Artificial Intelligence in

Gastroenterology

Artificial Intelligence in Gastroenterology

Contents

MINIREVIEWS

Contents

ABOUT COVER

Editorial Board Member of *Artificial Intelligence in Gastroenterology*, Michele Finotti, MD, Adjunct Professor, Department of Hepatobiliary Surgery, General Surgeon at the Treviso Hospital, Padova 35123, Italy. mi6le@libero.it

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastroenterology* (*AIG*, *Artif Intell Gastroenterol*) is to provide scholars and readers from various fields of artificial intelligence in gastroenterology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

   AIG mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastroenterology and covering a wide range of topics, including artificial intelligence in gastrointestinal cancer, liver cancer, pancreatic cancer, hepatitis B, hepatitis C, nonalcoholic fatty liver disease, inflammatory bowel disease, irritable bowel syndrome, and *Helicobacter pylori* infection.

INDEXING/ABSTRACTING

The *AIG* is now abstracted and indexed in Reference Citation Analysis, China Science and Technology Journal Database.

RESPONSIBLE EDITORS FOR THIS ISSUE

Liver surgery for colorectal metastasis: New paths and new goals with the help of artificial intelligence

Valeria Tonini, Gabriele Vigutto, Riccardo Donati

Valeria Tonini, Department of Medical and Surgical Sciences, Sant' Orsola Hospital University of Bologna, Bologna 40138, Italy

Gabriele Vigutto, Department of Medical and Surgical Sciences, St Orsola Hospital, University of Bologna, Bologna 40138, Italy

Riccardo Donati, Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi" (DEI), University of Bologna, Bologna 40138, Italy

Corresponding author: Valeria Tonini, MD, PhD, Professor, Surgeon, Department of Medical and Surgical Sciences, University of Bologna, Sant' Orsola Hospital, Via Massarenti 9, Bologna 40138, Italy. valeria.tonini@unibo.it

Abstract

Colorectal cancer is one of the most common neoplasia with an high risk to metastatic spread. Improving medical and surgical treatment is moving along with improving the precision of diagnosis and patient's assessment, the latter two aided more and more with the use of artificial intelligence (AI). The management of colorectal liver metastasis is multidisciplinary, and surgery is the main option. After the diagnosis, a surgical assessment of the patient is fundamental. Reaching a R0 resection with a proper remnant liver volume can be done using new techniques involving also artificial intelligence. Considering the recent application of artificial intelligence as a valid substitute for liver biopsy in chronic liver diseases, several authors tried to apply similar techniques to pre-operative imaging of liver metastasis. Radiomics showed good results in identifying structural changes in a unhealthy liver and in evaluating the prognosis after a liver resection. Recently deep learning has been successfully applied in estimating the remnant liver volume before surgery. Moreover AI techniques can help surgeons to perform an early diagnosis of neoplastic relapse or a better differentiation between a colorectal metastasis and a benign lesion. AI could be applied also in the histopathological diagnostic tool. Although AI implementation is still partially automatized, it appears faster and more precise than the usual diagnostic tools and, in the short future, could become the new gold standard in liver surgery.

Key Words: Colo-rectal cancer; Liver metastasis; Artificial intelligence; Radiomics; Deep learning

Core Tip: Colon cancer is one of the most frequent cancers that unfortunately has a high risk of metastatic spread especially to the liver. The treatment of liver metastases is multidisciplinary, but surgery remains undoubtedly the main act. The results in the treatment of liver metastases have improved significantly over the years, but we continue to seek further paths of improvement. A new path, to which we currently entrust many hopes, is that of artificial intelligence, which could bring revolutionary solutions both in the diagnosis of liver metastases, and as a useful guide for surgical techniques. The purpose of this article is to summarize the latest news reported in the literature and possible research developments on this topic.

INTRODUCTION

Nowadays, colorectal cancer is one of the most common neoplasia in Western countries and among the main causes of death for oncologic diseases[1,2]. Between 30% and 50% of patients with colorectal cancer will develop liver metastasis during their life and surgical resection remains a fundamental treatment[1,2]. The improvement of surgical techniques, along with the use of newer and better schemes of chemotherapy, will increase the chances of a longer disease free survival for these patients[3]. Meanwhile, artificial intelligence (AI) is infiltrating healthcare exponentially and it has already been applied to several fields related to gastroenterology and hepatology[4,5].

HEPATOBILIARY SURGERY FOR COLORECTAL METASTASIS

The treatment of colorectal metastasis is generally multidisciplinary, involving many professional figures and multiples pathways[1,2]. Discussing other therapies, such as chemotherapy or radiotherapy, is beyond the scope of this article.

Surgical treatment always goes with hepatic resection[1]. All metastatic patients need to undergo several pre-operative exams for a better definition of the disease and its extent: a thoraco-abdominal contrast-enhanced CT scan and/or a contrast-enhanced MRI[1,6]. The use of routine PET/CT scan remains controversial[1,7]. The main goals during the assessment are evaluating the extent of the hepatic disease and searching for any extra hepatic localization of disease, the latter one is an exclusion criteria for any kind of hepatic resection[1,8].

Once surgery is considered, the assessment becomes more operative: new main goals are estimating how complex is performing a R0 resection and evaluating the liver remnant volume[1]. Clearly, a R0 resection should be achieved to increase the disease free survival and the overall survival, but the well-known 1cm border of healthy tissue is now reconsidered due to the increasing effectiveness of chemotherapy and the complexity of the resection[1,9,10]. At the same time, the size of the remnant liver must be evaluated with a three dimensional CT volumetry and it should be more than 20% in a healthy liver, more than 30% in post- systemic chemotherapy liver and more than 40% in a cirrhotic liver[1,11]. In case of an insufficient liver remnant volume, a portal vein embolization can be considered to increase the size to the residual liver[1,12], while, in case of bilateral lesions with a majority of them in one lobe, a two-stage hepatectomy with or without contralateral limited resections can be done[1,13]. Finally, a mini invasive approach should be considered if the surgeon is experienced in these techniques, considering the well-known advantages of mini invasive approaches[14].

RADIOMICS AND ARTIFICIAL INTELLIGENCE APPLIED TO MEDICAL IMAGING

The recent advent of artificial intelligence has changed the paradigm in the field of medical imaging interpretation together with radiomics. Artificial intelligence is a discipline that aims at mimicking the function of human brain in solving complex problems using computers. Machine learning and deep learning are branches of AI in which machines are thought how to learn from data using analytical models and algorithms. While machine learning methods usually require less computation on the computer side and more human intervention, deep learning may involve a huge amount of information

(from which stems the adjective "deep") and thus requires high performance computers, but less or no human intervention.

Radiomics is a tool for extensive extraction of quantitative features from medical imaging[4] and can be applied to ultrasound (US), magnetic resonance imaging (MRI), positron emission tomography (PET) and computed tomography (CT). The science of radiomics has taken advantage of machine learning with great benefit for medicine in general. The large amount of information provided by radiomics together with the improvements in AI have given raise to new methods of reading and interpreting medical images. Experts in different domains have now the opportunity to make less challenging the hard task of interpreting images thanks to this machine-aided approach. As shown in Figure 1 the workflow of conventional radiomics and AI applied to medical imaging is split in image acquisition, preprocessing, segmentation, features extraction and selection, model construction and training, model testing and evaluation. In conventional radiomics, one of the main prerequisites during the phase of image acquisition and preprocessing is a certain degree of standardization of the processes, in order to obtain a database with images that have comparable characteristics. Images segmentation consists in locating lesions manually or with the aid of a computer, in order to identify the region of interest o volumes of interests. Feature extraction and selection is a crucial step in machine learning paradigms in order to obtain a subset of quantitative parameters that are given as inputs to train the analytical model. In case of radiomics, these can be shape-based features (*e.g.* size, shape, location), histogram features (or others first-order features like standard deviation and variance), textual features (*e.g.* tumor hetero-geneity) and other higher order features extracted with wavelet transforms or Laplacian filters. In the phase of model construction, it is important to choose the analytical engine that gives the best results in term of performance in relation to the selected features. To do so, several models can be chosen and then tested such as linear regression, support vector machines, decision tree, random forest, K-Means. The evaluation of the models and the assessment of their performance is inferred from indicators and methods such as the receiver operating characteristic, nomograms and the decision curve analysis.

