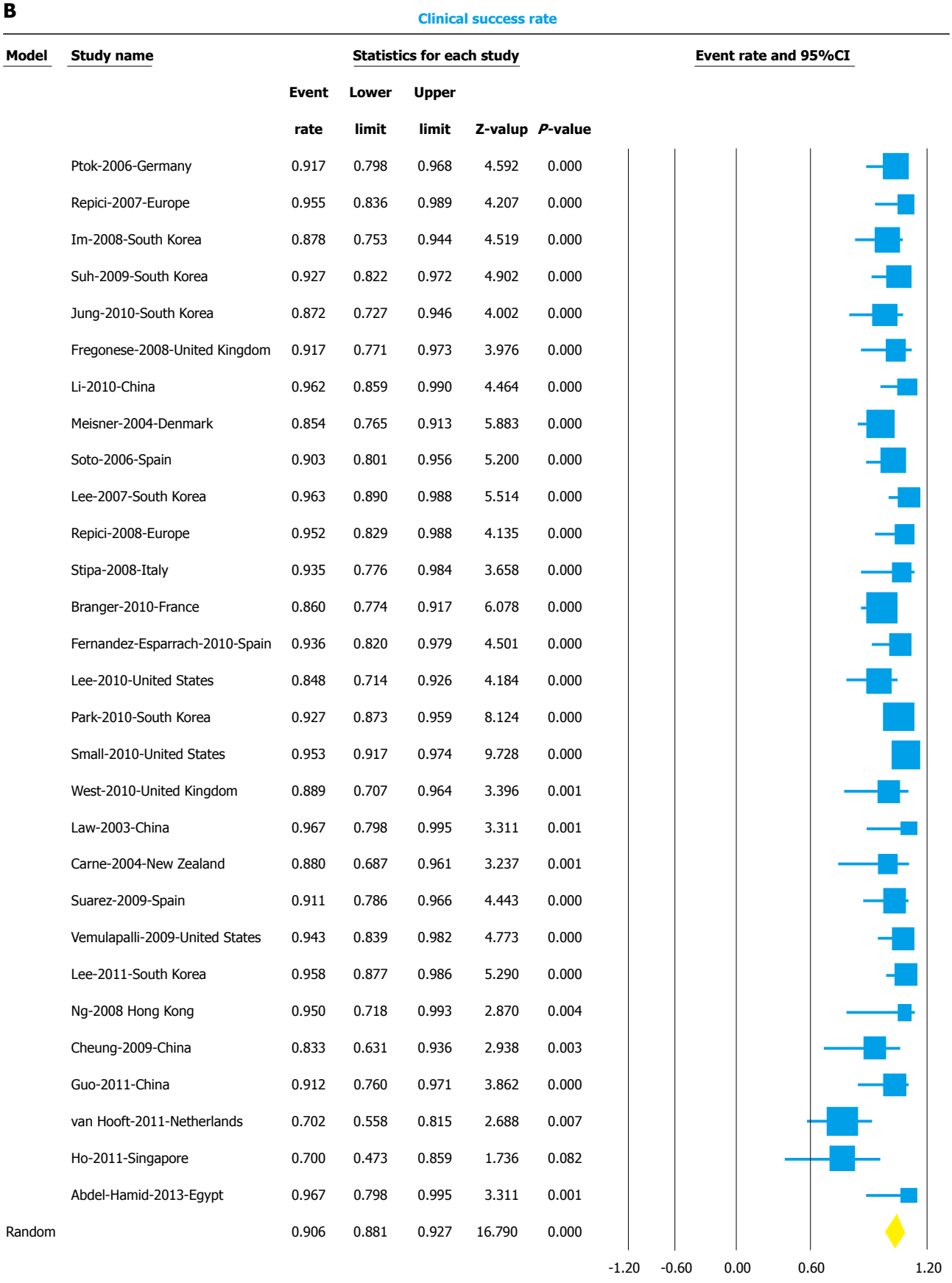


Figure 3 Forest plot for technical success (A) and clinical success (B) after endoscopic placement of self-expandable metal stent for malignant colorectal obstruction. The size of the each square is proportional to the sample size for each study, and the horizontal lines through the squares indicated the 95%CI for that study. For the pooled analysis, the diamond indicated the pooled value and the right and left ends of the vertical dashed bar indicated the 95%CI for the analysis.

Table 2 Efficacy of self-expandable metal stent placement

Category and subgroups		Studies	Event rate (%)	95%CI	*P* value	I^2 value
Technical success		30	94.0	91.8-95.6	0.000	58.07
Indication	Palliation	10	94.2	91.3-96.1	0.580	0.00
	BTS	8	89.4	79.5-94.8	0.003	67.14
Center	Single Center	24	94.1	92.1-95.6	0.06	32.74
	Multicenter	6	93.3	83.1-97.5	0.000	84.70
Design	Prospective	11	93.1	86.9-96.5	0.000	80.94
	Retrospective	18	94.4	92.7-95.7	0.846	0.00
Region	Asia	12	95.2	91.0-97.5	0.005	58.82
	Europe	13	92.8	89-95.4	0.001	65.40
	North America	3	96.0	93.2-97.6	0.679	0.00
Clinical success		29	90.6	88.1-92.7	0.001	50.58
Indication	Palliation	10	91.7	88.7-94	0.719	0.00
	BTS	8	87.9	78.1-93.7	0.004	66.35
Center	Single Center	23	90.8	88.4-92.7	0.042	36.52
	Multicenter	5	89.7	76.7-95.9	0.004	73.93
Design	Prospective	10	89.7	82.6-94.0	0.001	72.48
	Retrospective	18	91.0	88.7-92.3	0.238	18.04
Region	Asia	12	91.2	87.2-94.1	0.044	45.37
	Europe	12	89.1	84.5-92.3	0.029	48.73
	North America	3	92.5	83.9-96.7	0.040	69.03

BTS: Bridge to surgery.

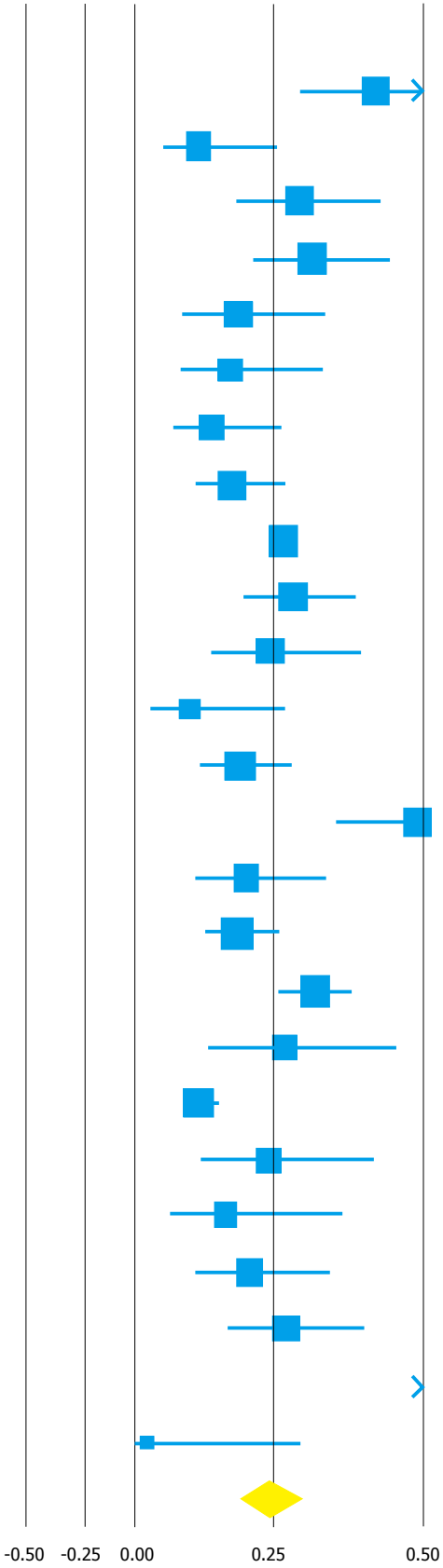
Table 3 Safety of self-expandable metal stent placement

Category and subgroups		Studies	Event rate (%)	95%CI	*P* value	I^2 value
Overall complication rate		25	23.1	18.5-28.6	0.000	81.86
Indication	Palliation	10	27.3	18.7-38	0.000	80.60
	BTS	3	13.8	8.3-22.2	0.371	0.00
Center	Single Center	20	25.2	20.1-31.1	0.000	78.20
	Multicenter	5	16.2	10.8-23.6	0.058	56.11
Design	Prospective	8	20.6	13.6-30.0	0.000	82.64
	Retrospective	16	25.0	19.1-31.8	0.000	79.08
Center	Asia	9	25.1	16.1-36.9	0.000	84.93
	Europe	12	21.3	15.1-29.1	0.000	81.28
	North America	3	28.0	21.9-35	1.000	0.00
Obstruction rate		25	8.3	6.0-11.4	0.000	68.79
Indication	Palliation	10	9.5	5.4-16.4	0.001	69.55
	BTS	3	1.9	0.5-7.1	0.959	0.00
Center	Single Center	20	8.7	6.2-12.1	0.000	63.23
	Multicenter	5	3.1	1.8-5.4	0.000	80.21
Design	Prospective	8	8.2	4.9-13.5	0.001	70.11
	Retrospective	16	8.7	5.7-13.0	0.000	67.79
Region	Asia	9	9.8	5.3-17.2	0.000	74.58
	Europe	12	7.7	4.8-12.1	0.001	63.77
	North America	3	7.5	5.1-10.9	0.961	0.00
Migration rate		29	7.6	5.7-10.0	0.000	55.48
Indication	Palliation	10	10.2	7.1-14.5	0.150	32.56
	BTS	7	4.1	2.0-8.1	0.687	0.00
Center	Single Center	23	9.7	7.7-12.1	0.061	33.51
	Multicenter	6	2.1	1.2-3.6	0.977	0.00
Design	Prospective	10	5.5	2.6-11.3	0.000	79.25
	Retrospective	18	9.0	7.3-11.1	0.389	5.55
Center	Asia	12	10.2	7.9-13.0	0.434	0.95
	Europe	13	5.4	2.9-10.0	0.000	75.80
	North America	3	7.8	5.4-11.3	0.972	0.00
Perforation rate		29	4.9	3.6-6.6	0.037	34.39
Indication	Palliation	10	5.4	2.9-9.8	0.054	46.00
	BTS	7	4.0	1.9-8.2	0.646	0.00
Center	Single Center	23	5.1	3.6-7.2	0.042	36.56
	Multicenter	6	4.0	2.6-4.7	0.586	0.00
Design	Prospective	10	3.1	2.1-4.7	0.722	0.00
	Retrospective	18	6.2	4.4-8.6	0.117	29.48
Center	Asia	12	3.6	1.8-7.0	0.024	50.02
	Europe	13	4.4	3.2-6.0	0.472	0.00
	North America	3	8.4	5.8-11.9	0.463	0.00

BTS: Bridge to surgery.

A **Overall complication rate**

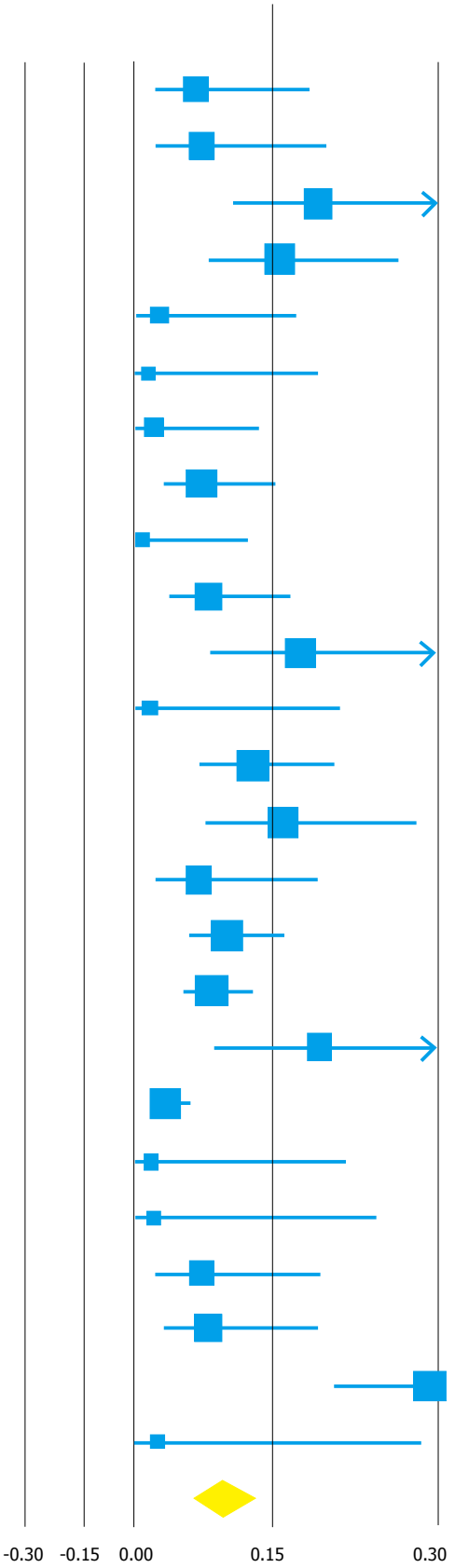
Model	Study name	Event rate	Lower limit	Upper limit	Z-valup	*P*-value
	Ptok-2006-Germany	0.417	0.287	0.559	-1.149	0.250
	Repici-2007-Europe	0.114	0.048	0.245	-4.324	0.000
	Im-2008-South Korea	0.286	0.177	0.426	-2.898	0.004
	Suh-2009-South Korea	0.309	0.202	0.442	-2.757	0.006
	Jung-2010-South Korea	0.179	0.088	0.331	-3.642	0.000
	Fregonese-2008-United Kingdom	0.167	0.077	0.325	-3.599	0.000
	Li-2010-China	0.135	0.066	0.256	-4.580	0.000
	Meisner-2004-Denmark	0.169	0.104	0.261	-5.636	0.000
	Soto-2006-Spain	0.258	0.165	0.381	-3.639	0.000
	Lee-2007-South Korea	0.275	0.188	0.383	-3.872	0.000
	Repici-2008-Europe	0.238	0.133	0.389	-3.211	0.001
	Stipa-2008-Italy	0.097	0.032	0.261	-3.677	0.000
	Branger-2010-France	0.183	0.117	0.275	-5.582	0.000
	Fernandez-Esparrach-2010-Spain	0.489	0.351	0.629	-0.146	0.884
	Lee-2010-United States	0.196	0.105	0.335	-3.804	0.000
	Park-2010-South Korea	0.179	0.126	0.248	-7.178	0.000
	Small-2010-United States	0.313	0.257	0.376	-5.556	0.000
	West-2010-United Kingdom	0.259	0.129	0.453	-2.391	0.017
	Meisner-2011-Demmark	0.113	0.085	0.148	-12.755	0.000
	Law-2003-China	0.233	0.116	0.415	-2.756	0.006
	Carne-2004-New Zealand	0.160	0.061	0.357	-3.040	0.002
	Suarez-2009-Spain	0.200	0.107	0.342	-3.720	0.000
	Vemulapalli-2009-United States	0.264	0.163	0.398	-3.288	0.001
	Lee-2011-South Korea	0.620	0.502	0.725	1.998	0.046
	Ho-2011-Singapore	0.024	0.001	0.287	-2.594	0.009
Random		0.231	0.185	0.286	-8.257	0.000



B

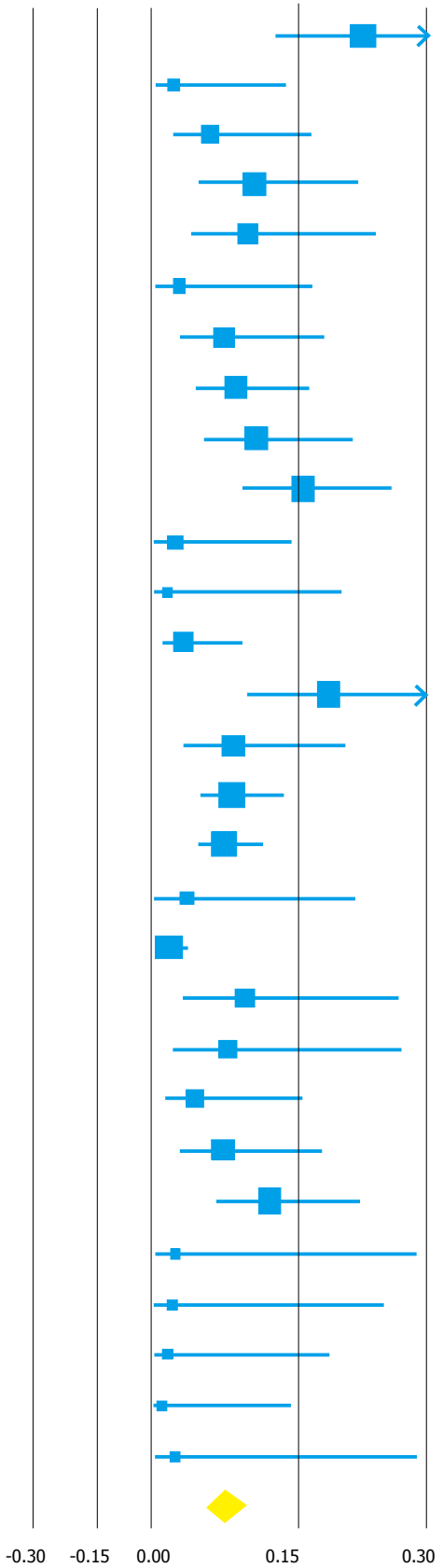
Stent obstruction rate

Model	Study name	Event rate	Lower limit	Upper limit	Z-valup	*P*-value
	Ptok-2006-Germany	0.063	0.020	0.177	-4.452	0.000
	Repici-2007-Europe	0.068	0.022	0.191	-4.372	0.000
	Im-2008-South Korea	0.184	0.098	0.317	-4.043	0.000
	Suh-2009-South Korea	0.145	0.074	0.265	-4.630	0.000
	Jung-2010-South Korea	0.026	0.004	0.161	-3.591	0.000
	Fregonese-2008-United Kingdom	0.014	0.001	0.182	-3.013	0.003
	Li-2010-China	0.019	0.003	0.124	-3.894	0.000
	Meisner-2004-Denmark	0.067	0.031	0.142	-6.214	0.000
	Soto-2006-Spain	0.008	0.000	0.115	-3.401	0.001
	Lee-2007-South Korea	0.075	0.034	0.157	-5.919	0.000
	Repici-2008-Europe	0.167	0.082	0.310	-3.887	0.000
	Stipa-2008-Italy	0.016	0.001	0.206	-2.907	0.004
	Branger-2010-France	0.118	0.067	0.201	-6.256	0.000
	Fernandez-Esparrach-2010-Spain	0.149	0.073	0.281	-4.254	0.000
	Lee-2010-United States	0.065	0.021	0.184	-4.459	0.000
	Park-2010-South Korea	0.093	0.056	0.150	-8.129	0.000
	Small-2010-United States	0.077	0.049	0.119	-10.108	0.000
	West-2010-United Kingdom	0.185	0.079	0.375	-2.991	0.003
	Meisner-2011-Demmark	0.031	0.018	0.054	-11.689	0.000
	Law-2003-China	0.016	0.001	0.211	-2.883	0.004
	Carne-2004-New Zealand	0.019	0.001	0.244	-2.753	0.006
	Suarez-2009-Spain	0.067	0.022	0.187	-4.416	0.000
	Vemulapalli-2009-United States	0.075	0.029	0.184	-4.818	0.000
	Lee-2011-South Korea	0.296	0.201	0.411	-3.336	0.001
	Ho-2011-Singapore	0.024	0.001	0.287	-2.594	0.009
Random		0.083	0.060	0.114	-13.311	0.000



C **Stent migration rate**

Model	Study name	Event rate	Lower limit	Upper limit	Z-valup	*P*-value
	Ptok-2006-Germany	0.229	0.132	0.368	-3.532	0.000
	Repici-2007-Europe	0.023	0.003	0.144	-3.718	0.000
	Im-2008-South Korea	0.061	0.020	0.173	-4.582	0.000
	Suh-2009-South Korea	0.109	0.050	0.222	-4.855	0.000
	Jung-2010-South Korea	0.103	0.039	0.243	-4.110	0.000
	Fregonese-2008-United Kingdom	0.028	0.004	0.173	-3.506	0.000
	Li-2010-China	0.077	0.029	0.188	-4.775	0.000
	Meisner-2004-Denmark	0.090	0.046	0.170	-6.247	0.000
	Soto-2006-Spain	0.113	0.055	0.218	-5.137	0.000
	Lee-2007-South Korea	0.163	0.097	0.260	-5.411	0.000
	Repici-2008-Europe	0.024	0.003	0.151	-3.669	0.000
	Stipa-2008-Italy	0.016	0.001	0.206	-2.907	0.004
	Branger-2010-France	0.032	0.010	0.095	-5.795	0.000
	Fernandez-Esparrach-2010-Spain	0.191	0.103	0.329	-3.885	0.000
	Lee-2010-United States	0.087	0.033	0.210	-4.494	0.000
	Park-2010-South Korea	0.086	0.051	0.143	-8.143	0.000
	Small-2010-United States	0.077	0.049	0.119	-10.108	0.000
	West-2010-United Kingdom	0.037	0.005	0.221	-3.197	0.001
	Meisner-2011-Demmark	0.018	0.009	0.038	-10.436	0.000
	Law-2003-China	0.100	0.033	0.268	-3.610	0.000
	Carne-2004-New Zealand	0.080	0.020	0.269	-3.313	0.001
	Suarez-2009-Spain	0.044	0.011	0.161	-4.241	0.000
	Vemulapalli-2009-United States	0.075	0.029	0.184	-4.818	0.000
	Lee-2011-South Korea	0.127	0.067	0.226	-5.410	0.000
	Ng-2008 Hong Kong	0.024	0.001	0.287	-2.594	0.009
	Cheung-2009-China	0.020	0.001	0.251	-2.724	0.006
	Guo-2011-China	0.014	0.001	0.191	-2.973	0.003
	van Hooft-2011-The Netherlands	0.010	0.001	0.146	-3.203	0.001
	Ho-2011-Singapore	0.024	0.001	0.287	-2.594	0.009
Random		0.076	0.057	0.100	-16.360	0.000



D Perforation rate

Model	Study name	Event rate	Lower limit	Upper limit	Z-valup	*P*-value
	Ptok-2006-Germany	0.010	0.001	0.143	-3.218	0.001
	Repici-2007-Europe	0.011	0.001	0.154	-3.156	0.002
	Im-2008-South Korea	0.041	0.010	0.149	-4.373	0.000
	Suh-2009-South Korea	0.018	0.003	0.118	-3.953	0.000
	Jung-2010-South Korea	0.051	0.013	0.183	-4.019	0.000
	Fregonese-2008-United Kingdom	0.083	0.027	0.229	-3.976	0.000
	Li-2010-China	0.009	0.001	0.134	-3.275	0.001
	Meisner-2004-Denmark	0.011	0.002	0.075	-4.452	0.000
	Soto-2006-Spain	0.048	0.016	0.140	-5.033	0.000
	Lee-2007-South Korea	0.025	0.006	0.094	-5.116	0.000
	Repici-2008-Europe	0.024	0.003	0.151	-3.669	0.000
	Stipa-2008-Italy	0.065	0.016	0.224	-3.658	0.000
	Branger-2010-France	0.032	0.010	0.095	-5.795	0.000
	Fernandez-Esparrach-2010-Spain	0.106	0.045	0.231	-4.499	0.000
	Lee-2010-United States	0.043	0.011	0.158	-4.275	0.000
	Park-2010-South Korea	0.003	0.000	0.050	-4.034	0.000
	Small-2010-United States	0.082	0.053	0.124	-10.116	0.000
	West-2010-United Kingdom	0.037	0.005	0.221	-3.197	0.001
	Meisner-2011-Demmark	0.039	0.024	0.064	-12.138	0.000
	Law-2003-China	0.033	0.005	0.202	-3.311	0.001
	Carne-2004-New Zealand	0.019	0.001	0.244	-2.753	0.006
	Suarez-2009-Spain	0.044	0.011	0.161	-4.241	0.000
	Vemulapalli-2009-United States	0.113	0.052	0.230	-4.748	0.000
	Lee-2011-South Korea	0.155	0.088	0.259	-5.172	0.000
	Ng-2008 Hong Kong	0.050	0.007	0.282	-2.870	0.004
	Cheung-2009-China	0.020	0.001	0.251	-2.724	0.006
	Guo-2011-China	0.029	0.004	0.181	-3.445	0.001
	van Hooft-2011-The Netherlands	0.010	0.001	0.146	-3.203	0.001
	Ho-2011-Singapore	0.024	0.001	0.287	-2.594	0.009
Random		0.049	0.036	0.066	-18.489	0.000

Event rate and 95%CI

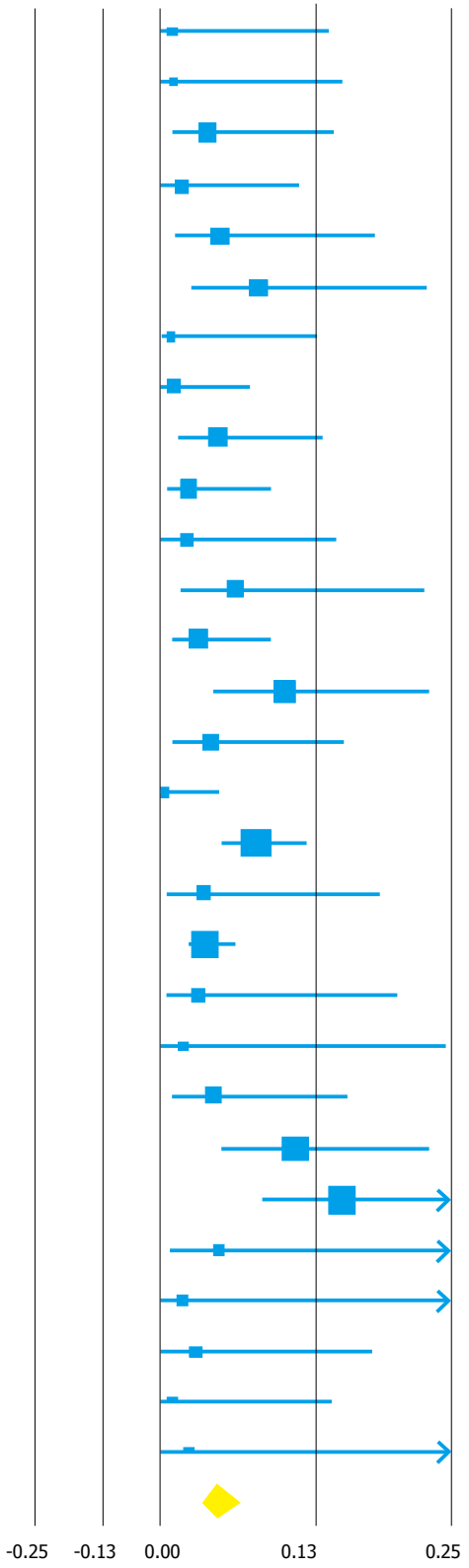
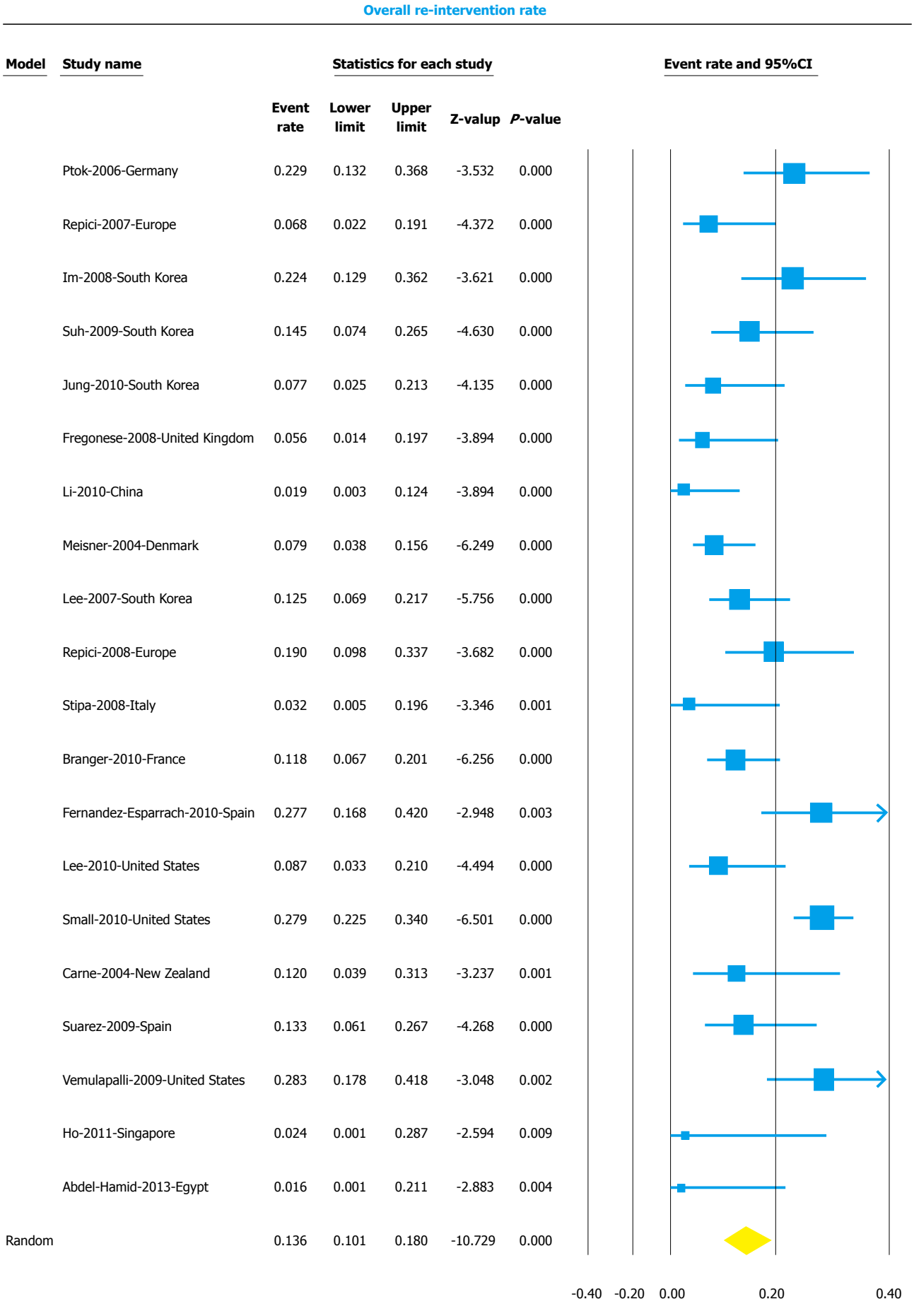
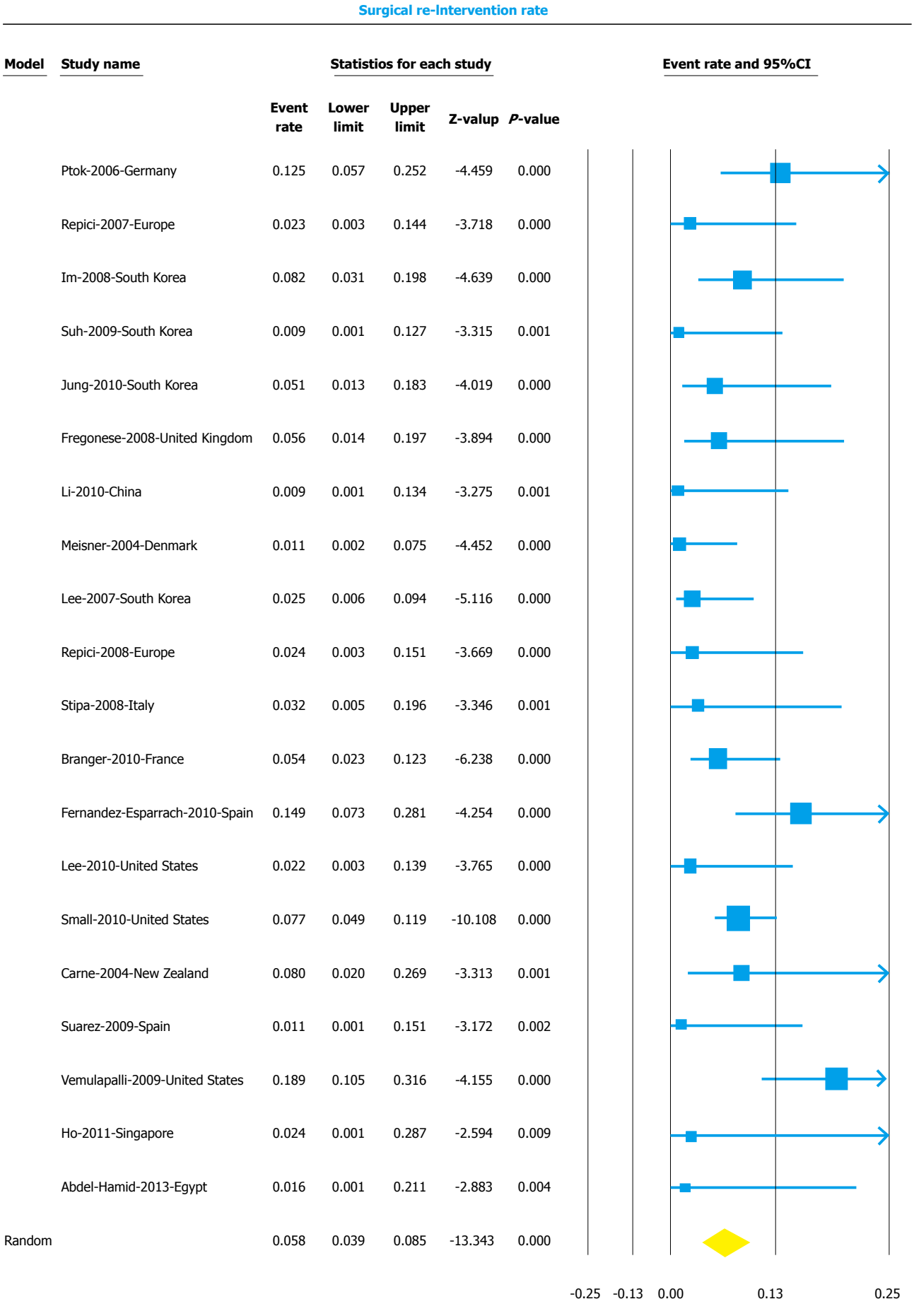


Figure 4 Forest plot for overall complication rate (A), stent occlusion rate (B), stent migration rate (C), and perforation rate (D) after endoscopic placement of self-expandable metal stent for malignant colorectal obstruction. The size of the each square is proportional to the sample size for each study, and the horizontal lines through the squares indicated the 95%CI for that study. For the pooled analysis, the diamond indicated the pooled value and the right and left ends of the vertical dashed bar indicated the 95%CI for the analysis.