Whereas conventional radiomics is still a widely used approach in medical image analysis, in recent years, deep learning has been introduced in the clinical practice thanks to its promising results[7]. This technique can reach high levels of performance while not requiring manual human intervention in the phases of image segmentation and features extraction (Figure 2). In this paradigm, features are in fact automatically selected by a neural network to maximize the performance of the algorithm (called "backpropagation algorithm"). However, a larger amount of data (*e.g.* of number of medical images) is commonly needed to train the neural network models using backpropagation. Among the most popular techniques are multilayer perceptron networks, convolutional neural network, long short-term memory recurrent neural networks. Such as in conventional radiomics, different deep learning techniques can be applied to the input data in order to obtain the best performance.

ARTIFICIAL INTELLIGENCE APPLIED TO LIVER SURGERY

Recently, artificial intelligence was applied to various fields in medicine, including general surgery and hepatology[4,5], as seen in Table 1. Decharatanachart *et al*[4] published a meta-analysis on AI supported imaging and standard liver biopsy. They showed a similar prediction rate for liver cirrhosis without the risk of complications of a biopsy and without the usual interpretation bias of ultrasonography. Meanwhile, Christou *et al*[5] focused more on the possibility of integrating diagnosis and management in several gastroenterological diseases, such as inflammatory bowel disease (IBD), *Helicobacter pylori* infection and gastric cancer, and several hepatic diseases, such as HCV infection and cirrhosis[5]. On one hand, they described how the use of machine learning and CAD can increase sensibility and specificity of a standard endoscopic or radiologic exam; on the other hand they describe the limitations of AI[5].

One the of the main application of AI in liver surgery is in the pre-operative imaging. Park *et al*[15] described the use of radiomics and deep learning in liver diseases. Radiomics appears to be an effective way to analyse the structural changes of an unhealthy liver, comparable to the standard techniques like biopsies[15,16]. Furthermore, radiomics is already in use for determining the prognosis after surgical resection or radiofrequency[17] for hepatocellular carcinoma, especially related to micro vascular invasion[15,18]. Deep learning finds its best application in liver segmentation, where it is fundamental in estimating the liver remnant volume and the fat ratio in post chemotherapy liver[15,19,20]. Fang *et al* [21] focused on the implementation of deep learning in CT-guided biopsy to obtain a better localization of the lesion. In addition they presented a basic algorithm that could offer good results. At the same time, Winkel *et al*[22] compared manual segmentation and automatic segmentation with the use of deep learning showing a similar efficacy of the automatic segmentation with a faster elaboration of the images.

Focusing on focal liver lesions, Zhou *et al*[23] illustrated a 5 categories classification based on dynamic contrast-enhanced CT scan with a deep learning software: applying this classification, the radiologist would be able to make a diagnosis between a carcinoma and a benign lesion without biopsy[23,24]. They reported the application of deep learning to a contrast-enhanced ultrasonography (CEUS) to better

Table 1 Main implementation of artificial intelligence in hepatology and liver surgery

Ref.	Type of paper	Main topic	AI implementation
Decharatanachart *et al*[4], 2021	Meta-analysis	Chronic liver diseases	Diagnosis and staging of liver fibrosis without biopsy
Christou *et al*[5], 2021	Review	IBD, GI bleeding and chronic liver diseases	Increasing accuracy of gold standard diagnostic exams
Park *et al*[15], 2020	Review	Liver diseases	Staging of liver disease and prognosis after liver resection or chemotherapy
Wang *et al*[16], 2012	Survey	Liver imaging	Diagnosis of structural changes in healthy liver
Shan *et al*[19], 2019	Research article	Liver imaging (CT)	Prediction of early recurrence after HCC resection/RF
Hu *et al*[18], 2019	Research article	Liver imaging (US)	Evaluating microvascular invasion in HCC
Iranmanesh *et al*[19], 2014	Research article	Liver imaging (CT)	Evaluating portal pressure without invasive methods
Wang *et al*[23], 2019	Research article	Liver imaging (CT/MRI)	Using liver segmentation to an automatized liver biometry
Fang *et al*[21], 2020	Research article	Liver imaging	Using liver segmentation to more accurate localization of a hepatic lesion
Winkel *et al*[22], 2020	Comparative study	Liver imaging	Comparing a fully automated liver segmentation to a manual one
Zhou *et al*[23], 2019	Review	Liver imaging	Detecting hepatic lesions, characterized them and evaluate a response after treatment
Yasaka *et al*[24], 2018	Retrospective study	Liver imaging (CT)	Differentiation between benign and malignant hepatic lesions
Guo *et al*[25], 2018	Research article	Liver imaging (US)	Differentiation between benign and malignant hepatic lesions
Schmauch *et al*[26], 2019	Research article	Liver imaging (US)	Differentiation between benign and malignant hepatic lesions
Tiyarattanachai *et al*[27], 2021	Retrospective study	Liver imaging (US)	Detect and diagnose hepatic lesions
Perez *et al*[28], 2020	Review	HCC	Improving diagnosis and evaluation after ancillary treatments
Vivanti *et al*[29], 2017	Research article	Liver neoplasia	Evaluating post chemotherapy response
Li *et al*[30], 2015	Research article	Liver imaging (CT)	Differentiation between benign and malignant hepatic lesions
Hamm *et al*[31], 2019	Research article	Liver imaging (MRI)	Differentiation between benign and malignant hepatic lesions
Zhang *et al*[32], 2018	Research article	HCC	Differentiation between healthy and tumoral tissue in patient's liver
Preis *et al*[33], 2011	Research article	Liver imaging (PET)	Differentiation between benign and malignant hepatic lesions
Chen *et al*[34], 2020	Review	Liver surgery	Implementation in pre and post operative care
Nakayama *et al*[35], 2017	Retrospective study	Liver surgery	Use of 3D modeling to improve hepatice resection
Zhang *et al*[36], 2018	Prospective study	Liver surgery	Diagnosis and treatment of perihilar CCC
Vorontsov *et al*[37], 2019	Retrospective study	Liver surgery	Improving CRM identification and segmentation
Chartrand *et al*[39], 2017	Comparative study	Liver imaging	Improving liver segmentation and volumetry
Cancian *et al*[40], 2021	Research article.	Liver pathology	Better assessment pf tumor microenvironment

AI: Artificial intelligence; CCC: Cholangiocarcinoma; CRM: Colo-rectal metastases; CT: Computed tomography; GI: Gastrointestinal; HCC: Hepatocellular carcinoma; IBD: inflammatory bowel disease; MRI: Magnetic resonance imaging; PET: Positron emission tomography; US: Ultrasound.

distinguish between a benign and malignant lesion of the liver, showing again a better performance using AI techniques compared to the conventional technique[23,25]. Schmauch *et al*[26] presented a glimpse of future implementations of the standard ultrasonography where the use of a deep learning technique could drastically improve the diagnostic value of a widespread imaging such as US. Similarly, Tiyarattanachai *et al*[27] implemented a deep learning software for the US reporting a better outcome both in prevention and diagnosis of a focal liver lesion. Closely related to our main topic, Perez *et al*[28]
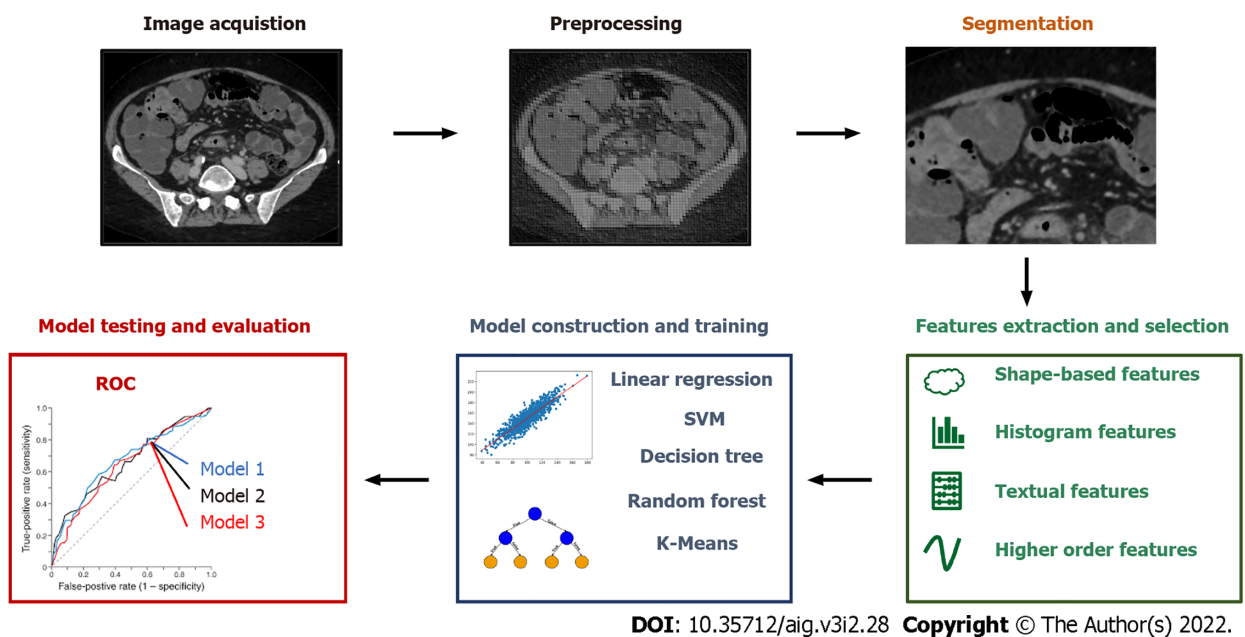
DOI: 10.35712/aig.v3i2.28 **Copyright** © The Author(s) 2022.

Figure 1 Workflow of conventional radiomics with machine learning.

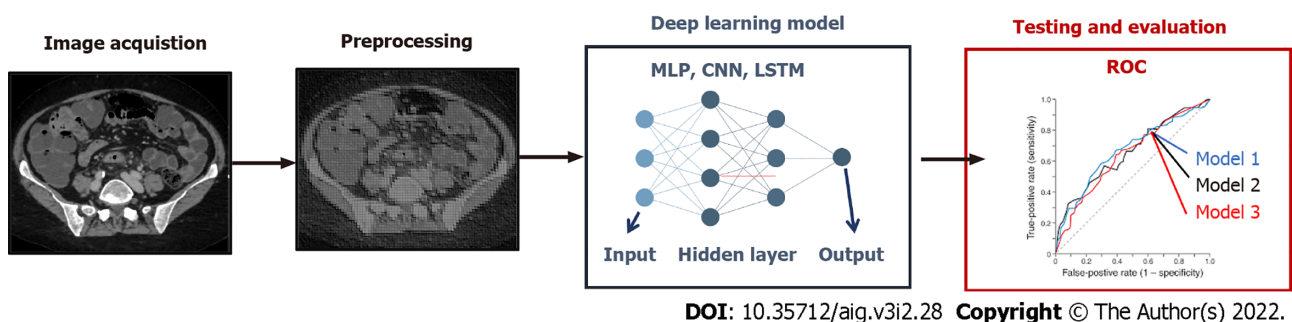**DOI**: 10.35712/aig.v3i2.28 **Copyright** © The Author(s) 2022.

**Figure 2 Deep learning techniques applied to radiomics.**

proposed a review on the management of hepatocellular carcinoma using AI for diagnosis, treatment and prognosis. Combining the US deep learning software[26] and the contrast-enhanced CT scan deep learning software[24,29,30], the clinician can reach a diagnosis on a focal liver lesion without the use of liver biopsy; in case of more doubts, a deep learning MRI software[31,32] and a deep learning PET software[33] are under external verification, but they appears promising.

Another main application of AI in liver surgery is the pre-operative patient assessment. The second part of the paper of Perez *et al*[28] described how the combined effort of US, CT, MRI scan and deep learning software increase the precision of the hepatic resection and the early recognition of a relapse. Beside the use of AI in the diagnosis, Chen *et al*[34] described the intra-operative advantages of using 3D rendering of the patient's liver to study and apply the best approach for a liver resection and, at the same time, to keep the same 3D model during the operation for a more intuitive way to reach the aforementioned R0 resection[34-36].

About colorectal liver metastasis, Voronstov *et al*[37] proposed a CT-based deep learning software to automatize and improve the recognition of metastasis rather than benign focal liver lesions. Detection performance of the software was still lower for lesion smaller than 10 mm, but it became more precise for lesions between 10 and 20 mm[37]. Manual liver segmentation was still more accurate for lesions smaller than 10mm, but it reached the same value for lesions greater than 10 mm and it was more efficient in lesions greater than 20 mm; the same results appeared considering lesion-volume estimation [37]. The authors also stated that all software calculations for an automatized or semi-automatized recognition and evaluation of metastasis is a significantly faster procedure than the usual manual one, as expected[37-39].

Within the same sphere, Cancian *et al*[40] focused on the analysis of the tumor microenvironment using a deep learning technique to evaluate the morphology of tumor associated macrophages. The same group recently described how different macrophages' morphologies are associated with different outcomes and therapeutic responses in colorectal liver metastasis[41], so they developed a pipeline
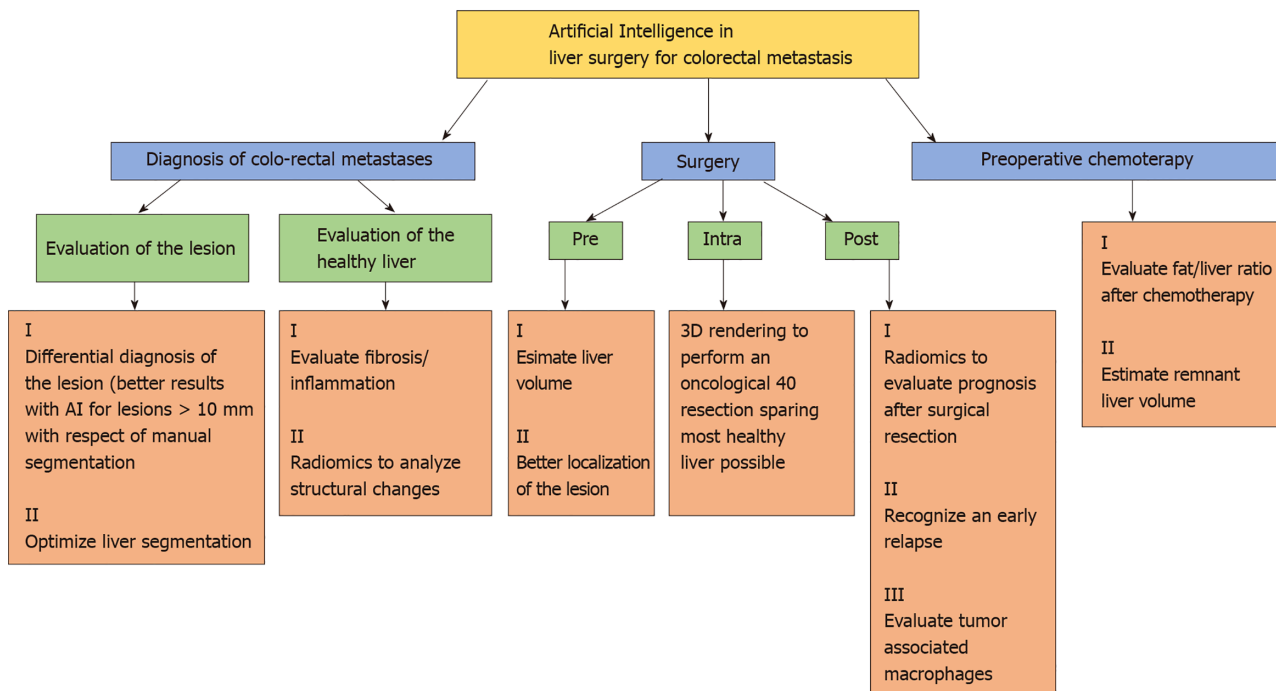
Figure 3 Main implementation of artificial intelligence in diagnosis and treatment of colo-rectal liver metastases.