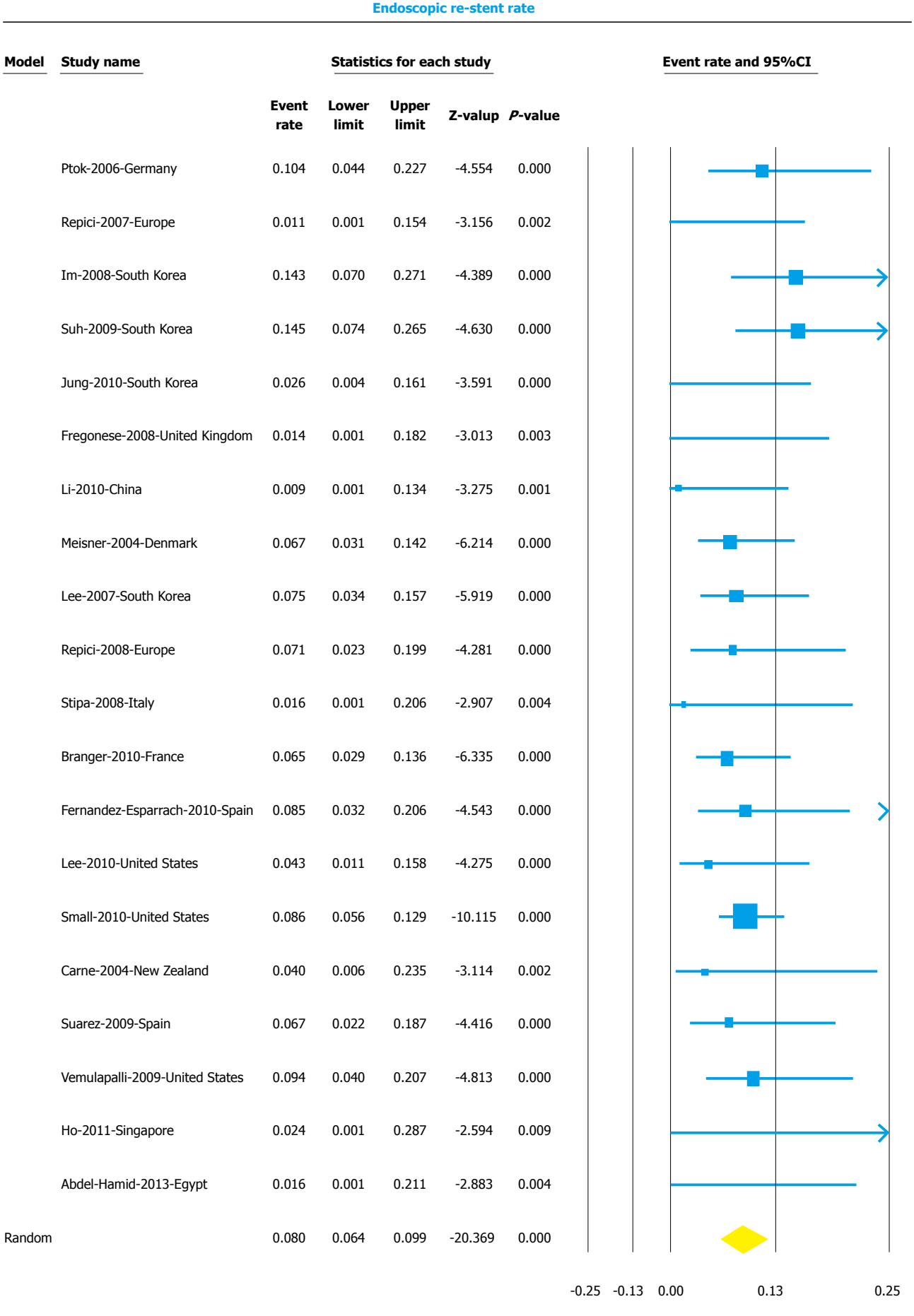
A


Overall re-intervention rate
Model
Study name
Statistics for each study
Event rate
Lower limit
Upper limit
Z-valup
P-value
Event rate and 95%CI
Ptok-2006-Germany 0.229 0.132 0.368 -3.532 0.000
Repici-2007-Europe 0.068 0.022 0.191 -4.372 0.000
Im-2008-South Korea 0.224 0.129 0.362 -3.621 0.000
Suh-2009-South Korea 0.145 0.074 0.265 -4.630 0.000
Jung-2010-South Korea 0.077 0.025 0.213 -4.135 0.000
Fregonese-2008-United Kingdom 0.056 0.014 0.197 -3.894 0.000
Li-2010-China 0.019 0.003 0.124 -3.894 0.000
Meisner-2004-Denmark 0.079 0.038 0.156 -6.249 0.000
Lee-2007-South Korea 0.125 0.069 0.217 -5.756 0.000
Repici-2008-Europe 0.190 0.098 0.337 -3.682 0.000
Stipa-2008-Italy 0.032 0.005 0.196 -3.346 0.001
Branger-2010-France 0.118 0.067 0.201 -6.256 0.000
Fernandez-Esparrach-2010-Spain 0.277 0.168 0.420 -2.948 0.003
Lee-2010-United States 0.087 0.033 0.210 -4.494 0.000
Small-2010-United States 0.279 0.225 0.340 -6.501 0.000
Carne-2004-New Zealand 0.120 0.039 0.313 -3.237 0.001
Suarez-2009-Spain 0.133 0.061 0.267 -4.268 0.000
Vemulapalli-2009-United States 0.283 0.178 0.418 -3.048 0.002
Ho-2011-Singapore 0.024 0.001 0.287 -2.594 0.009
Abdel-Hamid-2013-Egypt 0.016 0.001 0.211 -2.883 0.004
Random 0.136 0.101 0.180 -10.729 0.000
-0.40
-0.20
0.00
0.20
0.40

B


Surgical re-Intervention rate
Model
Study name
Statistios for each study
Event rate and 95%CI
Event rate
Lower limit
Upper limit
Z-valup
P-value
Ptok-2006-Germany 0.125 0.057 0.252 -4.459 0.000
Repici-2007-Europe 0.023 0.003 0.144 -3.718 0.000
Im-2008-South Korea 0.082 0.031 0.198 -4.639 0.000
Suh-2009-South Korea 0.009 0.001 0.127 -3.315 0.001
Jung-2010-South Korea 0.051 0.013 0.183 -4.019 0.000
Fregonese-2008-United Kingdom 0.056 0.014 0.197 -3.894 0.000
Li-2010-China 0.009 0.001 0.134 -3.275 0.001
Meisner-2004-Denmark 0.011 0.002 0.075 -4.452 0.000
Lee-2007-South Korea 0.025 0.006 0.094 -5.116 0.000
Repici-2008-Europe 0.024 0.003 0.151 -3.669 0.000
Stipa-2008-Italy 0.032 0.005 0.196 -3.346 0.001
Branger-2010-France 0.054 0.023 0.123 -6.238 0.000
Fernandez-Esparrach-2010-Spain 0.149 0.073 0.281 -4.254 0.000
Lee-2010-United States 0.022 0.003 0.139 -3.765 0.000
Small-2010-United States 0.077 0.049 0.119 -10.108 0.000
Carne-2004-New Zealand 0.080 0.020 0.269 -3.313 0.001
Suarez-2009-Spain 0.011 0.001 0.151 -3.172 0.002
Vemulapalli-2009-United States 0.189 0.105 0.316 -4.155 0.000
Ho-2011-Singapore 0.024 0.001 0.287 -2.594 0.009
Abdel-Hamid-2013-Egypt 0.016 0.001 0.211 -2.883 0.004
Random 0.058 0.039 0.085 -13.343 0.000
-0.25 -0.13 0.00 0.13 0.25

C


Endoscopic re-stent rate
Model
Study name
Statistics for each study
Event rate and 95%CI
Event rate
Lower limit
Upper limit
Z-valup
P-value
Ptok-2006-Germany 0.104 0.044 0.227 -4.554 0.000
Repici-2007-Europe 0.011 0.001 0.154 -3.156 0.002
Im-2008-South Korea 0.143 0.070 0.271 -4.389 0.000
Suh-2009-South Korea 0.145 0.074 0.265 -4.630 0.000
Jung-2010-South Korea 0.026 0.004 0.161 -3.591 0.000
Fregonese-2008-United Kingdom 0.014 0.001 0.182 -3.013 0.003
Li-2010-China 0.009 0.001 0.134 -3.275 0.001
Meisner-2004-Denmark 0.067 0.031 0.142 -6.214 0.000
Lee-2007-South Korea 0.075 0.034 0.157 -5.919 0.000
Repici-2008-Europe 0.071 0.023 0.199 -4.281 0.000
Stipa-2008-Italy 0.016 0.001 0.206 -2.907 0.004
Branger-2010-France 0.065 0.029 0.136 -6.335 0.000
Fernandez-Esparrach-2010-Spain 0.085 0.032 0.206 -4.543 0.000
Lee-2010-United States 0.043 0.011 0.158 -4.275 0.000
Small-2010-United States 0.086 0.056 0.129 -10.115 0.000
Carne-2004-New Zealand 0.040 0.006 0.235 -3.114 0.002
Suarez-2009-Spain 0.067 0.022 0.187 -4.416 0.000
Vemulapalli-2009-United States 0.094 0.040 0.207 -4.813 0.000
Ho-2011-Singapore 0.024 0.001 0.287 -2.594 0.009
Abdel-Hamid-2013-Egypt 0.016 0.001 0.211 -2.883 0.004
Random 0.080 0.064 0.099 -20.369 0.000
-0.25 -0.13 0.00 0.13 0.25

D

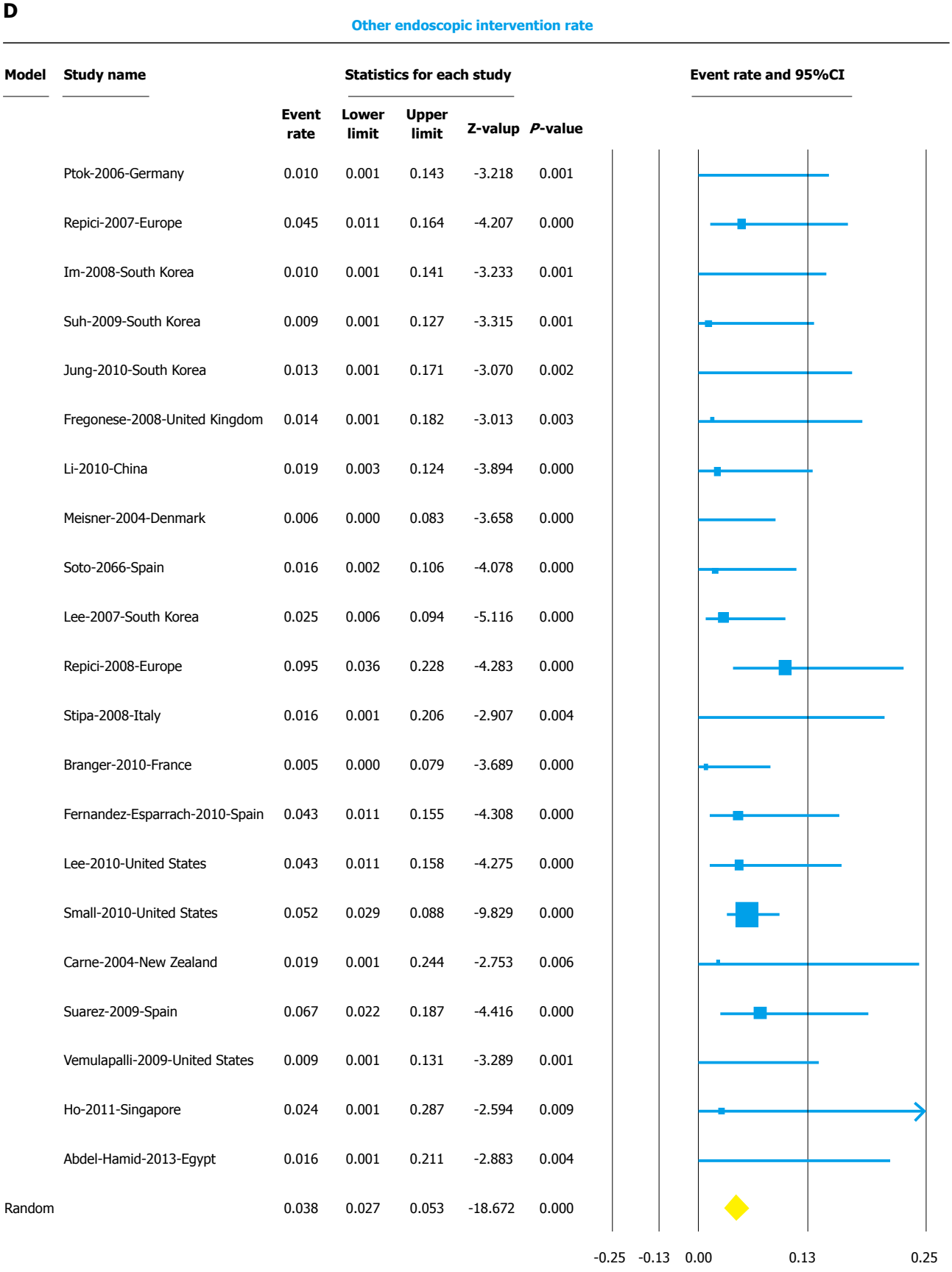


Figure 5 Forest plot for overall re-intervention rate (A), surgical intervention rate (B), endoscopic re-stent rate (C), and other endoscopic intervention rate (D) after endoscopic placement of self-expandable metal stent for malignant colorectal obstruction. The size of the each square is proportional to the sample size for each study, and the horizontal lines through the squares indicated the 95%CI for that study. For the pooled analysis, the diamond indicated the pooled value and the right and left ends of the vertical dashed bar indicated the 95%CI for the analysis.

Table 4 Re-intervention rates

Category and subgroups		Studies	Event rate (%)	95%CI	*P* value	I^2 value
Overall re-intervention rate		20	13.6	10.1-18.0	0.000	68.79
Indication	Palliation	8	16.7	11.8-22.9	0.077	45.30
	BTS	4	3.3	1.2-8.4	0.758	0.00
Center	Single Center	17	14.2	10.3-19.2	0.000	70.10
	Multicenter	3	10.3	4.3-22.6	0.114	53.94
Design	Prospective	7	14.8	9.5-22.4	0.069	48.81
	Retrospective	12	13.6	9.1-19.9	0.000	75.17
Region	Asia	6	11.6	6.8-19.3	0.066	51.74
	Europe	9	13.2	8.8-19.4	0.011	59.63
	North America	3	14.7	11.3-19.0	0.034	70.54
Surgical intervention rate		20	5.8	3.9-8.5	0.011	46.79
Indication	Palliation	8	7.6	4.1-13.8	0.048	50.62
	BTS	4	3.2	1.1-8.7	0.641	0.00
Center	Single Center	17	6.1	4.0-9.2	0.009	50.50
	Multicenter	3	3.6	1.3-9.2	0.675	0.00
Design	Prospective	7	5.1	2.6-9.8	0.193	30.76
	Retrospective	12	5.9	3.5-9.9	0.007	57.59
Region	Asia	6	4.2	2.2-7.9	0.411	0.81
	Europe	9	5.4	2.9-9.9	0.051	48.24
	North America	3	9.2	3.7-21.2	0.016	75.87
Endoscopic re-stent rate		20	8.0	6.4-9.9	0.439	2.15
Indication	Palliation	8	9.7	6.6-14.2	0.304	16.02
	BTS	4	1.5	0.4-5.7	0.973	0.00
Center	Single Center	17	8.3	6.6-10.2	0.486	0.00
	Multicenter	3	3.9	1.1-12.3	0.296	17.95
Design	Prospective	7	8.4	5.3-13.1	0.320	14.45
	Retrospective	12	7.9	6.1-10.1	0.486	0.00
Region	Asia	6	8.5	4.5-15.3	0.113	43.89
	Europe	9	6.8	4.8-9.7	0.706	0.00
	North America	3	8.3	5.7-11.8	0.600	0.00
Other endoscopic intervention rate		21	3.8	2.7-5.3	0.620	0.00
Indication	Palliation	8	3.1	1.5-6.1	0.601	0.00
	BTS	4	1.8	0.5-6.0	0.993	0.00
Center	Single Center	18	3.4	2.3-4.8	0.737	0.00
	Multicenter	3	6.3	2.7-13.7	0.333	8.99
Design	Prospective	7	4.2	2.3-7.7	0.412	1.59
	Retrospective	13	3.7	2.5-5.5	0.553	0.00
Region	Asia	6	1.8	0.8-4.3	0.980	0.00
	Europe	10	3.7	2.0-6.6	0.280	17.68
	North America	3	4.7	2.9-7.8	0.475	0.00

BTS: Bridge to surgery.

technical failure. Similarly no additional chemotherapy, dilation prior to stent placement, and extracolonic origin of the tumor were related to long-term clinical failure in the palliation group[48].

The overall adverse event rate in the SEMS group was 23.1%, with stent obstruction, migration, and perforation being the most common adverse events. Tenesmus, intra-abdominal abscesses and rectovesical fistulas were also reported as rare adverse events of SEMS in included studies. The overall rate of adverse event was greater when SEMSs were used for palliation (27.3%) when compared to their use as a BTS (13.8%); especially the frequency of stent obstruction (9.5% *vs* 1.9%) and stent migration (10.2% *vs* 4.1%) were higher in the palliation group, while the rate of perforation was comparable between both the groups (5.4% *vs* 4.0%). This increased adverse event rate in the palliation group was merely a function of the fact that the stent stayed for longer duration in palliation cases, so is more likely to eventually obstruct or migrate. Male sex, complete obstruction, stricture dilation during SEMS insertion, experience of the operator and stent diameter $\leqslant$ 22 mm, were identified as risk factors for adverse event related to SEMS placements[5,9]. Small *et al*[11] reported a threefold increase in the rate of perforation in patients receiving Bevacizumab™ therapy after SEMS placement. Patients who had received Bevacizumab™ therapy prior to SEMS placement showed no increased susceptibility to perforation[11]. Stool impaction, tumor ingrowth or tumor outgrowth were major reasons for stent obstruction[24,39,44].

Re-intervention measures including unplanned surgery 6.0% (95%CI: 4.0-8.9), endoscopic re-stenting 8.2% (95%CI: 6.6-10.1) and other endoscopic interventions 3.9% (95%CI: 2.8-5.4) were performed in 14.3% (95%CI: 10.7-18.9) of the patients. Re-interventions after SEMS placements were needed more often in the palliative group 16.7% (95%CI: 11.8-22.9)

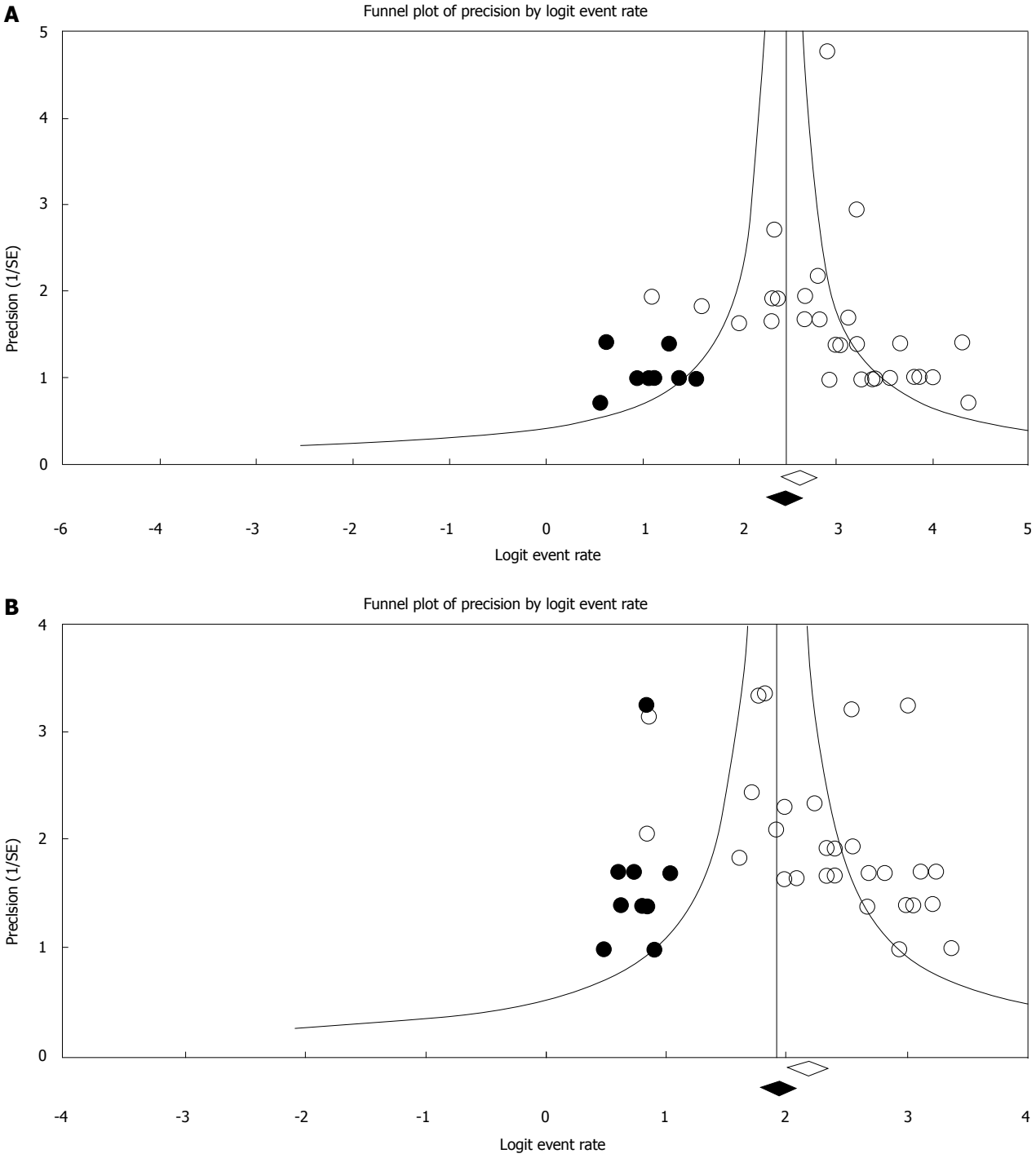


Figure 6 Funnel plots to depict publication bias with respect to technical success (A) and clinical success (B).

than in the BTS group 3.9% (95%CI: 1.3-11.4), which also might have attributed to higher incidence of long-term clinical failure in the palliation group. Adverse events such as SEMS migration and obstruction were chiefly managed endoscopically by procedures including reinsertion of stent, removal and replacement of stent, insertion of additional stents and trimming of the stent by argon plasma coagulation; whereas perforation was usually managed by surgical intervention.

Use of SEMS for decompression offers several advantages. In a palliative setting, it provides a better quality of life by avoiding a surgery with probable colostomy and significantly reduces hospital stay. In an emergency BTS setting, it allows time for bowel preparation, a full workup, and hyperalimentation of the patient prior to definitive surgery and thus increases likelihood of single stage colonic resection and anastomosis, with lower morbidity and mortality. To date, there are four RCTs[5,41,43,46] that evaluated endoscopic SEMS for BTS *vs* ES. There are no RCTs evaluating palliative SEMS *vs* palliative surgery. First RCT was a single center trial by Cheung *et al*[41] that evaluated SEMS for BTS *vs* ES

in 48 patients and reported that patients with SEMS as BTS followed by laparoscopic surgery had significantly less lower pain and cumulative blood loss, incidence of anastomotic leak, and infection in wound infection. Also, patients in SEMS group had significantly more successful one stage operation (16 *vs* 9, $P = 0.04$)[41]. The SEMS placements in the study were done by two dedicated endoscopists[41]. In contrast, more recently van Hooft *et al*[5] evaluated SEMS as BTS against ES in a multicenter (25 hospitals in Netherlands) RCT. Two successive interim analyses showed increased 30 d mortality in colonic stenting group and this study was stopped[12]. At final analysis of 98 patients (SEMS 47, ES 51), no difference was seen between two groups in 30 d mortality, morbidity, and stoma rate. This study had a relatively low technical success rate of only 70% and a high perforation rate of 20%, which are not the experience at most tertiary centers. This may be attributable to several factors. In order to increase recruitment, the study was not confined to tertiary care centers and endoscopists only had to have prior experience in deploying 10 colonic stents to participate in this study. Twenty-five centers performed 47 stent placements, with some presumably low volume centers performing only a single stent placement. The experience of the endoscopists at low volume centers may therefore have been a significant factor in modulating these results. Finally a higher than usual proportion (70%) of patients in this study had complete obstruction compared with most other series and this population can be expected to have a greater morbidity. In patients with complete obstruction, SEMS placement is difficult and it is a known risk factor for adverse events after SEMS placement[11]. Both additional RCTs[43,46] concluded that both SEMS and ES are feasible with trend towards lower post-operative morbidity and mortality with SEMS placement.

In conclusion, our meta-analysis showed that in CRC patients with malignant colonic obstruction, use of SEMS for palliation or as a BTS was quite efficacious with more than 90% of the cases achieving technical and clinical success. The overall adverse event rate was 23.5%, with the major adverse events being obstruction (8.3%), stent migration (7.6%), and perforation (4.9%). Also up to 15% of the patients required endoscopic and surgical re-interventions. Therefore, while choosing between ES and SEMS insertion it is advisable to individualize each decision after discussing the risks and benefits of both with the patients. Further prospective studies comparing the outcomes, in particular the adverse event rates between high and low volume centers for surgery and stent placements are desirable to provide further insights in the management of malignant colonic obstruction.

COMMENTS

Background

Malignant colonic obstruction is present in about 7% to 29% of the colorectal patients. Conventional treatment such as elective surgery is present, but has high chances of morbidity and mortality. Introduction of self-expandable metal stent (SEMS) is an initial therapy for malignant colonic obstruction provided an effective alternative, but long term outcomes are still warranted and under study.

Research frontiers

The meta-analysis evaluates the efficacy of SEMS over emergency surgery in colorectal patients having malignant colonic obstruction.

Innovations and breakthroughs

The key findings of the study indicate that use of SEMS significantly reduces the mortality and morbidity rates among colorectal cancer (CRC) patients. The authors found 91% of clinical success rate and 23% of complication rate when SEMS was used. Migration and perforation were among the most common complications observed when SEMS was used. Further interventions were required in about 14% of the patients.

Applications

SEMS is a safe device and its use can make it a more effective screening tool to prevent malignant colonic obstruction in CRC patients.

Terminology

A SEMS is a metallic tube used to hold and open a structure in the gastrointestinal tract. It helps in the passage of food, stool and other secretions. These stents are placed in gastrointestinal tract with the help of endoscopy either through mouth or through colon. Fluoroscopy is the other method used to place these stents.

Peer-review

This is a good paper.

REFERENCES

1 **Siegel R**, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin* 2012; **62**: 10-29 [PMID: 22237781 DOI: 10.3322/caac.20138]

2 **Thosani N**, Guha S, Singh H. Colonoscopy and colorectal cancer incidence and mortality. *Gastroenterol Clin North Am* 2013; **42**: 619-637 [PMID: 23931863 DOI: 10.1016/j.gtc.2013.05.011]

3 **Deans GT**, Krukowski ZH, Irwin ST. Malignant obstruction of the left colon. *Br J Surg* 1994; **81**: 1270-1276 [PMID: 7953385]

4 **Smothers L**, Hynan L, Fleming J, Turnage R, Simmang C, Anthony T. Emergency surgery for colon carcinoma. *Dis Colon Rectum* 2003; **46**: 24-30 [PMID: 12544518 DOI: 10.1097/01.DCR.0000044719.17980.4C]

5 **van Hooft JE**, Bemelman WA, Oldenburg B, Marinelli AW, Lutke Holzik MF, Grubben MJ, Sprangers MA, Dijkgraaf MG, Fockens P. Colonic stenting versus emergency surgery for acute left-sided malignant colonic obstruction: a multicentre randomised trial. *Lancet Oncol* 2011; **12**: 344-352 [PMID: 21398178 DOI: 10.1016/S1470-2045(11)70035-3]

6 **Spinelli P**, Dal Fante M, Mancini A. Self-expanding mesh stent for endoscopic palliation of rectal obstructing tumors: a preliminary report. *Surg Endosc* 1992; **6**: 72-74 [PMID: 1285349]

7 **Law WL**, Choi HK, Chu KW. Comparison of stenting with emergency surgery as palliative treatment for obstructing primary left-sided colorectal cancer. *Br J Surg* 2003; **90**: 1429-1433 [PMID: 14598426]

8 **Carne PW**, Frye JN, Robertson GM, Frizelle FA. Stents or open operation for palliation of colorectal cancer: a retrospective, cohort study of perioperative outcome and long-term survival. *Dis Colon Rectum* 2004; **47**: 1455-1461 [PMID: 15486741 DOI: 10.1007/s10350-004-0624-x]

9 **Sebastian S**, Johnston S, Geoghegan T, Torreggiani W, Buckley M. Pooled analysis of the efficacy and safety of self-expanding metal stenting in malignant colorectal obstruction. *Am J Gastroenterol* 2004; **99**: 2051-2057 [PMID: 15447772 DOI: 10.1111/j.1572-0241.

2004.40017.xAJG40017]
10 **Fernández-Esparrach G**, Bordas JM, Giráldez MD, Ginès A, Pellisé M, Sendino O, Martínez-Pallí G, Castells A, Llach J. Severe complications limit long-term clinical success of self-expanding metal stents in patients with obstructive colorectal cancer. *Am J Gastroenterol* 2010; **105**: 1087-1093 [PMID: 19935785 DOI: 10.1038/ajg.2009.660]
11 **Small AJ**, Coelho-Prabhu N, Baron TH. Endoscopic placement of self-expandable metal stents for malignant colonic obstruction: long-term outcomes and complication factors. *Gastrointest Endosc* 2010; **71**: 560-572 [PMID: 20189515 DOI: 10.1016/j.gie.2009.10.012]
12 **van Hooft JE**, Fockens P, Marinelli AW, Timmer R, van Berkel AM, Bossuyt PM, Bemelman WA. Early closure of a multicenter randomized clinical trial of endoscopic stenting versus surgery for stage IV left-sided colorectal cancer. *Endoscopy* 2008; **40**: 184-191 [PMID: 18322873 DOI: 10.1055/s-2007-995426]
13 **Moher D**, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535 [PMID: 19622551]
14 **Higgins JPT**, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2. The Cochrane Collaboration, 2009 [Updated September 2009]. Available from: URL: www.cochrane-handbook.org
15 **DerSimonian R**, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177-188 [PMID: 3802833]
16 **Higgins JP**, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539-1558 [PMID: 12111919 DOI: 10.1002/sim.1186]
17 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120 DOI: 10.1136/bmj.327.7414.557327/7414/557]
18 **Egger M**, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-634 [PMID: 9310563]
19 **Duval S**, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000; **56**: 455-463 [PMID: 10877304]
20 **Sterne JA**, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol* 2001; **54**: 1046-1055 [PMID: 11576817]
21 **Sterne JA**, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001; **323**: 101-105 [PMID: 11451790]
22 **Repici A**, Fregonese D, Costamagna G, Dumas R, Kähler G, Meisner S, Giovannini M, Freeman J, Petruziello L, Hervoso C, Comunale S, Faroux R. Ultraflex precision colonic stent placement for palliation of malignant colonic obstruction: a prospective multicenter study. *Gastrointest Endosc* 2007; **66**: 920-927 [PMID: 17904133 DOI: 10.1016/j.gie.2007.03.1042]
23 **Ptok H**, Meyer F, Marusch F, Steinert R, Gastinger I, Lippert H, Meyer L. Palliative stent implantation in the treatment of malignant colorectal obstruction. *Surg Endosc* 2006; **20**: 909-914 [PMID: 16738981 DOI: 10.1007/s00464-005-0594-7]
24 **Fregonese D**, Naspetti R, Ferrer S, Gallego J, Costamagna G, Dumas R, Campaioli M, Morante AL, Mambrini P, Meisner S, Repici A, Andreo L, Masci E, Mingo A, Barcenilla J, Petruzziello L. Ultraflex precision colonic stent placement as a bridge to surgery in patients with malignant colon obstruction. *Gastrointest Endosc* 2008; **67**: 68-73 [PMID: 18028916]
25 **Meisner S**, Hensler M, Knop FK, West F, Wille-Jørgensen P. Self-expanding metal stents for colonic obstruction: experiences from 104 procedures in a single center. *Dis Colon Rectum* 2004; **47**: 444-450 [PMID: 14994110 DOI: 10.1007/s10350-003-0081-y]
26 **Soto S**, López-Rosés L, González-Ramírez A, Lancho A, Santos A, Olivencia P. Endoscopic treatment of acute colorectal obstruction with self-expandable metallic stents: experience in a community hospital. *Surg Endosc* 2006; **20**: 1072-1076 [PMID: 16703437 DOI: 10.1007/s00464-005-0345-9]
27 **Repici A**, De Caro G, Luigiano C, Fabbri C, Pagano N, Preatoni P, Danese S, Fuccio L, Consolo P, Malesci A, D'Imperio N, Cennamo V. WallFlex colonic stent placement for management of malignant colonic obstruction: a prospective study at two centers. *Gastrointest Endosc* 2008; **67**: 77-84 [PMID: 18155427 DOI: 10.1016/j.gie.2007.08.019]
28 **Stipa F**, Pigazzi A, Bascone B, Cimitan A, Villotti G, Burza A, Vitale A. Management of obstructive colorectal cancer with endoscopic stenting followed by single-stage surgery: open or laparoscopic resection? *Surg Endosc* 2008; **22**: 1477-1481 [PMID: 18027039 DOI: 10.1007/s00464-007-9654-5]
29 **Branger F**, Thibaudeau E, Mucci-Hennekinne S, Métivier-Cesbron E, Vychnevskaia K, Hamy A, Arnaud JP. Management of acute malignant large-bowel obstruction with self-expanding metal stent. *Int J Colorectal Dis* 2010; **25**: 1481-1485 [PMID: 20607252 DOI: 10.1007/s00384-010-1003-9]
30 **West M**, Kiff R. Stenting of the colon in patients with malignant large bowel obstruction: a local experience. *J Gastrointest Cancer* 2011; **42**: 155-159 [PMID: 20596900 DOI: 10.1007/s12029-010-9178-4]
31 **Meisner S**, González-Huix F, Vandervoort JG, Goldberg P, Casellas JA, Roncero O, Grund KE, Alvarez A, García-Cano J, Vázquez-Astray E, Jiménez-Pérez J. Self-expandable metal stents for relieving malignant colorectal obstruction: short-term safety and efficacy within 30 days of stent procedure in 447 patients. *Gastrointest Endosc* 2011; **74**: 876-884 [PMID: 21855868 DOI: 10.1016/j.gie.2011.06.019]
32 **Súarez J**, Jiménez J, Vera R, Tarifa A, Balén E, Arrazubi V, Vila J, Lera JM. Stent or surgery for incurable obstructive colorectal cancer: an individualized decision. *Int J Colorectal Dis* 2010; **25**: 91-96 [PMID: 19859722 DOI: 10.1007/s00384-009-0814-z]
33 **Im JP**, Kim SG, Kang HW, Kim JS, Jung HC, Song IS. Clinical outcomes and patency of self-expanding metal stents in patients with malignant colorectal obstruction: a prospective single center study. *Int J Colorectal Dis* 2008; **23**: 789-794 [PMID: 18443807 DOI: 10.1007/s00384-008-0477-1]
34 **Suh JP**, Kim SW, Cho YK, Park JM, Lee IS, Choi MG, Chung IS, Kim HJ, Kang WK, Oh ST. Effectiveness of stent placement for palliative treatment in malignant colorectal obstruction and predictive factors for stent occlusion. *Surg Endosc* 2010; **24**: 400-406 [PMID: 19551432 DOI: 10.1007/s00464-009-0589-x]
35 **Jung MK**, Park SY, Jeon SW, Cho CM, Tak WY, Kweon YO, Kim SK, Choi YH, Kim GC, Ryeom HK. Factors associated with the long-term outcome of a self-expandable colon stent used for palliation of malignant colorectal obstruction. *Surg Endosc* 2010; **24**: 525-530 [PMID: 19597776 DOI: 10.1007/s00464-009-0604-2]
36 **Li YD**, Cheng YS, Li MH, Fan YB, Chen NW, Wang Y, Zhao JG. Management of acute malignant colorectal obstruction with a novel self-expanding metallic stent as a bridge to surgery. *Eur J Radiol* 2010; **73**: 566-571 [PMID: 19167177 DOI: 10.1016/j.ejrad.2008.12.004]
37 **Lee KM**, Shin SJ, Hwang JC, Cheong JY, Yoo BM, Lee KJ, Hahm KB, Kim JH, Cho SW. Comparison of uncovered stent with covered stent for treatment of malignant colorectal obstruction. *Gastrointest Endosc* 2007; **66**: 931-936 [PMID: 17767930 DOI: 10.1016/j.gie.2007.02.064]
38 **Park S**, Cheon JH, Park JJ, Moon CM, Hong SP, Lee SK, Kim TI, Kim WH. Comparison of efficacies between stents for malignant colorectal obstruction: a randomized, prospective study. *Gastrointest Endosc* 2010; **72**: 304-310 [PMID: 20561619 DOI: 10.1016/j.gie.2010.02.046]
39 **Lee HJ**, Hong SP, Cheon JH, Kim TI, Min BS, Kim NK, Kim WH. Long-term outcome of palliative therapy for malignant colorectal obstruction in patients with unresectable metastatic colorectal cancers: endoscopic stenting versus surgery. *Gastrointest Endosc* 2011; **73**: 535-542 [PMID: 21257165 DOI: 10.1016/j.gie.2010.10.052]
40 **Ng KC**, Law WL, Lee YM, Choi HK, Seto CL, Ho JW. Self-expanding metallic stent as a bridge to surgery versus emergency resection for obstructing left-sided colorectal cancer: a case-matched study. *J Gastrointest Surg* 2006; **10**: 798-803 [PMID: 16769535 DOI: 10.1016/j.gassur.2006.02.006]

41 **Cheung HY**, Chung CC, Tsang WW, Wong JC, Yau KK, Li MK. Endolaparoscopic approach vs conventional open surgery in the treatment of obstructing left-sided colon cancer: a randomized controlled trial. *Arch Surg* 2009; **144**: 1127-1132 [PMID: 20026830 DOI: 10.1001/archsurg.2009.216]

42 **Guo MG**, Feng Y, Zheng Q, Di JZ, Wang Y, Fan YB, Huang XY. Comparison of self-expanding metal stents and urgent surgery for left-sided malignant colonic obstruction in elderly patients. *Dig Dis Sci* 2011; **56**: 2706-2710 [PMID: 21442324 DOI: 10.1007/s10620-011-1648-4]

43 **Ho KS**, Quah HM, Lim JF, Tang CL, Eu KW. Endoscopic stenting and elective surgery versus emergency surgery for left-sided malignant colonic obstruction: a prospective randomized trial. *Int J Colorectal Dis* 2012; **27**: 355-362 [PMID: 22033810 DOI: 10.1007/s00384-011-1331-4]

44 **Lee JH**, Ross WA, Davila R, Chang G, Lin E, Dekovich A, Davila M. Self-expandable metal stents (SEMS) can serve as a bridge to surgery or as a definitive therapy in patients with an advanced stage of cancer: clinical experience of a tertiary cancer center. *Dig Dis Sci* 2010; **55**: 3530-3536 [PMID: 20721627 DOI: 10.1007/s10620-010-1370-7]

45 **Vemulapalli R**, Lara LF, Sreenarasimhaiah J, Harford WV, Siddiqui AA. A comparison of palliative stenting or emergent surgery for obstructing incurable colon cancer. *Dig Dis Sci* 2010; **55**: 1732-1737 [PMID: 19693667 DOI: 10.1007/s10620-009-0945-7]

46 **Ghazal AH**, El-Shazly WG, Bessa SS, El-Riwini MT, Hussein AM. Colonic endolumenal stenting devices and elective surgery versus emergency subtotal/total colectomy in the management of malignant obstructed left colon carcinoma. *J Gastrointest Surg* 2013; **17**: 1123-1129 [PMID: 23358847 DOI: 10.1007/s11605-013-2152-2]

47 **Pirlet IA**, Slim K, Kwiatkowski F, Michot F, Millat BL. Emergency preoperative stenting versus surgery for acute left-sided malignant colonic obstruction: a multicenter randomized controlled trial. *Surg Endosc* 2011; **25**: 1814-1821 [PMID: 21170659 DOI: 10.1007/s00464-010-1471-6]

48 **Yoon JY**, Jung YS, Hong SP, Kim TI, Kim WH, Cheon JH. Clinical outcomes and risk factors for technical and clinical failures of self-expandable metal stent insertion for malignant colorectal obstruction. *Gastrointest Endosc* 2011; **74**: 858-868 [PMID: 21862005 DOI: 10.1016/j.gie.2011.05.044]

P- Reviewer: Cheung HYS, Grund KE, Tsujikawa T
S- Editor: Ji FF **L- Editor**: A **E- Editor**: Jiao XK

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i6.254

World J Meta-Anal 2015 December 26; 3(6): 254-283
ISSN 2308-3840 (online)

META-ANALYSIS

Risk of pancreatic adverse events associated with the use of glucagon-like peptide-1 receptor agonist and dipeptidyl peptidase-4 inhibitor drugs: A systematic review and meta-analysis of randomized trials

Hasan M Shihab, Tokunbo Akande, Kacie Armstrong, Sonal Singh, Yoon K Loke

Hasan M Shihab, Division of Acute Care Surgery, Johns Hopkins University School of Medicine, Baltimore, MD 21205, United States

Tokunbo Akande, Department of Pediatrics, Bronx-Lebanon Hospital Center, Bronx, NY 10457, United States

Kacie Armstrong, Sonal Singh, Division of General Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, United States

Yoon K Loke, Norwich Medical School, University of East Anglia, Norwich NR4 7TJ, United Kingdom

Author contributions: Shihab HM designed study, extracted and analyzed data, and wrote manuscript; Akande T and Armstrong K contributed to study identification/extraction and edited manuscript; Singh S designed study, analysed data, contributed discussion and edited manuscript as corresponding author; Loke YK designed study, extracted and, analyzed data, and edited manuscript as corresponding author.

Conflict-of-interest statement: Sonal Singh had served as a consultant on the advisory board of Janssen Pharmaceuticals Inc. to comment on the safety of sodium glucose co-transporter inhibitor-2 (SGLT-2) canagliflozin. He was compensated for his time.

Data sharing statement: All available data has been presented in the manuscript.

Correspondence to: Yoon K Loke, MD, Professor, Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom. y.loke@uea.ac.uk
Telephone: +44-1603-591234
Fax: +44-1603-59752

Received: June 26, 2015
Peer-review started: June 28, 2015
First decision: September 17, 2015
Revised: October 13, 2015
Accepted: December 3, 2015
Article in press: December 4, 2015
Published online: December 26, 2015

Abstract

AIM: To systematically assess risk of pancreatic adverse events with glucagon-like peptide-1 (GLP-1) receptor agonist and dipeptidyl peptidase-4 (DPP-4) inhibitor drugs.

METHODS: We searched PubMed, Embase, CINAHL, Cochrane review of clinical trials, pharmaceutical company clinical trials register, United States Food and Drug Administration website, European Medicines Agency website and ClinicalTrials.gov for randomized controlled trials from inception to October 2013. Randomized control trial studies were selected for inclusion if they reported on pancreatic complication events and/or changes in pancreatic enzyme levels (serum amylase and serum lipase) as adverse events or as serious adverse events for patients who were on GLP-1 receptor agonist and DPP-4 inhibitor drugs. Two independent reviewers extracted data directly. We performed Peto odds ratio (OR) fixed effect meta-analysis of pancreatic adverse events a, and assessed heterogeneity with the I^2 statistic.

RESULTS: Sixty-eight randomized controlled trials were eligible. A total of 60720 patients were included in our analysis of the association of risk of pancreatic complication events with GLP-1 agents. A total of 89 pancreatic related adverse events occurred among the GLP-1 agents compared to 74 events among the controls. There was a statistically significant increased risk of elevation of pancreatic enzymes associated with GLP-1 agents compared with control (Peto OR = 3.15, 95%CI: 1.56-6.39, P = 0.001, I^2 = 0%). There was no statistically significant difference in the risk of pancreatic adverse event associated with GLP-1 agent compared with controls (Peto OR = 1.00, 95%CI: 0.73-1.37, P = 1.00, I^2 = 0%). There were a total of 71 pancreatitis events in patients on GLP-1 agents and 56 pancreatitis events occurred in the control patients. There were 36 reports of pancreatic cancer in these studies. Of these cases, 2 used linagliptin, 2 used alogliptin, 1 used vildagliptin, 7 used saxagliptin while 6 used sitagliptin. The remaining 18 cases occurred among controls.

CONCLUSION: Although GLP-1 based agents are associated with pancreatic enzyme elevation, we were unable to confirm a significant risk of pancreatitis or pancreatic cancer.

Key words: Diabetes mellitus; Pancreatitis; Glucagon-like peptide-1 agonists; Dipeptidyl peptidase-4 inhibitors; Meta-analysis

Core tip: There is conflicting data on the risk of pancreatitis with glucagon-like peptide-1 (GLP-1) receptor agonists and dipeptidyl peptidase-4 (DPP-4) inhibitors. We performed a meta-analysis of 68 randomized controlled trials of 11 different GLP-1 or DPP-4 targeted drugs. The incidence of pancreatic adverse events in the trials was generally low and we did not find any definitive evidence for pancreatitis or pancreatic cancer amongst the trials. However, we found a significantly raised risk of elevated pancreatic enzymes in a small number of trials that reported such enzyme elevations.

Shihab HM, Akande T, Armstrong K, Singh S, Loke YK. Risk of pancreatic adverse events associated with the use of glucagon-like peptide-1 receptor agonist and dipeptidyl peptidase-4 inhibitor drugs: A systematic review and meta-analysis of randomized trials. *World J Meta-Anal* 2015; 3(6): 254-283 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i6/254.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i6.254

INTRODUCTION

Glucagon-like peptide-1 (GLP-1) is a naturally occurring gut hormone that is mainly secreted by the intestinal L cell. It is a potent antihyperglycemic hormone, inducing glucose-dependent stimulation of insulin secretion while suppressing glucagon secretion. Once in the circulation, GLP-1 has a half-life of less than 2 min, due to the rapid degradation by the enzyme dipeptidyl peptidase-4 (DPP-4). The GLP-1 based therapies include GLP-1 receptor agonists and DPP-4 inhibitors. As GLP-1 is a gut hormone, it is possible that patients may experience adverse effects on the gastrointestinal system such as nausea or abdominal pain.

There are already several GLP-1 receptor agonists and DPP-4 inhibitor drugs approved by the Food and Drug Administration (FDA) or the European Medicines Agency (EMA), and we are aware of additional agents in development. However, sitagliptin and exenatide have been shown to cause acute pancreatitis in rodent models *via* amplification of ductal replication and induction of acinar to ductal metaplasia[1-4]. A recent case-control study showed a significant increased risk of hospitalization for acute pancreatitis associated with the use of sitagliptin or exenatide among adult patients with type-2 diabetes mellitus[5]. A meta-analysis of clinical trials reported no difference for sitagliptin use compared with placebo or other oral hypoglycemic in the incidence rates of pancreatitis[6]. Although complications involving the pancreas (acute pancreatitis, chronic pancreatitis and pancreatic cancer) are potentially serious adverse effects of GLP-1 receptor agonist drugs, there is a paucity of data available to clinicians regarding these effects of GLP-1 receptor agonist drugs. A recent meta-analysis[7] suggested that neither exenatide nor liraglutide increases the risk for acute pancreatitis when used in the treatment of type-2 diabetes mellitus. This analysis, however, was based on small studies, non-clinical evaluation of pancreatitis in the included RCTs and residual confounding in the observational studies that were included. None of the previous studies have adequately evaluated the role of pancreatic enzyme elevations. These studies have not evaluated the occurrence of reports of pancreatic cancer in these trials. Finally, the risk of pancreatic complication associated with individual therapies has not been evaluated.

Our objective was to conduct a systematic review to ascertain the risk of pancreatic complications (acute and chronic pancreatitis and pancreatic cancer) and pancreatic enzyme elevations associated with GLP-1 based therapies, as compared to placebo or other oral hypoglycemic drugs in randomized controlled trials of GLP-1 based therapies.

MATERIALS AND METHODS

Methods

We defined study aims and procedures in the study protocol registered with PROSPERO register of systematic reviews[8].

Data sources and searches

We searched MEDLINE, EMBASE, CINAHL and the Cochrane database from inception to October 2013 using the search terms: (drug name OR chemical com-

pound OR drug class) AND ["Pancreatic Neoplasms" (Mesh) OR "Pancreatitis"(Mesh) OR "pancreas"(tiab) OR "pancreatitis"(tiab) OR "pancreatic"(tiab) OR "pancreatic cancer"(tiab) OR "serum amylase"(tiab) OR "serum lipase"(tiab) OR "Islet Cell Adenoma" (tiab) OR "Insulinoma"(tiab) OR "Islet Cell Carcinoma" (tiab) OR "Gastrinoma"(tiab) OR "Glucagonoma"(tiab) OR "Somatostatinoma"(tiab) OR "Vipoma"(tiab) OR "Pancreatic Ductal Carcinoma"(tiab) OR "Islet Cell Adenomas"(tiab) OR "Insulinomas"(tiab) OR "Islet Cell Carcinomas"(tiab) OR "Gastrinomas"(tiab) OR "Glucagonomas"(tiab) OR "Somatostatinomas"(tiab) OR "Vipomas"(tiab) OR "Pancreatic Ductal Carcinomas" (tiab)] AND English(lang) NOT ["Animals"(Mesh)] NOT ["Animals"(Mesh) AND "Humans"(Mesh)].

We did not specify any language or population restrictions. To identify any unpublished studies, we keyed in the names of specific drug compounds into the search boxes of all GLP-1 agent pharmaceutical companies, three of which had publicly available clinical trials, these were Boehringer Ingelheim clinical trials register, Novartis clinical trials register and Takeda Pharmaceuticals register. We also searched the FDA, the EMA and ClinicalTrials.gov up to August 2013. Bibliographies of included studies and recent review articles were checked for additional relevant studies.

Study selection

We selected randomized controlled trials that enrolled participants using GLP-1 agonist and DPP-4 inhibitor drugs and reported on the risk of pancreatic complications either as adverse events or as serious adverse events. We included studies that reported on the use of FDA approved GLP-1 receptor agonists such as Exenatide (Byetta, Bydureon), Liraglutide (Victoza) and Albiglutide (Tanzeum). Other GLP-1 receptor agonists that were studied but have not yet been approved by FDA included Taspoglutide, Lixisenatide (Lyxumia), Dulaglutide and Semaglutide were included. Studies that also used FDA approved DPP-4 inhibitors such as Vidagliptin (Eucreas, Galvus, Icandra, Jalra, Xiliarx, Zomarist), Sitagliptin (Efficib, Januvia, Janumet, Ristaben, Ristfor, Tesavel, Velmetia, Xelevia), Saxagliptin (Komboglyze, Onglyza), Linagliptin (Jentadueto, Trajenta) and Alogliptin (Nesina) were included. Other DPP-4 inhibitors in development were included in our search. These include Septagliptin, Anagliptin, Biseglitpin, Carmegliptin, Denagliptin, Dutogliptin, Gosogliptin, Isoleucine Thiazolidide, Valine pyrrolidide, Evogliptin, Gemigliptin, Melogliptin, Omarigliptin, Tenegliptin and Trelagliptin. We did not restrict studies by healthcare settings, methods of diagnosing pancreatitis or by indication for the drug.

Data extraction and quality assessment

Two reviewers (HMS and TA) evaluated all titles and abstracts for studies that met the inclusion criteria, and excluded any articles that clearly did not meet the selection criteria. Full reports of potentially relevant studies were retrieved and independently checked for eligibility. Data from the included studies were then extracted independently by two reviewers (HMS and TA) who collected information on study design, study location, study population description, drug exposure, pancreatic complication (acute pancreatitis, chronic pancreatitis, pancreatic cancer) events, pancreatic enzyme derangement (elevated serum pancreatic amylase and/or pancreatic serum lipase) data, mortality from pancreatic events, how the pancreatic events were defined and monitored, confounders for pancreatic events and characteristics of participants onto a pre-formatted spreadsheet. Another reviewer (YKL or SS) then checked the data. Any uncertainties or discrepancies were resolved through rechecking against the source papers, and through discussion with a third reviewer.

We used a pre-specified spreadsheet to record the location and duration of the randomized controlled trials (in years), dose and frequency of GLP-1 agonist drug and DPP-4 inhibitor drug and placebo or alternative hypoglycemic agent, mean age and sex of participants, number of pancreatic complication events and confounders.

The Cochrane toolkit was used for the assessment of bias in evaluating each trial for the reporting of randomization, allocation concealment, the use of blinding of participants and staff, and information on loss to follow-up or withdrawal rates[9]. In accordance with the Cochrane handbook of systematic reviews, we assessed the quality of data on adverse events by recording how they were monitored and recorded by the investigators[10]. We aimed to generate funnel plots to assess the possibility of publication bias, provided that there were > 10 studies available in the meta-analysis, with no evidence of substantial statistical heterogeneity[11].

Statistical analysis

We used RevMan[12] 5.3 to conduct meta-analysis based on the summary statistic of Peto Odds Ratios, which is the recommended approach for rare events[9]. We assumed similarity between the risk ratio and OR because the incidence of adverse outcomes was low[13]. We evaluated both adjusted and unadjusted data from primary studies, although we preferentially used adjusted data where available.

Statistical heterogeneity was assessed using I^2 statistic[14], with I^2 values of 30%-60% representing a moderate level of heterogeneity. Pre-specified subgroup analysis was performed by evaluating the effect of study design, study setting and outcome ascertainment.

The statistical methods of this study were reviewed by Yoon K Loke, convenor of the Cochrane Adverse Effects Methods Group.

RESULTS

After a review of 3583 citations, we identified 68 randomized controlled trials (Figure 1) with a total of 60811

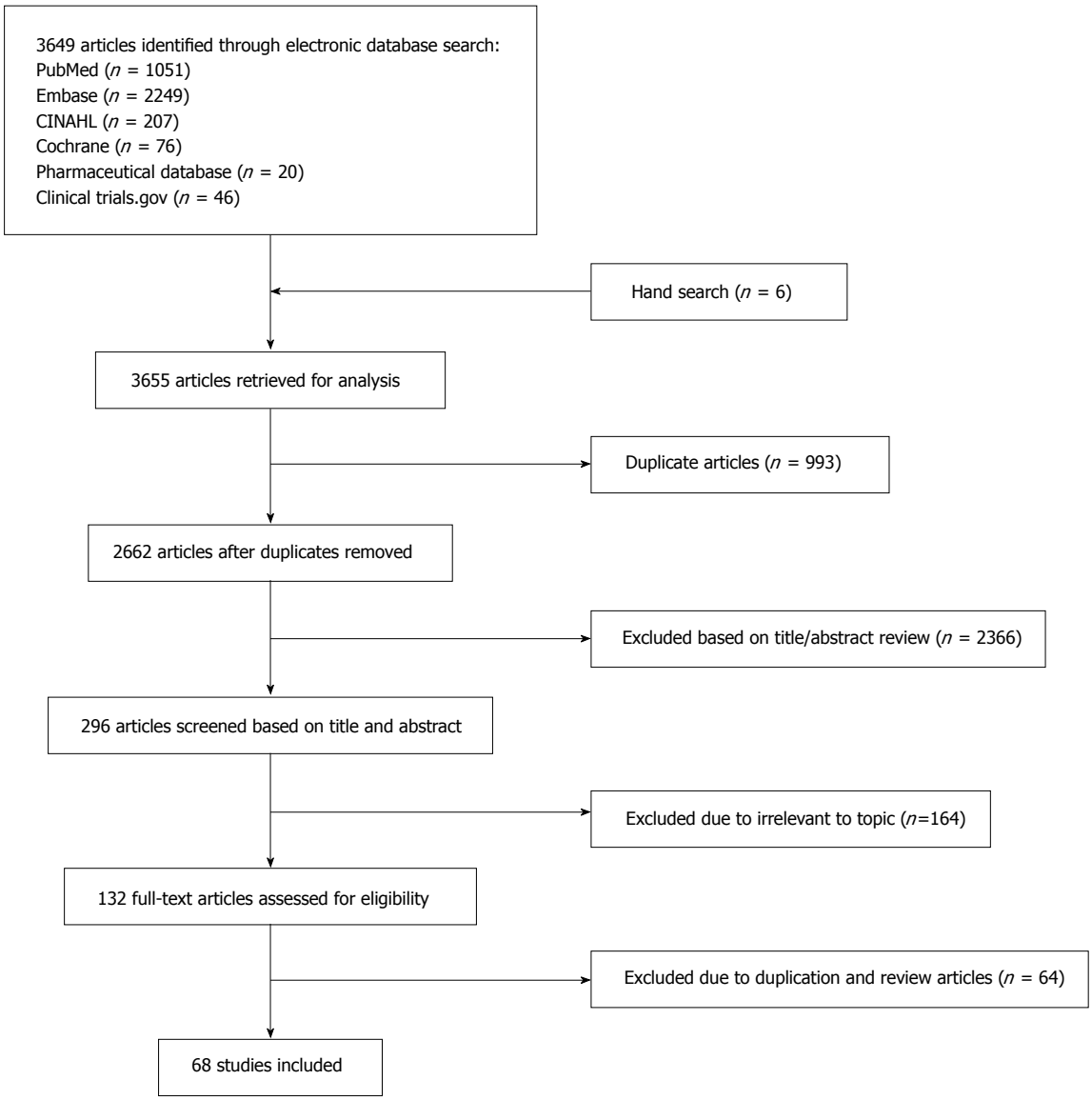


Figure 1 Flow diagram of studies identified and selected.

patients for inclusion in our analysis of the association of risk of pancreatic complication events with the use of GLP-1 agonist and DPP-4 inhibitor drugs.

Description of studies

The study characteristics are listed in Table 1 and quality assessment of the trials in Table 2.

Of these 69 studies, data was abstracted from 28 published reports, 32 studies from clinicaltrials.gov and 9 studies were abstracted from pharmaceutical company databases (Boehringer Ingelheim, Novartis and Takeda Pharmaceutical Company). Almost all the trials were multicenter or multinational studies in patients with type Ⅱ diabetes mellitus.

The majority of studies did not report on the method of generating the random sequence, or on the means of concealing allocation. However, most of the trials (*n* = 55) were double-blinded, thus reducing risk of bias in the diagnosis of pancreatic adverse events. We found that most of the trials, except for two, did not specific pancreatitis as a part of their safety monitoring protocol. As such, there is a strong possibility that pancreatic adverse events may have been missed or wrongly diagnosed. Moreover, the included studies did not specify whether they applied similar criteria in defining cases of pancreatitis.

Overall

Total pancreatic related adverse events: With a total of 89 pancreatic related adverse events among the 34340 number of patients receiving GLP-1 agents and 74 events among 26380 patients receiving the control agents, there was no statistically significant difference in the risk of pancreatic adverse event associated with GLP-1 agent compared with controls (Peto OR = 0.99, 95%CI: 0.72-1.36, $P = 0.96$; $I^2 = 0\%$) (Figure 2).

Pancreatitis: There were a total of 71 pancreatitis

Table 1 Characteristics of glucagon-like peptide-1 based agents in randomized controlled trials included in analysis of pancreatic events

Ref.	Location (No. of centers)	Year of study completion	Total duration (wk)	Duration of GLP-1 exposure (wk)	Participant disease	Arms	No. of participants	Mean age, yr (SD)	Female, *n* (%)
Ross *et al*[21]	Multi-national (84 centers in 9 countries)	2010	43	12	Type 2 diabetes	Linagliptin 2.5 mg bid	223	58.7 (9.9)	85 (38.1)
						Linagliptin 5 mg qd	224	58.4 (10.6)	103 (46.0)
						Placebo	44	59.9 (10.7)	23 (52.3)
Haak *et al*[22]	Multi-national (133 clinics in 14 countries)	2010	73	24	Type 2 diabetes	Linagliptin 5 mg qd	142	56.2 (10.8)	62 (43.7)
						Metformin 500 mg bid	144	52.9 (10.4)	62 (43.1)
						Metformin 1000 mg bid	147	55.2 (10.6)	69 (46.9)
						Linagliptin 2.5 mg qd + Metformin 500 mg bid	143	55.6 (11.2)	73 (49.0)
						Linagliptin 2.5 mg qd + Metformin 1000 mg bid	143	56.4 (10.7)	66 (46.2)
						Placebo	72	55.7 (11.0)	36 (50)
NCT00328172[23]	Multi-national (71 sites in 6 countries)	2007	65	12	Type 2 diabetes	Linagliptin 0.5 mg	58	58.0 (9.4)	13 (22.4)
						Linagliptin 2.5 mg	57	59.8 (10.3)	30 (52.6)
						Linagliptin 5.0 mg	55	56.6 (9.6)	24 (43.6)
						Metformin	65	53.7 (10.7)	26 (40.0)
						Placebo	67	58.6 (8.9)	34 (50.7)
Yki-Jarvinen *et al*[24,25]	Multi-national (169 sites in 19 countries)	2011	108	52	Type 2 diabetes	Linagliptin 5.0 mg	631	59.7 (9.9)	302 (47.9)
						Placebo	630	60.4 (10.0)	301 (47.8)
NCT00654381[26]	Japan	2010	91	12	Type 2 diabetes	Linagliptin 5.0 mg	159	60.3 (9.4)	48 (30.2)
						Linagliptin 10.0 mg	160	61.3 (10.0)	48 (30.0)
						Voglibose	162	58.5 (9.9)	47 (29.0)
						Placebo	80	59.7 (8.9)	23 (28.7)
NCT00622284[27]	Multi-national (221 sites in 16 countries)	2010	146	104	Type 2 diabetes	Linagliptin	776	59.8 (9.4)	314 (40.5)
						Glimepiride	775	59.8 (9.4)	304 (39.2)
BI Trial No: 1218.15/ U09-2519-01[28]	Multi-national (43 sites in 7 countries)	2009	61	24	Type 2 diabetes	Linagliptin 5 mg + Pioglitazone 30 mg	259	NR	NR
						Pioglitazone 30 mg + Placebo	130	NR	NR
BI Trial No: 1218.52/ U11-1782-01[29]	Multi-national (101 sites in 14 countries)	2011	102	54	Type 2 diabetes	Linagliptin 2.5 mg + Metformin (500 mg and 1000 mg bid)	396	NR	NR
						Metformin 1000 mg bid	170	NR	NR
BI Trial No: 1218.63/ U11-1781-02[30]	Multi-national (33 sites in 5 countries)	2011	67	24	Type 2 diabetes	Linagliptin 5 mg	162	NR	46 (28.4)
						Placebo	79	NR	30 (38.0)
BI Trial No: 1218.75/ U12-3204-01[31]	Multi-center study (Black/African American patients only)	2011	55	24	Type 2 diabetes	Linagliptin 5 mg	106	NR	NR
						Placebo	120	NR	NR
BI Trial No: 1218.61/ U13-3124-01[32]	Multi-national study (4 countries)	2012	123	24	Type 2 diabetes	Linagliptin 5 mg	183	NR	NR
						Placebo	89	NR	NR
BI Trial No: 1218.65/ U12-2143-01[33]	Multi-national study (19 sites in 3 countries)	2012	74	24	Type 2 diabetes	Linagliptin 5 mg	205	82%(< 65 yr)	NR
						Placebo	100	83% (< 65 yr)	NR
BI Trial No: 1218.64/ U13-1283-01[34]	Multi-national study (52 sites in 9 countries)	2012	117	52	Type 2 diabetes	Linagliptin 5 mg	113	NR	43 (38.1)
						Placebo (first 12 wk)/ Glimepiride (next 40 wk)	122	NR	43 (35.2)
BI Trial No: 1218.66/ U12-2076-01[35]	Multi-national study (19 sites in 3 countries	2012	80	24	Type 2 diabetes	Linagliptin 5 mg	200	84.0% (< 65 yr)	NR
						Placebo	99	89.9% (< 65 yr)	NR
Rosenstock *et al*[36]	Multi-national study (110 sites in 13 countries)	2007	65	26	Type 2 diabetes	Alogliptin 12.5 mg	131	55.4 (9.8)	76 (58)
						Alogliptin 25 mg	129	55.9 (10.2)	85 (66)
						Placebo	130	55.0 (10.6)	68 (52)
White *et al*[37]	Multi-national study (898 centers in 49 countries)	2013	193	76 (median)	Type 2 diabetes	Alogliptin	2701	36.0% (≥ 65 yr)	873 (32.3)
						Placebo	2679	34.9% (≥ 65 yr)	856 (32.0)

NCT01318135[38]	Japan (58 sites)	2010	52	52	Type 2 diabetes	Alogliptin 12.5 mg qd + Glimepiride 1-6 mg qd or bid	150	38.0% (≥ 65 yr)	53 (35.3)
						Alogliptin 25 mg qd + Glimepiride 1-6 mg qd or bid	152	30.9% (≥ 65 yr)	52 (34.2)
NCT01289119[39]	Multi-national study (21 sites in 3 countries)	2011	52	16	Type 2 diabetes	Alogliptin monotherapy	92	51.6 (10.41)	37 (40.2)
						Metformin	98	53.2 (9.46)	50 (51.0)
						Metformin + Alogliptin Add-on Therapy	99	53.0 (9.88)	48 (48.5)
						Pioglitazone	63	51.8 (10.37)	24 (38.1)
						Pioglitazone + Alogliptin Add-on Therapy	61	52.6 (9.44)	33 (54.1)
						Placebo	93	53.1 (8.88)	39 (41.9)
NCT01263496[40]	Japan (58 sites)	2008	72	52	Type 2 diabetes	Alogliptin 6.25 mg qd	96	28.1 (≥ 65 yr)	26 (27.1)
						Alogliptin 12.5 mg qd	101	33.7 (≥ 6 5 yr)	29 (28.7)
						Alogliptin 25 mg qd	97	34.0 (≥ 65 yr)	22 (22.7)
						Alogliptin 50 mg qd	97	32.9 (≥ 65 yr)	29 (29.9)
						Voglibose 0.2 mg tid	83	38.6 (≥ 65 yr)	27 (32.5)
NCT00328627[41]	Multi-national study (90 sites in 19 countries)	2008	93	26	Type 2 diabetes	Alogliptin 12.5 mg + Placebo	128	53.1 (9.59)	61 (47.6)
						Alogliptin 25 mg + Placebo	129	53.7 (9.31)	79 (61.2)
						Placebo	129	55.2 (9.89)	68 (52.7)
NCT00395512[42]	Multi-national study (268 sites in 23 countries)	2008	67	26	Type 2 diabetes	Alogliptin 25 mg + Pioglitazone 30 mg	164	52.8 (11.01)	91 (55.5)
						Alogliptin 12.5 mg + Pioglitazone 30 mg	164	53.5 (11.37)	83 (50.6)
						Pioglitazone 30 mg	163	51.5 (10.72)	73 (44.8)
Kikuchi *et al*[43]	Japan (26 sites)	2007	52	12	Type 2 diabetes	Vildagliptin 50 mg bid + glimepiride	102	59.2 (9.8)	27 (26.5)
						Placebo + glimepiride	100	60.3 (10.1)	31 (31.0)
Lukashevich *et al*[44]	Multi-national study (12 countries)	2010	291	24	Type 2 diabetes	Vildagliptin 50 mg qd (moderate RI)	165	67.7 (8.8)	69 (41.8)
						Placebo (moderate RI)	129	69.7 (7.3)	49 (38.0)
						Vildagliptin 50 mg qd (severe RI)	124	64.1 (9.2)	59 (47.6)
						Placebo (severe RI)	97	64.5 (10.8)	44 (45.4)
Strain *et al*[45]	Multi-national study (45 centers in 7 countries)	2012	64	24	Type 2 diabetes	Vildagliptin	139	75.1 (4.3)	66 (47.5)
						Placebo	139	74.4 (4.0)	86 (61.9)
NCT00106340[46] (CLAF237A2308)	Multi-national study (402 centers in 25 countries)	2008	166	104	Type 2 diabetes	Vildagliptin 50 mg bid + Metformin	1562	57.5 (9.07)	733 (46.9)
						Glimepiride up to 6 mg qd + Metformin	1556	57.5 (9.19)	718 (46.1)
NCT00300287[47] (CLAF237A2307)	Multi-national study (69 centers in 6 countries)	2006	85	52	Type 2 diabetes	Vildagliptin 50 mg qd	156	63.27 (10.18)	63 (40.4)
						Placebo	150	62.84 (11.03)	61 (40.7)
CLAF237A1301[48]	Japan (51 centers)	2007	44	12	Type 2 diabetes	Vildagliptin 50 mg bid	188	60.3 (10.48)	67 (35.6)
						Voglibose 0.2 mg tid	192	58.0 (9.32)	62 (32.3)
CLAF237A23119[49]	United States (796 centers)	2007	53	12	Type 2 diabetes	Vildagliptin 100 mg + Metformin	1776	55.3	864 (48.6)
						Thiazolinedione + Metformin	888	56.2	467 (52.6)
NCT00110240[50] (CLAF237A2323)	Multi-national study (31 centers in 3 countries)	2006	87	24	Type 2 diabetes	Vildagliptin 50 mg bid	441	51.79 (10.13)	176 (40.0)
						Acarbose up to 100 mg tid	220	51.93 (10.34)	81 (37.0)
NCT00327015[51]	Multi-national study (211 sites in 13 countries)	2007	78	24	Type 2 diabetes	Saxagliptin 5 mg + Metformin 500 mg	320	51.95 (10.43)	155 (48.4)

						Saxagliptin 10 mg + Metformin 500 mg	323	52.08 (11.59)	177 (54.8)
						Metformin 500 mg + Placebo	328	51.83 (10.74)	165 (50.3)
Hollander *et al*[52] (NCT00295633)	Multi-national study (133 sites in 7 countries)	2007	82	24	Type 2 diabetes	Saxagliptin 2.5 mg + TZD	195	54.9 (9.7)	89 (45.6)
						Saxagliptin 5 mg + TZD	186	53.2 (10.6)	97 (52.2)
						Placebo + TZD	184	54.0 (10.1)	99 (53.8)
NCT00757588[53]	Multi-national study (80 sites in 10 countries)	2010	73	24	Type 2 diabetes	Saxagliptin 5 mg + Insulin	304	57.2 (9.4)	184 (60.5)
						Placebo + Insulin	151	57.3 (9.3)	83 (54.9)
Scirica *et al*[54]	Multi-national study (788 sites in 26 countries)	2013	156	109	Type 2 diabetes	Saxagliptin	8280	65.1 (8.5)	2768 (33.4)
						Placebo	8212	65 (8.6)	2687 (32.7)
Göke *et al*[55]	Multi-national study (130 sites in 11 countries)	2010	139	104	Type 2 diabetes	Saxagliptin + Metformin	428	57.5 (10.26)	216 (50.5)
						Glipizide + Metformin	430	57.59 (10.37)	198 (46.1)
NCT00316082[56]	Multi-national study (74 sites in 4 countries)	2007	74	24	Type 2 diabetes	Saxagliptin 2.5/5 mg QAM	71	54.28 (10.93)	34 (47.9)
						Saxagliptin 2.5 mg QAM	74	55.24 (10.44)	49 (66.2)
						Saxagliptin 5 mg QAM	74	54.66 (9.71)	36 (48.6)
						Saxagliptin 5 mg QPM	72	55.11 (10.35)	39 (54.2)
						Placebo	74	55.57 (10.32)	39 (52.7)
NCT00614939[57]	Multi-national study (74 sites in 14 countries)	2009	74	12	Type 2 diabetes	Saxagliptin	85	66.8 (8.3)	53 (62.4)
						Placebo	85	66.2 (9.1)	44 (51.8)
Chan *et al*[58,59]	Multi-national study (30 sites in 13 countries)	2006	NR	54	Type 2 diabetes	Sitagliptin 50 mg or 25 mg once daily	65	68.9 (9.8)	34 (52.3)
						Glipizide	26	65.3 (9.7)	10 (38.5)
Kojima *et al*[60]	Japan (Japanese Red Cross Medical Center)	2011	65	12	Type 2 diabetes	Sitagliptin	20	63.85 (12.92)	5 (0.25)
						Nateglinide	16	66.44 (9.02)	4 (0.25)
NCT00509262 (Arjona Ferreira *et al*[61,62])	Multi-national study	2011	178	54	Type 2 diabetes	Sitagliptin	211	64.2 (10.7)	80 (37.9)
						Glipizide	212	64.2 (9.4)	90 (42.5)
Henry *et al*[63,64]	Multi-national study	2010	108	54	Type 2 diabetes	Sitagliptin 100 mg/ Pioglitazone 15 mg	230	NR	112 (48.7)
						Sitagliptin 100 mg/ Pioglitazone 30 mg	231	NR	96 (41.6)
						Sitagliptin 100 mg/ Pioglitazone 45 mg	230	NR	95 (41.3)
						Pioglitazone 15 mg	230	NR	82 (35.7)
						Pioglitazone 30 mg	233	NR	106 (45.5)
						Pioglitazone 45 mg	230	NR	117 (50.9)
Raz *et al*[65,66]	Multi-national study (30 sites in 13 countries)	2007	47	30	Type 2 diabetes	Sitagliptin 100 mg	96	53.6 (9.5)	47 (48.9)
						Placebo	94	56.1 (9.5)	55 (58.5)
NCT01131182[67]	NR	2010	22	4	Type 2 diabetes	Sitagliptin	507	55.0 (11.0)	238 (46.9)
						Sulfonylurea	514	55.0 (11.0)	259 (50.4)
Goldstein *et al*[68,69]	Multi-national study	2006	69	54	Type 2 diabetes	Sitagliptin 50 mg bid + Metformin 500 mg bid	190	54.1 (10.0)	85 (44.7)
						Sitagliptin 50 mg bid + Metformin 1000 mg bid	182	53.3 (9.6)	105 (57.7)
						Sitagliptin 50 mg bid + Metformin 1000 mg bid (Open Label Cohort)	117	52.6 (10.0)	50 (42.7)
						Metformin 500 mg bid	182	53.4 (10.2)	93 (51.1)
						Metformin 1000 mg bid	182	53.2 (9.6)	100 (54.9)
						Placebo/Metformin 1000 mg bid	176	53.6 (10.0)	83 (47.2)
Arechavaleta *et al*[70,71]	Multi-national study	2009	74	30	Type 2 diabetes	Sitagliptin	516	56.3 (9.7)	232 (44.9)
						Glimepiride	519	56.2 (10.1)	240 (46.2)

NCT00086515 *et al*[72,73]	Multi-national study	2007	135	24	Type 2 diabetes	Sitagliptin 100 mg	464	54.4 (10.4)	205 (44.2)
						Placebo/Glipizide 5 mg	237	54.7 (9.7)	96 (40.5)
Bergenstal *et al*[74,75]	Multi-national study (62 sites in 3 countries)	2009	56	26	Type 2 diabetes	Exenatide once weekly	160	52.4 (10.41)	71 (44.4)
						Sitagliptin	166	52.2 (10.54)	80 (48.2)
						Pioglitazone	165	53.0 (9.92)	86 (52.1)
NCT00094757[76]	Multi-national study	2006	78	54	Type 2 diabetes	Sitagliptin 100 mg	205	54.5 (10.0)	95 (46.3)
						Sitagliptin 200 mg	206	55.4 (9.2)	102 (49.5)
						Placebo/Pioglitazone	110	55.5 (10.1)	41 (37.3)
NCT00094770[77]	Multi-national study (173 sites in 27 countries)	2006	139	104	Type 2 diabetes	Sitagliptin 100 mg	588	56.8 (9.3)	252 (42.8)
						Glipizide	584	56.6 (9.8)	226 (38.7)
NCT01137812[78,79]	Multi-national study (182 sites in 17 countries)	2012	87	52	Type 2 diabetes	Sitagliptin 100 mg	378	56.6 (9.33)	163 (43.1)
						Canagliflozin 300 mg	377	56.5 (9.62)	170 (45.1)
NCT00482729[80]	Multi-national study (209 sites in United States)	2008	74	44	Type 2 diabetes	Sitagliptin/ Meformin-Fixed Dose Combination	625	49.5 (10.5)	272 (43.5)
						Metformin	621	50.0 (10.5)	266 (42.8)
Bunck *et al*[81]	Multi-national study (3 sites in 3 countries)	2007	154	52	Type 2 diabetes	Exenatide	36	58.4 (1.4)	13 (36.1)
						Insulin glargine	33	58.3 (1.3)	11 (33.3)
Diamant *et al*[82]	Multi-national study (72 sites in 7 countries)	2009	53	26	Type 2 diabetes	Exenatide	233	58.0 (10.0)	113 (48.0)
						Insulin glargine	223	58 (9.0)	100 (45.0)
Inagaki *et al*[83]	Japan (22 sites)	2010	61	26	Type 2 diabetes	Exenatide once weekly	215	57.07 (10.44)	73 (34.0)
						Insulin glargine once daily	212	56.44 (11.16)	64 (30.2)
Russell-Jones *et al*[84]	Multi-national study (106 sites in 22 countries)	2010	82	26	Type 2 diabetes	Exenatide 2 mg once weekly + Oral placebo	248	53.7 (10.91)	109 (43.9)
						Sitagliptin 100 mg/d + SC placebo	163	52.3 (11.05)	69 (42.3)
						Metformin starting at 1000 mg/d + SC placebo	246	53.7 (11.08)	92 (37.4)
						Pioglitazone starting at 30 mg/d+ SC placebo	163	55.3 (10.96)	66 (40.5)
NCT01003184[85]	34 sites in Ireland and United Kingdom	2011	91	26	Type 2 diabetes	Exenatide once weekly	111	59.2 (9.86)	40 (36.04)
						Insulin Detemir twice daily	105	57.8 (9.48)	33 (31.4)
Astrup *et al*[86]	Multi-national study (19 sites in 8 European countries)	2009	117	104	Type 2 diabetes	Liraglutide 1.2 mg	95	47.18 (9.72)	73 (76.8)
						Liraglutide 1.8 mg	90	45.53 (10.9)	68 (75.6)
						Liraglutide 2.4 mg	93	45.01 (11.09)	71 (76.3)
						Liraglutide 3.0 mg	93	45.91 (10.71)	70 (75.3)
						Placebo	98	45.86 (10.28)	74 (75.5)
Garber *et al*[87]	126 sites in United States and 12 sites in Mexico	2007	91	52	Type 2 diabetes	Liraglutide 1.2 mg	251	53.7 (11.0)	134 (53.4)
						Liraglutide 1.8 mg	247	52.0 (10.8)	126 (51.0)
						Glimepiride 8 mg	248	53.4 (10.9)	115 (46.4)
Nauck *et al*[88]	Multi-national study (170 sites in 21 countries)	2007	52	26	Type 2 diabetes	Once daily Liraglutide (0.6 mg)	242	56.0 (11.0)	91 (37.6)
						Once daily Liraglutide (1.2 mg)	241	57 (9.0)	111 (46.1)
						Once daily Liraglutide (1.8 mg)	242	57 (9.0)	100 (41.3)
						Once daily Glimepiride (4 mg)	244	57 (9.0)	103 (42.2)
						Placebo	122	56 (9.0)	49 (40.2)
Marre *et al*[89]	Multi-national study (116 sites in 21 countries)	NR	NR	26	Type 2 diabetes	Liraglutide 0.6 mg	233	55.7 (9.9)	107 (46.0)
						Liraglutide 1.2 mg	228	57.7 (9.0)	125 (55.0)
						Liraglutide 1.8 mg	234	55.6 (10.0)	110 (47.0)
						Placebo	114	54.7 (10.0)	60 (53.0)
Zinman *et al*[90]	90 sites in United States and Canada	2007	65	26	Type 2 diabetes	Liraglutide 1.2 mg	178	55.0 (10.0)	77 (43.0)
						Liraglutide 1.8 mg	178	55.0 (11.0)	87 (49.0)
						Placebo	177	55.0 (10.0)	67 (38.0)

Raz *et al*[91]	Multi-national study (53 centers in 11 countries)	2011	134	24	Type 2 diabetes	Taspoglutide 10 mg	116	NR	NR
						Taspoglutide 20 mg	129	NR	NR
						Placebo	123	NR	NR
Rosenstock *et al*[92]	Multi-national study (118 sites in 4 countries)	2008	56	16	Type 2 diabetes	Albiglutide 4 mg weekly	35	50.4 (10.3)	20 (57.1)
						Albiglutide 15 mg weekly	35	55.5 (10.5)	17 (48.6)
						Albiglutide 30 mg weekly	31	54.2 (9.7)	23 (74.2)
						Albiglutide 15 mg biweekly	33	52.5 (9.6)	19 (57.6)
						Albiglutide 30 mg biweekly	32	55.5 (9.9)	16 (50.0)
						Albiglutide 50 mg biweekly	35	51.1 (10.3)	16 (45.7)
						Albiglutide 50 mg monthly	35	54.1 (11.3)	18 (51.4)
						Albiglutide 100 mg monthly	34	54.4 (9.9)	15 (44.1)
						Placebo	51	54.0 (10.6)	23 (45.1)
Seino *et al*[93]	Multi-national study (57 centers in 4 Asian countries)	NR	NR	24	Type 2 diabetes	Lixisenatide (10 ug for 1 wk, 15 mg for 1 wk, then 20 mg-maintenance dose)	154	58.7 (10.2)	85 (55.2)
						Placebo	157	58.0 (10.1)	77 (49.0)
Umpierrez *et al*[94]	36 sites in United States and 3 in Puerto Rico	2008	39	16	Type 2 diabetes	LY2189265 (LY 0.5/1.0)	66	59.0 (12.0)	31 (47.0)
						LY2189265 (LY 1.0/1.0)	65	57.0 (12.0)	30 (46.0)
						LY2189265 (LY 1.0/2.0)	65	54.0 (11.0)	31 (48.0)
						Placebo	66	56.0 (12.0)	37 (56.0)

Table 2 Quality assessment of glucagon-like peptide-1 based agents in randomized controlled trials included in analysis of pancreatic events

Ref.	Sequence generation	Blinding	Allocation concealment	Was Pancreatitis an AE or SAE?	Adverse event monitoring	Arms	Withdrawal rate (%)	Loss to follow- up (%)
Ross *et al*[21]	Central computer based; randomization: block in a 5:5:1 ratio	Double blind	Adequate	AE	Safety and tolerability end-points were the incidence of adverse events (including adverse changes observed during physical examinations or ECGs), protocol-specified significant AEs, hypoglycemia and changes from baseline in vital signs, clinical laboratory parameters and body weight	Linagliptin 2.5 mg bid	7.2	0
						Linagliptin 5 mg qd	4.5	0
						Placebo	2.3	0
Haak *et al*[22]	NR	Double blind	Adequate	AE	Incidence of AEs, serious AEs, discontinuation due to AEs,12-lead ECGs, vital signs and clinical laboratory parameters. The causal relationships between study medications and AEs were evaluated by the investigators at the site	Linagliptin 5 mg qd	14.8	2.1
						Metformin 500 mg bid	11.8	2.1
						Metformin 1000 mg bid	14.3	2.7
						Linagliptin 2.5 mg qd + Metformin 500 mg bid	11.2	2.8
						Linagliptin 2.5 mg qd + Metformin 1000 mg bid	7.7	0
						Placebo	25.0	1.4