using a CAD tool to process faster the histopathological slides. Although the pipeline is still under verification for a fully automatic application, a combined use of a manual and automatic approach showed a better and faster identification of macrophages' morphologies[40,41]. In Figure 3 are shown in a schematic manner the main tools of AI in diagnosis and treatment of colo-rectal liver metastases.

CONCLUSION

Artificial intelligence and deep learning offer new hopes in diagnosis and therapy of the liver metastasis. Therefore new promising research directions open up in this field, that must be confirmed with larger studies in the future.

FOOTNOTES

REFERENCES

1 **Dhir M**, Sasson AR. Surgical Management of Liver Metastases From Colorectal Cancer. *J Oncol Pract* 2016; **12**: 33-39 [PMID: 26759464 DOI: 10.1200/JOP.2015.009407]

2 **Manfredi S**, Lepage C, Hatem C, Coatmeur O, Faivre J, Bouvier AM. Epidemiology and management of liver metastases from colorectal cancer. *Ann Surg* 2006; **244**: 254-259 [PMID: 16858188 DOI: 10.1097/01.sla.0000217629.94941.cf]

3 **Folprecht G**. Liver Metastases in Colorectal Cancer. *Am Soc Clin Oncol Educ Book* 2016; **35**: e186-e192 [PMID: 27249722 DOI: 10.1200/EDBK_159185]

4 **Decharatanachart P**, Chaiteerakij R, Tiyarattanachai T, Treeprasertsuk S. Application of artificial intelligence in chronic liver diseases: a systematic review and meta-analysis. *BMC Gastroenterol* 2021; **21**: 10 [PMID: 33407169 DOI: 10.1186/s12876-020-01585-5]

5 **Christou CD**, Tsoulfas G. Challenges and opportunities in the application of artificial intelligence in gastroenterology and hepatology. *World J Gastroenterol* 2021; **27**: 6191-6223 [PMID: 34712027 DOI: 10.3748/wjg.v27.i37.6191]

6 **van Kessel CS**, Buckens CF, van den Bosch MA, van Leeuwen MS, van Hillegersberg R, Verkooijen HM. Preoperative imaging of colorectal liver metastases after neoadjuvant chemotherapy: a meta-analysis. *Ann Surg Oncol* 2012; **19**: 2805-2813 [PMID: 22396005 DOI: 10.1245/s10434-012-2300-z]

7 **Moulton CA**, Gu CS, Law CH, Tandan VR, Hart R, Quan D, Fairfull Smith RJ, Jalink DW, Husien M, Serrano PE, Hendler AL, Haider MA, Ruo L, Gulenchyn KY, Finch T, Julian JA, Levine MN, Gallinger S. Effect of PET before liver resection on surgical management for colorectal adenocarcinoma metastases: a randomized clinical trial. *JAMA* 2014; **311**: 1863-1869 [PMID: 24825641 DOI: 10.1001/jama.2014.3740]

8 **Ekberg H**, Tranberg KG, Andersson R, Lundstedt C, Hägerstrand I, Ranstam J, Bengmark S. Determinants of survival in liver resection for colorectal secondaries. *Br J Surg* 1986; **73**: 727-731 [PMID: 3756436 DOI: 10.1002/bjs.1800730917]

9 **Dhir M**, Lyden ER, Wang A, Smith LM, Ullrich F, Are C. Influence of margins on overall survival after hepatic resection for colorectal metastasis: a meta-analysis. *Ann Surg* 2011; **254**: 234-242 [PMID: 21694583 DOI: 10.1097/SLA.0b013e318223c609]

10 **Pawlik TM**, Scoggins CR, Zorzi D, Abdalla EK, Andres A, Eng C, Curley SA, Loyer EM, Muratore A, Mentha G, Capussotti L, Vauthey JN. Effect of surgical margin status on survival and site of recurrence after hepatic resection for colorectal metastases. *Ann Surg* 2005; **241**: 715-722, discussion 722 [PMID: 15849507 DOI: 10.1097/01.sla.0000160703.75808.7d]

11 **Ribero D**, Chun YS, Vauthey JN. Standardized liver volumetry for portal vein embolization. *Semin Intervent Radiol* 2008; **25**: 104-109 [PMID: 21326551 DOI: 10.1055/s-2008-1076681]

12 **van Lienden KP**, van den Esschert JW, de Graaf W, Bipat S, Lameris JS, van Gulik TM, van Delden OM. Portal vein embolization before liver resection: a systematic review. *Cardiovasc Intervent Radiol* 2013; **36**: 25-34 [PMID: 22806245 DOI: 10.1007/s00270-012-0440-y]

13 **Lam VW**, Laurence JM, Johnston E, Hollands MJ, Pleass HC, Richardson AJ. A systematic review of two-stage hepatectomy in patients with initially unresectable colorectal liver metastases. *HPB (Oxford)* 2013; **15**: 483-491 [PMID: 23750490 DOI: 10.1111/j.1477-2574.2012.00607.x]

14 **Ye SP**, Qiu H, Liao SJ, Ai JH, Shi J. Mini-invasive *vs* open resection of colorectal cancer and liver metastases: A meta-analysis. *World J Gastroenterol* 2019; **25**: 2819-2832 [PMID: 31236004 DOI: 10.3748/wjg.v25.i22.2819]

15 **Park HJ**, Park B, Lee SS. Radiomics and Deep Learning: Hepatic Applications. *Korean J Radiol* 2020; **21**: 387-401 [PMID: 32193887 DOI: 10.3348/kjr.2019.0752]

16 **Wang S**, Summers RM. Machine learning and radiology. *Med Image Anal* 2012; **16**: 933-951 [PMID: 22465077 DOI: 10.1016/j.media.2012.02.005]

17 **Shan QY**, Hu HT, Feng ST, Peng ZP, Chen SL, Zhou Q, Li X, Xie XY, Lu MD, Wang W, Kuang M. CT-based peritumoral radiomics signatures to predict early recurrence in hepatocellular carcinoma after curative tumor resection or ablation. *Cancer Imaging* 2019; **19**: 11 [PMID: 30813956 DOI: 10.1186/s40644-019-0197-5]

18 **Hu HT**, Wang Z, Huang XW, Chen SL, Zheng X, Ruan SM, Xie XY, Lu MD, Yu J, Tian J, Liang P, Wang W, Kuang M. Ultrasound-based radiomics score: a potential biomarker for the prediction of microvascular invasion in hepatocellular carcinoma. *Eur Radiol* 2019; **29**: 2890-2901 [PMID: 30421015 DOI: 10.1007/s00330-018-5797-0]