NCT00328172[23]	NR	Double blind	NR	SAE	NR	Linagliptin 0.5 mg	24.1	1.7
						Linagliptin 2.5 mg	17.5	3.5
						Linagliptin 5.0 mg	23.6	1.8
						Metformin	7.7	1.5
						Placebo	32.8	1.5
Yki-Jarvinen *et al*[24,25]	NR	Double blind	NR	SAE	NR	Linagliptin 5.0 mg	13.9	2.2
						Placebo	17.5	1.3
NCT00654381[26]	NR	Double blind	NR	SAE	NR	Linagliptin 5.0 mg	1.89	0
						Linagliptin 10.0 mg	3.13	0
						Voglibose	2.5	0
						Placebo	7.5	0
NCT00622284[27]	NR	Double blind	NR	SAE	NR	Linagliptin	24.4	1.4
						Glimepiride	22.1	1.7
BI Trial No: 1218.15/ U09-2519-01[28]	Randomized into 1;2 ratio to receive either placebo or linagliptin	Double blind	Adequate	SAE	Incidence and intensity of AEs, withdrawals due to AEs, physical examination, 12-lead ECG, vital signs, clinical laboratory parameters	Linagliptin 5 mg + Pioglitazone 30 mg	5.8	NR
						Pioglitazone 30 mg + Placebo	14.6	NR
BI Trial No: 1218.52/ U11-1782-01[29]	NR	Double blind	NR	SAE	Safety endpoints were the incidence and intensity of AEs, withdrawals due to AEs, clinically relevant new or worsening findings in physical examination, 12-lead ECG, vital signs and clinical laboratory parameters	Linagliptin 2.5 mg + Metformin (500 mg and 1000 mg bid)	0.0	NR
						Metformin 1000 mg bid	0.6	NR
BI Trial No: 1218.63/ U11-1781-02[30]	NR	Double blind	NR	SAE	Incidence and intensity of AEs, withdrawals due to AEs, physical examination, 12-lead ECG, vital signs, clinical laboratory parameters	Linagliptin 5 mg	1.23	NR
						Placebo	1.26	NR
BI Trial No: 1218.75/ U12-3204-01[31]	NR	Double blind	NR	AE	Incidence and intensity of AEs, withdrawals due to AEs, clinically relevant changes from baseline in vital signs (blood pressure and pulse rate), clinically relevant new or worsening findings in 12-lead ECG as reported as AEs, clinically relevant changes from baseline in clinical laboratory assessments, cardiac and cerebrovascular events adjudicated CEC	Linagliptin 5 mg	12.3	NR
						Placebo	12.5	NR
BI Trial No: 1218.61/ U13-3124-01[32]	NR	Double blind	NR	AE	Incidence and intensity of AEs, primarily based on spontaneous AEs; withdrawal due to AEs; clinically relevant new or worsening findings in physical examination reported as AEs; changes from baseline in vital signs (BP and pulse); clinically relevant new or worsening findings in 12 lead ECG reported as AEs; changes from baseline in clinical lab assessments; and hypoglycemic events	Linagliptin 5 mg	2.2	NR
						Placebo	0.0	NR

BI Trial No: 1218.65/ U12-2143-01[33]	NR	Double blind	NR	SAE	Incidence and intensity of adverse events, withdrawals due to AEs, physical examination, ECGs, change from baseline in clinical lab parameters and cardiovascular events (Clinical Event Committee adjudication results)	Linagliptin 5 mg Placebo	0.98 3.0	NR NR
BI Trial No: 1218.64/ U13-1283-01[34]	NR	Double blind	NR	AE	Incidence and intensity of adverse events (AEs), withdrawals due to AEs, physical examination, vital signs, 12 lead ECG, change from baseline in clinical lab parameters	Linagliptin 5 mg Placebo (first 12 wk)/ Glimepiride (next 40 wk)	0.0 1.64	NR NR
BI Trial No: 1218.66/ U12-2076-01[35]	NR	Double blind	NR	SAE	Incidence and intensity of adverse events, withdrawals due to AEs, physical examination and vital signs, 12-lead ECG, clinical laboratory assessments	Linagliptin 5 mg Placebo	5.1 2.0	NR NR
Rosenstock *et al*[36]	Automated interactive voice response system using a randomization schedule	Double blind	NR	SAE	During the treatment period, patients were reviewed for adverse event evaluations. Further safety assessments included clinical examination of skin and digits. Hematology, serum chemistry, vital signs, physical exam and ECG parameters were done	Alogliptin 12.5 mg Alogliptin 25 mg Placebo	36.6 40.3 57.7	3.05 2.33 1.54
White *et al*[37]	NR	Double blind	NR	SAE	The principal secondary safety end point was the primary composite end point with the addition of urgent revascularization due to unstable angina within 24 h after hospital admission. Additional safety end points included angioedema, hypoglycemia, pancreatitis, cancer, and the results of laboratory testing	Alogliptin Placebo	NR NR	NR NR
NCT01318135[38]	NR	Open Label	Inadequate	SAE (Pancreatic cancer only)		Alogliptin 12.5 mg qd + Glimepiride 1-6 mg qd or bid Alogliptin 25 mg qd + Glimepiride 1-6 mg qd or bid	NR NR	NR NR
NCT01289119[39]	NR	Double blind	NR	SAE	TEAE were defined as any adverse events that started on or after the date of the first dose of double-blind study drug and within 14 d after the date of the last dose of double-blind study drug	Alogliptin monotherapy Metformin Metformin + Alogliptin Add-on Therapy Pioglitazone Pioglitazone + Alogliptin Add-on Therapy Placebo	9.78 9.18 6.06 7.94 6.56 9.78	3.26 0 0 0 1.64 0
NCT01263496[40]	NR	Open Label	Inadequate	SAE	A TEAE is defined as an adverse event with an onset that occurs after receiving study drug and within 30 d after receiving the last dose of study drug	Alogliptin 6.25 mg qd Alogliptin 12.5 mg qd Alogliptin 25 mg qd Alogliptin 50 mg qd Voglibose 0.2 mg tid	NR NR NR NR NR	NR NR NR NR NR

NCT00328627[41]	NR	Double blind	NR	SAE	NR	Alogliptin 12.5 mg + Placebo	24.2	1.56
						Alogliptin 25 mg + Placebo	21.7	1.55
						Placebo	45.7	3.1
NCT00395512[42]	NR	Double blind	Adequate	SAE	NR	Alogliptin 25 mg + Pioglitazone 30 mg	17.1	3.05
						Alogliptin 12.5 mg + Pioglitazone 30 mg	23.2	3.05
						Pioglitazone 30 mg	22.7	3.68
Kikuchi *et al*[43]	Dynamic randomization	Double blind	NR	SAE	Adverse events were recorded at each visit, and these AEs were assessed for severity and suspected relationship to the study drug. Hematology, biochemistry and urinalysis were performed at each scheduled visit. All laboratory assessments were processed at a central testing to ensure consistency	Vildagliptin 50 mg bid + glimepiride	2.9	NR
						Placebo + glimepiride	4	NR
Lukashevich *et al*[44]	NR	Double blind	NR	SAE	All treatment emergent AEs were recorded and assessed by the investigator as to severity and potential relationship to study drug. Particular attention was paid to hepatic, infections, skin, pancreatitis as well as edema and cardiovascular safety	Vildagliptin 50 mg qd (moderate RI)	10.3	2.4
						Placebo (moderate RI)	10.9	1.6
						Vildagliptin 50 mg qd (severe RI)	13.7	1.6
						Placebo (severe RI)	13.4	2.1
Strain *et al*[45]	Validated automated system	Double blind	Adequate	AE	All AEs and their severity, serious AEs, and their presumed relation with the study drug were monitored and recorded at each study visit	Vildagliptin	5.8	0.72
						Placebo	5.8	0
NCT00106340[46] (CLAF237A2308)	NR	Double blind	NR	SAE	Safety assessments included monitoring and recording all AEs, SAEs and pregnancies; regular monitoring of hematology, blood chemistry, and urine (performed at a central lab); and regular assessments of vital signs, ECG, physical condition and body weight. Severity and relationship to study drug were recorded for all AEs and SAEs	Vildagliptin 50 mg bid + Metformin	36.4	0
						Glimepiride up to 6 mg qd + Metformin	38.8	0
NCT00300287[47]	NR	Double blind	NR	SAE	Safety assessments included monitoring and recording all AEs, SAEs with their severity and presumed relationship to study drug and pregnancies, recording of hypoglycemic events, the regular monitoring of hematology, blood chemistry and urine, and regular assessments of vital signs, physical condition, body weight, and ECGs	Vildagliptin 50 mg qd	14.7	0.6
						Placebo	12.7	0.7

(CLAF237A2307) CLAF237A1301[48]	NR	Double blind	NR	AE (elevated pancreatic enzymes)	Safety assessments included monitoring and recording all AEs, SAEs with their severity and presumed relationship to study drug and pregnancies, recording of hypoglycemic events, the regular monitoring of hematology, blood chemistry and urine, and regular assessments of vital signs, physical condition, body weight, and ECGs	Vildagliptin	4.8	NR
						50 mg bid Voglibose 0.2 mg tid	5.2	NR
CLAF237A23119[49]	NR	Open Label	NA	SAE	Safety assessments included monitoring and recording all AEs, SAEs with their severity and presumed relationship to study drug and pregnancies, recording of hypoglycemic events, the regular monitoring of hematology, blood chemistry and urine, and regular assessments of vital signs, physical condition, body weight, and ECGs	Vildagliptin 100 mg + Metformin	10.4	2.5
						Thiazolinedione + Metformin	11.8	2.1
NCT00110240[50] (CLAF237A2323)	NR	Double Blind	NR	SAE	Safety assessments included adverse events, hypoglycemic events and serious adverse events, physical examination, vital signs, laboratory evaluations, and ECGs	Vildagliptin 50 mg bid	9.5	1.6
						Acarbose up to 100 mg tid	12.7	1.4
NCT00327015[51]	NR	Double Blind	NR	SAE	Safety and tolerability end-points included incidence of AEs, SAEs, discontinuation due to AEs, physical and ECG examinations, vital signs and results of clinical laboratory tests	Saxagliptin 5 mg + Metformin 500 mg	28.4	6.9
						Saxagliptin 10 mg + Metformin 500 mg	28.5	7.1
						Metformin 500 mg + Placebo	33.2	6.7
Hollander *et al*[52] (NCT00295633)	NR	Double Blind	NR	SAE	Safety assessments included incidence of AEs, SAEs and discontinuation due to AEs, changes from baseline lab parameters; changes from baseline vital signs; and incidence of marked clinical laboratory abnormalities	Saxagliptin 2.5 mg + TZD	31.8	NR
						Saxagliptin 5 mg + TZD	36	NR
						Placebo + TZD	41.3	NR
NCT00757588[53]	Interactive voice response system	Double Blind	NR	SAE	Safety end points included AEs, hypoglycemia and weight gain	Saxagliptin 5 mg + Insulin	11.8	0.98
						Placebo + Insulin	11.3	3.31
Scirica *et al*[54]	Central computerized telephone or web based system	Double Blind	NR	NR (Safety End Point)	A clinical events committee comprising specialists in cardiovascular and pancreatic medicine, all of whom were unaware of the study group assignments, adjudicated	Saxagliptin	NR	NR
						Placebo	NR	NR
Goke *et al*[55]	NR	Double Blind	NR	SAE	Safety and tolerability assessments included AEs and SAEs, lab measurements, vital signs, physical examination and ECG testing	Saxagliptin + Metformin	61.4	0.23
						Glipizide + Metformin	65.8	0.69
NCT00316082[56]	NR	Double Blind	NR	SAE	NR	Saxagliptin 2.5/5 mg QAM	38.0	9.9
						Saxagliptin 2.5 mg QAM	44.6	9.5
						Saxagliptin 5 mg QAM	29.7	8.1
						Saxagliptin 5 mg QPM	36.1	11.1
						Placebo	35.1	8.1

NCT00614939[57]	Interactive voice response system	Double Blind	NR	SAE	Safety and tolerability assessments included AEs, SAEs, treatment-related AEs, discontinuations of randomized study medication because of AEs, deaths, AEs of special interest and hypoglycemic events	Saxagliptin Placebo	71.8 80.0	NR NR
Chan *et al*[58,59]	Computer generated randomization schedule	Double Blind	Adequate	SAE	Assessment of safety and tolerability included evaluation of the data from physical examinations, vital signs and ECGs collected at specified study visits. All adverse experiences were rated by the investigators for intensity and relationship to study drug	Sitagliptin 50 mg or 25 mg once daily Placebo/ Glipizide	29.2 23.1	NR NR
Kojima *et al*[60]	Random allocation sequence performed centrally	Open label	NA	AE	NR	Sitagliptin Nateglinide	NR NR	NR NR
NCT00509262 (Arjona Ferreira JC *et al*[61,62])	Computer generated randomization schedule	Double	NR	SAE	Safety measurements included evaluation of AEs, physical exam and vital signs, and ECG. Lab safety studies included serum chemistry, hematology and urinalysis. All AEs were rated by the investigator for intensity and relationship to study drug	Sitagliptin Glipizide	210 212	
Henry RR *et al*[63,64]	NR	Blind Double blind	NR	SAE	Safety and tolerability were evaluated throughout the study by physical examination, monitoring of vital signs and safety lab measurements that included serum chemistry, hematology and urinalysis. AEs were monitored and evaluated by the investigators for intensity (severity), duration, outcome and relationship to study drug	Sitagliptin 100 mg/ Pioglitazone 15 mg Sitagliptin 100 mg/ Pioglitazone 30 mg Sitagliptin 100 mg/ Pioglitazone 45 mg Pioglitazone 15 mg Pioglitazone 30 mg Pioglitazone 45 mg	20.9 22.9 22.2 31.3 27.9 27.4	3.5 6.9 5.7 6.1 9 5.7
Raz I *et al*[65,66]	Computer generated schedule	Double blind	NR	SAE	Safety and tolerability were evaluated by physical examination, vital signs and lab measurements that included routine serum chemistry, hematology, urinalysis and pregnancy testing. AEs were monitored through the study for intensity, duration, outcome, relationship to study drug and level of severity	Sitagliptin 100 mg Placebo	17.7 14.9	3.13 3.19
NCT01131182[67]	NR	Open label	NA	SAE	NR	Sitagliptin Sulfonylurea	NR NR	NR NR

Goldstein *et al*[68,69]	NR	Double blind	NR	SAE	Data were collected regarding AEs, physical exam, vital signs, ECGs and body weight throughout the study. All AEs were rated by investigators for intensity and relationship to study drug	Sitagliptin 50 mg bid + Metformin 500 mg bid	22.1	2.6
						Sitagliptin 50 mg bid + Metformin 1000 mg bid	22.5	5.5
						Sitagliptin 50 mg bid + Metformin 1000 mg bid (OLC)	32.5	2.6
						Metformin 500 mg bid	30.8	2.2
						Metformin 1000 mg bid	25.8	3.8
						Placebo/ Metformin 1000 mg bid	34.7	5.1
Arechavaleta *et al*[70,71]	Concealed computer-generated allocation schedule	Double blind	Adequate	SAE	Safety and tolerability were assessed by a review of all safety parameters including adverse experiences, laboratory safety parameters, body weight and vital signs	Sitagliptin	9.3	1.7
						Glimepiride	9.8	1.7
NCT00086515 *et al*[72,73]	NR	Double blind	NR	SAE	Safety and tolerability were assessed throughout the study. Monitoring for adverse experiences, physical examinations, vital signs, body weight, 12-lead ECGs (read at a central reading laboratory), and safety laboratory measurements comprising routine hematology, serum chemistry, and urinalysis were performed	Sitagliptin 100 mg	10.6	0.86
						Placebo/ Glipizide 5 mg	18.9	2.11
Bergenstal *et al*[74,75]	Interactive voice response system	Double blind	Adequate	SAE	NR	Exenatide once weekly	26.9	5
						Sitagliptin	16.9	5.4
						Pioglitazone	24.8	7.8
NCT00094757[76]	NR	Double blind	NR	SAE	Data for adverse experiences, physical examinations, vital signs, ECGs, and body weight were collected throughout the study	Sitagliptin 100 mg	25.8	1.5
						Sitagliptin 200 mg	30.1	2.4
						Placebo/ Pioglitazone	27.3	5.4
NCT00094770[77]	NR	Double blind	NR	SAE	Data on adverse experiences, physical examinations, vital signs, ECGs and body weight were collected throughout the study. All adverse experiences were rated by the study site investigators for intensity and relationship to study drug. Laboratory safety evaluations included blood chemistry, haematology and urinalysis	Sitagliptin 100 mg	34.4	3.2
						Glipizide	29.5	1.7
NCT01137812[78,79]	Interactive Voice Response System/ Interactive Web Response System	Double blind	Adequate	SAE	Safety evaluations included AEs, clinical laboratory tests, vital sign measurements, physical examinations, self-monitored blood glucose, 12-lead electrocardiograms, and documentation of hypoglycemic episodes	Sitagliptin 100 mg	44.4	2.1
						Canagliflozin 300 mg	32.6	1.6

NCT00482729[80]	NR	Double blind	NR	SAE	NR	Sitagliptin/ Meformin-Fixed Dose Combination	34.7 (217/626)	13.7 (86/626)
						Metformin	34.9 (218/624)	10.6 (66/624)
Bunck *et al*[81]	NR	Open label	NA	SAE	NR	Exenatide	16.7	0
						Insulin glargine	9.1	3.03
Diamant *et al*[82]	Computer generated randomization sequence	Open label	NA	SAE	Safety endpoints were adverse events, clinical lab assessments, vital signs, and hypoglycemia. We defined adverse events as those occurring at or after randomization or worsening during the study	Exenatide	10.3	0.86
						Insulin glargine	6.3	0.45
Inagaki *et al*[83]	Computer generated randomization sequence	Open label	NA	AE	Safety profile end points included AEs and hypoglycemia	Exenatide once weekly	10.2	0.47
						Insulin glargine once daily	5.2	0
Russell-Jones *et al*[84]	Computer generated randomization sequence	Double blind	Adequate	SAE	Safety end points were adverse events, clinical lab assessments, vital signs, hypoglycemia and antibodies to exenatide. Treatment emergent adverse events were defined as those occurring or worsening after the first dose of study drug	Exenatide 2 mg once weekly + Oral placebo	15.3	1.6
						Sitagliptin 100 mg/d + SC placebo	14.1	2.4
						Metformin starting at 1000 mg/d + SC placebo	13.4	0.4
						Pioglitazone starting at 30 mg/d + SC placebo	1.8	1.8
NCT01003184[85]	NR	Open label	NR	SAE	NR	Exenatide once weekly	17.1	0.9
						Insulin Detemir twice daily	11.4	0
Astrup *et al*[86]	NR	Double blind (first 20 wk) Weeks 20-104: Open label	NR	SAE	Safety assessments included adverse events, recorded at every visit, standard lab tests and serum liraglutide antibodies. A safety committee for data surveillance was established	Liraglutide 1.2 mg	10.5	0
						Liraglutide 1.8 mg	17.8	0
						Liraglutide 2.4 mg	21.5	0
						Liraglutide 3.0 mg	11.8	0
						Placebo	19.4	0
Garber *et al*[87]	Telephone based or web-based systems	Double blind	Adequate	SAE	Key safety assessments were tolerability (including nausea and other gastrointestinal adverse events), serum calcitonin and hypoglycemic episodes	Liraglutide 1.2 mg	35.5	NR
						Liraglutide 1.8 mg	29.7	NR
						Glimepiride 8 mg	38.7	NR
Nauck *et al*[88]	Telephone based or web-based randomization systems	Double blind	Adequate	SAE	Safety variables included adverse events, vital signs, ECG, biochemical and hematology measures and subject reported hypoglycemic episodes	Once daily Liraglutide (0.6 mg)	14.0	0
						Once daily Liraglutide (1.2 mg)	18.0	0.4
						Once daily Liraglutide (1.8 mg)	21.0	0
						Once daily Glimepiride (4 mg)	14.0	0
						Placebo	39.0	0

Marre *et al*[89]	NR	Double blind	NR	SAE	Safety variables included hypoglycemic episodes, liraglutide antibodies, tolerability (gastrointestinal complaints) and pulse. AEs, vital signs, ECG, biochemical and hematology measures including calcitonin were also monitored	Liraglutide 0.6 mg	10.7	NR
						Liraglutide 1.2 mg	14.0	NR
						Liraglutide 1.8 mg	8.9	NR
						Placebo	27.2	NR
Zinman *et al*[90]	Telephone based or web-based randomization systems	Double blind	Adequate	SAE	Safety variables included AEs, vital signs, ECG, biochemical and hematology measures and subject reported hypoglycemic episodes	Liraglutide 1.2 mg	14.0	NR
						Liraglutide 1.8 mg	25.0	NR
						Placebo	32.0	NR
Raz *et al*[91]	NR	Double blind	NR	SAE	Safety assessments included AEs, vital signs, physical examinations, clinical lab tests, ECG and hypoglycemia	Taspoglutide 10 mg	11.2	NR
						Taspoglutide 20 mg	13.2	NR
						Placebo	3.3	NR
Rosenstock *et al*[92]	NR	Double blind	NR	SAE	Adverse event assessments and safety analyses were conducted throughout the study	Albiglutide 4 mg weekly	48.6	5.7
						Albiglutide 15 mg weekly	31.4	8.6
						Albiglutide 30 mg weekly	32.3	3.2
						Albiglutide 15 mg biweekly	45.5	9.1
						Albiglutide 30 mg biweekly	24.2	0
						Albiglutide 50 mg biweekly	42.9	2.9
						Albiglutide 50 mg monthly	14.3	2.9
						Albiglutide 100 mg monthly	44.1	2.9
						Placebo	23.5	0
Seino *et al*[93]	Interactive voice response system	Double blind	Adequate	SAE	Safety and tolerability included reported AEs and other safety information such as symptomatic hypoglycemia	Lixisenatide (10 ug for 1 wk, 15 ug for 1 wk, then 20 ug-maintenance dose)	NR	NR
						Placebo	NR	NR
Umpierrez *et al*[94]	Computer generated random sequence	Double blind	Adequate	SAE	Safety measures included AEs, vital signs, hypoglycemia events and lab tests	LY2189265 (LY 0.5/1.0)	12.1	1.5
						LY2189265 (LY 1.0/1.0)	10.8	1.5
						LY2189265 (LY 1.0/2.0)	13.8	1.5
						Placebo	9.1	1.5

NR: Not reported; AE: Adverse event; SAE: Serious adverse event; ECGs: Electrocardiograms; TEAE: Treatment-emergent adverse events; CEC: Clinical events committee.

events in patients on GLP-1 agents and 56 pancreatitis events occurred in the control patients. There was no statistically significant difference in the risk of pancreatic adverse event associated with GLP-1 agent compared with controls (Peto OR = 1.07, 95%CI: 0.75-1.52, P = 0.72, I^2 = 0%) (Figure 3).

Elevated pancreatic enzymes: Eight studies reported on elevated pancreatic enzymes. There was a statistically significant increased risk of elevation of pancreatic enzymes associated with GLP-1 agents compared with control (Peto OR = 3.15, 95%CI: 1.56-6.39, P = 0.001, I^2 = 0%) (Figure 4).

Pancreatic cancer: Eighteen studies reported on pancreatic cancer (Table 3). There were a total of 35 cases of pancreatic cancer reported from studies that used GLP-1 agents. Seventeen cases of pancreatic occurred among 18259 patients taking GLP-1 agents compared to 18 cases among 15785 controls. Of these cases, 2 used linagliptin, 2 used alogliptin, 1 used vildagliptin, 7 used saxagliptin while 5 used sitagliptin. The remaining 18 cases occurred among controls.

Individual GLP-1 agents

DPP-4 inhibitors: (1) Linagliptin: Fifteen studies that used Linagliptin had a total of 7263 patients. There

A

Study or subgroup	GLP1 agent Events	Total	Control Events	Total	Weight	Peto odds ratio Peto, Fixed, 95%CI
Active control						
1218.52/U11-1782-01	0	396	0	170		Not estimable
1218.64/U13-1283-01	0	113	0	122		Not estimable
Bunk 2009	1	36	0	33	0.6%	6.80 [0.13, 343.88]
Chan 2008	1	65	0	26	0.5%	4.06 [0.05, 310.62]
CLAF237A23119	2	1756	2	871	2.3%	0.47 [0.06, 3.75]
Diamant 2010	1	233	0	223	0.6%	7.08 [0.14, 357.08]
Garber 2009	2	498	0	248	1.1%	4.48 [0.24, 85.11]
Goke 2013	0	428	2	430	1.3%	0.14 [0.01, 2.17]
Haak LM v M 2012	0	286	0	291		Not estimable
Inagaki 2012	0	215	0	212		Not estimable
Kojima 2013	0	20	0	16		Not estimable
Nauck LvG 2009	1	725	1	244	1.0%	0.27 [0.01, 6.53]
NCT00086515	1	464	0	237	0.6%	4.53 [0.07, 285.42]
NCT00094757	0	411	2	110	0.9%	0.01 [0.00, 0.26]
NCT00094770	1	588	0	584	0.6%	7.34 [0.15, 369.87]
NCT00103857	1	489	2	540	1.9%	0.57 [0.06, 5.46]
NCT00106340	3	1553	2	1546	3.2%	1.49 [0.26, 8.59]
NCT00110240	1	440	0	220	0.6%	4.48 [0.07, 286.49]
NCT00327015	1	643	1	328	1.2%	0.48 [0.03, 9.08]
NCT00328627	1	257	0	129	0.6%	4.49 [0.07, 286.28]
NCT00395512	1	328	0	163	0.6%	4.47 [0.07, 286.84]
NCT00482729	1	625	2	621	1.9%	0.51 [0.05, 4.91]
NCT00509262	1	210	0	212	0.6%	7.46 [0.15, 375.96]
NCT00622284	2	776	2	775	2.6%	1.00 [0.14, 7.10]
NCT00637273 E v P	0	160	1	82	0.6%	0.05 [0.00, 3.29]
NCT00637273 SvP	0	166	1	83	0.6%	0.05 [0.00, 3.18]
NCT00701090	1	516	0	518	0.6%	7.42 [0.15, 373.83]
NCT00722371	1	691	1	693	1.3%	1.00 [0.06, 16.05]
NCT01131182	0	507	1	514	0.6%	0.14 [0.00, 6.91]
NCT01137812	0	378	1	377	0.6%	0.13 [0.00, 6.80]
NCT01263496	2	391	0	83	0.7%	3.37 [0.09, 129.71]
NCT01289119 A + M	0	99	0	98		Not estimable
NCT01289119 A + P	0	61	0	63		Not estimable
NCT01318135	1	142	0	145	0.6%	7.55 [0.15, 380.42]
Russell-Jones E v P 2012	0	242	0	254		Not estimable
Russell-Jones S v P 2012	1	163	0	255	0.6%	12.99 [0.23, 722.62]
Subtotal (95%CI)		15071		11516	29.1%	0.96 [0.53, 1.70]
Total events	28		21			

Heterogeneity: χ^2 = 28.87, df = 27 (P = 0.37); I^2 = 6%

Test for overall effect: Z = 0.18 (P = 0.86)

B

Study or subgroup	GLP1 agent Events	Total	Control Events	Total	Weight	Peto odds ratio Peto, Fixed, 95%CI
Placebo control						
1218.15/U09-2519-01	0	259	1	130	0.6%	0.05 [0.00, 3.20]
1218.61/U13-3124-01	0	183	0	89		Not estimable
1218.63/U11-1781-02	0	162	0	79		Not estimable
1218.65/U12-2143-01	0	205	0	100		Not estimable
1218.66/U12-2076-01	0	200	0	99		Not estimable
1218.75/U12-3204-01	0	106	0	120		Not estimable
Astrup 2012	1	371	0	98	0.4%	3.54 [0.03, 439.21]
Haak LvP 2012	0	142	0	72		Not estimable
Hollander 2011	1	381	0	184	0.6%	4.41 [0.07, 288.68]
Kikuchi 2010	0	102	0	100		Not estimable
Lukashevich 2011	0	289	0	226		Not estimable
Marre 2009	1	695	0	114	0.3%	3.20 [0.01, 895.32]
Nauck LvP 2009	1	725	0	122	0.3%	3.22 [0.01, 854.21]
NCT00300287	0	156	1	150	0.6%	0.13 [0.00, 6.56]
NCT00316082	1	291	0	74	0.4%	3.51 [0.03, 459.13]
NCT00328172	1	170	0	67	0.5%	4.03 [0.05, 313.12]
NCT00328627	1	257	0	129	0.6%	4.49 [0.07, 286.28]
NCT00337610	0	96	1	94	0.6%	0.13 [0.00, 6.68]
NCT00614939	0	85	1	85	0.6%	0.14 [0.00, 6.82]
NCT00654381	1	319	0	80	0.4%	349 [0.03, 466.89]
NCT00757588	1	304	1	151	1.1%	0.47 [0.02, 8.90]
NCT00954447	2	631	1	630	1.9%	1.95 [0.20, 18.75]
NCT01289119 A alone	0	92	1	93	0.6%	0.14 [0.00, 6.89]
Raz 2012	0	245	0	123		Not estimable
Rosenstock Albi 2009	0	270	0	51		Not estimable
Rosenstock Aloglip 2009	2	260	0	130	1.1%	4.50 [0.24, 85.42]
Ross 2012	0	447	0	44		Not estimable
Scirica 2013	29	8280	34	8212	40.4%	0.85 [0.52, 1.39]
Seino 2012	0	154	0	157		Not estimable
Strain 2013	0	139	0	139		Not estimable
Umpierrez 2011	2	196	0	66	1.0%	3.83 [0.16, 93.74]
White 2013	17	2701	12	2679	18.6%	1.40 [0.68, 2.91]
Zinman 2009	0	356	0	177		Not estimable
Subtotal (95%CI)		19269		14864	70.9%	1.01 [0.70, 1.47]
Total events	61		53			
Heterogeneity: χ^2 = 12.05, df = 18 (P = 0.84); I^2 = 0%						
Test for overall effect: Z = 0.06 (P = 0.95)						
Total (95%CI)		34340		26380	100.0%	0.99 [0.72, 1.36]
Total events	89		74			
Heterogeneity: χ^2 = 40.95, df = 46 (P = 0.68); I^2 = 0%						
Test for overall effect: Z = 0.05 (P = 0.96)						
Test for subqroup differences: χ^2 = 0.03, df = 1 (P = 0.85); I^2 = 0%						

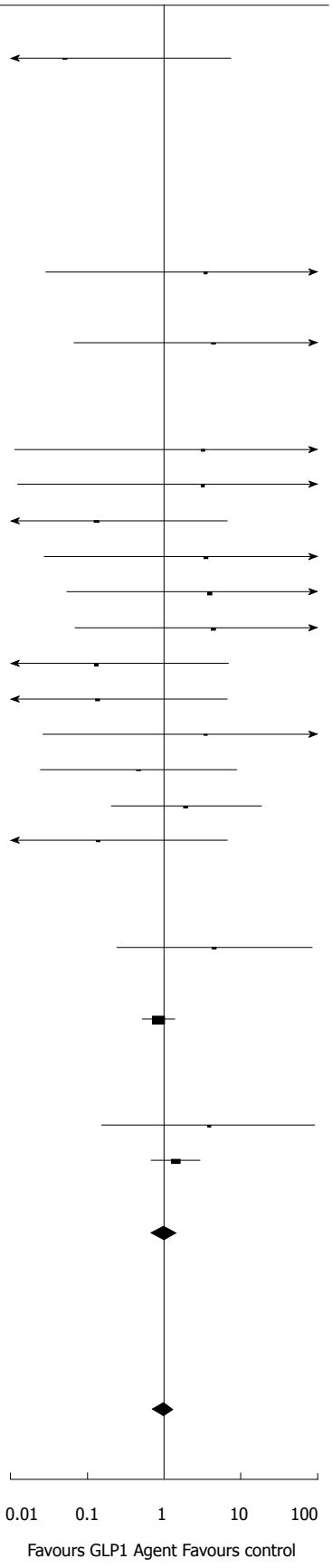


Figure 2 Risk of pancreatic adverse events in patients treated with glucagon-like peptide-1 based therapies (A and B).

A

Study or subgroup	GLP1 agent Events	Total	Control Events	Total	Weight	Peto odds ratio Peto, Fixed, 95%CI
Active control						
1218.52/U11-1782-01	0	396	0	170		Not estimable
1218.64/U13-1283-01	0	113	0	122		Not estimable
Bunk 2009	1	36	0	33	0.8%	6.80 [0.13, 343.88]
Chan 2008	0	65	0	26		Not estimable
CLAF237A23119	1	1756	2	871	2.2%	0.22 [0.02, 2.44]
Diamant 2010	1	233	0	223	0.8%	7.08 [0.14, 357.08]
Garber 2009	2	498	0	248	1.5%	4.48 [0.24, 85.11]
Goke 2013	0	428	2	430	1.6%	0.14 [0.01, 2.17]
Haak LM v M 2012	0	286	0	291		Not estimable
Inagaki 2012	0	215	0	212		Not estimable
Kojima 2013	0	20	0	16		Not estimable
Nauck LvG 2009	1	725	1	244	1.2%	0.27 [0.01, 6.53]
NCT00086515	0	464	0	237		Not estimable
NCT00094757	0	411	1	110	0.5%	0.01 [0.00, 1.07]
NCT00094770	1	588	0	584	0.8%	7.34 [0.15, 369.87]
NCT00103857	1	489	1	540	1.6%	1.10 [0.07, 17.74]
NCT00106340	3	1553	2	1546	4.1%	1.49 [0.26, 8.59]
NCT00110240	1	440	0	220	0.7%	4.48 [0.07, 286.49]
NCT00327015	1	643	1	328	1.5%	0.48 [0.03, 9.08]
NCT00395512	1	328	0	163	0.7%	4.47 [0.07, 286.84]
NCT00482729	1	625	2	621	2.5%	0.51 [0.05, 4.91]
NCT00509262	0	210	0	212		Not estimable
NCT00622284	1	776	0	775	0.8%	7.38 [0.15, 371.91]
NCT00637273 E v P	0	160	1	82	0.7%	0.05 [0.00, 3.29]
NCT00637273 SvP	0	166	1	83	0.7%	0.05 [0.00, 3.18]
NCT00701090	0	516	0	518		Not estimable
NCT00722371	0	691	1	693	0.8%	0.14 [0.00, 6.84]
NCT01131182	0	507	1	514	0.8%	0.14 [0.00, 6.91]
NCT01137812	0	378	1	377	0.8%	0.13 [0.00, 6.80]
NCT01263496	1	391	0	83	0.5%	3.36 [0.02, 583.65]
NCT01289119 A + M	0	99	0	98		Not estimable
NCT01289119 A + P	0	61	0	63		Not estimable
NCT01318135	0	142	0	145	0.6%	Not estimable
Russell-Jones E v P 2012	0	242	0	254		Not estimable
Russell-Jones S v P 2012	1	163	0	255	0.8%	12.99 [0.23, 722.62]
Subtotal (95%CI)		14814		11387	26.7%	0.75 [0.38, 1.50]
Total events	18		17			

Heterogeneity: χ^2 = 22.59 df =21 (P = 0.37); I^2 = 7%

Test for overall effect: Z = 0.81 (P = 0.42)

B

Study or subgroup	GLP1 agent Events	GLP1 agent Total	Control Events	Control Total	Weight	Peto odds ratio Peto, Fixed, 95%CI
Placebo control						
1218.15/U09-2519-01	0	259	0	130		Not estimable
1218.61/U13-3124-01	0	183	0	89		Not estimable
1218.63/U11-1781-02	0	162	0	79		Not estimable
1218.65/U12-2143-01	0	205	0	100		Not estimable
1218.66/U12-2076-01	0	200	0	99		Not estimable
1218.75/U12-3204-01	0	106	0	120		Not estimable
Astrup 2012	1	371	0	98	0.5%	3.54 [0.03, 439.21]
Haak LvP 2012	0	142	0	72		Not estimable
Hollander 2011	1	381	0	184	0.7%	4.41 [0.07, 288.68]
Kikuchi 2010	0	102	0	100		Not estimable
Lukashevich 2011	0	289	0	226		Not estimable
Marre 2009	1	695	0	114	0.4%	3.20 [0.01, 895.32]
Nauck LvP 2009	1	725	0	122	0.4%	3.22 [0.01, 854.21]
NCT00300287	0	156	1	150	0.8%	0.13 [0.00, 6.56]
NCT00316082	0	291	0	74		Not estimable
NCT00328172	1	170	0	67	0.7%	4.03 [0.05, 313.12]
NCT00328627	1	257	0	129	0.7%	4.49 [0.07, 286.28]
NCT00337610	0	96	0	94		Not estimable
NCT00614939	0	85	1	85	0.8%	0.14 [0.00, 6.82]
NCT00654381	0	319	0	80		Not estimable
NCT00757588	0	304	1	151	0.7%	0.05 [0.00, 3.16]
NCT00954447	2	631	1	630	2.5%	1.95 [0.20, 18.75]
NCT01289119 A alone	0	92	1	93	0.8%	0.14 [0.00, 6.89]
Raz 2012	0	245	0	123		Not estimable
Rosenstock Aloglip 2009	2	260	0	130	1.5%	4.50 [0.24, 85.42]
Ross 2012	0	447	0	44		Not estimable
Scirica 2013	24	8280	22	8212	37.7%	1.08 [0.61, 1.93]
Seino 2012	0	154	0	157		Not estimable
Strain 2013	0	139	0	139		Not estimable
Umpierrez 2011	2	196	0	66	1.2%	3.83 [0.16, 93.74]
White 2013	17	2701	12	2679	23.7%	1.40 [0.68, 2.91]
Zinman 2009	0	356	0	177		Not estimable
Subtotal (95%CI)		18999		14813	73.3%	1.21 [0.80, 1.84]
Total events	53		39			
Heterogeneity: χ^2 = 9.11, df = 14 (P = 0.82); I^2 = 0%						
Test for overall effect: Z = 0.06 (P = 0.36)						
Total (95%CI)		33813		26200	100.0%	0.99 [0.72, 1.36]
Total events	71		56			
Heterogeneity: χ^2 = 33.06, df = 36 (P = 0.61); I^2 = 0%						
Test for overall effect: Z = 0.05 (P = 0.72)						
Test for subqroup differences: χ^2 = 1.36, df = 1 (P = 0.24); I^2 = 26.3%						

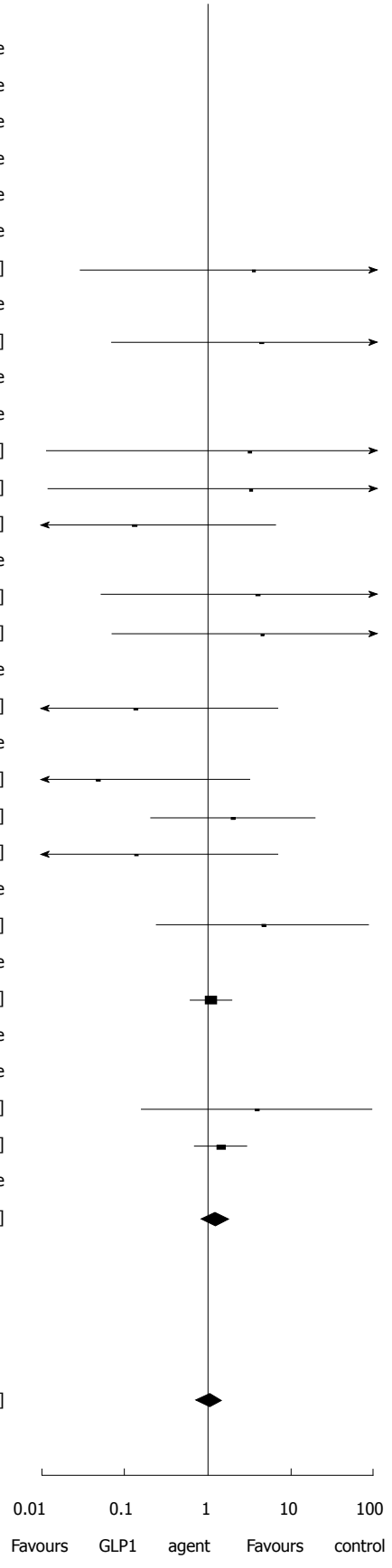


Figure 3 Risk of pancreatitis in patients treated with glucagon-like peptide-1 based therapies (A and B).