19 **Iranmanesh P**, Vazquez O, Terraz S, Majno P, Spahr L, Poncet A, Morel P, Mentha G, Toso C. Accurate computed tomography-based portal pressure assessment in patients with hepatocellular carcinoma. *J Hepatol* 2014; **60**: 969-974 [PMID: 24362073 DOI: 10.1016/j.jhep.2013.12.015]

20 **Wang K**, Mamidipalli A, Retson T, Bahrami N, Hasenstab K, Blansit K, Bass E, Delgado T, Cunha G, Middleton MS, Loomba R, Neuschwander-Tetri BA, Sirlin CB, Hsiao A; members of the NASH Clinical Research Network. Automated CT and MRI Liver Segmentation and Biometry Using a Generalized Convolutional Neural Network. *Radiol Artif Intell* 2019; **1** [PMID: 32582883 DOI: 10.1148/ryai.2019180022]

21 **Fang X**, Xu S, Wood BJ, Yan P. Deep learning-based liver segmentation for fusion-guided intervention. *Int J Comput Assist Radiol Surg* 2020; **15**: 963-972 [PMID: 32314228 DOI: 10.1007/s11548-020-02147-6]

22 **Winkel DJ**, Weikert TJ, Breit HC, Chabin G, Gibson E, Heye TJ, Comaniciu D, Boll DT. Validation of a fully automated liver segmentation algorithm using multi-scale deep reinforcement learning and comparison versus manual segmentation. *Eur J Radiol* 2020; **126**: 108918 [PMID: 32171914 DOI: 10.1016/j.ejrad.2020.108918]

23 **Zhou LQ**, Wang JY, Yu SY, Wu GG, Wei Q, Deng YB, Wu XL, Cui XW, Dietrich CF. Artificial intelligence in medical imaging of the liver. World J Gastroenterol 2019; 25: 672-682 [PMID: 30783371 DOI: 10.3748/wjg.v25.i6.672]

24 **Yasaka K**, Akai H, Abe O, Kiryu S. Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. *Radiology* 2018; **286**: 887-896 [PMID: 29059036 DOI: 10.1148/radiol.2017170706]

25 **Guo LH**, Wang D, Qian YY, Zheng X, Zhao CK, Li XL, Bo XW, Yue WW, Zhang Q, Shi J, Xu HX. A two-stage multi-view learning framework based computer-aided diagnosis of liver tumors with contrast enhanced ultrasound images. *Clin Hemorheol Microcirc* 2018; **69**: 343-354 [PMID: 29630528 DOI: 10.3233/CH-170275]

26   **Schmauch B**, Herent P, Jehanno P, Dehaene O, Saillard C, Aubé C, Luciani A, Lassau N, Jégou S. Diagnosis of focal liver lesions from ultrasound using deep learning. *Diagn Interv Imaging* 2019; **100**: 227-233 [PMID: 30926443 DOI: 10.1016/j.diii.2019.02.009]

27   **Tiyarattanachai T**, Apiparakoon T, Marukatat S, Sukcharoen S, Geratikornsupuk N, Anukulkarnkusol N, Mekaroonkamol P, Tanpowpong N, Sarakul P, Rerknimitr R, Chaiteerakij R. Development and validation of artificial intelligence to detect and diagnose liver lesions from ultrasound images. *PLoS One* 2021; **16**: e0252882 [PMID: 34101764 DOI: 10.1371/journal.pone.0252882]

28   **Jiménez Pérez M**, Grande RG. Application of artificial intelligence in the diagnosis and treatment of hepatocellular carcinoma: A review. *World J Gastroenterol* 2020; **26**: 5617-5628 [PMID: 33088156 DOI: 10.3748/wjg.v26.i37.5617]

29   **Vivanti R**, Szeskin A, Lev-Cohain N, Sosna J, Joskowicz L. Automatic detection of new tumors and tumor burden evaluation in longitudinal liver CT scan studies. *Int J Comput Assist Radiol Surg* 2017; **12**: 1945-1957 [PMID: 28856515 DOI: 10.1007/s11548-017-1660-z]

30   **Li W**, Jia F, Hu Q. Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks. *J Comput Commun* 2015; **3**: 146-151

31   **Hamm CA**, Wang CJ, Savic LJ, Ferrante M, Schobert I, Schlachter T, Lin M, Duncan JS, Weinreb JC, Chapiro J, Letzen B. Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. *Eur Radiol* 2019; **29**: 3338-3347 [PMID: 31016442 DOI: 10.1007/s00330-019-06205-9]

32   **Zhang F**, Yang J, Nezami N, Laage-Gaupp F, Chapiro J, De Lin M, Duncan J. Liver Tissue Classification Using an Auto-context-based Deep Neural Network with a Multi-phase Training Framework. *Patch Based Tech Med Imaging (2018)* 2018; **11075**: 59-66 [PMID: 32432233 DOI: 10.1007/978-3-030-00500-9_7]

33   **Preis O**, Blake MA, Scott JA. Neural network evaluation of PET scans of the liver: a potentially useful adjunct in clinical interpretation. *Radiology* 2011; **258**: 714-721 [PMID: 21339347 DOI: 10.1148/radiol.10100547]

34   **Chen H**, He Y, Jia W. Precise hepatectomy in the intelligent digital era. *Int J Biol Sci* 2020; **16**: 365-373 [PMID: 32015674 DOI: 10.7150/ijbs.39387]

35   **Nakayama K**, Oshiro Y, Miyamoto R, Kohno K, Fukunaga K, Ohkohchi N. The Effect of Three-Dimensional Preoperative Simulation on Liver Surgery. *World J Surg* 2017; **41**: 1840-1847 [PMID: 28271263 DOI: 10.1007/s00268-017-3933-7]

36   **Zhang J**, Qiao QL, Guo XC, Zhao JX. Application of three-dimensional visualization technique in preoperative planning of progressive hilar cholangiocarcinoma. *Am J Transl Res* 2018; **10**: 1730-1735 [PMID: 30018714]

37   **Vorontsov E**, Cerny M, Régnier P, Di Jorio L, Pal CJ, Lapointe R, Vandenbroucke-Menu F, Turcotte S, Kadoury S, Tang A. Deep Learning for Automated Segmentation of Liver Lesions at CT in Patients with Colorectal Cancer Liver Metastases. *Radiol Artif Intell* 2019; **1**: 180014 [PMID: 33937787 DOI: 10.1148/ryai.2019180014]

38   **Boykov YY**, Jolly MP.   Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. International Conference on Computer Vision. 2001; 105–112

39   **Chartrand G**, Cresson T, Chav R, Gotra A, Tang A, De Guise JA. Liver Segmentation on CT and MR Using Laplacian Mesh Optimization. *IEEE Trans Biomed Eng* 2017; **64**: 2110-2121 [PMID: 27893375 DOI: 10.1109/TBME.2016.2631139]

40   **Cancian P**, Cortese N, Donadon M, Di Maio M, Soldani C, Marchesi F, Savevski V, Santambrogio MD, Cerina L, Laino ME, Torzilli G, Mantovani A, Terracciano L, Roncalli M, Di Tommaso L. Development of a Deep-Learning Pipeline to Recognize and Characterize Macrophages in Colo-Rectal Liver Metastasis. *Cancers (Basel)* 2021; **13** [PMID: 34282750 DOI: 10.3390/cancers13133313]

41   **Donadon M**, Torzilli G, Cortese N, Soldani C, Di Tommaso L, Franceschini B, Carriero R, Barbagallo M, Rigamonti A, Anselmo A, Colombo FS, Maggi G, Lleo A, Cibella J, Peano C, Kunderfranco P, Roncalli M, Mantovani A, Marchesi F. Macrophage morphology correlates with single-cell diversity and prognosis in colorectal liver metastasis. *J Exp Med* 2020; **217** [PMID: 32785653 DOI: 10.1084/jem.20191847]