Table 3 Pancreatic cancer events in randomized controlled trials of glucagon-like peptide-1 agents

Ref.	Duration of GLP-1 exposure (wk)	Arms	No. of participants	No. of cases
NCT00654381[26]	52	Linagliptin 5 mg	159	0
		Linagliptin 10 mg	160	1
		Voglibose	162	0
		Placebo	80	0
NCT00622284[27]	104	Linagliptin	776	1
		Glimepiride	775	2
BI Trial No: 1218.15/ U09-2519-01[28]	24	Linagliptin 5 mg + Pioglitazone 30 mg	130	0
		Pioglitazone 30 mg + Placebo	259	1
White *et al*[37]	76	Alogliptin	2701	0
		Placebo	2679	0
NCT01318135[38]	52	Alogliptin 12.5 mg qd + Metformin 500 mg bid or 750 mg tid	142	1
		Metformin 500 mg bid or 750 mg tid	145	0
NCT01263496[40]	52	Alogliptin 6.25 mg qd	96	0
		Alogliptin 12.5 mg qd	101	0
		Alogliptin 25 mg qd	97	1
		Alogliptin 50 mg qd	97	0
		Voglibose 0.2 mg tid	83	0
CLAF237A23119[49]	12	Vildagliptin 100 mg + Metformin	1756	1
		Thiazolinedione + Metformin	871	NR
NCT00757588[53]	52	Saxagliptin 5 mg + Insulin	304	1
		Placebo + Insulin	151	0
Scirica *et al*[54]	109	Saxagliptin	8280	5
		Placebo	8212	12
NCT00316082[56]	24	Saxagliptin 2.5/5 mg QAM	71	1
		Saxagliptin 2.5 mg QAM	74	0
		Saxagliptin 5 mg QAM	74	0
		Saxagliptin 5 mg QPM	72	0
		Placebo	74	0
Chan *et al*[58,59]	54	Sitagliptin 50 mg or 25 mg once daily	65	1
		Placebo/Glipizide	26	0
Ferreira *et al*[61,62]	54	Sitagliptin	210	1
		Glipizide	212	0
Henry *et al*[63,64]	54	Pioglitazone 15 mg	230	0
		Pioglitazone 30 mg	233	0
		Pioglitazone 45 mg	230	0
		Sitagliptin 100 mg/Pioglitazone 15 mg	230	0
		Sitagliptin 100 mg/Pioglitazone 30 mg	231	1
		Sitagliptin 100 mg/Pioglitazone 45 mg	230	0
Raz *et al*[65,66]	30	Sitagliptin 100 mg	96	0
		Placebo	94	1
Goldstein *et al*[68,69]	104	Metformin 500 mg bid	182	0
		Metformin 1000 mg bid	182	0
		Sitagliptin 50 mg bid + Metformin 500 mg bid	190	0
		Sitagliptin 50 mg bid + Metformin 1000 mg bid	182	0
		Sitagliptin 50 mg bid + Metformin 1000 mg bid	117	0
		Placebo/Metformin 1000 mg bid	176	1
Arechavaleta *et al*[70,71]	30	Sitagliptin	516	1
		Glimepiride	518	0
Charbonnel *et al*[72,73]	104	Sitagliptin 100 mg	464	1
		Placebo/Glipizide 5 mg	237	0
NCT00094757[76]	54	Sitagliptin 100 mg	205	0
		Sitagliptin 200 mg	206	0
		Placebo/Pioglitazone	110	1

GLP-1: Glucagon-like peptide-1.

was no statistically significant difference in the risk of pancreatic adverse event (Peto OR = 1.14, 95%CI: 0.32-4.13) or pancreatitis (Peto OR = 2.90, 95%CI: 0.49-17.36) associated with linagliptin compared with controls; (2) Alogliptin: Nine studies that used Alogliptin had a total of 7914 patients. In comparison with control, there was no increased risk of having a pancreatic adverse event (Peto OR = 1.59, 95%CI: 0.82-3.07) or pancreatitis (Peto OR = 1.50, 95%CI: 0.77-2.94) with alogiptin; (3) Vildagliptin: Seven studies that used Vildagliptin had a total of 7687 patients. In comparison with control, there was no statistically significant difference in the risk of pancreatic adverse event (Peto OR = 0.87, 95%CI: 0.26-2.94) or pancreatitis (Peto OR = 0.75, 95%CI: 0.21-2.67) with vildagliptin; (4) Saxagliptin: Seven studies that used Saxagliptin had

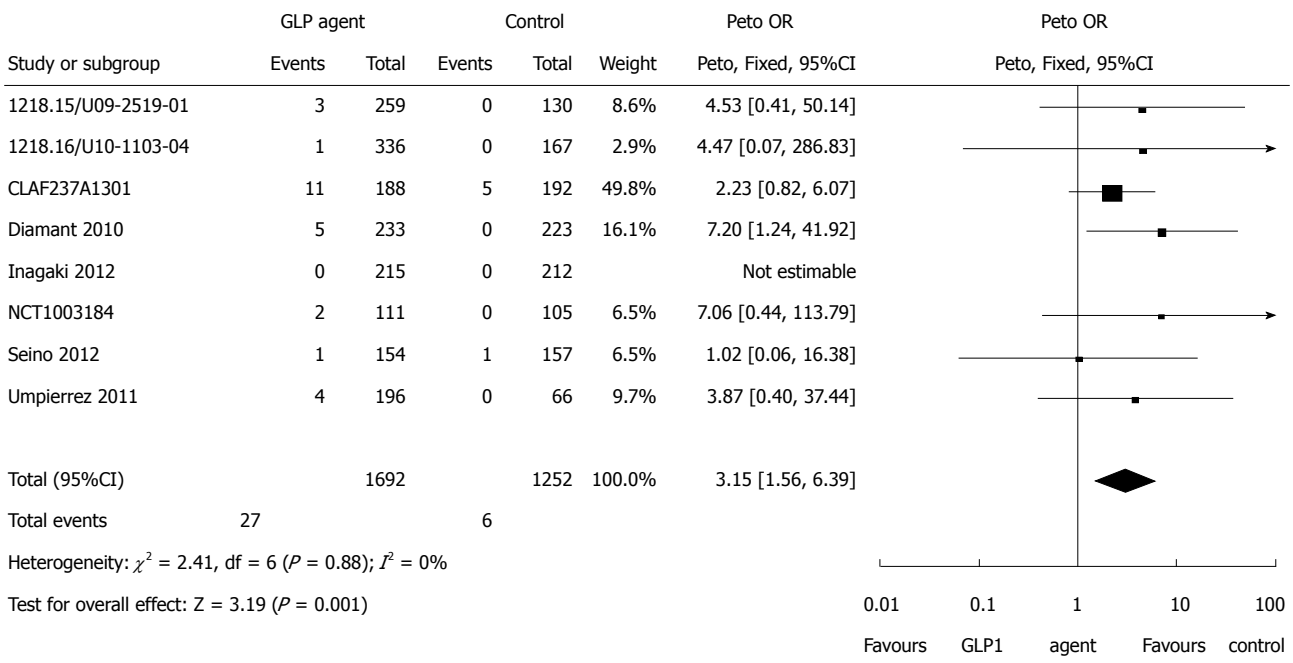


Figure 4 Risk of elevated pancreatic enzymes for glucagon-like peptide-1 based agents.

a total of 19876 patients. In comparison with control, there was no statistically significant difference in the risk of pancreatic adverse event (Peto OR = 0.79, 95%CI: 0.49-1.25) or pancreatitis (Peto OR 0.91, 95%CI: 0.53-1.56) with saxaglipitin; and (5) sitagliptin: Sixteen studies that used Sitagliptin had a total of 10360 patients. In comparison with control, there was no statistically significant difference in the risk of pancreatic adverse event (Peto OR = 0.66, 95%CI: 0.27-1.63) or pancreatitis (Peto OR = 0.45, 95%CI: 0.14-1.43) with sitagliptin.

GLP-1 receptor agonists

Exenatide: Five studies that used Exenatide had a total of 1690 patients. In comparison with control, there was no statistically significant difference in the risk of pancreatic adverse event (Peto OR = 1.53, 95%CI: 0.15-15.29) or pancreatitis (Peto OR = 1.53, 95%CI: 0.15-15.29) with exenatide.

Liraglutide: Six studies that used Liraglutide had a total of 4373 patients. In comparison with control, there was no statistically significant difference in the risk of pancreatic adverse event (Peto OR = 1.71, 95%CI: 0.29-10.04) or pancreatitis (Peto OR = 1.71, 95%CI: 0.29-10.04 with liraglutide.

Dulaglutide: One study that used Dulaglutide had 262 patients. In comparison with control, there was no statistically significant difference in the risk of pancreatic adverse event (Peto OR = 3.83, 95%CI: 0.16-93.74) or pancreatitis (Peto OR = 3.83, 95%CI: 0.16-93.74) with dulaglutide.

Taspoglutide, albiglutide and lixisenatide: Taspoglutide, Albiglutide and Lixisenatide all had 1 study each with 368, 321 and 311 patients each. The effect estimates were not estimable due to the small number of events.

In a post-hoc analysis, we examined whether there was any difference between DPP-4 inhibitors and GLP-1 based therapies. The results showed that neither the DPP-4 inhibitors nor the GLP- 1 based therapies were associated with a risk of pancreatic complications (Figure 5).

Publication bias

We did not detect any publication bias in the funnel plot (Figure 6).

DISCUSSION

Summary of results

Our study showed a significantly increased risk of pancreatic enzyme elevation with GLP-1 based therapies**.** However, the use of GLP-1 based therapies was not associated with a statistically significant increased risk of pancreatic complication events in patients with type 2 diabetes in randomized controlled trials. Additionally, when we examined individual agents, none of the DPP-4 inhibitors or GLP-1 agonists was associated with a statistically significant increased risk of pancreatitis (Figure 3). Despite the lack of statistical significance the upper bounds of the CI in several analyses, particularly for the GLP-1 receptor agonists (exenatide, liraglutide and albiglutide) exceeded 1 and could not rule out a clinically significant hazard. There were an insufficient number

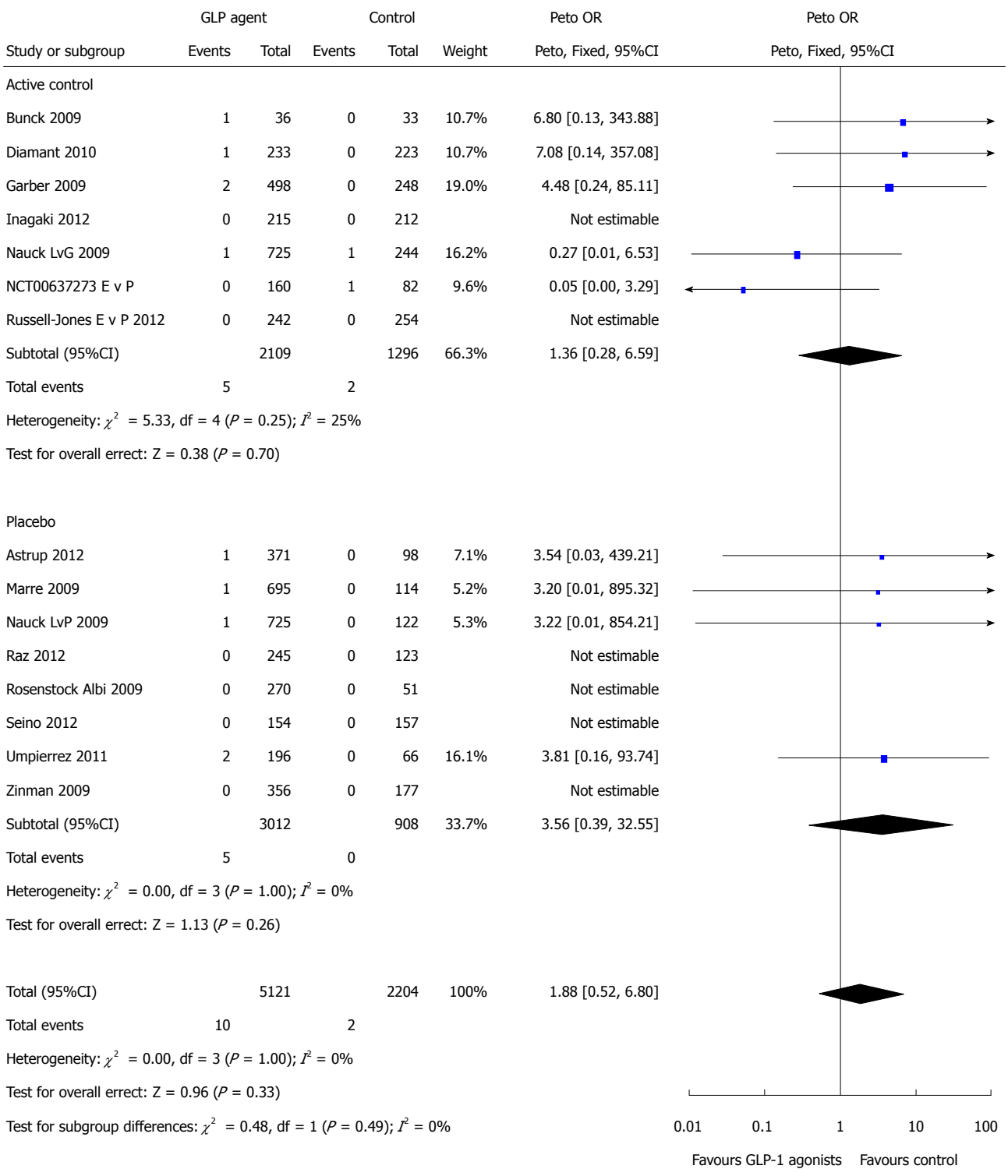


Figure 5 Risk of pancreatic events for glucagon-like peptide-1 receptor agonist drugs only.

of cases of pancreatic cancer to allow for the estimation of meaningful differences between GLP-1 based agents and controls.

Explanations

These discordant results-no significant effect on the outcome of acute pancreatitis but significant increase in the risk of pancreatic enzyme elevation associated with GLP-1 based therapies in a small number of studies may have two alternative explanations.

These could indicate that injury with GLP-1 based therapies is sub-threshold and result in pancreatic inflammation that may not reach the level of acute pancreatitis. Alternatively, the ascertainment of pancreatic adverse events/complications may have been more complete in this subset of studies showing an elevation in pancreatic

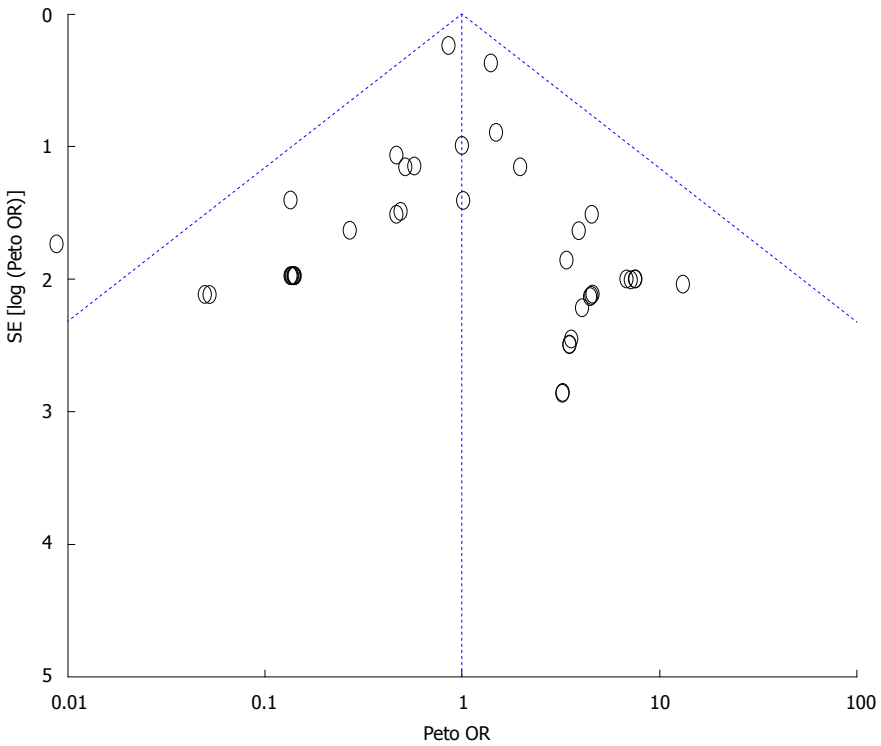


Figure 6 Funnel plot for risk of pancreatic adverse events.

enzymes. It was not clear whether pancreatitis adverse events were rigorously defined or captured in an objective rather than subjective manner across the trials, potentially biasing towards the null due to misclassification. In contrast, measurement of elevated pancreatic enzymes is a more objective measure, serial enzyme measurements should be regularly checked in trial participants on GLP-1 agents who present with gastrointestinal symptoms. Lack of awareness for the need to assess pancreatic enzymes could lead to under-ascertainment of pancreatic adverse events in patients presenting with upper abdominal symptoms. Among patients with type 2 diabetes, one previous study reported an increase in enzyme associated with DPP-4 inhibitors compared to controls (36% *vs* 18%), suggesting that this adverse reaction deserves further investigation[15].

Our meta-analysis should be seen in the light of other recent studies. A recent review reported a slightly increased trend for reporting of acute pancreatitis associated with GLP-1 receptor agonists but not with DPP-Ⅳ inhibitors[16]. Two other systematic reviews reported no increased risk of acute pancreatitis, but with very wide confidence intervals that could not rule out a significant increase[6,17]. However, one such meta-analysis included observational studies, which may be prone to confounding[17]. The difference in meta-analysis should reflect differences in inclusion of trials and ascertainment of events. Importantly none of the previous meta-analysis have reported on elevations in pancreatic enyzmes associated with GLP-1 based therapies. However, the CIs were wide in all meta-analyses and we could not rule out a significant increase in the risk of pancreatitis with GLP-1 based therapies. The lack of statistical significance may reflect incomplete ascertainment of pancreatic adverse events in clinical trials of GLP-1 based therapies or inadequate statistical power to detect rare but serious complication such as pancreatitis. Observational studies have also shown inconsistent results between GLP-1 based therapies and acute pancreatitis due to incomplete ascertainment of covariates, or poor performance of the diagnostic codes for acute pancreatitis[5,18-20]. It is also unclear whether the inflammatory process from recurrent or chronic pancreatitis is a predisposing factor to subsequent development of pancreatic cancer.

Limitations

Our study has some limitations. We limited our analysis to published RCTs. However; there may be unpublished studies that report on this outcome. We did not have access to data to conduct individual patient data meta-analysis and ascertain time to the occurrence of pancreatic enzyme elevations. Importantly, clinical trials may not have ascertained the occurrence of pancreatitis on participants who withdrew from the trial (as a result of the complication). This may bias our estimates towards the null. The availability of sponsors of individual patient data to independent investigators may allow for further analyses.

Our meta-analysis shows a three-fold increased risk of pancreatic enzyme elevation with GLP-1 based agents compared to controls, without an a significant increased risk of pancreatitis or pancreatic cancer due to small number of cases. Future adequately powered observational studies with well validated codes for pancreatitis

and pancreatic cancer and careful control of confounding are needed to evaluate the risk of pancreatic enzyme elevation, pancreatitis and pancreatic cancer with GLP-1 based therapies.

COMMENTS

Background

Recent developments have led to an increasingly wide range of glucose lowering drugs being trialed for treatment of type II diabetes mellitus. However, a variety of concerns have been raised regarding the safety of these new agents for long-term chronic use. This has led to tightening of the regulatory landscape and closer scrutiny of data regarding serious rare adverse events.

Research frontiers

Many trials have been conducted to demonstrate the efficacy of glucagon-like-peptide-1 (GLP-1) agonists and dipeptidyl peptidase-4 (DPP-4) inhibitors in reducing blood glucose levels. However, there have been suggestions of a potential increase in risk of pancreatic adverse events with these drugs due to a postulated proliferative effect on pancreatic cells. The existing evidence base is conflicting, and difficult to interpret due to the very low incidence of pancreatic adverse events.

Innovations and breakthroughs

The findings of this meta-analysis are that risks of pancreatitis or pancreatic cancer have not been definitively established with any of the GLP-1 agonists or DPP-4 inhibitors. However, there is a signal suggesting increased risk of elevated pancreatic enzymes, which has not previously been described in other systematic reviews.

Applications

GLP-1 agonists or DPP-4 inhibitors may have some relationship with elevations in the pancreatic enzyme levels. Further large scale studies are needed to determine if these elevations may or may not be associated with adverse clinical outcomes.

Terminology

GLP-1 belongs to the incretin group of hormones which act to stimulate insulin secretion dependent on glucose levels. GLP-1 receptor agonists are drugs developed as incretin-mimetics. DPP-4 is an enzyme that breaks down GLP-1, thus causing GLP-1 to have a short half-life. Drugs that inhibit DPP-4 would be expected to increase the availability of endogenous GLP-1.

Peer-review

This manuscript has a great collecting data about this topic.

REFERENCES

1 **Butler PC**, Dry S, Elashoff R. GLP-1-based therapy for diabetes: what you do not know can hurt you. *Diabetes Care* 2010; **33**: 453-455 [PMID: 20103562 DOI: 10.2337/dc09-1902]

2 **Drucker DJ**, Sherman SI, Gorelick FS, Bergenstal RM, Sherwin RS, Buse JB. Incretin-based therapies for the treatment of type 2 diabetes: evaluation of the risks and benefits. *Diabetes Care* 2010; **33**: 428-433 [PMID: 20103558]

3 **Nachnani JS**, Bulchandani DG, Nookala A, Herndon B, Molteni A, Pandya P, Taylor R, Quinn T, Weide L, Alba LM. Biochemical and histological effects of exendin-4 (exenatide) on the rat pancreas. *Diabetologia* 2010; **53**: 153-159 [PMID: 19756486 DOI: 10.1007/s00125-009-1515-4]

4 **Matveyenko AV**, Dry S, Cox HI, Moshtaghian A, Gurlo T, Galasso R, Butler AE, Butler PC. Beneficial endocrine but adverse exocrine effects of sitagliptin in the human islet amyloid polypeptide transgenic rat model of type 2 diabetes: interactions with metformin. *Diabetes* 2009; **58**: 1604-1615 [PMID: 19403868 DOI: 10.2337/db09-0058]

5 **Singh S**, Chang HY, Richards TM, Weiner JP, Clark JM, Segal JB. Glucagonlike peptide 1-based therapies and risk of hospitalization for acute pancreatitis in type 2 diabetes mellitus: a population-based matched case-control study. *JAMA Intern Med* 2013; **173**: 534-539 [PMID: 23440284 DOI: 10.1001/jamainternmed.2013.2720]

6 **Monami M**, Dicembrini I, Martelli D, Mannucci E. Safety of dipeptidyl peptidase-4 inhibitors: a meta-analysis of randomized clinical trials. *Curr Med Res Opin* 2011; **27** Suppl 3: 57-64 [PMID: 22106978 DOI: 10.1185/03007995.2011.602964]

7 **Alves C**, Batel-Marques F, Macedo AF. A meta-analysis of serious adverse events reported with exenatide and liraglutide: acute pancreatitis and cancer. *Diabetes Res Clin Pract* 2012; **98**: 271-284 [PMID: 23010561]

8 **Shihab HM,** Akande T, Loke YK, Singh S. Risk of pancreatic complication events associated with the use of GLP-1 receptor agonist and DPP-4 inhibitor drugs: A systematic review and meta-analysis. PROSPERO: International prospective register of systematic reviews. Available from: URL: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013004742. Accessed June 23, 2014

9 **Higgins JPT,** Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration; 2011

10 **Loke YK,** Price D, Herxheimer A. Chapter 14: Adverse Effects. In: Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Chichester: John Wiley and Sons, 2008

11 **Ioannidis JP**, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *CMAJ* 2007; **176**: 1091-1096 [PMID: 17420491 DOI: 10.1503/cmaj.060410]

12 Review Manager (RevMan) [computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014

13 **Davies HT**, Crombie IK, Tavakoli M. When can odds ratios mislead? *BMJ* 1998; **316**: 989-991 [PMID: 9550961]

14 **Higgins JP**, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539-1558 [PMID: 12111919 DOI: 10.1002/sim.1186]

15 **Lando HM**, Alattar M, Dua AP. Elevated amylase and lipase levels in patients using glucagonlike peptide-1 receptor agonists or dipeptidyl-peptidase-4 inhibitors in the outpatient setting. *Endocr Pract* 2012; **18**: 472-477 [PMID: 22440997]

16 **Meier JJ**, Nauck MA. Risk of pancreatitis in patients treated with incretin-based therapies. *Diabetologia* 2014; **57**: 1320-1324 [PMID: 24723174 DOI: 10.1007/s00125-014-3231-y]

17 **Li L**, Shen J, Bala MM, Busse JW, Ebrahim S, Vandvik PO, Rios LP, Malaga G, Wong E, Sohani Z, Guyatt GH, Sun X. Incretin treatment and risk of pancreatitis in patients with type 2 diabetes mellitus: systematic review and meta-analysis of randomised and non-randomised studies. *BMJ* 2014; **348**: g2366 [PMID: 24736555 DOI: 10.1136/bmj.g2366]

18 **Garg R**, Chen W, Pendergrass M. Acute pancreatitis in type 2 diabetes treated with exenatide or sitagliptin: a retrospective observational pharmacy claims analysis. *Diabetes Care* 2010; **33**: 2349-2354 [PMID: 20682680 DOI: 10.2337/dc10-0482]

19 **Dore DD,** Hussein M, Hoffman C, Pelletier EM, Smith DB, Seeger JD. A pooled analysis of exenatide use and risk of acute pancreatitis. *Curr Med Res Opin* 2013; **29**: 1577-1586 [PMID: 23981106 DOI: 10.1185/03007995.2013.838550]

20 **Chou HC**, Chen WW, Hsiao FY. Acute pancreatitis in patients with type 2 diabetes mellitus treated with dipeptidyl peptidase-4 inhibitors: a population-based nested case-control study. *Drug Saf* 2014; **37**: 521-528 [PMID: 24859164 DOI: 10.1007/s40264-014-0171-x]

21 **Ross SA**, Rafeiro E, Meinicke T, Toorawa R, Weber-Born S, Woerle HJ. Efficacy and safety of linagliptin 2.5 mg twice daily versus 5 mg once daily in patients with type 2 diabetes inadequately controlled on metformin: a randomised, double-blind, placebo-controlled trial. *Curr Med Res Opin* 2012; **28**: 1465-1474 [PMID: 22816729 DOI: 10.1185/03007995.2012.714360]

22 **Haak T**, Meinicke T, Jones R, Weber S, von Eynatten M, Woerle HJ. Initial combination of linagliptin and metformin in patients with type 2 diabetes: efficacy and safety in a randomised, double-blind

1-year extension study. *Int J Clin Pract* 2013; **67**: 1283-1293 [PMID: 24118640 DOI: 10.1111/j.1463-1326.2012.01590.x]

23 **Boehringer Ingelheim Pharmaceuticals.** Efficacy and safety of 3 doses of BI1356 (linagliptin) in type 2 diabetes patients. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 8]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00328172 NLM Identifier: NCT00328172

24 **Yki-Järvinen H**, Rosenstock J, Durán-Garcia S, Pinnetti S, Bhattacharya S, Thiemann S, Patel S, Woerle HJ. Effects of adding linagliptin to basal insulin regimen for inadequately controlled type 2 diabetes: a ≥52-week randomized, double-blind study. *Diabetes Care* 2013; **36**: 3875-3881 [PMID: 24062327 DOI: 10.2337/dc12-2718]

25 **Boehringer Ingelheim Pharmaceuticals, Eli Lilly and Company.** Efficacy and safety of linagliptin in combination with insulin in patients with type 2 diabetes. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 8]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00954447 NLM Identifier: NCT00954447

26 **Boehringer Ingelheim Pharmaceuticals.** Japanese P III vs voglibose and placebo. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 8]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00654381 NLM Identifier: NCT00654381

27 **Boehringer Ingelheim Pharmaceuticals.** Efficacy and safety of BI 1356 in combination with metformin in patients with type 2 diabetes. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 8]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00622284 NLM Identifier: NCT00622284

28 **Boehringer Ingelheim Pharmaceuticals.** A randomised, double-blind, placebo controlled, parallel group 24 week study to assess the efficacy and safety of linagliptin (5mg) in combination with 30mg pioglitazone (both administered orally once daily), compared to 30mg pioglitazone plus placebo in drug-naive or previously treated type 2 diabetic patients with insufficient glycaemic control. [accessed 2014 Jun 8]. Available from: URL: http://trials.boehringer-ingelheim.com/content/dam/internet/opu/clinicaltrial/com_EN/results/1218/1218.15_U09-2519.pdf

29 **Boehringer Ingelheim Pharmaceuticals.** A phase III randomised, double-blind parallel group extension study to investigate the safety and efficacy of twice daily administration of the free combination of linagliptin 2.5 mg metformin 500 mg or of linagliptin 2.5 mg metformin 1000mg versus monotherapy with metformin 1000 mg twice daily over 54 weeks in type 2 diabetic patients previously completing the double-blind part of study 1218.46. [accessed 2014 Jun 8]. Available from: URL: http://trials.boehringer-ingelheim.com/content/dam/internet/opu/clinicaltrial/com_EN/results/1218/1218.52_U11-1782-01.pdf

30 **Boehringer Ingelheim Pharmaceuticals.** A phase III randomised, double-blind, placebo-controlled, parallel group, efficacy and safety study of linagliptin (5mg), administered orally once daily over 24 weeks in type 2 diabetic patients (age 70 years) with insufficient glycaemic control (HbA1c 7.0%) despite metformin and/or sulphonylurea and/or insulin therapy. [accessed 2014 Jun 8]. Available from: URL: http://trials.boehringer-ingelheim.com/content/dam/internet/opu/clinicaltrial/com_EN/results/1218/1218.63_U11-1781-02.pdf

31 **Boehringer Ingelheim Pharmaceuticals.** A phase IIIb, 24 week, randomised, placebo-controlled, double-blinded, efficacy and safety study of linagliptin in black/african american patients with type 2 diabetes with a MTT sub-study. [accessed 2014 Jun 8]. Available from: URL: http://trials.boehringer-ingelheim.com/content/dam/internet/opu/clinicaltrial/com_EN/results/1218/1218.75_U12-3204-01.pdf

32 **Boehringer Ingelheim Pharmaceuticals.** A phase III, randomised, double-blind, placebo-controlled parallel group efficacy and safety study of linagliptin 5mg administered orally once daily over 24 weeks in type 2 diabetic patients with insufficient glycaemic control despite a therapy of metformin in combination with pioglitazone. [accessed 2014 Jun 8]. Available from: URL: http://trials.boehringer-ingelheim.com/content/dam/internet/opu/clinicaltrial/com_EN/results/1218/1218.61_U13-3124-01.pdf

33 **Boehringer Ingelheim Pharmaceuticals.** A randomised, double-blind, placebo-controlled parallel group efficacy and safety study of linagliptin (5mg administered orally once daily) over 24 weeks in type 2 diabetic patients with insufficient glycaemic control despite metformin therapy in asian population. [accessed 2014 Jun 8]. Available from: URL: http://trials.boehringer-ingelheim.com/content/dam/internet/opu/clinicaltrial/com_EN/results/1218/1218.65_U12-2143-01.pdf

34 **Boehringer Ingelheim Pharmaceuticals.** A phase III, randomised, double-blind, placebo-controlled parallel group safety and efficacy study of linagliptin (5mg administered orally once daily) over 12 weeks followed by a 40 week double-blind extension period (placebo patients switched to glimepiride) in drug naive or previously treated type 2 diabetic patients with moderate to severe renal impairment and insufficient glycaemic control. [accessed 2014 Jun 8]. Available from: URL: http://trials.boehringer-ingelheim.com/content/dam/internet/opu/clinicaltrial/com_EN/results/1218/1218.64_U13-1283-01-DS.pdf

35 **Boehringer Ingelheim Pharmaceuticals.** A randomised, double-blind, placebo-controlled parallel group, efficacy and safety study of linagliptin (5mg administered orally once daily) over 24 weeks, in drug naive or previously treated type 2 diabetic patients with insufficient glycaemic control. [accessed 2014 Jun 8]. Available from: URL: http://trials.boehringer-ingelheim.com/content/dam/internet/opu/clinicaltrial/com_EN/results/1218/1218.66_U12-2076-01.pdf

36 **Rosenstock J**, Rendell MS, Gross JL, Fleck PR, Wilson CA, Mekki Q. Alogliptin added to insulin therapy in patients with type 2 diabetes reduces HbA(1C) without causing weight gain or increased hypoglycaemia. *Diabetes Obes Metab* 2009; **11**: 1145-1152 [PMID: 19758359 DOI: 10.1111/j.1463-1326.2009.01124.x]

37 **White WB**, Cannon CP, Heller SR, Nissen SE, Bergenstal RM, Bakris GL, Perez AT, Fleck PR, Mehta CR, Kupfer S, Wilson C, Cushman WC, Zannad F. Alogliptin after acute coronary syndrome in patients with type 2 diabetes. *N Engl J Med* 2013; **369**: 1327-1335 [PMID: 23992602 DOI: 10.1056/NEJMoa1305889]

38 **Takeda.** Long-term safety study of alogliptin used in combination with sulfonylurea or metformin in participants with type 2 diabetes in japan. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT01318135 NLM Identifier: NCT01318135

39 **Takeda.** Efficacy and safety of alogliptin in participants with type 2 diabetes. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT01289119 NLM Identifier: NCT01289119

40 **Takeda**. Long-term safety study of alogliptin in participants with type 2 diabetes in japan. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/results?term=NCT01263496 NLM Identifier: NCT01263496

41 **Takeda.** Efficacy and safety of alogliptin combined with pioglitazone in treating subjects with type 2 diabetes mellitus. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00328627 NLM Identifier: NCT00328627

42 **Takeda**. Efficacy of alogliptin with pioglitazone (actos) in subjects with type 2 diabetes mellitus. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00395512 NLM Identifier: NCT00395512

43 **Kikuchi M**, Haneda M, Koya D, Tobe K, Onishi Y, Couturier A, Mimori N, Inaba Y, Goodman M. Efficacy and tolerability of vildagliptin as an add-on to glimepiride in Japanese patients with Type 2 diabetes mellitus. *Diabetes Res Clin Pract* 2010; **89**:

216-223 [PMID: 20537746 DOI: 10.1016/j.diabres.2010.04.017]
44 **Lukashevich V**, Schweizer A, Shao Q, Groop PH, Kothny W. Safety and efficacy of vildagliptin versus placebo in patients with type 2 diabetes and moderate or severe renal impairment: a prospective 24-week randomized placebo-controlled trial. *Diabetes Obes Metab* 2011; **13**: 947-954 [PMID: 21733061 DOI: 10.1111/j.1463-1326.2011.01467.x]
45 **Strain WD**, Lukashevich V, Kothny W, Hoellinger MJ, Paldánius PM. Individualised treatment targets for elderly patients with type 2 diabetes using vildagliptin add-on or lone therapy (INTERVAL): a 24 week, randomised, double-blind, placebo-controlled study. *Lancet* 2013; **382**: 409-416 [PMID: 23706759 DOI: 10.1016/S0140-6736(13)60995-2]
46 **Novartis Pharmaceuticals.** Vildagliptin compared to glimepiride in combination with metformin in patients with type 2 diabetes. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00106340 NLM Identifier: NCT00106340
47 **Novartis Pharmaceuticals**. A 56-week extension to a clinical study to assess the efficacy and safety of vildagliptin compared to placebo in drug naive patients with type 2 diabetes and mild hyperglycemia. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00300287 NLM Identifier: NCT00300287
48 **Novartis Pharmaceuticals.** A multicenter, randomized, double-blind, active-controlled study to compare the effects of 12 weeks treatment with vildagliptin 50 mg b.i.d. to voglibose 0.2 mg t.i.d. in patients with type 2 diabetes. [accessed 2014 Jun 22]. Available from: URL: http://www.novctrd.com/ctrdWebApp/clinicaltrialrepository/displayFile.do?trialResult=2524
49 **Novartis Pharmaceuticals.** A multi-center, randomized, open-label, active controlled, parallel arm study to compare the efficacy of 12 weeks of treatment with vildagliptin 100 mg, once daily (qd) to thiazolidinedione (TZD) as add-on therapy in patients with type 2 diabetes inadequately controlled with metformin monotherapy in a community-based practice setting. [accessed 2014 Jun 22]. Available from: URL: http://www.novctrd.com/ctrdWebApp/clinicaltrialrepository/displayFile.do?trialResult=2567.
50 **Novartis Pharmaceuticals.** Efficacy and safety of vildagliptin compared to acarbose in drug naive patients with type 2 diabetes. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: https://clinicaltrials.gov/ct2/show/NCT00110240 NLM Identifier: NCT00110240
51 **AstraZeneca**. A phase 3 study of BMS-477118 in combination with metformin in subjects with type 2 diabetes who are not controlled with diet and exercise. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00327015 NLM Identifier: NCT00327015
52 **Hollander PL**, Li J, Frederich R, Allen E, Chen R. Safety and efficacy of saxagliptin added to thiazolidinedione over 76 weeks in patients with type 2 diabetes mellitus. *Diab Vasc Dis Res* 2011; **8**: 125-135 [PMID: 21562064 DOI: 10.1177/1479164111404575]
53 **AstraZeneca.** Safety and efficacy of saxagliptin plus insulin with or without metformin. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00757588 NLM Identifier: NCT00757588
54 **Scirica BM**, Bhatt DL, Braunwald E, Steg PG, Davidson J, Hirshberg B, Ohman P, Frederich R, Wiviott SD, Hoffman EB, Cavender MA, Udell JA, Desai NR, Mosenzon O, McGuire DK, Ray KK, Leiter LA, Raz I. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. *N Engl J Med* 2013; **369**: 1317-1326 [PMID: 23992601 DOI: 10.1056/NEJMoa1307684]
55 **Göke B**, Gallwitz B, Eriksson JG, Hellqvist Å, Gause-Nilsson I. Saxagliptin vs. glipizide as add-on therapy in patients with type 2 diabetes mellitus inadequately controlled on metformin alone: long-term (52-week) extension of a 52-week randomised controlled trial. *Int J Clin Pract* 2013; **67**: 307-316 [PMID: 23638466 DOI: 10.1111/ijcp.12119]
56 **AstraZeneca.** Study of BMS-477118 as monotherapy with titration in subjects with type 2 diabetes who are not controlled with diet and exercise. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: https://clinicaltrials.gov/ct2/show/NCT00316082 NLM Identifier: NCT00316082
57 **AstraZeneca**. Treatment effect of saxagliptin compared with placebo in patients with type 2 diabetes and renal impairment. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00614939 NLM Identifier: NCT00614939
58 **Merck Sharp and Dohme Corp.** An investigational drug in patients with type 2 diabetes mellitus and chronic renal insufficiency. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00095056 NLM Identifier: NCT00095056
59 **Chan JC**, Scott R, Arjona Ferreira JC, Sheng D, Gonzalez E, Davies MJ, Stein PP, Kaufman KD, Amatruda JM, Williams-Herman D. Safety and efficacy of sitagliptin in patients with type 2 diabetes and chronic renal insufficiency. *Diabetes Obes Metab* 2008; **10**: 545-555 [PMID: 18518892 DOI: 10.1111/j.1463-1326.2008.00914.x]
60 **Kojima Y**, Kaga H, Hayashi S, Kitazawa T, Iimura Y, Ohno M, Yoshitsugu M, Fujiwara M, Hiyoshi T. Comparison between sitagliptin and nateglinide on postprandial lipid levels: The STANDARD study. *World J Diabetes* 2013; **4**: 8-13 [PMID: 23493856 DOI: 10.4239/wjd.v4.i1.8]
61 **Arjona Ferreira JC**, Marre M, Barzilai N, Guo H, Golm GT, Sisk CM, Kaufman KD, Goldstein BJ. Efficacy and safety of sitagliptin versus glipizide in patients with type 2 diabetes and moderate-to-severe chronic renal insufficiency. *Diabetes Care* 2013; **36**: 1067-1073 [PMID: 23248197 DOI: 10.2337/dc12-1365]
62 **Merck Sharp and Dohme Corp**. Sitagliptin versus glipizide in participants with type 2 diabetes mellitus and chronic renal insufficiency (MK-0431-063 AM1). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00509262 NLM Identifier: NCT00509262
63 **Merck Sharp and Dohme Corp**. MK0431 and pioglitazone co-administration factorial study in patients with type 2 diabetes mellitus (0431-102 AM2). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00722371 NLM Identifier: NCT00722371
64 **Henry RR**, Staels B, Fonseca VA, Chou MZ, Teng R, Golm GT, Langdon RB, Kaufman KD, Steinberg H, Goldstein BJ. Efficacy and safety of initial combination treatment with sitagliptin and pioglitazone--a factorial study. *Diabetes Obes Metab* 2014; **16**: 223-230 [PMID: 23909985 DOI: 10.1111/dom.12194]
65 **Merck Sharp and Dohme Corp.** Sitagliptin metformin add-on study in patients with TYpe 2 diabetes mellitus. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00337610 NLM Identifier: NCT00337610
66 **Raz I**, Chen Y, Wu M, Hussain S, Kaufman KD, Amatruda JM, Langdon RB, Stein PP, Alba M. Efficacy and safety of sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes. *Curr Med Res Opin* 2008; **24**: 537-550 [PMID: 18194595 DOI: 10.1185/030079908X260925]
67 **Merck Sharp and Dohme Corp.** Study of sitagliptin treatment in patients with type 2 diabetes during ramadhan (0431-263). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT01131182 NLM Identifier: NCT01131182

68 **Goldstein BJ**, Feinglos MN, Lunceford JK, Johnson J, Williams-Herman DE. Effect of initial combination therapy with sitagliptin, a dipeptidyl peptidase-4 inhibitor, and metformin on glycemic control in patients with type 2 diabetes. *Diabetes Care* 2007; **30**: 1979-1987 [PMID: 17485570]

69 **Merck Sharp and Dohme Corp.** MK0431 (sitagliptin) and metformin co-administration factorial study in patients with type 2 diabetes mellitus (0431-036). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00103857 NLM Identifier: NCT00103857

70 **Merck Sharp and Dohme Corp**. A study to test the safety and efficacy of sitagliptin compared to glimepiride in patients with type 2 diabetes on a stable dose of metformin (0431-803). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00701090 NLM Identifier: NCT00701090

71 **Arechavaleta R**, Seck T, Chen Y, Krobot KJ, O'Neill EA, Duran L, Kaufman KD, Williams-Herman D, Goldstein BJ. Efficacy and safety of treatment with sitagliptin or glimepiride in patients with type 2 diabetes inadequately controlled on metformin monotherapy: a randomized, double-blind, non-inferiority trial. *Diabetes Obes Metab* 2011; **13**: 160-168 [PMID: 21199268 DOI: 10.1111/j.1463-1326.2010.01334.x]

72 **Merck Sharp and Dohme Corp**. Metformin add-on study in patients with type 2 diabetes mellitus (0431-020). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00086515 NLM Identifier: NCT00086515

73 **Charbonnel B**, Karasik A, Liu J, Wu M, Meininger G. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes inadequately controlled with metformin alone. *Diabetes Care* 2006; **29**: 2638-2643 [PMID: 17130197 DOI: 10.2337/dc06-0706]

74 **Merck Sharp and Dohme Corp**. A study to compare the glycemic effects, safety, and tolerability of exenatide once weekly to those of sitagliptin and pioglitazone in subjects with type 2 diabetes treated with metformin (DURATION-2). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00637273 NLM Identifier: NCT00637273

75 **Bergenstal RM**, Wysham C, Macconell L, Malloy J, Walsh B, Yan P, Wilhelm K, Malone J, Porter LE. Efficacy and safety of exenatide once weekly versus sitagliptin or pioglitazone as an adjunct to metformin for treatment of type 2 diabetes (DURATION-2): a randomised trial. *Lancet* 2010; **376**: 431-439 [PMID: 20580422 DOI: 10.1016/S0140-6736(10)60590-9]

76 **Merck Sharp and Dohme Corp.** An investigational drug study in patients with type 2 diabetes mellitus (MK0431-023). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00094757 NLM Identifier: NCT00094757

77 **Merck Sharp and Dohme Corp.** An investigational drug study in patients with type 2 diabetes mellitus (0431-024). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00094770 NLM Identifier: NCT00094770

78 **Janssen Research and Development, LLC.** The CANTATA-D2 trial (CANagliflozin treatment and trial analysis-DPP-4 inhibitor second comparator trial). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT01137812 NLM Identifier: NCT01137812

79 **Schernthaner G**, Gross JL, Rosenstock J, Guarisco M, Fu M, Yee J, Kawaguchi M, Canovatchel W, Meininger G. Canagliflozin compared with sitagliptin for patients with type 2 diabetes who do not have adequate glycemic control with metformin plus sulfonylurea: a 52-week randomized trial. *Diabetes Care* 2013; **36**: 2508-2515 [PMID: 23564919 DOI: 10.2337/dc12-2491]

80 **Merck Sharp and Dohme Corp.** MK0431 A comparative study in patients with type 2 diabetes (0431A-079). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT00482729 NLM Identifier: NCT00482729

81 **Bunck MC**, Diamant M, Cornér A, Eliasson B, Malloy JL, Shaginian RM, Deng W, Kendall DM, Taskinen MR, Smith U, Yki-Järvinen H, Heine RJ. One-year treatment with exenatide improves beta-cell function, compared with insulin glargine, in metformin-treated type 2 diabetic patients: a randomized, controlled trial. *Diabetes Care* 2009; **32**: 762-768 [PMID: 19196887 DOI: 10.2337/dc08-1797]

82 **Diamant M**, Van Gaal L, Stranks S, Northrup J, Cao D, Taylor K, Trautmann M. Once weekly exenatide compared with insulin glargine titrated to target in patients with type 2 diabetes (DURATION-3): an open-label randomised trial. *Lancet* 2010; **375**: 2234-2243 [PMID: 20609969 DOI: 10.1016/S0140-6736(10)60406-0]

83 **Inagaki N**, Atsumi Y, Oura T, Saito H, Imaoka T. Efficacy and safety profile of exenatide once weekly compared with insulin once daily in Japanese patients with type 2 diabetes treated with oral antidiabetes drug(s): results from a 26-week, randomized, open-label, parallel-group, multicenter, noninferiority study. *Clin Ther* 2012; **34**: 1892-1908.e1 [PMID: 22884767 DOI: 10.1016/j.clinthera.2012.07.007]

84 **Russell-Jones D**, Cuddihy RM, Hanefeld M, Kumar A, González JG, Chan M, Wolka AM, Boardman MK. Efficacy and safety of exenatide once weekly versus metformin, pioglitazone, and sitagliptin used as monotherapy in drug-naive patients with type 2 diabetes (DURATION-4): a 26-week double-blind study. *Diabetes Care* 2012; **35**: 252-258 [PMID: 22210563 DOI: 10.2337/dc11-1107]

85 **AstraZeneca.** Efficacy of once-weekly exenatide versus once or twice daily insulin detemir in patients with type 2 diabetes. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). [accessed 2014 Jun 22]. Available from: URL: https://clinicaltrials.gov/ct2/show/NCT01003184 NLM Identifier: NCT01003184

86 **Astrup A**, Carraro R, Finer N, Harper A, Kunesova M, Lean ME, Niskanen L, Rasmussen MF, Rissanen A, Rössner S, Savolainen MJ, Van Gaal L. Safety, tolerability and sustained weight loss over 2 years with the once-daily human GLP-1 analog, liraglutide. *Int J Obes* (Lond) 2012; **36**: 843-854 [PMID: 21844879 DOI: 10.1038/ijo.2011.158]

87 **Garber A**, Henry R, Ratner R, Garcia-Hernandez PA, Rodriguez-Pattzi H, Olvera-Alvarez I, Hale PM, Zdravkovic M, Bode B; LEAD-3 (Mono) Study Group. Liraglutide versus glimepiride monotherapy for type 2 diabetes (LEAD-3 mono): A randomised, 52-week, phase III, double-blind, parallel-treatment trial. *Lancet* 2009; **373**: 473-481 [PMID: 18819705 DOI: 10.1016/S0140-6736(08)61246-5]

88 **Nauck M**, Frid A, Hermansen K, Shah NS, Tankova T, Mitha IH, Zdravkovic M, Düring M, Matthews DR. Efficacy and safety comparison of liraglutide, glimepiride, and placebo, all in combination with metformin, in type 2 diabetes: the LEAD (liraglutide effect and action in diabetes)-2 study. *Diabetes Care* 2009; **32**: 84-90 [PMID: 18931095 DOI: 10.2337/dc08-1355]

89 **Marre M**, Shaw J, Brändle M, Bebakar WM, Kamaruddin NA, Strand J, Zdravkovic M, Le Thi TD, Colagiuri S. Liraglutide, a once-daily human GLP-1 analogue, added to a sulphonylurea over 26 weeks produces greater improvements in glycaemic and weight control compared with adding rosiglitazone or placebo in subjects with Type 2 diabetes (LEAD-1 SU). *Diabet Med* 2009; **26**: 268-278 [PMID: 19317822 DOI: 10.1111/j.1464-5491.2009.02666.x]

90 **Zinman B**, Gerich J, Buse JB, Lewin A, Schwartz S, Raskin P, Hale PM, Zdravkovic M, Blonde L. Efficacy and safety of the human glucagon-like peptide-1 analog liraglutide in combination with metformin and thiazolidinedione in patients with type 2 diabetes (LEAD-4 Met+TZD). *Diabetes Care* 2009; **32**: 1224-1230 [PMID: 19289857 DOI: 10.2337/dc08-2124]

91 **Raz I**, Fonseca V, Kipnes M, Durrwell L, Hoekstra J, Boldrin M,

Balena R. Efficacy and safety of taspoglutide monotherapy in drug-naive type 2 diabetic patients after 24 weeks of treatment: results of a randomized, double-blind, placebo-controlled phase 3 study (T-emerge 1). *Diabetes Care* 2012; **35**: 485-487 [PMID: 22301126 DOI: 10.2337/dc11-1942]

92 **Rosenstock J**, Reusch J, Bush M, Yang F, Stewart M. Potential of albiglutide, a long-acting GLP-1 receptor agonist, in type 2 diabetes: a randomized controlled trial exploring weekly, biweekly, and monthly dosing. *Diabetes Care* 2009; **32**: 1880-1886 [PMID: 19592625 DOI: 10.2337/dc09-0366]

93 **Seino Y**, Min KW, Niemoeller E, Takami A. Randomized, double-blind, placebo-controlled trial of the once-daily GLP-1 receptor agonist lixisenatide in Asian patients with type 2 diabetes insufficiently controlled on basal insulin with or without a sulfonylurea (GetGoal-L-Asia). *Diabetes Obes Metab* 2012; **14**: 910-917 [PMID: 22564709 DOI: 10.1111/j.1463-1326.2012.01618.x]

94 **Umpierrez GE**, Blevins T, Rosenstock J, Cheng C, Anderson JH, Bastyr EJ. The effects of LY2189265, a long-acting glucagon-like peptide-1 analogue, in a randomized, placebo-controlled, double-blind study of overweight/obese patients with type 2 diabetes: the EGO study. *Diabetes Obes Metab* 2011; **13**: 418-425 [PMID: 21251180 DOI: 10.1111/j.1463-1326.2011.01366.x]

P- Reviewer: Rungsakulkij N
S- Editor: Qi Y **L- Editor**: A **E- Editor**: Jiao XK

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i6.284

World J Meta-Anal 2015 December 26; 3(6): 284-294
ISSN 2308-3840 (online)

META-ANALYSIS

Antireflux surgery *vs* medical treatment for gastroesophageal reflux disease: A meta-analysis

Ya Jiang, Wen-Xia Cui, Ying Wang, Ding Heng, Jia-Cheng Tan, Lin Lin

Ya Jiang, Wen-Xia Cui, Ying Wang, Ding Heng, Jia-Cheng Tan, Lin Lin, Department of Gastroenterology, First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, Jiangsu Province, China

Author contributions: Jiang Y and Cui WX acquisition of data, analysis and interpretation of data, drafting the article, final approval; Wang Y, Heng D and Tan JC interpretation of data, revising the article, final approval; Lin L conception and design of the study, critical revision, final approval.

Supported by The Project of Jiangsu Province Department of Health, No. H201328; the Project of Jiangsu Provincial Administration of traditional Chinese Medicine, No. LZ13225; the Postgraduates' Innovation Program supported by the Education Department of Jiangsu Province, No. Jx22013279.

Conflict-of-interest statement: The authors deny any conflict of interest.

Data sharing statement: No additional data are available.

Correspondence to: Lin Lin, MD, PhD, Department of Gastroenterology, First Affiliated Hospital of Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, Jiangsu Province, China. lin9100@aliyun.com
Telephone: +86-25-68136920
Fax: +86-25-83674636

Received: June 9, 2015
Peer-review started: June 11, 2015
First decision: August 4, 2015
Revised: September 29, 2015
Accepted: October 23, 2015
Article in press: October 27, 2015

Published online: December 26, 2015

Abstract

AIM: To compare the effect of antireflux surgery with medicine in treating gastroesophageal reflux disease (GERD) patients using meta- analysis.

METHODS: MEDLINE, Embase and the Cochrane Library were searched. We only included randomized controlled trials (RCTs) comparing the effect of surgical intervention with medical therapy for GERD. Statistical analyses were performed using RevMan 5.2 and STATA 12.0 software. RevMan 5.2 was used to assess the risk of bias and calculate the pooled effect size, while Stata 12.0 was used to evaluate publication bias and for sensitivity analysis. We evaluated the primary outcomes with GERD-/health-related quality of life in short (one to three years) and long (three to twelve years) periods of follow-up. Secondary outcomes evaluated were DeMeester scores and the percentage of time that pH < 4 to evaluate the degree of acid exposure.

RESULTS: This meta-analysis included 7 studies with 1972 patients. It showed a positive effect of antireflux surgery compared with medical treatment in terms of health-related quality of life [standardized mean difference (SMD) = 0.18; 95%CI: 0.01 to 0.34] and GERD-related quality of life (SMD = 0.35; 95%CI: 0.11 to 0.59). We also conducted the subgroup analyses based on follow-up periods and found that surgery remained more effective than medicine over the short to medium follow-up time, but the advantage of antireflux surgery probably not maintained for long time. GERD-related quality of life in the surgical group was significantly higher than medical group for the < 3 years follow-up (SMD = 0.45; 95%CI: 0.23 to 0.66); the difference was not statistically significant when the follow-up time was ⩾ 3 years (SMD = 0.30; 95%CI: -0.10 to 0.69). Meta-analysis showed a statistically significant difference between the

surgical group and medical group in the percentage of time that pH < 4 (SMD = 0.38; 95%CI: 0.14 to 0.61). Meta-analysis indicated a positive effect of antireflux surgery compared with medical treatment concerning DeMeester scores (SMD = 0.32; 95%CI: 0.00 to 0.65).

CONCLUSION: Although both were effective, in some respects surgical intervention was more effective than medical therapy to treat GERD when follow-up time was up to three years.

Key words: Gastroesophageal reflux disease; Antireflux surgery; Medical treatment; Meta-analysis

Core tip: The advantage of our study is its systematic approach to identifying all randomized controlled trials comparing surgical intervention with medical therapy for gastroesophageal reflux disease (GERD). We conducted the subgroup analyses based on follow-up periods and concluded that antireflux surgery may in some ways be more effective than medicine over short to medium follow-up periods, but its effect declined over time. We also completed analyses for the percentage of time pH < 4 and DeMeester scores. Both analyses favored surgery over medical treatment. However, long-term studies are still required to determine whether surgical intervention is better than medicine in treating GERD patients.

Jiang Y, Cui WX, Wang Y, Heng D, Tan JC, Lin L. Antireflux surgery *vs* medical treatment for gastroesophageal reflux disease: A meta-analysis. *World J Meta-Anal* 2015; 3(6): 284-294 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i6/284.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i6.284

INTRODUCTION

Gastroesophageal reflux disease (GERD) is a highly prevalent chronic disease of the digestive system[1,2], which is resulted from the reflux of gastric contents into the esophagus. The disease is chronic and relapsing that can lead to a marked reduction in patients' health-related quality of life[3].

There are two available treatments for patients to choose: medical treatment and antireflux surgery. The medical treatment of GERD usually involves long-term administration of antacid agents; proton pump inhibitors (PPIs) are now the mainstay of pharmacological treatment[4]. Although PPIs are effective in the treatment of GERD, some patients (4%) continue to experience abnormal acid reflux according to 24-h pH testing in some studies[5], and up to 37% experience a relapse of symptoms during a 5-year follow-up period[6]. With these treatments, the risks of hip fracture and enteric infections are increased[7].

Antireflux surgery is an alternative to medical therapy, but has traditionally been reserved for patients with persistent symptoms despite medication[8]. Since the introduction of the laparoscopic nissen fundoplication (LNF) procedure in the early 1990s, the operation has been found to be effective and is now widely applied throughout the world[9]. In some studies fundoplication produced a therapy of reflux symptoms in up to 90% of patients. A Cochrane review from 2010 suggested that laparoscopic fundoplication surgery was more effective than medical treatment for GERD in the short follow-up period[10]. However, some risk may come along with surgery, and the long-term benefits of it remained uncertain. Since then, several randomized controlled trials (RCTs) comparing medicine with surgery over a longer period of time have been accomplished. The present meta-analysis aimed to systematically summarize and quantify the effects of antireflux surgery and medical treatment for GERD in RCTs on health-related quality of life and GERD-related quality of life.

MATERIALS AND METHODS

Search strategy

To identify all relevant studies, computerized searches of MEDLINE (January 1966-April 2015), Embase (January 1980-April 2015), and the Cochrane Library (January 1970-April 2015) were performed regardless of language. All relevant RCTs comparing antireflux surgery (open or laparoscopic) with medical treatment (PPIs or H2RAs) for patients with GERD were identified. In addition, reference lists in selected papers were searched manually. We also contacted authors to obtain additional data.

Study selection

Two reviewers independently screened the retrieved database files and the full texts of potentially eligible studies for relevance. We used different combinations of the keywords: "GERD" "antireflux surgery, open surgery, laparoscopic fundoplication, Nissen fundoplication" and "medical treatment, PPIs, antacid, histamine receptor antagonists". Disagreements were resolved by consensus.

Eligibility criteria and exclusion criteria

Studies were eligible if they met the following criteria: RCTs; included individuals with GERD (over 14 years of age), whether subjectively or objectively defined, who were judged to be suitable for either medical and/or surgical management; investigated currently used laparoscopic or open antireflux surgery techniques; investigated PPIs and histamine receptor antagonists (H2RAs) as comparative medical treatment for GERD; and reported health-related quality of life with PGWBI, EQ-5D or SF-36, GERD-related quality of life with GI well-being score, REFLUX quality-of life-score, GRESS, GSRS or GRACI scores, GERD-related symptoms and

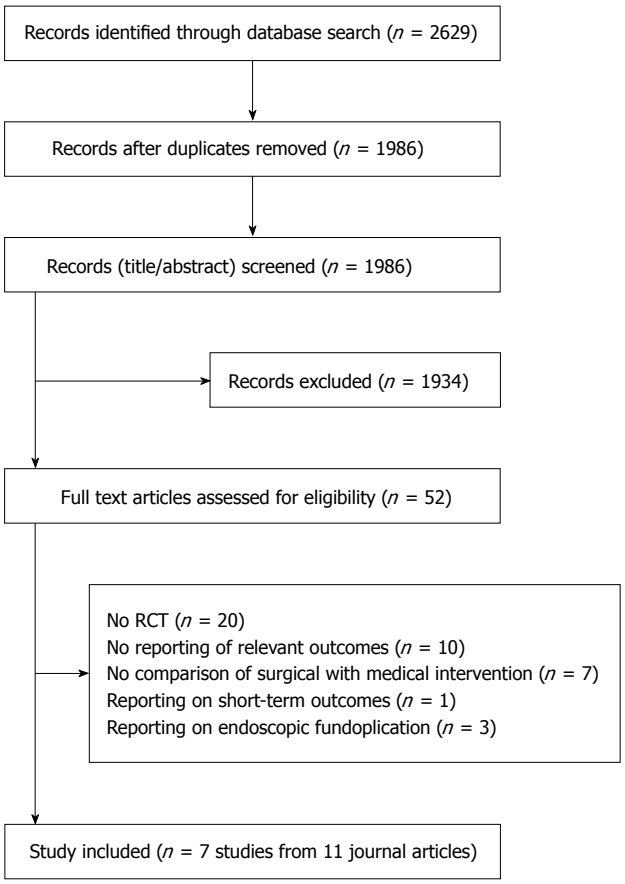


Figure 1 Study selection.

adverse events.

Studies were excluded for the following criteria: Age below 14 years old; duplicate publication; language other than English; not a RCT; unable to extract data from the original literature; no reporting of relevant outcomes; no comparison of surgical with medical intervention; reporting on short-term outcomes or reporting on endoscopic fundoplication.

Outcomes

One of the primary measurements was the standardized mean differences (SMDs) of patients' health-related quality of life. Another primary indicator was the SMDs of patients' GERD-related quality of life. Secondary outcomes included DeMeester scores and the percentage of time that pH < 4 to evaluate the degree of acid exposure.

Data extraction and risk of bias assessment

From each article, we extracted the details of authors, year of publication, country of origin, study population, sex, inclusion and exclusion criteria, type of medication, type of surgery, length of follow-up, and outcomes mentioned above independently in each arm. We assessed the risk of bias for each included study at the level of the selected outcomes. We applied Begg's test to evaluate publication bias risk in our study[11].

Data analysis

Statistical analyses were performed using RevMan 5.2 and STATA 12.0 software. RevMan 5.2 was used to assess the risk of bias and calculate the pooled effect size, while Stata 12.0 was used to assess publication bias and for sensitivity analysis.

SMDs were calculated for continuous data. Considering the different types of scores the included studies used for reporting health- or GERD-related quality of life, we calculated SMDs for pooling effect size of outcomes. The SMD was calculated by dividing the mean difference in each trial by the standard deviation. An SMD greater than 0 corresponded to a favorable effect of the surgical intervention compared with medical management; While an SMD less than 0 meant a negative effect of surgery. In some questionnaires in which higher scores equaled lower GERD-related quality of life (GSRS, GRACI, GRESS), the mean values were multiplied by -1 before calculation of the SMD. Likewise, higher DeMeester scores and the percentage of time that pH < 4 indicate worse acid control, so these two items were calculated in the same way as described above.

Outcome measures were quantitatively summarized, so we used the I^2 index to analyze heterogeneity. I^2 < 25% indicated that there was low heterogeneity and a fixed effect model was applied to pool effect size. If the I^2 value was > 25% and < 50%, it suggested that there was moderate heterogeneity. When I^2 was > 50%, it showed that there was significant heterogeneity[12]. To achieve more conservative results, a random effect model was applied in the last two situations[13]. Furthermore, subgroup analyses were performed. To identify sources of significant heterogeneity, sensitivity analysis was applied by removing individual study from the data set followed by analysis of the effect of their removal on the overall results.

RESULTS

From a total of 1763 records, 11 journal articles reporting 7 studies (*n* = 1972) were qualified for the criteria and included in this meta-analysis[6,14-23] (Figure 1).

General study characteristics

The 7 studies were respectively completed in the United Kingdom (3), Canada (1), and the United States (1), with multicenter investigations in 11 European countries (2) (Table 1). The studies were conducted on adults (mean age range 42-58 years across studies). All but one study[19,20] included more males than females (female/male ratio 1:0.7-1:3). The follow-up period ranged from 1-12 years in the 11 included articles. Among them, three had a follow-up period of 1 year[14,16,19], two had three years of follow-up[20,22], and the remaining studies had a follow-up period longer than 3 years. The studies included one 3-arm study and six 2-arm studies; five of them evaluated the effects of laparoscopic antireflux surgery[14-17,19,20,22,23] while two evaluated open

Table 1 Description of patients at baseline from the 11 articles included in meta-analysis

Ref.	Group	No. of patients randomized	Mean age (yr)	Female/male	Mean follow-up years	No. of patients at follow-up	Country
Mahon *et al*[14]	LNF	109	48	1:1.9	1	106	United Kingdom
	PPIs	108	47	1:2.6		97	
Mehta *et al*[15]	LNF	91	47	1:2.0	6.9	91	United Kingdom
	PPIs	92	47	1:2.5		38	
Grant *et al*[16]	LF	178	46.7	1:1.9	1	145	United Kingdom
	Best medical management(mainly PPIs)	179	45.9	1:2.2		154	
Grant *et al*[17]	LF	178	46.7	1:1.9	5	127	United Kingdom
	Best medical management(mainly PPIs)	179	45.9	1:2.2		119	
Lundell *et al*[6]	Open antireflux surgery	155	51	1:3.2	5	79	Sweden
	omeprazole	155	55	1:3.0		104	
Lundell *et al*[18]	Open antireflux surgery	155	51	1:3.2	12	53	Sweden
	omeprazole	155	55	1:3.0		71	
Anvari *et al*[19]	LNF	52	42.9	1:0.7	1	48	Canada
	PPIs	52	42.1	1:1		48	
Anvari *et al*[20]	LNF	52	42.9	1:0.7	3	49	Canada
	PPIs	52	42.1	1:1		44	
Spechler *et al*[21]	Open Nissen fundoplication	82	58	NR	9.1	38	United States
	Antacid and ranitidin	77	58	NR	10.6	91	
Lundell *et al*[22]	Laparoscopic antireflux surgery	288	44.8	1:2.2	3	204	11 European countries
	Esomeprazole	266	45.4	1:3.0		208	
Galmiche *et al*[23]	Laparoscopic antireflux surgery	288	44.8	1:2.2	5	180	11 European countries
	Esomeprazole	266	45.4	1:3.0		192	

In the 11 articles, Ref. [16,17] were one study of different follow-up times, so were Ref. [6,18-20,22,23]. LNF: Laparoscopic Nissen fundoplication; PPIs: Proton pump inhibitors; LF: Laparoscopic fundoplication.

Table 2 Summary of inclusion criteria and relevant outcomes for the 7 studies included in meta-analysis

Ref.	Inclusion criteria	Relevant outcomes
Mahon *et al*[14]	Symptoms of GERD for at least 6 mo and pathologic reflux	PGWBI, GI well-being score, DeMeester score, pH testing
Mehta *et al*[15]	Symptoms of GERD for at least 6 mo and pathologic reflux pathologic reflux	DeMeester symptom score
Grant *et al*[16,17]	Symptoms of GERD requiring maintenance treatment with PPI (or other) for at least 12 mo, pathologic reflux	REFLUX quality-of life-score, EQ-5D, SF-36
Lundell *et al*[6,18]	Chronic GERD with concomitant esophagitis and response to omeprazole	PGWBI, GSRS
Anvari *et al*[19,20]	Chronic GERD, PPI for at least 1 yr, pathologic reflux, symptom control	GERSS, 24 h pH monitoring, EQ-5D
Spechler *et al*[21]	Complicated GERD(abnormal acid reflux at least 1 of the 4 disorders: Barrett esophagus, peptic esophageal ulcer, peptic esophageal stricture, or erosive esophagitis)	GRACI scores, 24 h pH monitoring, SF-36
Lundell *et al*[22] and Galmiche *et al*[23]	Chronic GERD, assessed through endoscopy, 24 h pH-metry and response to esomeprazole. Esophagitis no more than LA grade B, GERD symptoms no more than mild	GSRS

GERD: Gastroesophageal reflux disease; PPI: Proton pump inhibitor.

surgery[6,18,21]. The GERD criteria in the included studies ranged from mild symptoms[22,23] to complicated GERD (Table 2). In all but one study, PPIs were used. In two studies, the response to PPI treatment was an inclusion criterion[6,18]. Participation in two additional studies required treatment with PPI for at least 1 year[17,19].

Risk of bias assessment

The results of the risk of bias assessment are presented in Table 3. The sequence generation for randomization was adequate in all but two studies[6,18,22,23]. Concealment of group allocation was not clear in four[6,14,15,18,22,23]. Healthcare providers and patients were not blinded in any study. Observer bias could not be excluded because none of the studies reported blinded outcome assessment. Four studies used intention-to-treat analysis to avoid attrition bias[6,15-18,21-23].

In the 7 studies, the individuals were analyzed in the groups to which they were randomized. Across these trials, the percentage of patients not analyzed ranged from 6% to 42% in the medical group and from 0% to 34% in the surgical group. Two studies[14,19,20] did not address the missing continuous outcome data appropriately. Figure 2 shows the summary of the risk

Table 3 Summary of risk of bias assessment for the 7 studies included in meta-analysis

Ref.	Sequence generation adequate	Allocation concealed	Blinding		Intention-to-treat analysis performed	Exact data available
			Patients	Healthcare providers		
Mahon *et al*[14]	Yes	Unclear	No	No	No	Yes
Mehta *et al*[15]	Yes	Unclear	No	No	Yes	Yes
Grant *et al*[16,17]	Yes	Yes	No	No	Yes	Yes
Lundell *et al*[6,18]	Unclear	Unclear	No	No	Yes	Yes
Anvari *et al*[19,20]	Yes	Yes	No	No	No	Yes
Spechler *et al*[21]	Yes	Yes	No	No	Yes	Yes
Lundell *et al*[22] and Galmiche *et al*[23]	Unclear	Unclear	No	No	Yes	Yes

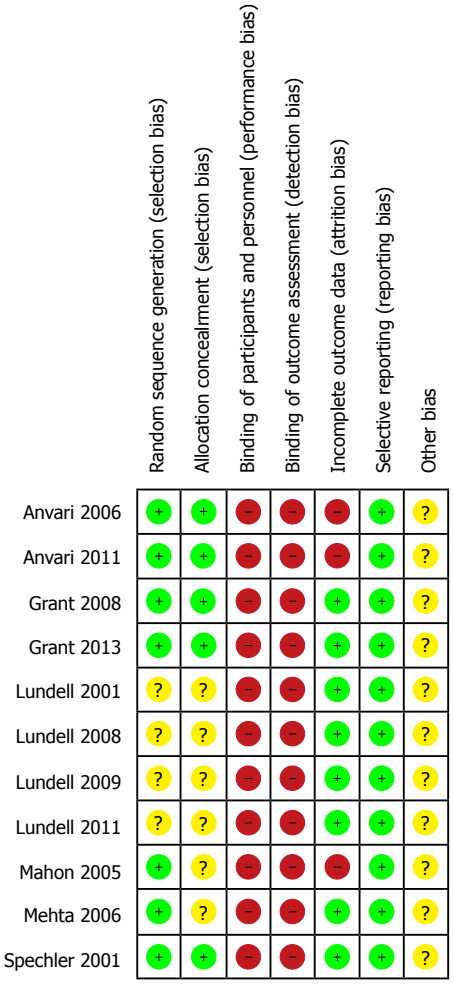


Figure 2 Risk of bias assessment.

of bias assessment for each study.

Health-related quality of life

Health-related quality of life was reported in five studies. Four of them reported higher health-related quality of life in the surgical group than in the medical group, with the differences being statistically significant in 2 studies. One reported a slightly negative effect of surgery on health-related quality of life[6]. Meta-analysis favored antireflux surgery over medical treatment concerning health-related quality of life (SMD = 0.18; 95%CI: 0.01 to 0.34), with no significant heterogeneity observed (I^2 = 36%; P = 0.17) (Figure 3). It is obvious that the general well-being of patients who received antireflux surgery improved.

Subgroup analyses: Health-related quality of life was reported to be higher in the surgical group than in the medical group when the follow-up period was < 3 years; but this was not significantly different (SMD = 0.16; 95%CI: -0.04 to 0.36) with no obvious heterogeneity observed (I^2 = 25%; P = 0.26). Likewise, no significant difference was found between the surgical and medical group when the follow-up period was ⩾ 3 years (SMD = 0.23; 95%CI: -0.10 to 0.56); in this situation the heterogeneity was high (I^2 = 61%; P = 0.08) (Figure 3).

GERD-related quality of life

GERD-related quality of life was reported in six studies. All of them reported higher GERD-related quality of life in the surgical group than in the medical group and the differences were statistically significant in 2 studies. Meta-analysis revealed a statistically significant difference in the two groups (SMD = 0.35; 95%CI: 0.11 to 0.59), but the heterogeneity was high (I^2 = 76.0%; P = 0.001) (Figure 4). GERD-related quality of life was found to improve more in the patients who underwent surgery.

Subgroup analyses: GERD-related quality of life was significantly different between the surgical and medical groups when the follow-up period was < 3 years (SMD = 0.45; 95%CI: 0.23 to 0.66), with no significant heterogeneity observed (I^2 = 26.0%; P = 0.25). The difference was not as significant when the follow-up period was ⩾ 3 years (SMD = 0.30; 95%CI: -0.10 to 0.69) with significant heterogeneity observed (I^2 = 84.0%; P < 0.001) (Figure 4).

DeMeester scores and the percentage of time that pH < 4

The percentage of time that pH < 4 was reported in three studies[14,19,21]. All of them reported a higher percentage of time that pH < 4 in the medical group, with one being statistically significant. Meta-analysis revealed a statistically significant difference in this parameter (SMD = 0.38; 95%CI: 0.14 to 0.61) with no heterogeneity observed (I^2 = 0%; P = 0.91) (Figure 5). It showed that antireflux surgery was better than

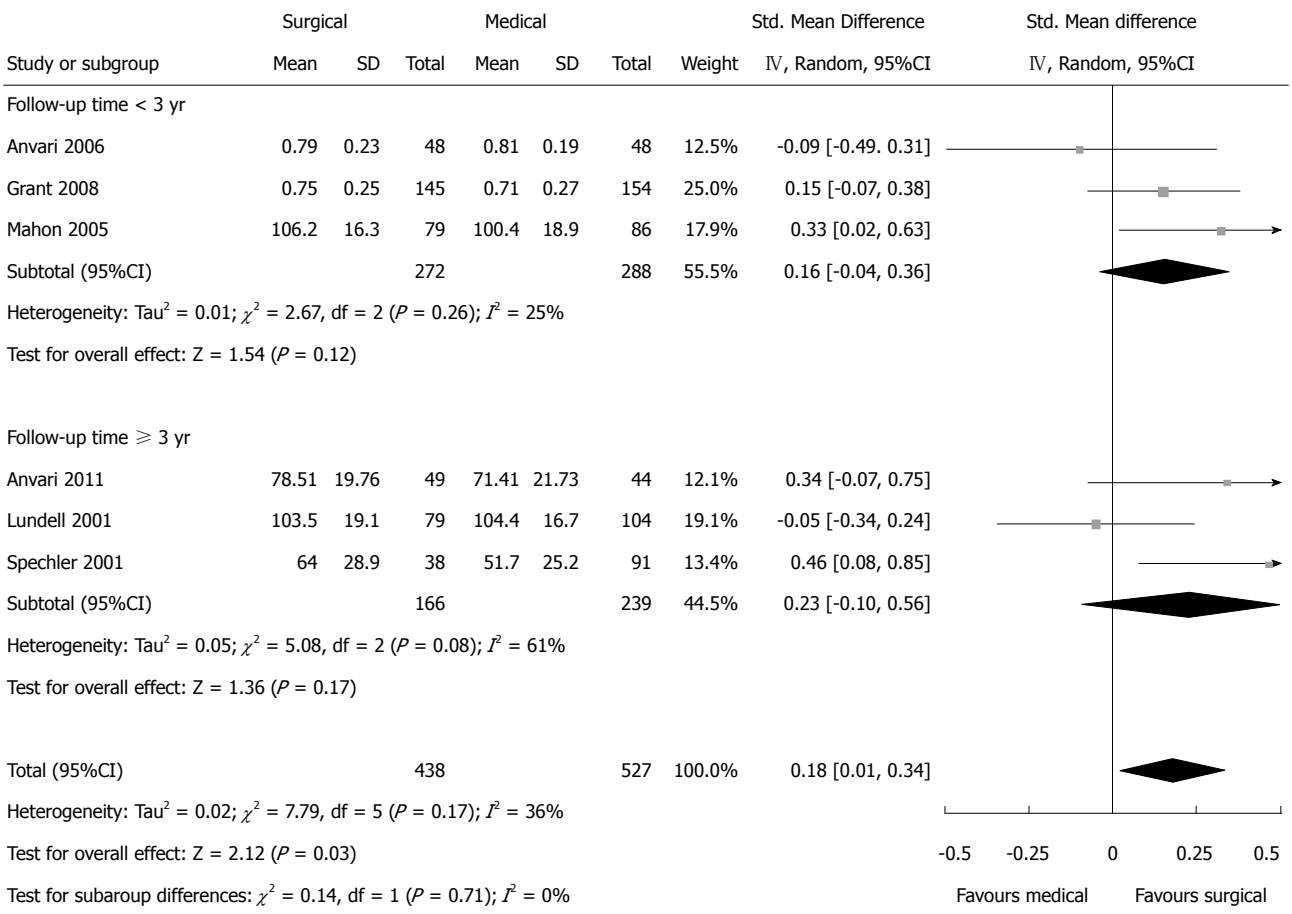


Figure 3 Meta-analysis Forest plot concerning health-related quality of life.

medical treatment in controlling pH. DeMeester scores were reported in two studies[14,15]. Both of them reported lower DeMeester scores in the surgical group than in the medical group. Meta-analysis showed a better effect of antireflux surgery than medical treatment on DeMeester scores (SMD = 0.32; 95%CI: 0.00 to 0.65), the heterogeneity was not significant (I^2 = 44%; P = 0.18) (Figure 6). This outcome showed the positive role of antireflux surgery in alleviating acid exposure.

Sensitivity analysis

To assess whether a single study could influence the main results, we excluded each study individually to evaluate its influence on the pooled effect size and the heterogeneity of the main results. When analyzing GERD-related quality of life, exclusion of the study of Grant *et al*[16] and Lundell *et al*[22] resulted in a loss of the statistical significance of the pooled SMD (Figure 7). Removing any individual study from the data set did not substantially influence the general health-related quality of life (Figure 8).

Publication bias

Funnel plot was used to assess the outcome of health-related quality of life. The *P*-value calculated by Begg's test was 0.707, indicating no serious publication bias (Figure 9). However, given that the number of studies was limited, publication bias was judged as not clear[24].

DISCUSSION

This meta-analysis indicates that antireflux surgery may be more effective than medical therapy for GERD in improving health-related quality of life and GERD-related quality of life when the follow-up period was < 3 years. Antireflux surgery could also result in a lower percentage of time that pH < 4 and lower DeMeester scores.

The difference in the GERD-related quality of life is obvious for one to three years, but it declines over time. The difference was not significant in the health-related quality of life between the surgical and medical group over longer time. Therefore, the long-term sustainability of the surgical benefits is still unclear.

No perioperative or surgery-associated deaths have been reported during the follow-up period. However, one study observed a decrease in survival among the patients randomized to surgery[21], while another reported more deaths attributed to cardiac causes in the medical group[18]. The complication of dysphagia may increase after fundoplication as well, and the symptoms associated with antireflux surgery like bloating, inability to belch or vomit, and abdominal fullness were conflicting.