Colorectal cancer: Artificial intelligence and its role in surgical decision making

Nalini Kanta Ghosh, Ashok Kumar

Nalini Kanta Ghosh, Ashok Kumar, Department of Surgical Gastroenterology, Sanjay Gandhi Post Graduate Institute of Medical Sciences, Lucknow 226014, UP, India

Corresponding author: Ashok Kumar, BSc, FASCRS, FRCS (Ire), FRCS (Ed), FRCS (Eng), MBBS, M. Ch, MS, Professor, Department of Surgical Gastroenterology, Sanjay Gandhi Post Graduate Institute of Medical Sciences, Raebareli Road, Lucknow 226014, UP, India. doc.ashokgupta@gmail.com

Abstract

Despite several advances in the oncological management of colorectal cancer (CRC), there still remains a lacuna in the treatment strategy, which differs from center to center and on the philosophy of the treating clinician that is not without bias. Personalized treatment is essential for the treatment of CRC to achieve better long-term outcomes and to reduce morbidity. Surgery has an important role to play in the treatment. Surgical treatment of CRC is decided based on clinical parameters and investigations and hence likely to have judgmental errors. Artificial intelligence has been reported to be useful in the surveillance, diagnosis, treatment, and follow-up with accuracy in several malignancies. However, it is still evolving and yet to be established in surgical decision making in CRC. It is not only useful preoperatively but also intraoperatively. Artificial intelligence helps to rectify the human surgical decision when clinical data and radiological and laboratory parameters are fed into the computer and may guide correct surgical treatment.

Key Words: Artificial Intelligence; Colorectal cancer; Clinical implications; Treatment strategy; Surgical treatment

Core Tip: Treatment decision making in colorectal cancer significantly affects the outcome, which is a multidisciplinary team approach and is not without bias. Surgery plays a significant role in the treatment. Whether artificial intelligence may improve the outcome of surgery in colorectal cancer is not known. The present review focuses on its current role in surgical decision making and future impact.

INTRODUCTION

Mr. Alan Turing in 1950 hypothesized that a machine can also think like a human being in his book entitled "Computing Machinery and Intelligence"[1]. The term "artificial intelligence (AI)" was later coined by John McCarthy in a summer workshop[1,2]. AI has evolved from simple tasks to more complex tasks similar to a human brain[1].

AI has proven its worth in various day-to-day life and human requirements, including health care (health tracking devices)[3], automobiles (autopilot)[4], banking and finances (chatbots, robotraders)[5], surveillance (CCTV cameras), social media, entertainment, education, space exploration, industries (aluminum, dairy)[6-8], and disaster management[9,10]. One recent example is the efficient production of facemasks during the coronavirus disease 2019 pandemic[11] (Table 1). Its potential has been exploited in various fields of medicine, including online appointment scheduling, online check-in at hospitals, digitization of medical records, follow-up and immunization reminder, drug dosage algorithm, and adverse effect warnings during the prescription of multidrug combinations. Besides this, its application in the field of oncology is immense. AI is assisting in generating new approaches for cancer detection, screening of healthy subjects, diagnosis, classification of cancers using genomics, tumor microenvironment analysis, prognostication, follow-up, and new drug discovery[12-15].

Colorectal cancer (CRC) is one of the most common types of gastrointestinal (GI) tract malignancy and is the fourth most leading cause of cancer death globally[16,17]. AI has been used to facilitate screening, diagnosis (colonoscopy, advanced endoscopic modalities, imaging), genetic testing, and treatment (chemotherapy, radiotherapy, robotic assisted surgery)[18]. New research and developments are required for better patient management to improve the outcome.

In the past decade, several developments have taken place in the management of CRC, *e.g.*, revised anatomy of the rectum and concept of total mesorectal excision by Heald *et al*[19], concept of complete mesocolic excision and central vascular ligation by Hohenberger[20] for colon cancer, imaging and staging techniques, introduction of staplers[21], newer chemotherapeutic agents and biologicals, radiation therapy, and mode of surgery (laparoscopic and robotic surgery)[22,23] have significantly improved the outcome and sphincter preservation. However, there still remain numerous challenging issues like accurate preoperative diagnosis, staging, individualized and personalized treatment planning, and intraoperative challenges to minimize complications and improve the surgical outcome. Newer tools of AI have been used in various fields of medicine, including drug development, health monitoring, managing medical data, disease diagnostics, digital consultations, personalized treatment, analysis of health plans, and medical and surgical treatment[24] and is quickly finding a role in surgery and surgical decision making.

Two common fields of the AI used in medicine are: virtual and physical[25]. Virtual field is commonly used in medical imaging, clinical diagnosis, treatment, and drug research and development. Surgical and nursing robots are the part of physical fields. Because of ongoing innovations in AI, it is being used widely in medicine, both for diagnosis and management of tumors. AI has played a significant role in CRC at various stages and is reported to have improved the 5-year survival. The subsection of AI used in medicine is deep learning, which is responsible for widespread application of AI[26]. This method encompasses all the concepts of AI and is based on artificial neural networks (ANN), which is inspired by the neurons in a biological brain. Deep learning involves application of training a specific task on a larger data set, extracting information from them, and using them for future predictions about these tasks through flexible adaptation to the new data. Recently, deep learning has been used to predict cardiovascular risk based on retinal images[27], classification of skin lesions[28], mammogram-based breast cancer detection[29], and esophageal carcinoma[30]. However, application of AI in surgery is challenging, as unlike the use of static images, surgery includes dynamic procedural data like the patient clinical parameters, different devices used, and knowledge of clinical guidelines and from the experiences[31]. The uses and applications of various branches of AI in medicine as well as other fields are shown in Table 1.

In 2007, IBM began development of Deep QA technology (Watson). In 2017, Artery's medical imaging platform was the first Food and Drug Administration approved cloud-based deep learning application in healthcare for cardiac disorders, which was faster in giving results as compared to the professionals(15 s *vs* 30 s)[32]. The Food and Drug Administration-approved "GI genius" in the year 2019 is the first device based on machine learning to aid clinicians in detecting polyps or tumors during colonoscopy.

Table 1 Subfields of artificial intelligence and its application in day-to-day human life

S. No	Fields of AI	Description
1	Machine learning	Pattern identification and analysis where machine can help to improve based on past experiences provided from the given data set
2	Deep learning	Consists of multilayered neural networks called artificial neural network, which enables the computer to learn and make decisions on its own
3	Natural language processing	Ability of the computer to extract data from human language and make decisions
4	Computer vision	Potential to obtain information from a series of images or videos
5	Mixed-integer linear programming model[11]	It is helpful in finding the locational, supply, production, distribution, collection, quarantine, recycling, reuse, and disposal decisions within a multiperiod multiechelon multiproduct supply chain
6	Covering tour approach[9]	Optimizing the distribution and allocation of resources among individuals. It is useful at the time of crisis
7	Mixed-integer linear mathematical model[6]	This model optimizes economic, social, and environmental objectives simultaneously
8	Neural network with runner root algorithm[8]	Minimizing risk and maximizing return in industrial production
9	Meta-heuristic algorithms[7]	A comprehensive framework to predict the demand for dairy products
10	Hybrid shapley value and multimoora method[10]	An intelligent performance evaluation system for different supply chains in industries

AI: Artificial intelligence; S. No: Serial number.

This paper reviews the current status of AI in CRC surgical decision making and its future implications.

USES OF AI IN GASTROINTESTINAL DISORDERS AND COLORECTAL CANCER

AI is progressively being used in the understanding of GI diseases[33-35]. Imaging such as X-ray, computed tomography scanning, magnetic resonance imaging, or endoscopic imaging is being used for diagnosis[36-39]. The application of AI has led to early detection of intestinal malignancies or premalignant lesions, and inflammatory or other non-malignant diseases or lesions[40].