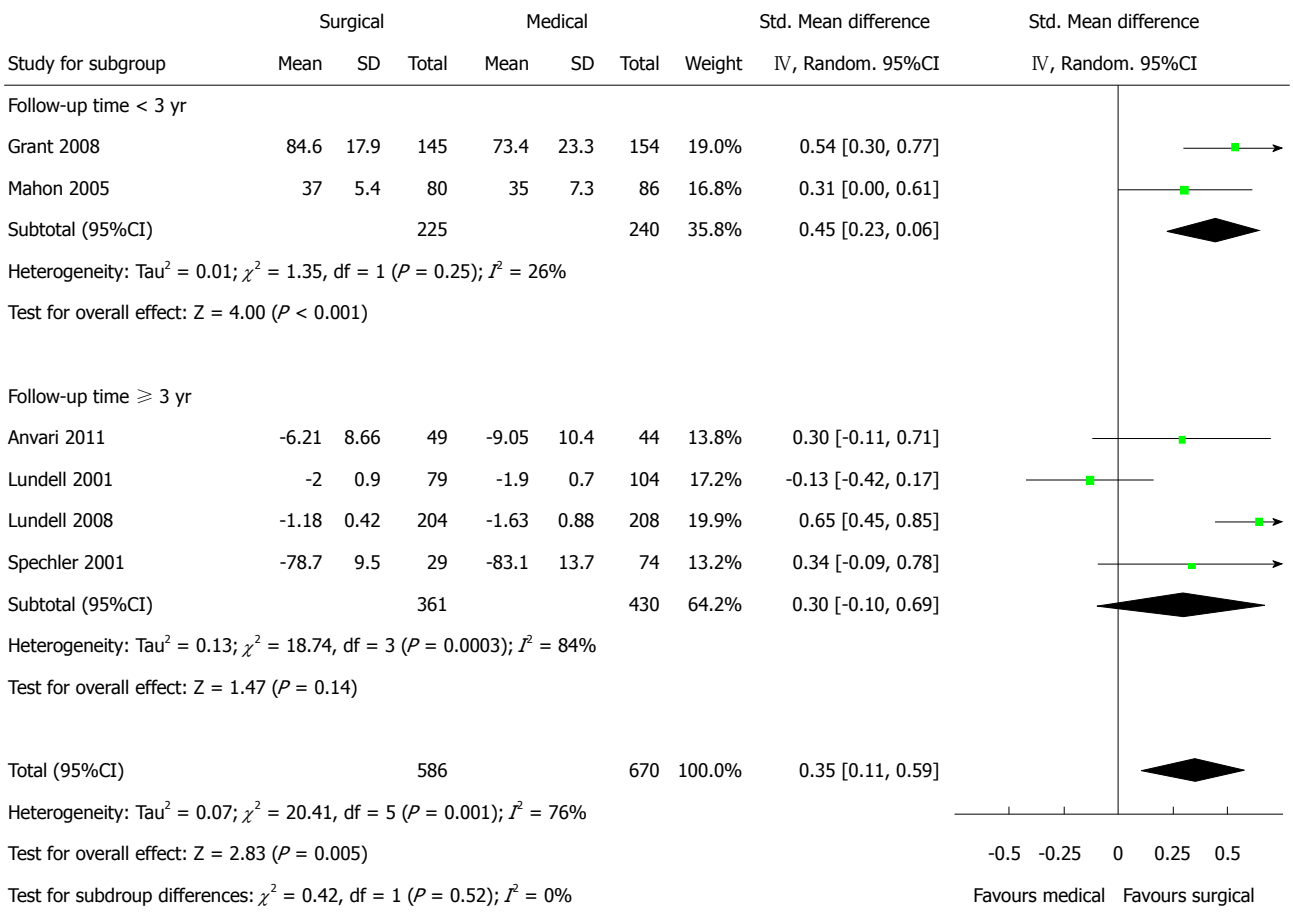


Figure 4 Meta-analysis Forest plot concerning gastroesophageal reflux disease-related quality of life.

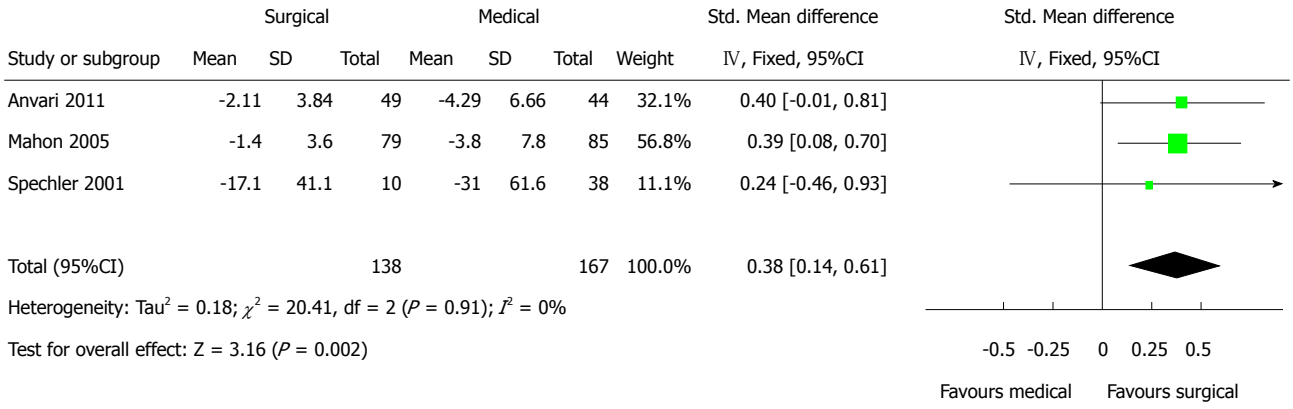


Figure 5 Meta-analysis Forest plot concerning % of time pH < 4.

The advantage of this meta-analysis is its comprehensive method to identifying all RCTs comparing antireflux surgery with medical treatment for GERD. The outcome of health-related quality of life proved to be robust through a variety of sensitivity analyses.

We also conducted analyses for the percentage of time that pH < 4 and DeMeester scores. Both analyses favored surgery over medical treatment. This might indicate that the main symptoms of GERD (heartburn and regurgitation) were less frequent after surgical intervention than medical management.

Through subgroup analyses, antireflux surgery in some ways may be more effective than medicine over short to medium follow-up time, but its effect declined over time. A published meta-analysis comparing surgery with medical management using PPI or H2 receptor antagonists[25] concluded that surgical treatment is more effective than medical management, and that surgery may be considered as a treatment alternative to long-term medical therapy. However, the review did not perform subgroup analyses based on follow-up time. We did not find evidence that the advantage

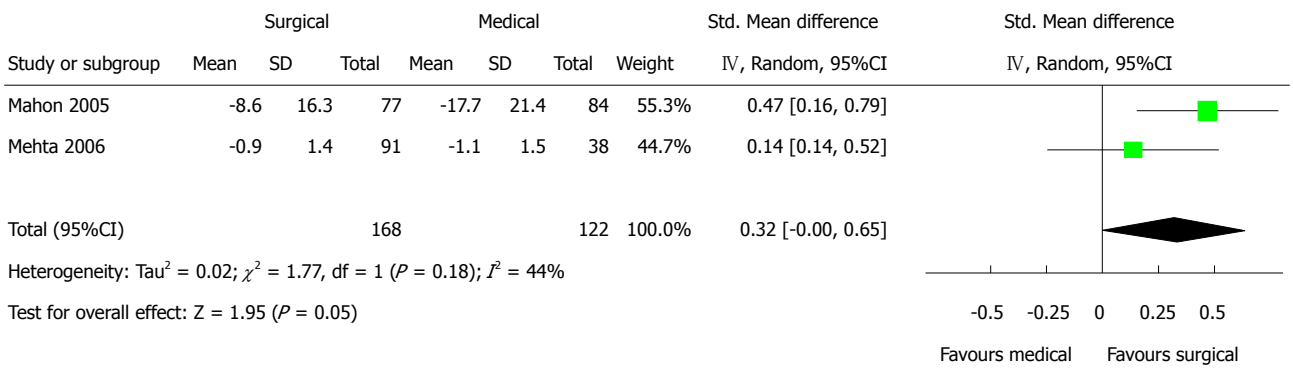


Figure 6 Meta-analysis Forest plot concerning DeMeester scores.

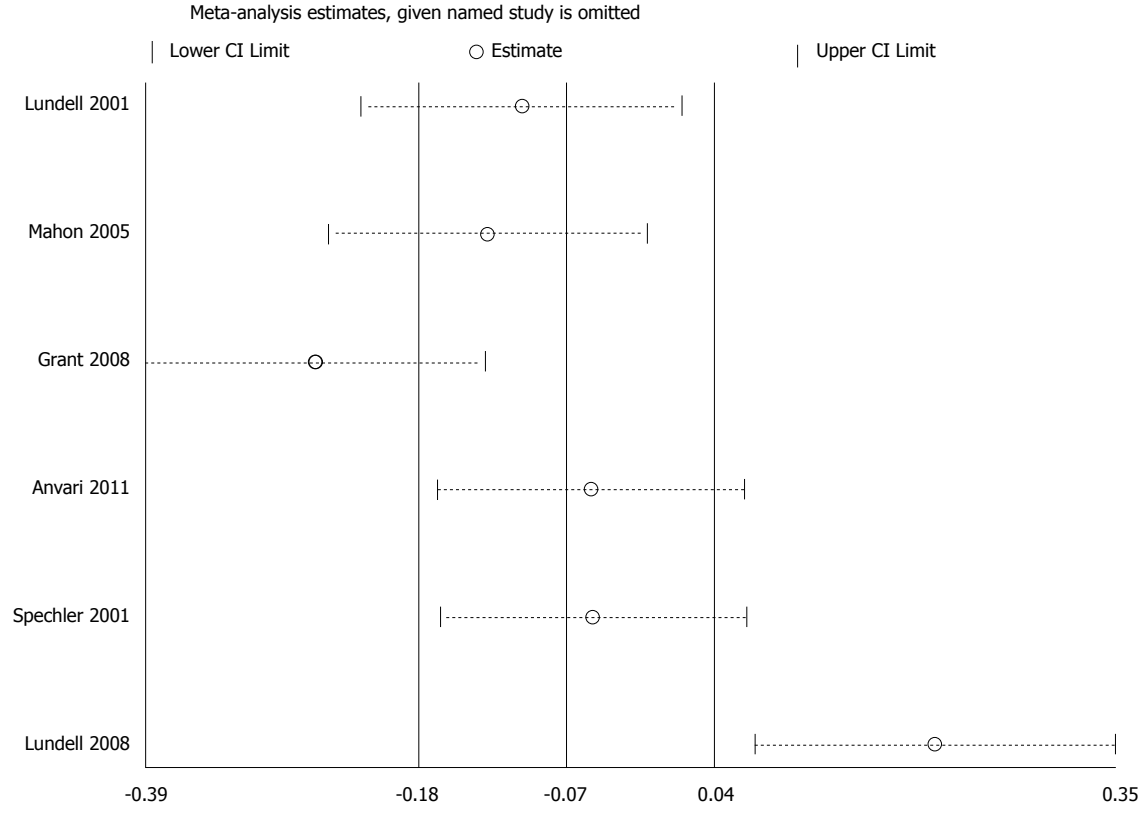


Figure 7 Sensitivity analysis concerning gastroesophageal reflux disease-related quality of life.

of antireflux surgery in effectiveness remained for long time. Because large majority of GERD patients do not experience erosive disease, given the possible adverse events associated with antireflux surgery, it should be considered only in selected GERD patients.

There are also some limitations in our study. The pooled effect sizes were just based on 7 studies. In addition, the methodological quality of 4 studies was judged as unclear allocation concealment. Two trials did not report their recruitment strategy adequately, making it possible that only a certain subgroup of the eligible patient population was included in these studies[6,18,22,23].

Our main results were based on questionnaires, which are sometimes affected by subjective factors. We could not obtain subjective values such as 24 h - pH testing, LES, and endoscopic results. The inclusion criteria of GERD in our included studies varied from mild symptoms to complicated GERD, making us unable to analyze subgroups based on the severity of GERD. Nor could we analyze GERD-related symptoms or patient satisfaction because the methods each study used were different.

Long-term acid suppression may result in some complications[26,27]. However, nowadays no specific serious adverse events have been judged to be attributable to acid suppressive treatment alone. This might be because the follow-up period was insufficient, and therefore, we couldn't draw any conclusion concerning adverse events except for the longer period.

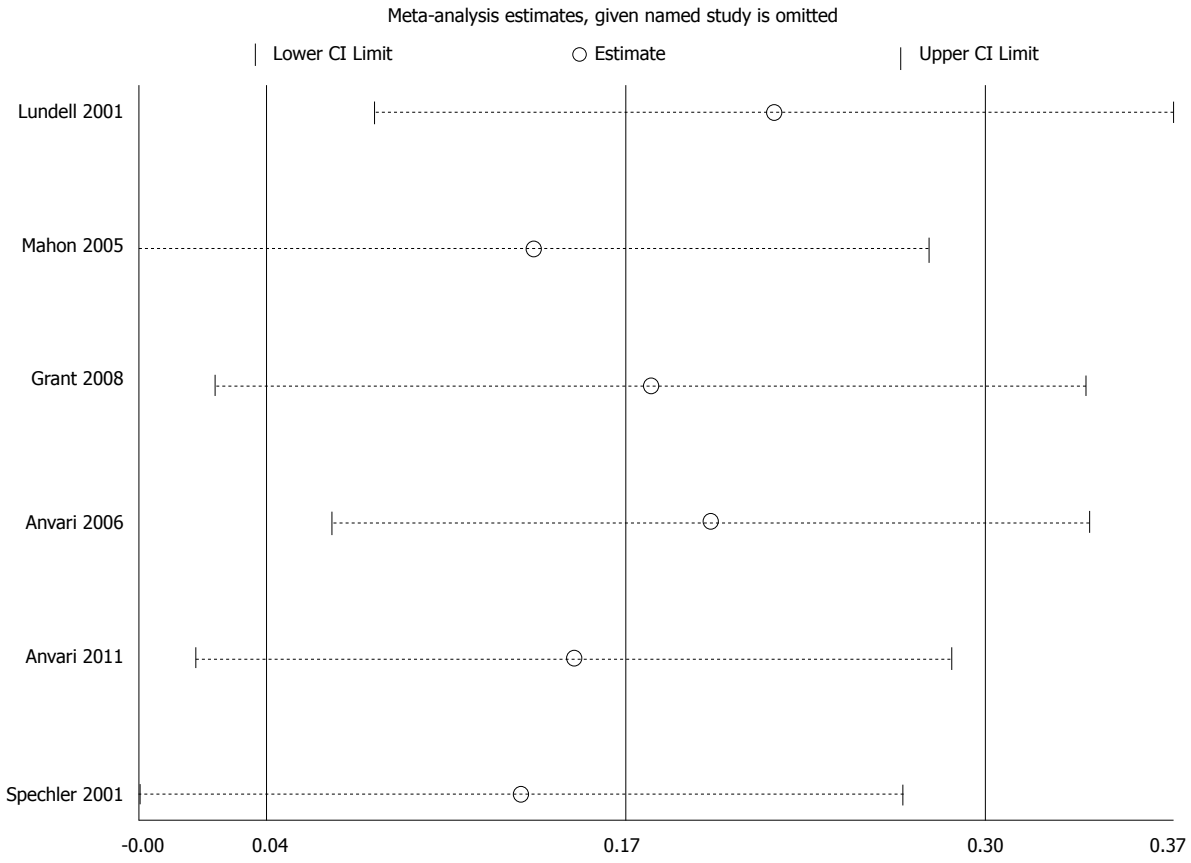


Figure 8 Sensitivity analysis concerning health-related quality of life.

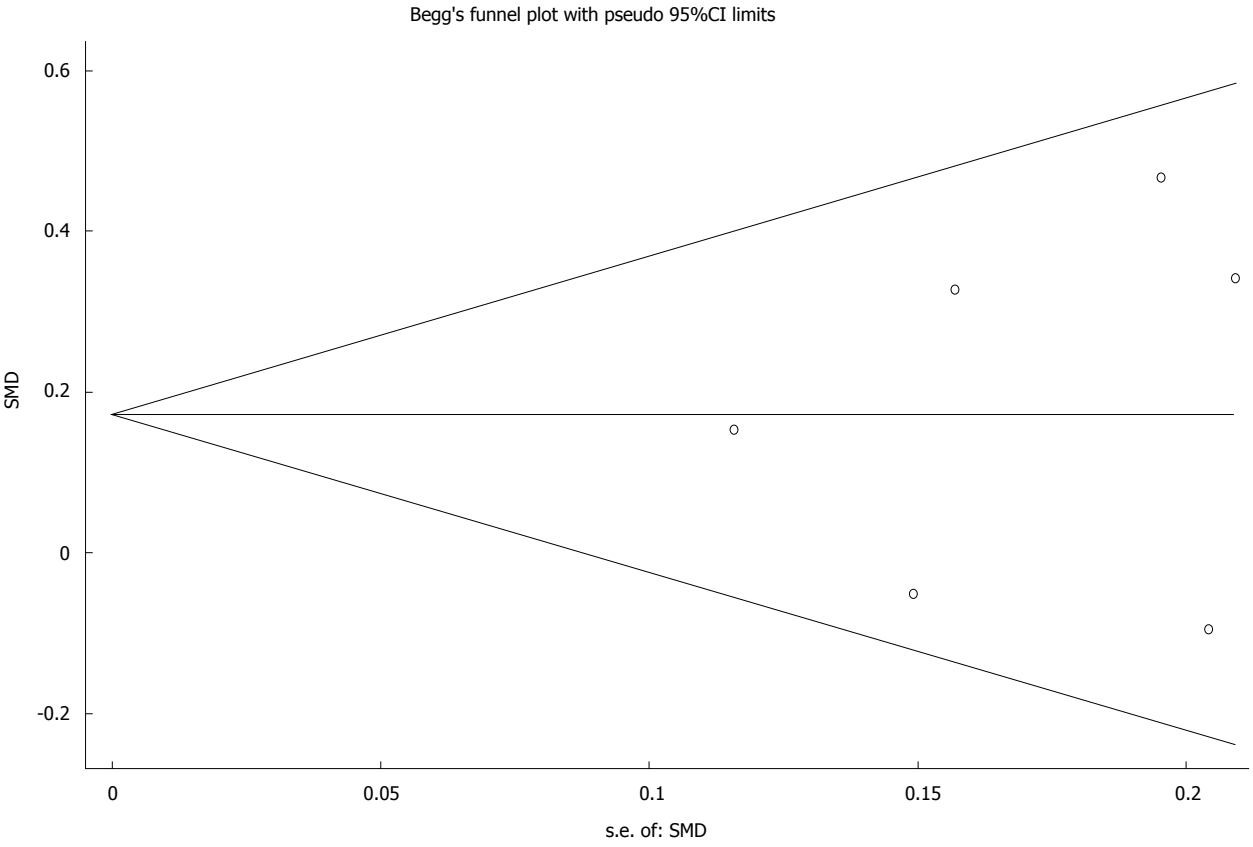


Figure 9 Funnel plot of health-related quality of life. SMD: Standardized mean difference.

In two of the included studies, only PPI responders were enrolled. Due to the limited data on PPI responders and patients who initially were partially or completely refractory to PPI therapy, we presented no results stratified by it. Nevertheless, these poor responders were a heterogeneous group with lots of underlying reasons for their non-responsiveness to acid suppressive therapy[28]. The most common reason was the absence of actual reflux disease, with symptoms resulting from non-reflux conditions.

Surgery also appears to be a relatively safe option if performed by an experienced surgeon. There were few cases of intraoperative morbidity and postoperative complications reported in the studies. Recently, surgical therapy has become a option for long-term treatment in GERD patients, which is generally not recommended for patients resistant to PPI therapy[29]. In contrast, the National Institute for Health and Care Excellence does not suggest surgery to the patients with persistent GERD as the routine management[30].

The outcomes of our study apply to a general population of GERD individuals over 14 years of age. In the 7 studies, the proportion of females ranged from 25% to 59%. No information was available on body mass index. The results might be different for specific subgroups.

In summary, this meta-analysis shows comprehensive evidence that open or laparoscopic antireflux surgery might be more effective than medical treatment for GERD in terms of GERD - and health-related quality of life, especially for 3 years of follow-up with GERD-related quality of life. However, the advantage of antireflux surgery in effectiveness may not remain for long time. In addition, the differences in long term adverse events between surgery and medical management remain unclear. Overall, we conclude that for most GERD patients, medical treatment may be the first choice, and that surgical intervention should be considered only after critical evaluation by experienced doctors. Further studies are warranted on this topic.

COMMENTS

Background

Gastroesophageal reflux disease (GERD) is a highly prevalent chronic disorder of the gastrointestinal tract. The main treatments for GERD include medical treatment and antireflux surgery. However, the effect of medical treatment and antireflux surgery on GERD symptoms remains unclear.

Research frontiers

Current treatment options for GERD include medical treatment and antireflux surgery. Several randomized controlled trials (RCTs) comparing pharmaceutical treatment with surgery over a longer term have been completed. Worldwide research is directed at treatment options that maximize safety and effectiveness.

Innovations and breakthroughs

In the present study, the authors analyzed the effect of antireflux surgery and medical treatment for GERD by pooling results from different RCTs. They conducted subgroup analyses to investigate GERD and health-related quality of life to systematically and comprehensively evaluate the effect of the two treatments.

Applications

The present meta-analysis allows comprehensive understanding of the roles of antireflux surgery and medical treatment for GERD.

Peer-review

It is a good job, up to date results of medical and surgical treatment of gastroesophageal reflux.

REFERENCES

1 **Kahrilas PJ**. Gastroesophageal reflux disease. *JAMA* 1996; **276**: 983-988 [PMID: 8805734]

2 **Orlando RC**. Reflux esophagitis: overview. *Scand J Gastroenterol Suppl* 1995; **210**: 36-37 [PMID: 8578203]

3 **Camilleri M**, Dubois D, Coulie B, Jones M, Kahrilas PJ, Rentz AM, Sonnenberg A, Stanghellini V, Stewart WF, Tack J, Talley NJ, Whitehead W, Revicki DA. Prevalence and socioeconomic impact of upper gastrointestinal disorders in the United States: results of the US Upper Gastrointestinal Study. *Clin Gastroenterol Hepatol* 2005; **3**: 543-552 [PMID: 15952096]

4 **Katz PO**, Gerson LB, Vela MF. Guidelines for the diagnosis and management of gastroesophageal reflux disease. *Am J Gastroenterol* 2013; **108**: 308-328; quiz 329 [PMID: 23419381 DOI: 10.1038/ajg.2012.444]

5 **Gerson LB**, Boparai V, Ullah N, Triadafilopoulos G. Oesophageal and gastric pH profiles in patients with gastro-oesophageal reflux disease and Barrett's oesophagus treated with proton pump inhibitors. *Aliment Pharmacol Ther* 2004; **20**: 637-643 [PMID: 15352912]

6 **Lundell L**, Miettinen P, Myrvold HE, Pedersen SA, Liedman B, Hatlebakk JG, Julkonen R, Levander K, Carlsson J, Lamm M, Wiklund I. Continued (5-year) followup of a randomized clinical study comparing antireflux surgery and omeprazole in gastroesophageal reflux disease. *J Am Coll Surg* 2001; **192**: 172-179; discussion 172-179 [PMID: 11220717]

7 **Khalili H**, Huang ES, Jacobson BC, Camargo CA, Feskanich D, Chan AT. Use of proton pump inhibitors and risk of hip fracture in relation to dietary and lifestyle factors: a prospective cohort study. *BMJ* 2012; **344**: e372 [PMID: 22294756 DOI: 10.1136/bmj.e372]

8 Guidelines for surgical treatment of gastroesophageal reflux disease (GERD). Society of American Gastrointestinal Endoscopic Surgeons (SAGES). *Surg Endosc* 1998; **12**: 186-188 [PMID: 9479742]

9 **Zacharoulis D**, O'Boyle CJ, Sedman PC, Brough WA, Royston CM. Laparoscopic fundoplication: a 10-year learning curve. *Surg Endosc* 2006; **20**: 1662-1670 [PMID: 17024541]

10 **Wileman SM**, McCann S, Grant AM, Krukowski ZH, Bruce J. Medical versus surgical management for gastro-oesophageal reflux disease (GORD) in adults. *Cochrane Database Syst Rev* 2010; **17**: CD003243 [PMID: 20238321 DOI: 10.1002/14651858.CD003243.pub2]

11 **Begg CB**, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994; **50**: 1088-1101 [PMID: 7786990]

12 **Higgins JP**, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557-560 [PMID: 12958120]

13 **DerSimonian R**, Levine RJ. Resolving discrepancies between a meta-analysis and a subsequent large controlled trial. *JAMA* 1999; **282**: 664-670 [PMID: 10517720]

14 **Mahon D**, Rhodes M, Decadt B, Hindmarsh A, Lowndes R, Beckingham I, Koo B, Newcombe RG. Randomized clinical trial of laparoscopic Nissen fundoplication compared with proton-pump inhibitors for treatment of chronic gastro-oesophageal reflux. *Br J Surg* 2005; **92**: 695-699 [PMID: 15898130]

15 **Mehta S**, Bennett J, Mahon D, Rhodes M. Prospective trial of laparoscopic nissen fundoplication versus proton pump inhibitor therapy for gastroesophageal reflux disease: Seven-year follow-up. *J Gastrointest Surg* 2006; **10**: 1312-1316; discussion 1316-1317

[PMID: 17114017]
16 **Grant AM**, Wileman SM, Ramsay CR, Mowat NA, Krukowski ZH, Heading RC, Thursz MR, Campbell MK. Minimal access surgery compared with medical management for chronic gastro-oesophageal reflux disease: UK collaborative randomised trial. *BMJ* 2008; **337**: a2664 [PMID: 19074946 DOI: 10.1136/bmj.a2664]
17 **Grant AM**, Cotton SC, Boachie C, Ramsay CR, Krukowski ZH, Heading RC, Campbell MK. Minimal access surgery compared with medical management for gastro-oesophageal reflux disease: five year follow-up of a randomised controlled trial (REFLUX). *BMJ* 2013; **346**: f1908 [PMID: 23599318 DOI: 10.1136/bmj.f1908]
18 **Lundell L**, Miettinen P, Myrvold HE, Hatlebakk JG, Wallin L, Engström C, Julkunen R, Montgomery M, Malm A, Lind T, Walan A. Comparison of outcomes twelve years after antireflux surgery or omeprazole maintenance therapy for reflux esophagitis. *Clin Gastroenterol Hepatol* 2009; **7**: 1292-1298; quiz 1260 [PMID: 19490952 DOI: 10.1016/j.cgh.2009.05.021]
19 **Anvari M**, Allen C, Marshall J, Armstrong D, Goeree R, Ungar W, Goldsmith C. A randomized controlled trial of laparoscopic nissen fundoplication versus proton pump inhibitors for treatment of patients with chronic gastroesophageal reflux disease: One-year follow-up. *Surg Innov* 2006; **13**: 238-249 [PMID: 17227922]
20 **Anvari M**, Allen C, Marshall J, Armstrong D, Goeree R, Ungar W, Goldsmith C. A randomized controlled trial of laparoscopic Nissen fundoplication versus proton pump inhibitors for the treatment of patients with chronic gastroesophageal reflux disease (GERD): 3-year outcomes. *Surg Endosc* 2011; **25**: 2547-2554 [PMID: 21512887 DOI: 10.1007/s00464-011-1585-5]
21 **Spechler SJ**, Lee E, Ahnen D, Goyal RK, Hirano I, Ramirez F, Raufman JP, Sampliner R, Schnell T, Sontag S, Vlahcevic ZR, Young R, Williford W. Long-term outcome of medical and surgical therapies for gastroesophageal reflux disease: follow-up of a randomized controlled trial. *JAMA* 2001; **285**: 2331-2338 [PMID: 11343480]
22 **Lundell L**, Attwood S, Ell C, Fiocca R, Galmiche JP, Hatlebakk J, Lind T, Junghard O. Comparing laparoscopic antireflux surgery with esomeprazole in the management of patients with chronic gastro-oesophageal reflux disease: a 3-year interim analysis of the LOTUS trial. *Gut* 2008; **57**: 1207-1213 [PMID: 18469091 DOI: 10.1136/gut.2008.148833]
23 **Galmiche JP**, Hatlebakk J, Attwood S, Ell C, Fiocca R, Eklund S, Långström G, Lind T, Lundell L. Laparoscopic antireflux surgery vs esomeprazole treatment for chronic GERD: the LOTUS randomized clinical trial. *JAMA* 2011; **305**: 1969-1977 [PMID: 21586712 DOI: 10.1001/jama.2011.626]
24 **Egger M**, Smith GD, Altman DG. Systematic reviews in health care: meta-analysis in context, second edition. London: BMJ Books, 2001 [DOI: 10.1002/9780470693926]
25 **Rickenbacher N**, Kötter T, Kochen MM, Scherer M, Blozik E. Fundoplication versus medical management of gastroesophageal reflux disease: systematic review and meta-analysis. *Surg Endosc* 2014; **28**: 143-155 [PMID: 24018760 DOI: 10.1007/s00464-013-3140-z]
26 **Moayyedi P**, Cranney A. Hip fracture and proton pump inhibitor therapy: balancing the evidence for benefit and harm. *Am J Gastroenterol* 2008; **103**: 2428-2431 [PMID: 18855852 DOI: 10.1111/j.1572-0241.2008.02031.x]
27 **Lodato F**, Azzaroli F, Turco L, Mazzella N, Buonfiglioli F, Zoli M, Mazzella G. Adverse effects of proton pump inhibitors. *Best Pract Res Clin Gastroenterol* 2010; **24**: 193-201 [PMID: 20227032 DOI: 10.1016/j.bpg.2009.11.004]
28 **Zaninotto G**, Attwood SE. Surgical management of refractory gastro-oesophageal reflux. *Br J Surg* 2010; **97**: 139-140 [PMID: 20069606 DOI: 10.1002/bjs.6863]
29 **Lundell L**. Borderline indications and selection of gastroesophageal reflux disease patients: 'Is surgery better than medical therapy'? *Dig Dis* 2014; **32**: 152-155 [PMID: 24603401 DOI: 10.1159/000357182]
30 **North of England Dyspepsia Guideline Development Group (UK)**. Dyspepsia: managing dyspepsia in adults in primary care. National Institute for Health and Clinical Excellence: Guidance. 2004 [PMID: 21678625]

P- Reviewer: Lara FJP, Skrypnyk IN
S- Editor: Ji FF **L- Editor**: A **E- Editor**: Jiao XK

World Journal of Meta-Analysis

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.13105/wjma.v3.i6.295

World J Meta-Anal 2015 December 26; 3(6): 295-303
ISSN 2308-3840 (online)

META-ANALYSIS

Radiofrequency ablation with or without transarterial chemoembolization for hepatocellular carcinoma: A systematic review and meta-analysis

Ming-Zheng Hu, Shao-Fang Li

Ming-Zheng Hu, Shao-Fang Li, Institute of Hepatopancreatobiliary Surgery, Yichang Central People's Hospital, China Three Gorges University, Yichang 443000, Hubei Province, China

Shao-Fang Li, Department of Gastroenterology, Zhijiang People's Hospital, Yichang 443200, Hubei Province, China

Author contributions: Hu MZ and Li SF conceived and designed the study, acquired, analyzed and interpreted data, and drafted the article; both authors have approved the final version to be published

Supported by Hubei Provincial Health and Family Planning Scientific Research Project, No. QJX2012-42.

Conflict-of-interest statement: The authors declared that there is no conflict of interest in this study.

Data sharing statement: No additinal data are available.

Correspondence to: Dr. Ming-Zheng Hu, Institute of Hepatopancreatobiliary Surgery, Yichang Central People's Hospital, China Three Gorges University, No. 183, Yiling Road, Yichang 443000, Hubei Province, China. hmz08@sina.com.cn
Telephone: +86-71-76485275
Fax: +86-71-76482302

Received: May 14, 2015
Peer-review started: May 15, 2015
First decision: June 19, 2015
Revised: August 27, 2015
Accepted: November 30, 2015
Article in press: December 3, 2015

Published online: December 26, 2015

Abstract

AIM: To determine whether combined transarterial chemoembolization (TACE) and radiofrequency ablation (RFA) improve overall and recurrence-free survival (RFS) compared with RFA alone.

METHODS: We reviewed randomized clinical trials (RCTs) comparing overall survival rate as well as recurrence-free rate for hepatocellular carcinoma (HCC) between TACE-RFA therapy and RFA alone published before April 2015 by conducting a systematic review and meta-analysis. Eligible studies were identified by searching PubMed and EMBASE up to April 2015. Additional studies were retrieved *via* China Medical Collections, Google Scholar or a hand review of the reference lists of the retrieved articles. The summarized relative risks (RRs) with their 95%CIs were estimated using random-effects model. I^2 statistic was calculated to measure the heterogeneity of RRs across studies and Cochran's Q test was used to test the statistical significance accordingly. Publication bias was assessed primarily based on visual assessment using a funnel plot, and secondly by using Egger's regression asymmetry test or Begg's rank correlation test as appropriate. Meta-regression was implemented to examine potential effect modifiers.

RESULTS: Nine single-center RCTs conducted in China and Japan were included, with a total of 618 patients with HCC; 321 of whom (51.9%) received TACE/RFA therapy and 297 received RFA alone. The pooled RRs with corresponding CIs comparing combined TACE/RFA to RFA alone were 1.12 (1.004-1.26) and 1.20 (1.02-1.41) for 1-year and 3-year survival rates, respectively. Similar positive associations were found for 1-year (1.19; 1.05-1.35) and 3-year (1.44; 1.00-2.07) RFS. The

beneficial effect was more evident in patients with medium-sized (3-5 cm) tumors and among the Chinese population.

CONCLUSION: Combined TACE/RFA has a beneficial effect on survival and recurrence rates compared with RFA alone, especially for medium-sized HCC and among Chinese patients.

Key words: Radiofrequency ablation; Hepatocellular carcinoma; Transarterial chemoembolization; Meta-analysis; Randomized clinical trial

Core tip: A systematic review and meta-analysis were conducted to determine whether combined transarterial chemoembolization (TACE) and radiofrequency ablation (RFA) improve overall and recurrence-free survival compared with RFA alone. Nine single-center randomized controlled trials were included, with a total of 618 patients with hepatocellular carcinoma (HCC), 321 of whom (51.9%) received TACE/RFA and 297 received RFA alone. We found that combined TACE/RFA has a beneficial effect on survival and recurrence rates compared with RFA alone, especially for medium-sized (3-5 cm) HCC and among Chinese patients.

Hu MZ, Li SF. Radiofrequency ablation with or without transarterial chemoembolization for hepatocellular carcinoma: A systematic review and meta-analysis. *World J Meta-Anal* 2015; 3(6): 295-303 Available from: URL: http://www.wjgnet.com/2308-3840/full/v3/i6/295.htm DOI: http://dx.doi.org/10.13105/wjma.v3.i6.295

INTRODUCTION

Hepatocellular carcinoma (HCC) is the fifth most common cancer, and one of the leading causes of cancer-related deaths throughout the world. More than 700000 incident cases are diagnosed and > 600000 deaths are attributed to HCC each year[1]. Surgical resection is the first-line therapeutic option for HCC patients with small solitary nodules without underlying cirrhosis[2-4]; however, its role in treating HCC is limited by strict inclusion criteria. Consequently, various non-surgical, liver-directed, locoregional therapies, such as radiofrequency ablation (RFA)[5], one of the therapies for HCC, and transcatheter arterial chemoembolization (TACE)[6], which involves administration of chemotherapy directly to the liver tumor *via* a catheter, have been developed as alternatives, particularly for patients with nonresectable HCC. Although local therapies may have less invasiveness, shorter hospital stay and lower associated mortality, their higher recurrence rates and lower disease-free survival rates are still major concerns. In the past decade, some evidence[7,8] has suggested that combining RFA with TACE improves overall survival (OS) rate and reduces recurrence rate, while other studies have not shown these[9-15].

One recent review found that combination of RFA with TACE increased 1-year survival rate by 114% and 5-year survival rate by 170%[16], which might have been overestimates of the true effect, due to the combination of retrospective cohort studies and randomized clinical trials (RCTs). Three other meta-analyses used odds ratio (OR) as a measure of effect size[17-19], which always exaggerates any "effect", especially when the survival rate in the control group is very high[20], in addition to its hard interpretability. Of note, RFA is effective in the treatment of small but not surgically resectable HCC (< 3 cm in diameter)[21], so combining RFA with TACE for treating small HCCs may not improve the efficacy compared with using RFA alone. It would be important for clinical practice to establish whether combination therapy improves the survival rate of patients with medium-sized (3-5 cm in diameter) or even large (> 5 cm in diameter) lesions. None of the previous studies has answered this question, and none has discussed whether combined therapy improves the recurrence-free survival (RFS) rate.

Therefore, in this study we investigated whether combined TACE/RFA improved OS and RFS rates, compared with RFA alone, especially for patients with medium-sized (3-5 cm) or large (> 5 cm) HCCs, by conducting a systematic review as well as meta-analysis of RCTs.

MATERIALS AND METHODS

Data sources and searches

This meta-analysis was based on a pre-specified protocol - the PRISMA Statement[22]. PubMed and Embase were systematically searched up to April 2015, with the following terms: "Carcinoma, Hepatocellular", "Liver Tumor", "Liver Cell Carcinoma", "Radiofrequency Ablation", "Transcatheter Arterial Chemoembolization" and "Clinical Trials". Studies published in Chinese were searched in Wanfang China Medical Collections (1990 to April 2015) using the corresponding Chinese terms. The references of the retrieved articles were also reviewed. In addition, Google scholar was used to give confirmation of the literature search.

Study selection

The titles and abstracts of all relevant studies were scanned independently by two authors (Hu MZ and Li SF). Reviews, case reports and letters to the editor were excluded. Studies were included if they met the following criteria: (1) RCTs that involved liver cancer patients; (2) patients were treated with RFA alone or combined with TACE; and (3) OS and/or RFS rate was reported in each group, or these data could be derived from the presented results.

Data abstraction

Two authors (Hu MZ and Li SF) independently reviewed the literature and extracted the following data from the included studies: last name of the first author; publication year; country where the study was conducted; number of participants; number of male patients; age; Child-Pugh class; tumor size and number; follow-up time; and 1-, 3- and/or 5-year OS or RFS rates after surgery. If the required data were not available in the primary article, we contacted the authors and requested *de novo* data. Disagreement on data extraction was resolved by group discussion.

Statistical analysis

The estimate of the principal effect was defined as the relative risk (RR) of OS rate or RFS rate, comparing the patients that were assigned to combined TACE/RFA therapy to those who received RFA alone. RR > 1 indicates that the combined therapy can benefit the survival/recurrence-free rate. I^2 statistic was calculated and used to define low, moderate, and high degrees of heterogeneity with 50% and 75% as the cutoffs. Cochran's Q test was used to test the statistical heterogeneity of RRs across studies, using 0.10 as the significance level. The summarized RRs with their 95%CIs were estimated using a random-effects model since high heterogeneity was shown in most of the pooled analyses[23].

Potential publication bias was assessed primarily based on visual assessment using a funnel plot and secondarily using Egger's regression asymmetry test (when the number of studies pooled was ⩾ 3) or Begg's rank correlation test (when the number of studies pooled = 2). Meta-regression was implemented to examine potential effect modifiers that may have affected the observed effects, including tumor size (< 3, 3-5 and > 5 cm) or study location (China or Japan)[24]. Sensitivity analyses were also conducted to evaluate the influence of each study included in the meta-analysis by omitting one study at each time, and statistical model selection in which we repeated the analyses with I^2 < 50% and P > 0.10 using a fixed-effects model.

The meta-analysis was performed with STATA statistical software version 13.0 (Stata Corporation, College Station, TX, United States). All statistical tests were two-sided and $P \leqslant 0.05$ was considered statistically significant, unless otherwise specified.

RESULTS

Flow of the included studies

Our literature search resulted in an initial set of 1494 publications from the PubMed database. Of these, 1366 were excluded at the title screening due to at least one of the following reasons: (1) not human studies (n = 37); (2) not original studies, for example, reviews, meta-analyses, letters to editors, or abstracts (n = 467); or (3) not clinical trials (n = 862). Among the remaining 128 studies, 115 were excluded after abstract review because: (1) they were *in vitro* studies, feasibility studies, or studies of diagnosis with computed tomography, magnetic resonance imaging or ultrasound (n = 22); (2) they were not RCTs (n = 29); (3) they did not compare survival or recurrence rates between RFA plus TACE and RFA alone (n = 58); or (4) they focused on metastases or complications (n = 6). Among the remaining 13 studies, the following were excluded after full-text review: five that did not compare efficacy; two that did not have a control; and one that was retracted by the journal. Four additional studies were found in other resources: one from the references of the relevant articles; two from China Medical Collections ("Wanfang" in Chinese); and one from Embase (Figure 1). Finally, a total of nine RCTs were included in the meta-analysis[7-15].

Features of RCTs

The main features of the trials included in the meta-analysis are shown in Table 1. Of the nine included studies, seven were published in English[7,8,10,12-15] and two in Chinese[9,11]. All the studies were conducted in Asia: six in China[7-11,13], and three in Japan[12,14,15]. The nine RCTs included 618 HCC patients; Three hundred and twenty-one of whom (51.9%) received combined TACE/RFA therapy and 297 RFA alone. The average follow-up time was 47.1 mo (range: 24-60 mo) among the nine studies.

All the studies were single-center trials. The majority of patients (74.1%) were male, ranging from 23.5%[10] to 84.8%[11]. There was no difference in the average proportion of male patients between the two groups (75.1% for the TACE/RFA group *vs* 73.1% for the RFQ alone group). The average age was 60.0 (range: 50.7-3.0) years, with no difference between the two groups (59.8 years for the TACE/RFA group *vs* 60.3 years for the RFA alone group), either.

According to the available data, the distribution of Child-Pugh class A, B and C for liver function was 80.6%, 18.9% and 0.5%, respectively. There was no significant difference in this distribution (80.0%/19.6%/0.4% *vs* 81.1%/18.2%/0.7%) across two comparison groups. About 23.0% of patients had more than two tumors (25.1% for the TACE/RFA group *vs* 21.0% for the RFA alone group). Patients in three studies each had average tumor size > 5 cm, 3-5 cm, and < 3 cm. The average tumor size among the nine studies was similar between the two groups: approximately 4.0 cm in the combined TACE/RFA group and 3.7 cm in the RFA alone group.

OS rate

Eight studies reported 1-year OS rate[7-11,13-15]; six presented 3-year OS rate[7,8,10,11,14,15]; and one had 5-year OS rate[8]. The pooled RR (95% CI) for 1-year survival rate comparing combined TACE/RFA to RFA alone was 1.12 (1.004-1.26), with high heterogeneity among eight studies (I^2 = 74.6% and P < 0.01) (Figure 2), which could be explained by the fact that one study reported

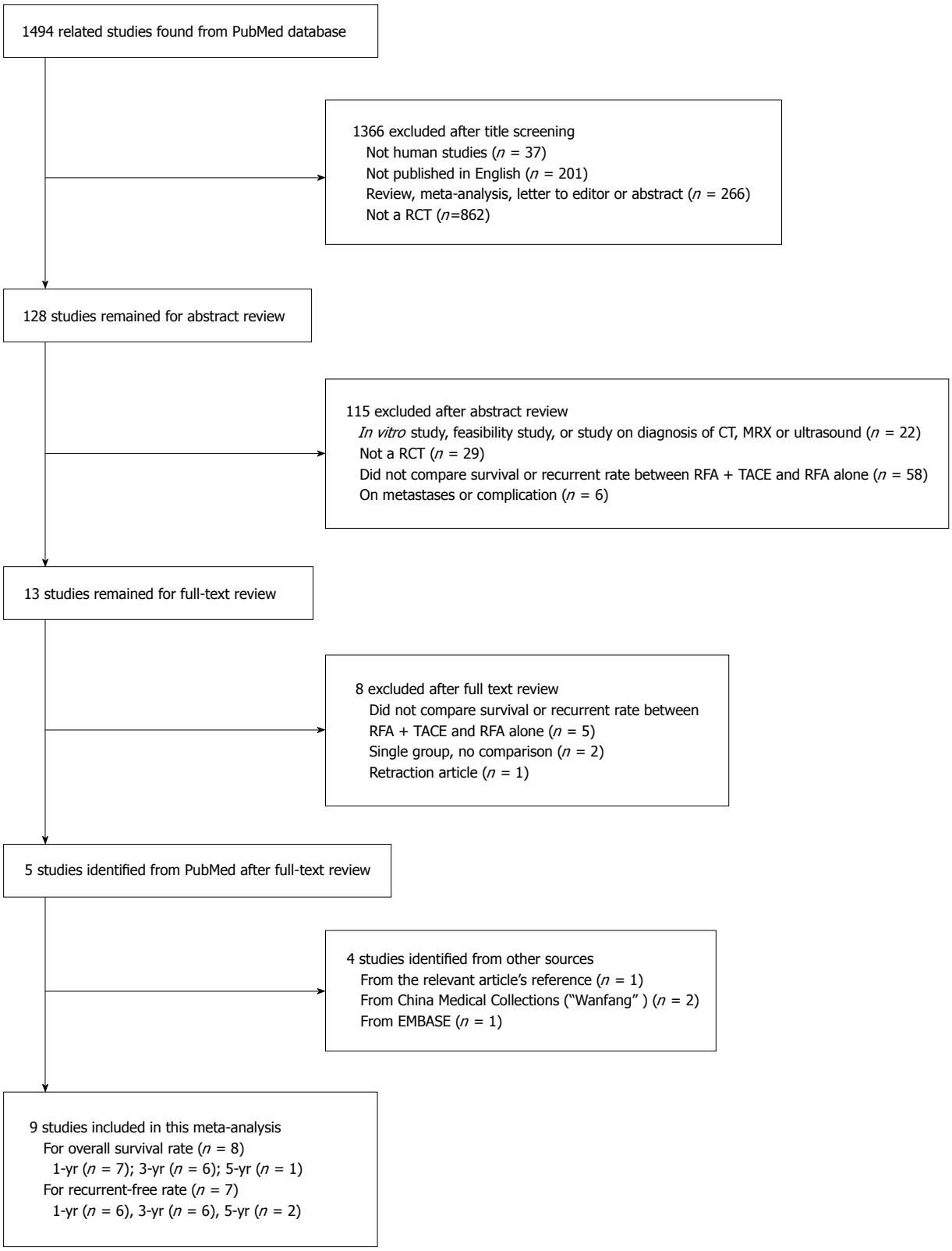


Figure 1 Flow chart of study screening and selection. RCT: Randomized clinical trial; TACE: Transarterial chemoembolization; RFA: Radiofrequency ablation; CT: Computed tomography; MRX: Magnetic resonance, soft spectrum coupled X-ray laser.

1-year survival rate of 100% for both groups[14]. After excluding this study, the heterogeneity disappeared among the other seven studies (I^2 = 0.0% and P = 0.82). A similar positive association (RR = 1.20; 95%CI: 1.02-1.41) was found for 3-year survival rate without heterogeneity (I^2 = 37.5% and P = 0.16). No evidence on publication bias was found (Egger's test: P = 0.40 for 1-year survival rate; and P = 0.09 for 3-year survival

Table 1 Characteristics of the included studies in this meta-analysis by treatment arms

Ref.	Arms	No. of patients	Male gender	Age (yr)	Child–Pugh Class (A/B/C)	Tumor size, cm	No. of tumors (1 *vs* ⩾ 2)	Follow-up (mo)	OS rate (%)			RFS rate, %		
									1-yr	3-yr	5-yr	1-yr	3-yr	5-yr
Zhang *et al*[9]	RFA + TACE	15	12	57.8 ± 11.0 (39-72)	--	4.6 ± 1.3 (2.3-7.1)	--	24	100	NA	NA	92.3	NA	NA
	RFA	15	13	58.3 ± 12.7 (38-78)	--	4.2 ± 1.1 (2.4-6.0)	--	24	80	NA	NA	--	NA	NA
Shen *et al*[10]	IRFAPA	18	5	52.7 (20-72)	4/14/0	5.6 (2.2-15.8)	0/18	28.3 (6-38)	87.5	52.2	NA	63.9	50.0	NA
	RFA	16	3	56.1 (36-75)	6/10/0	5.0 (2.3-12.3)	3/13	19.3 (5-36)	73.3	20.4	NA	30.0	18.7	NA
Kang *et al*[11]	RRA + TACE	19	14	52.2	12/7/0	6.7 ± 1.1	--	36	84.2	36.8	NA	--	--	NA
	RFA	18	14	50.7	12/6/0	6.2 ± 1.2	--	36	61.1	16.7	NA	--	--	NA
Kobayashi *et al*[12]	RFA + AO	10	7	67 (50-76)	4/5/1	1.7 (1.0-2.4)	--	48	--	--	NA	87.5	25	NA
	RFA	10	8	63 (51-75)	3/5/2	2.3 (1.0-2.6)	--	48	--	--	NA	70.0	20	NA
Yang *et al*[13]	RFA + TACE	31	23	60.3 ± 10.9	--	6.5 ± 0.8	--	--	81.2	--	--	--	82.2	--
	RFA	12	8	61.0 ± 10.4	--	5.2 ± 0.4	--	--	57.6	--	--	--	65.3	--
Shibata *et al*[14]	RFA + TACE	46	31	67.2 ± 8.9 (45-83)	32/14/0	1.7 ± 0.6 (0.9-3.0)	43/3	60	100	84.8	--	71.3	48.8	--
	RFA	43	33	69.8 ± 8.0 (44-87)	33/10/0	1.6 ± 0.5 (0.8-2.6)	42/1	60	100	84.5	--	74.3	29.7	--
Morimoto *et al*[15]	RFA + TACE	19	15	70 (57-78)	18/1/0	3.7 ± 0.6	--	30 (12-46)	100	93	NA	67	--	9
	RFA	18	12	73 (48-84)	16/2/0	3.6 ± 0.7	--	32 (15-46)	89	80	NA	56	--	28
Peng *et al*[8]	RFA + TACE	69	59	57.5 ± 10.0 (19-75)	60/9/0	2.1 ± 0.5 (0.8-5.0)	65/4	39.2 (5.0-95.0)	94	69	46	80	45	40
	RFA	70	55	55.1 ± 9.5 (22-75)	59/11/0	2.1 ± 0.4 (0.9-5.0)	65/5	33.6 (2.0-87.0)	82	47	36	64	18	18
Peng *et al*[7]	RFA + TACE	94	75	53.3 ± 11.0	90/4/0	3.5 ± 1.4	62/32	60	92.6	66.6	--	79.4	66.7	--
	RFA	95	71	55.3 ± 13.3	90/5/0	3.4 ± 1.4	67/28	60	85.3	59.0	--	60.6	44.2	--

AO: Arterial occlusion; IRFAPA: Intraoperative radiofrequency thermal ablation with portal vein infusion chemotherapy and transarterial chemoembolization; NA: Not applicable; TACE: Transarterial chemoembolization; RFA: Radiofrequency ablation.

rate).

RFS rate

Six studies reported 1-year RFS rate[7,8,10,12,14,15]; six presented 3-year RFS rate[7,8,10,12,14]; and two reported 5-year RFS rate[8,15]. One study reported 1-year RFS rate in the combination treatment group[9], and was not included in the pooling analysis for this outcome. The pooled RR was 1.19 (1.05-1.35) for 1-year RFS rate, 1.44 (1.0-2.07) for 3-year RFS rate, and 1.05 (0.19-5.75) for 5-year RFS rate (Figure 3).

There was no heterogeneity among six studies for 1-year RFS rate (I^2 = 19.5% and P = 0.29). High heterogeneity existed among six studies for 3-year RFS rate and in two studies for 5-year RFS rate (I^2 = 70.7% and P < 0.01 for 3-year RFS; I^2 = 78.2% and P = 0.03 for 5-year RFS). No evidence of publication bias was found for all the three pooled analyses.

Subgroup analysis

Table 2 presents the results from the subgroup analysis stratified by two predetermined factors: tumor size (< 3, 3-5, > 5 cm in diameter) and study location (China *vs* Japan). The results were more evident in those with medium-sized tumors or among Chinese population.

Sensitivity analysis

First, we replaced the random-effects model with the fixed-effects model for the pooled analyses. The findings were generally consistent. Second, we omitted one study each time from the pooled analyses and found that no single study substantially influenced the pooled results in the main analyses.

DISCUSSION

Our meta-analysis showed that RFA with TACE had a beneficial effect on 1- or 3-year OS rate as well as RFS rate, compared with RFA alone, especially for patients with medium-sized (3-5 cm) tumors. Based on the available evidence, there was no difference between the RFA and the TACE + RFA groups in terms of 5-year RFS rate. Whether or not there is difference between groups in 5-year OS rate remains unclear because only a few studies have reported this. A beneficial effect of combined TACE/RFA therapy on RFS rate was more evident among studies conducted in China.

So far, several reviews have been published on this topic. All of them had some flaws in their statistical design. One of them had a combined mixed cohort and RCT design and yielded a biased estimation of the

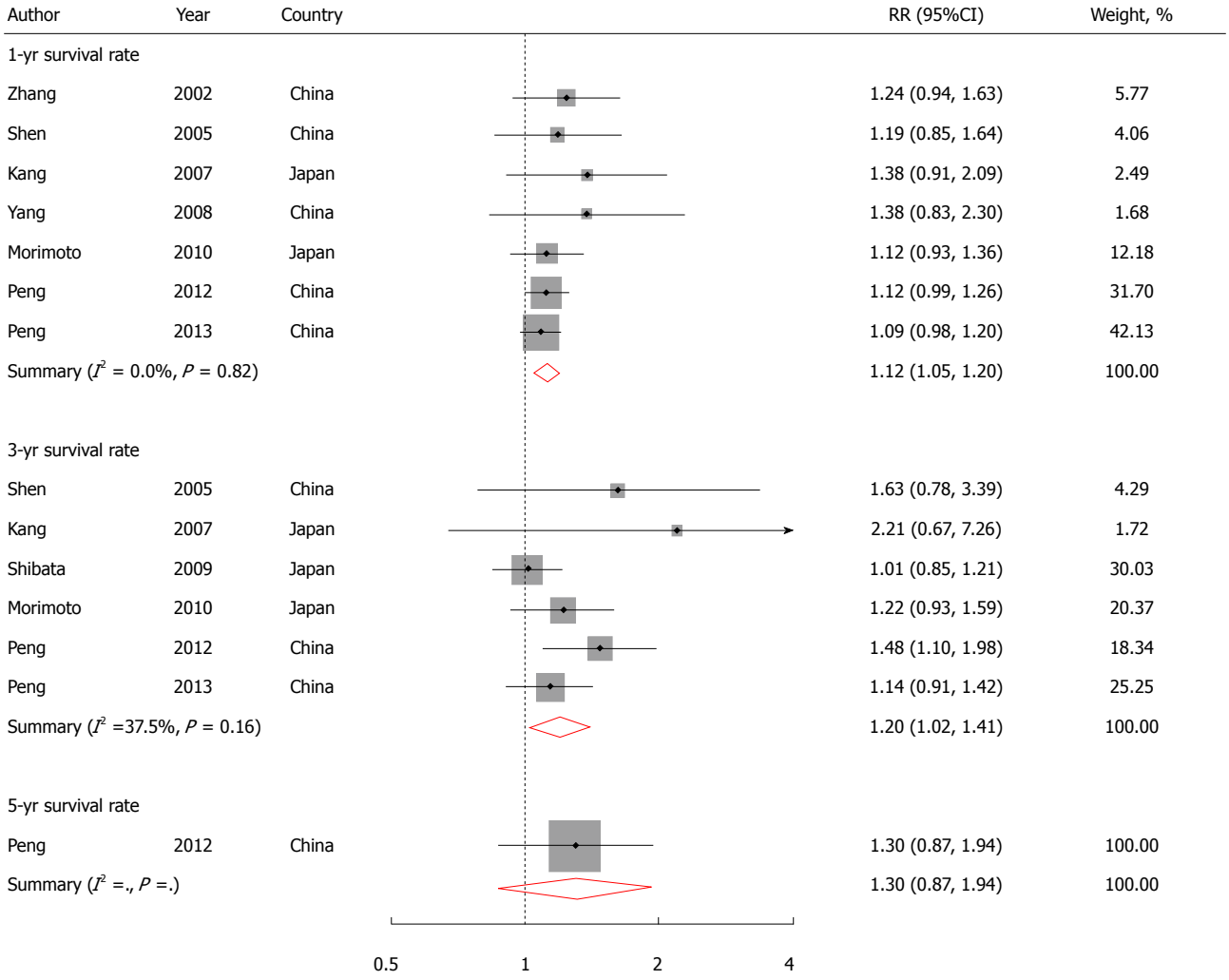


Figure 2 Relative risks and 95%CIs for risk of overall survival rate of hepatocellular carcinoma comparing combined therapy (radiofrequency ablation/transarterial chemoembolization) with radiofrequency ablation alone. The overall estimates were obtained by using a random-effects model. The dots indicate the RRs. The size of the shaded square is proportional to the weight of each study. The horizontal lines represent 95%CIs. The diamond data markers indicate the summary RRs. RRs: Relative risks.

true effects of combined therapy[16]. Three other meta-analyses gave misleading estimation of effect using OR instead of RR[17-19]. The other one pooled seven RCTs and found that RFA with TACE improved survival in patients with HCC > 3 cm[25]. Unfortunately, this review included one article that had already been retracted by the journal due to scientific misconduct[26], in addition to including an abstract[27]. None of these reviews gave a stratified analysis by tumor size and location, or reported results on RFS rate.

Several strengths of this meta-analysis should be mentioned. First, this was a meta-analysis of multiple RCTs, which provided strong evidence for casual inference. Second, this meta-analysis answered whether combined therapy is beneficial to medium-sized tumors in addition to small-sized ones, which is more meaningful in clinical practice. In addition, we found that combined therapy benefited the OS and RFS rates of HCC.

This meta-analysis had some limitations. First, we only included articles published in English and Chinese. Any bias caused by excluding studies published in other languages could not be ruled out. Second, our results were based on unadjusted RRs that were calculated according to the data derived from the original studies. The potential confounding effect could not be completely excluded. However, this meta-analysis included only RCTs, which substantially alleviated this concern. Indeed, there were no significant differences found in age, sex ratio, Child-Pugh class, and tumor size between TACE/RFA and RFA alone. In addition, there was insufficient information to conduct a quality assessment of the included studies. Any inherit limitation in the original studies may have biased our results. Moreover, due to the small number of studies included, we could not comment on the benefit of the combined therapy on 5-year OS and RFS rates.

Surgical resection, liver transplantation and local ablation are currently considered to be the three best and curative treatments for early-stage HCC[28]. However, only 10%-30% of early-stage HCC is suitable for surgery due to poor liver reserve, comorbidity, and shortage of liver donors. Therefore, local ablation plays an important

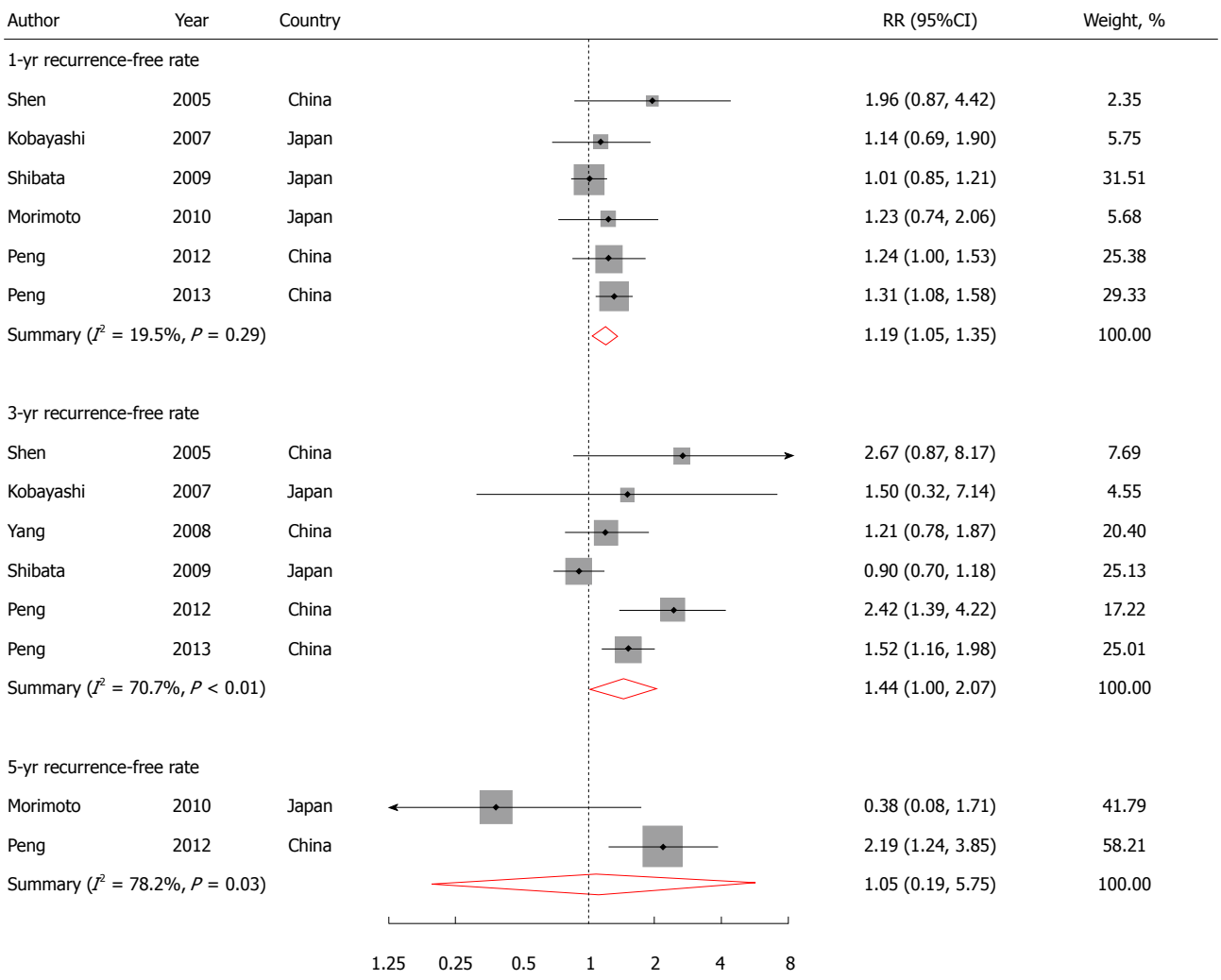


Figure 3 Relative risks and 95%CIs for risk of radiofrequency ablation rate of hepatocellular carcinoma comparing combined therapy (radiofrequency ablation/transarterial chemoembolization) with radiofrequency ablation alone. The overall estimates were obtained by using a random-effects model. The dots indicate the RRs. The size of the shaded square is proportional to the weight of each study. The horizontal lines represent 95%CIs. The diamond data markers indicate the RRs. RRs: Relative risks.

role in the treatment of unresectable or resectable early-stage HCC. Among the various local ablative modalities, RFA is a curative treatment with minimal invasiveness and high efficacy for small HCC that is generally defined as maximal diameter no larger than 3 cm[29]. Efficacy of RFA is reduced as the tumor size increases, which could be due to incomplete ablation and increased blood flow in larger lesions, resulting in heat loss. Therefore, as to the treatment for medium- or large-sized tumors, RFA alone may not be a first choice. Then, combining RFA with TACE may overcome the limitations of each of them used alone and provide better local control of HCCs > 3 cm. In particular, RFA is more efficient when the blood flow is reduced or cells are dying because of chemoembolization and the retained iodized oil after TACE can transfer heat fast.

No strictly designed crossover RCT has been conducted to validate which should be used first in the combination of RFA with TACE. However, it is hypothesized that using TACE before RFA may enhance subsequent RFA because: (1) TACE partially kills some tumor cells through chemotherapy and hypoxic injury, which could reduce tumor size; and (2) TACE decreases arterial blood flow and reduces or even eliminates the heat loss mediated by tissue perfusion[30], which will enlarge ablation and heat coagulation zone. In addition, whether an additional TACE regimen should be used to consolidate the treatment effect after combined TACE/RFA therapy deserves further research.

In this meta-analysis, the beneficial effect of combined TACE/RFA therapy was found to be more pronounced in studies conducted in China than those studies conducted in Japan. This phenomenon might be explained by the fact that HCC etiology in China is different from that in Japan, although the detailed mechanism is not clear. HCC in China is mostly related to hepatitis B virus, while HCC in Japan is associated with alcoholic fatty liver disease.

In conclusion, our meta-analysis found a 12% and 20% increase in 1-year and 3-year OS rate, as well as a 19% and 44% increase in 1-year and 3-year recurrence rate, respectively. The beneficial effect is more evident

Table 2 Stratified analysis of the pooled associations for overall survival and radiofrequency ablation rates comparing transarterial chemoembolization/radiofrequency ablation with radiofrequency ablation alone

Outcome	Potential modifiers		No. of studies	No. of events/patients		Heterogeneity test	RR (95%CI)
				RFA + TACE	RFA		
1-yr OS rate	Tumor size, cm	< 3	1	65/69	59/70	NA	1.12 (0.99, 1.26)
		3–5	3	121/128	109/128	I^2 = 0.0%, P = 0.66	1.11 (1.02, 1.21)[1]
		> 5	3	57/68	30/46	I^2 = 0.0%, P = 0.80	1.28 (1.02, 1.61)[1]
	Study location	China	5	208/227	171/208	I^2 = 0.0%, P = 0.77	1.12 (1.04, 1.20)[1]
		Japan	2	35/38	27/36	I^2 = 11.2%, P = 0.29	1.17 (0.96, 1.43)
3-yr OS rate	Tumor size, cm	< 3	2	87/115	69/113	I^2 = 83.2%, P = 0.02	1.21 (0.79, 1.83)
		3–5	2	81/113	70/113	I^2 = 0.0%, P = 0.68	1.17 (0.99, 1.39)[1]
		> 5	2	18/37	9/34	I^2 = 0.0%, P = 0.66	1.77 (0.95, 3.31)
	Study location	China	3	122/181	95/181	I^2 = 18.7%, P = 0.29	1.29 (1.05, 1.58)[1]
		Japan	3	64/84	53/79	I^2 = 33.4%, P = 0.22	1.12 (0.90, 1.39)
1-yr RFS rate	Tumor size, cm	< 3	3	102/125	88/123	I^2 = 10.7%, P = 0.33	1.11 (0.96, 1.28)
		3–5	2	88/113	68/113	I^2 = 0.0%, P = 0.83	1.30 (1.09, 1.55)[1]
		> 5	1	11/18	5/16	NA	1.96 (0.87, 4.42)
	Study location	China	3	141/181	108/181	I^2 = 0.0%, P = 0.55	1.29 (1.12, 1.49)[1]
		Japan	3	60/75	53/71	I^2 = 0.0%, P = 0.70	1.04 (0.89, 1.23)
3-year RFS rate	Tumor size, cm	< 3	3	67/128	47/123	I^2 = 84.3%, P < 0.01	1.45 (0.59, 3.54)
		3–5	1	63/94	42/95	NA	1.52 (1.16, 1.98)[1]
		> 5	2	34/49	11/28	I^2 = 51.0%, P = 0.15	1.56 (0.70, 3.47)
	Study location	China	4	128/212	66/193	I^2 = 41.2%, P = 0.16	1.63 (1.19, 2.24)[1]
		Japan	2	36/59	34/53	I^2 = 0.0%, P = 0.52	0.90 (0.71, 1.19)

[1]Denotes significant or marginally significant. TACE: Transarterial chemoembolization; RFA: Radiofrequency ablation.

in medium-sized tumors and among the Chinese population. Further RCTs with large tumors, long-term follow-up, and assessment of the combination model of RFA and TACE, for example, TACE-RFA-TACE, are needed.

ACKNOWLEDGMENTS

We thank Dr. Peng-cheng Xun for his help in verifying our analysis.

COMMENTS

Background

Surgical resection is the first-line therapeutic option for hepatocellular carcinoma (HCC) patients with small solitary nodules without underlying cirrhosis; however, its role in treating HCC is limited by strict inclusion criteria. Radiofrequency ablation (RFA), and transcatheter arterial chemoembolization (TACE), therapies for HCC, which involves administration of chemotherapy directly to the liver tumor *via* a catheter, have been developed as alternatives, particularly for patients with nonresectable HCC. Although local therapies may have less invasiveness, shorter hospital stay and lower associated mortality, their higher recurrence rates and lower disease-free survival rates are still major concerns. In the past decade, some evidence has suggested that combining RFA with TACE improves overall survival (OS) rate and reduces recurrence rate, while other studies have not shown these.

Innovations and breakthroughs

To investigate whether combined TACE/RFA improved OS and recurrence-free survival rates, compared with RFA alone, especially for patients with medium-sized (3-5 cm) or large (> 5 cm) HCCs, by conducting a systematic review as well as meta-analysis of RCTs.

Applications

Combined TACE/RFA has a beneficial effect on survival and recurrence rates compared with RFA alone, especially for medium-sized (3-5 cm) HCC and among Chinese patients.

Peer-review

This is a well-written article. The literature search was thorough. The statistical analysis was appropriate. The included studies did not reflect publication bias. The possible limitations of the meta analyses were identified.

REFERENCES

1 **Dhanasekaran R**, Limaye A, Cabrera R. Hepatocellular carcinoma: current trends in worldwide epidemiology, risk factors, diagnosis, and therapeutics. *Hepat Med* 2012; **4**: 19-37 [PMID: 24367230 DOI: 10.2147/HMER.S16316]

2 **Shirabe K**, Kanematsu T, Matsumata T, Adachi E, Akazawa K, Sugimachi K. Factors linked to early recurrence of small hepatocellular carcinoma after hepatectomy: univariate and multivariate analyses. *Hepatology* 1991; **14**: 802-805 [PMID: 1657754 DOI: 10.1002/hep.1840140510]

3 **Nagasue N**, Uchida M, Makino Y, Takemoto Y, Yamanoi A, Hayashi T, Chang YC, Kohno H, Nakamura T, Yukaya H. Incidence and factors associated with intrahepatic recurrence following resection of hepatocellular carcinoma. *Gastroenterology* 1993; **105**: 488-494 [PMID: 8392955]

4 **Izumi R**, Shimizu K, Ii T, Yagi M, Matsui O, Nonomura A, Miyazaki I. Prognostic factors of hepatocellular carcinoma in patients undergoing hepatic resection. *Gastroenterology* 1994; **106**: 720-727 [PMID: 8119543]

5 **Livraghi T**, Goldberg SN, Lazzaroni S, Meloni F, Solbiati L, Gazelle GS. Small hepatocellular carcinoma: treatment with radiofrequency ablation versus ethanol injection. *Radiology* 1999; **210**: 655-661 [PMID: 10207464 DOI: 10.1148/radiology.210.3.r99fe40655]

6 **Lo CM**, Ngan H, Tso WK, Liu CL, Lam CM, Poon RT, Fan ST, Wong J. Randomized controlled trial of transarterial lipiodol chemoembolization for unresectable hepatocellular carcinoma. *Hepatology* 2002; **35**: 1164-1171 [PMID: 11981766 DOI: 10.1053/jhep.2002.33156]

7 **Peng ZW**, Zhang YJ, Chen MS, Xu L, Liang HH, Lin XJ, Guo RP, Zhang YQ, Lau WY. Radiofrequency ablation with or without transcatheter arterial chemoembolization in the treatment of hepatocellular carcinoma: a prospective randomized trial. *J Clin Oncol* 2013; **31**: 426-432 [PMID: 23269991 DOI: 10.1200/

JCO.2012.42.9936]
8 **Peng ZW**, Zhang YJ, Liang HH, Lin XJ, Guo RP, Chen MS. Recurrent hepatocellular carcinoma treated with sequential transcatheter arterial chemoembolization and RF ablation versus RF ablation alone: a prospective randomized trial. *Radiology* 2012; **262**: 689-700 [PMID: 22157201 DOI: 10.1148/radiol.11110637]
9 **Zhang ZJ**, Wu MC, Chen H, Chen D, He J. Percutaneous radiofrequency ablation combined with transcatheter arterial chemoembolization for hepatocellular carcinoma. *Zhonghua Wai Ke Za Zhi* 2002; **40**: 826-829
10 **Shen SQ**, Xiang JJ, Xiong CL, Wu SM, Zhu SS. Intraoperative radiofrequency thermal ablation combined with portal vein infusion chemotherapy and transarterial chemoembolization for unresectable HCC. *Hepatogastroenterology* 2005; **52**: 1403-1407 [PMID: 16201083]
11 **Kang CB**, Xu HB, Wang SL, Rui B. Treatment of large hepatoma by TACE in combination with RFA. *Zhonghua Gandan Waike Zazhi* 2007; **13**: 828-830
12 **Kobayashi M**, Ikeda K, Kawamura Y, Hosaka T, Sezaki H, Yatsuji H, Akuta N, Suzuki F, Suzuki Y, Arase Y, Kumada H. Randomized controlled trial for the efficacy of hepatic arterial occlusion during radiofrequency ablation for small hepatocellular carcinoma--direct ablative effects and a long-term outcome. *Liver Int* 2007; **27**: 353-359 [PMID: 17355457 DOI: 10.1111/j.1478-3231.2006.01434.x]
13 **Yang P**, Liang M, Zhang Y, Shen B. Clinical application of a combination therapy of lentinan, multi-electrode RFA and TACE in HCC. *Adv Ther* 2008; **25**: 787-794 [PMID: 18670743 DOI: 10.1007/s12325-008-0079-x]
14 **Shibata T**, Isoda H, Hirokawa Y, Arizono S, Shimada K, Togashi K. Small hepatocellular carcinoma: is radiofrequency ablation combined with transcatheter arterial chemoembolization more effective than radiofrequency ablation alone for treatment? *Radiology* 2009; **252**: 905-913 [PMID: 19567647 DOI: 10.1148/radiol.2523081676]
15 **Morimoto M**, Numata K, Kondou M, Nozaki A, Morita S, Tanaka K. Midterm outcomes in patients with intermediate-sized hepatocellular carcinoma: a randomized controlled trial for determining the efficacy of radiofrequency ablation combined with transcatheter arterial chemoembolization. *Cancer* 2010; **116**: 5452-5460 [PMID: 20672352 DOI: 10.1002/cncr.25314]
16 **Kong QF**, Jiao JB, Chen QQ, Li L, Wang DG, Lv B. Comparative effectiveness of radiofrequency ablation with or without transarterial chemoembolization for hepatocellular carcinoma. *Tumour Biol* 2014; **35**: 2655-2659 [PMID: 24197985 DOI: 10.1007/s13277-013-1349-z]
17 **Ni JY**, Liu SS, Xu LF, Sun HL, Chen YT. Meta-analysis of radiofrequency ablation in combination with transarterial chemoembolization for hepatocellular carcinoma. *World J Gastroenterol* 2013; **19**: 3872-3882 [PMID: 23840128 DOI: 10.3748/wjg.v19.i24.3872]
18 **Ni JY**, Liu SS, Xu LF, Sun HL, Chen YT. Transarterial chemoembolization combined with percutaneous radiofrequency ablation versus TACE and PRFA monotherapy in the treatment for hepatocellular carcinoma: a meta-analysis. *J Cancer Res Clin Oncol* 2013; **139**: 653-659 [PMID: 23292073 DOI: 10.1007/s00432-012-1369-x]
19 **Yan S**, Xu D, Sun B. Combination of radiofrequency ablation with transarterial chemoembolization for hepatocellular carcinoma: a meta-analysis. *Dig Dis Sci* 2012; **57**: 3026-3031 [PMID: 22585384 DOI: 10.1007/s10620-012-2212-6]
20 **Davies HT**, Crombie IK, Tavakoli M. When can odds ratios mislead? *BMJ* 1998; **316**: 989-991 [PMID: 9550961]
21 **Lin SM**, Lin CJ, Lin CC, Hsu CW, Chen YC. Radiofrequency ablation improves prognosis compared with ethanol injection for hepatocellular carcinoma < or =4 cm. *Gastroenterology* 2004; **127**: 1714-1723 [PMID: 15578509]
22 **Moher D**, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the Quality of Reports of Meta-Analyses of Randomised Controlled Trials: The QUOROM Statement. *Onkologie* 2000; **23**: 597-602 [PMID: 11441269]
23 **Jackson D**, White IR, Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Stat Med* 2010; **29**: 1282-1297 [PMID: 19408255 DOI: 10.1002/sim.3602]
24 **Knapp G**, Hartung J. Improved tests for a random effects meta-regression with a single covariate. *Stat Med* 2003; **22**: 2693-2710 [PMID: 12939780 DOI: 10.1002/sim.1482]
25 **Lu Z**, Wen F, Guo Q, Liang H, Mao X, Sun H. Radiofrequency ablation plus chemoembolization versus radiofrequency ablation alone for hepatocellular carcinoma: a meta-analysis of randomized-controlled trials. *Eur J Gastroenterol Hepatol* 2013; **25**: 187-194 [PMID: 23134976 DOI: 10.1097/MEG.0b013e32835a0a07]
26 **Cheng BQ**, Jia CQ, Liu CT, Fan W, Wang QL, Zhang ZL, Yi CH. Chemoembolization combined with radiofrequency ablation for patients with hepatocellular carcinoma larger than 3 cm: a randomized controlled trial. *JAMA* 2008; **299**: 1669-1677 [PMID: 18398079 DOI: 10.1001/jama.299.14.1669]
27 **Aikata H**, Shirakawa H, Takaki S. Radiofrequency ablation combined with transcatheter arterial chemoembolizaiton for small hepatocelluar carcinoma. *Hepatology* 2006; **44**: A487 [DOI: 10.1002/hep.21398]
28 **Lau WY**, Lai EC. Hepatocellular carcinoma: current management and recent advances. *Hepatobiliary Pancreat Dis Int* 2008; **7**: 237-257 [PMID: 18522878]
29 **Kitamoto M**, Imagawa M, Yamada H, Watanabe C, Sumioka M, Satoh O, Shimamoto M, Kodama M, Kimura S, Kishimoto K, Okamoto Y, Fukuda Y, Dohi K. Radiofrequency ablation in the treatment of small hepatocellular carcinomas: comparison of the radiofrequency effect with and without chemoembolization. *AJR Am J Roentgenol* 2003; **181**: 997-1003 [PMID: 14500217 DOI: 10.2214/ajr.181.4.1810997]
30 **Yamasaki T**, Kurokawa F, Shirahashi H, Kusano N, Hironaka K, Okita K. Percutaneous radiofrequency ablation therapy for patients with hepatocellular carcinoma during occlusion of hepatic blood flow. Comparison with standard percutaneous radiofrequency ablation therapy. *Cancer* 2002; **95**: 2353-2360 [PMID: 12436442 DOI: 10.1002/cncr.10966]

P- Reviewer: Pouponneau P, Sollano JDD, Wang DS, Xu WG
S- Editor: Yu J **L- Editor**: A **E- Editor**: Jiao XK

Published by **Baishideng Publishing Group Inc**
8226 Regency Drive, Pleasanton, CA 94588, USA
Telephone: +1-925-223-8242
Fax: +1-925-223-8243
E-mail: bpgoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com