With IBM Watson for oncology (WFO), AI has found its increasing role in oncology therapy. It has been used in several malignancies like breast carcinoma, lung carcinoma, gastric cancer, colon and rectal cancer, *etc.* Initially, Memorial Sloan Kettering Cancer Center (New York, United States) started the use of WFO machine learning. WFO uses natural language processing and clinical data from multiple resources (treatment guidelines, expert opinions, literature, and medical records) to formulate treatment recommendations[41]. A recent meta-analysis[42] had shown the highest concordance between WFO and Mass Detection Tool in breast carcinoma and the lowest in stomach carcinoma. The Manipal Comprehensive Cancer Centre (Bangalore, India) has implemented WFO for treatment in 250 CRC patients[43]. There was a concordance in 92.7% of rectal and 81.0% of colon cancer patients between WFO and Mass Detection Tool recommendations[43].

AI IN COLORECTAL CANCER

AI is used in the diagnosis and treatment of colorectal polyps and cancer. In colorectal cancer, it helps in diagnosis, staging (lymph node or liver metastasis), preoperative treatment planning, response to treatment assessment, intraoperative assistance, postoperative prognostic information, *etc*[44-46].

AI in preoperative surgical decision making: staging and planning

After diagnosis of CRC is made, the most important consideration is staging to determine a further plan of management, whether upfront surgery, neoadjuvant treatment, or palliative treatment.

In locally advanced rectal cancer, preoperative chemoradiotherapy is known to reduce the local recurrence. However, selection of patients is essential to avoid unnecessary complications due to overtreatment. Therefore, there is a need for a system that can differentiate between T2 and T3 rectal cancers. Kim *et al*[47] used convolutional neural network models to distinguish T2 from T3 lesions from magnetic resonance imaging with an accuracy of 94%. Similarly, Wu *et al*[48] also used convolutional

neural network to stage rectal cancers.

In addition to its role in preoperative imaging, AI provides faster interpretation compared to radiologists (20 s *vs* 600 s) in the detection of lymph node metastasis in rectal cancer[49]. Preoperatively, positron emission tomography/computed tomography is commonly used in the case of indeterminate lesions on contrast-enhanced computed tomography to potentially find curable M1 disease (National Comprehensive Cancer Network guidelines version 3.2021). Recently, application of AI has improved the sensitivity and specificity of detection of pulmonary nodules[50]. AI can also be used to reconstruct the area of interest from two-dimensional data obtained from imaging and endoscopic findings to generate a three-dimensional structure for better delineation of the tumor in relation to the surrounding vital structures, which may be useful in preoperative surgical planning[51]. This is extremely useful in determining which patient will require a pelvic exenteration or which patient will require a lateral pelvic lymph node dissection. This is also useful to safeguard the important surrounding structures during surgery to reduce the postoperative morbidity and mortality related to it.

In colon cancer, clinical evidence of bulky nodal disease or T4b lesion entails neoadjuvant therapy (National Comprehensive Cancer Network guidelines version 3.2021). It is also recommended that the presence of nodal involvement in T1 cancer requires colectomy and lymphadenectomy. Kudo *et al*[52] applied machine learning ANN in 3134 patients with T1 CRC based on the patient's data on age, gender, tumor size, location, morphology, lymphatic and vascular invasion, and histologic grade to predict nodal involvement. ANN model was significantly better in lymph node metastasis detection compared to guidelines (area under the curve: 0.83 *vs* area under the curve: 0.73, *P* value = 0.005). Therefore, these patients can be subjected to upfront surgery and lymphadenectomy instead of endoscopic treatment. A meta-analysis by Bedrikovetski *et al*[53] using 17 studies (12 used radiomics models and 5 used deep learning models) concluded that AI was more efficient than radiologists in predicting lymph node metastasis. Similarly, AI was found to be better in detecting metastatic nodes as compared to conventional positron emission tomography/computed tomography imaging[54].

AI in intraoperative surgical decision making

Execution of a surgery depends upon the operating skill and ability of decision making. In 1978, Dr. Spencer[55], a cardiovascular surgeon, mentioned that "a skilfully performed operation is about 75% decision making and 25% dexterity." The decision making can be both technical or non-technical, which impacts patient outcome. Studies of surgical errors have shown that over half of the adverse events are due to cognitive errors[56]. But surgical training is more focused on skill training rather than decision making as it is a challenging task to train[57]. Decision-making skills may vary with experience of operating surgeons[58]. Thus, improving the quality of surgical decision making could help to improve the outcome of surgery.

Decision making is a three-step process, *i.e.* assessment of the situation, action-taking, and re-evaluation of the action's consequences. AI has been used as a decision making aid in a variety of fields, both in medicine and in surgery[59,60]. AI can help surgeons to assess a given situation (*e.g.*, retrieving better data about a clinical situation), the types of actions taken (*e.g.*, through decision suggestion), and the process of re-evaluating the impact of the decision taken. Therefore, it can be achieved in three different ways: (1) Retrieving data and experience from similar clinical scenarios and to supplement sensory input during minimal access surgery, which are not available compared to open surgery; (2) Intraoperative pathology assessment, tumor margin mapping, tumor classification, and tissue identification; and (3) Suggestion of steps of surgery.

Identification of surrounding structures: Harangi *et al*[61] used an ANN model to distinguish ureter from uterine artery during laparoscopic hysterectomy with 94.2% accuracy. Similarly, Quellec *et al*[62] applied a system of retrieving related videos of retinal surgery, and subsequent steps were followed during surgery to minimize the risk of injury. AI made it possible to define dissection planes in the robotic gastrectomy and to identify the recurrent laryngeal nerve during thyroidectomy[63,64]. Various studies have shown improved detection of vital structures during laparoscopic cholecystectomy to prevent bile duct injury using AI (Madani *et al*[65], Mascagni *et al*[66], Tokuyasu *et al*[67]). Table 2 highlights the studies where AI was used for identification of vital structures.

In CRC surgery, AI can be used to detect nearby vital structures (nerve plexus, presacral venous plexus, ureter, bladder, urethra, prostate, seminal vesicles), lymph node metastasis (lateral pelvic nodes, nodes near the root of inferior mesenteric artery), determination of the margin of resection, vascularity, and adequacy of anastomosis.

Augmented reality augments surgeons' intraoperative vision by providing a semi-transparent overlay of preoperative imaging on the area of interest[68]. It has been used in several GI surgical procedures like laparoscopic splenectomy[69] and pancreaticoduodenectomy[70]. Augmented reality can be applied to CRC surgeries to identify and preserve the nearby vital structures.

Deciding the level of resection: In CRC surgery, determination of margin status is important to decide the level of resection and consideration for the feasibility of an anastomosis or the creation of a stoma. Margin status can be obtained with "optical biopsy" (*in vivo* diagnostic imaging), which can avoid time-consuming resection and frozen section analysis. Fluorescence-guided surgery is evolving, and it has

Table 2 Studies having found the role of artificial intelligence in identification of vital structures in surgery

S. No	Primary aim	AI method used	Ref.
1	Recognition of ureter and uterine artery	Convolutional neural network	Harangi *et al*[61], 2017
2	Recognition of surgical steps of retinal surgery	Content-based video retrieval system	Quellec *et al*[62], 2011
3	To define safe dissection plane in robot assisted gastrectomy	Deep learning model based on U-net	Kumazu *et al*[63], 2021
4	Recurrent laryngeal nerve detection during thyroidectomy	Deep learning computer vision algorithm	Gong *et al*[64], 2021

AI: Artificial intelligence; S. No: Serial number.

shown promising results in determination of liver or peritoneal metastasis, anastomotic perfusion, detection of sentinel nodes, ureter, and nerves, and intraoperative detection of primary and recurrent lesions during colorectal cancer surgery[71]. Such a concept can be extrapolated on to AI for more efficient performance. Modalities used for intraoperative optical biopsy are confocal laser endomicroscopy, hyperspectral imaging, optical coherence tomography, and contrast-enhanced ultrasonography. There are several studies where these modalities have been used to distinguish abnormal epithelium from normal with the help of AI (Table 3). Using hyperspectral imaging, Jansen-Winkeln *et al*[72] reported 94% accuracy in distinguishing carcinoma from adenoma and healthy mucosa using ANN on post-resection of colonic lesions during surgery. A couple of experimental studies have shown that laparoscopic hyperspectral imaging can be used to distinguish malignant tissue in CRC from normal tissue. These modalities can be used to help in surgical decision making in CRC as revisional surgery can be done intraoperatively rather than waiting for frozen sections or final histology avoiding another surgery[73,74]. AI has been effective in differentiating glioblastoma, parathyroid gland, and malignant lesions of the colon from adjacent normal tissues[75-77].

Deciding the site of anastomosis: Studies have shown the incidence of colocolic and colorectal anastomosis leak to be 3.3% and 8.6%, respectively[78] and has adverse clinical outcomes and economic burden[79]. It can lead to anastomotic site stricture, recurrence of malignancy, and poor evacuatory function. The literature has shown poor predictive value of surgeons' perceptions of possible anastomotic site leaks that led to investigating other methods like the use of indocyanine green[80]. The robotic platform provides an inbuilt near infrared camera for assessment of vascularity at the resection margin and to reduce anastomotic site leakage[81]. A study by Mazaki *et al*[82], where auto-artificial intelligence was used to develop a predictive model for anastomotic leakage, showed that triple-row staplers can decrease the leak rate. There is an ongoing study by Taha *et al*[83] known as the PANIC study (The Prediction of Anastomotic Insufficiency risk after Colorectal surgery), which utilizes machine learning principles to formulate an algorithm for prediction of anastomotic leak following colonic (PANIC-C) or colorectal (PANIC-R) anastomosis. The results of the study are expected to be available by December 2022.

Helping in operative step suggestion: Operative step suggestion in CRC is at a developmental stage. In the literature, AI has been used in cataract surgery and spinal cord surgery with satisfactory results. Tian *et al*[84] developed VeBIRD (Video-Based Intelligent Recognition and Decision system) to track and classify the cataract grade on videos of phacoemulsification surgeries. It helped to decide the amount of ultrasonic energy needed to emulsify a cataract based on the grade. Therefore, a less experienced surgeon can perform the procedure with as much efficiency as that of an experienced surgeon. Somatosensory evoked potential is used during spinal cord surgeries to detect spinal cord injury. A decrease in somatosensory evoked potential value needs to be confirmed with awakening the patient and checking spinal cord function and this decrease in somatosensory evoked potential can be due to the effect of anesthesia. Fan *et al*[85] applied support vector regression and multi-support vector regression to distinguish spinal cord injury from anesthetic effect. Similarly, in CRC surgery such methods can help to find the area of interest to formulate standardized resection and differentiate intraoperative lymphorrhea from ureter or bladder injury using AI.

Colorectal cancer surgery requires accurate and judicious preoperative decisions to optimize the outcome of surgery (personalized treatment). The decision can be augmented by the use of AI, which is expected to be precise and without errors. It can assist in imaging, tissue diagnosis, and staging before surgery. It can be used preoperatively to choose patients for neoadjuvant therapy and those requiring upfront surgeries. Intraoperatively, it helps in the identification of tumor tissue (to determine the margin of resection), metastatic lymph nodes (for the extent of lymphadenectomy), and important surrounding structures. Its assistance is also useful in assessing the adequate vascularity at the anastomotic site that can decrease the postoperative anastomotic leak and thereby reduce the morbidity and mortality.

Table 3 Studies of artificial intelligence differentiating normal epithelium from abnormal or malignant cells

S. No	Modality used	Primary aim of study	AI method used	Ref.
1	CEUS	To differentiate glioblastoma from normal tissue	Support vector machines	Ritschel *et al*[75], 2015
2	OCT	To distinguish parathyroid tissue from thyroid, lymph node, and adipose tissue	Texture feature analysis and back propagation artificial neural network	Hou *et al*[76], 2017
3	CLE	Normal colonic mucosa from malignant lesion	Fractal analysis and neural network modelling	Ştefănescu *et al*[77], 2016
4	Hyperspectral imaging	Differentiation of colonic carcinoma from adenoma and healthy mucosa	Artificial neural network	Jansen-Winkeln *et al* [72], 2021

AI: Artificial intelligence; CEUS: Contrast-enhanced ultrasonography; OCT: Optical coherence tomography; CLE: Confocal laser endomicroscopy; S. No: Serial number.

Like the application of AI in several domains of medicine and health, it may play a significant role in surgical decision making, enhancing the outcome, in addition to diagnosis (imaging, endoscopy, tissue diagnosis).

FUTURE IMPLICATIONS

The future is promising, where AI is likely to play a significant role in reducing the bias of the Mass Detection Tool in deciding the treatment strategy and reducing the diagnosis and planning time with uniformity and with no or minimum error. The day is not far when the surgical world may be able to find a personalized surgical treatment for each and every patient of CRC, with improved intraoperative technical execution and reduced complications. The overall time taken in the management of CRC will be reduced, the treatment will be standardized, and the outcome will be maximized.

CONCLUSION

The role of AI in CRC is currently limited to preoperative staging and assessment of surgical resection margins and anastomotic sites. Its application to surgical decision making is still evolving, and the literature is very limited. However, the future is promising.

FOOTNOTES

ORCID number: Nalini Kanta Ghosh 0000-0003-1213-0235; Ashok Kumar 0000-0003-3959-075X.

REFERENCES

1   **Kaul V**, Enslin S, Gross SA. History of artificial intelligence in medicine. *Gastrointest Endosc* 2020; **92**: 807-812 [PMID: 32565184 DOI: 10.1016/j.gie.2020.06.040]

2   **Hamamoto R**, Suvarna K, Yamada M, Kobayashi K, Shinkai N, Miyake M, Takahashi M, Jinnai S, Shimoyama R, Sakai A, Takasawa K, Bolatkan A, Shozu K, Dozen A, Machino H, Takahashi S, Asada K, Komatsu M, Sese J, Kaneko S. Application of Artificial Intelligence Technology in Oncology: Towards the Establishment of Precision Medicine. *Cancers (Basel)* 2020; **12** [PMID: 33256107 DOI: 10.3390/cancers12123532]

3   **Meng Y**, Speier W, Shufelt C, Joung S, E Van Eyk J, Bairey Merz CN, Lopez M, Spiegel B, Arnold CW. A Machine Learning Approach to Classifying Self-Reported Health Status in a Cohort of Patients With Heart Disease Using Activity Tracker Data. *IEEE J Biomed Health Inform* 2020; **24**: 878-884 [PMID: 31199276 DOI: 10.1109/JBHI.2019.2922178]

4   **Kirsch D**. Autopilot and algorithms: accidents, errors, and the current need for human oversight. *J Clin Sleep Med* 2020; **16**: 1651-1652 [PMID: 32844741 DOI: 10.5664/jcsm.8762]

5   **Bredt S**. Artificial Intelligence (AI) in the Financial Sector-Potential and Public Strategies. *Front Artif Intell* 2019; **2**: 16 [PMID: 33733105 DOI: 10.3389/frai.2019.00016]

6   **Pahlevan SM**, Hosseini SMS, Goli A. Sustainable supply chain network design using products' life cycle in the aluminum industry. *Environ Sci Pollut Res Int* 2021 [PMID: 33474670 DOI: 10.1007/s11356-020-12150-8]

7   **Goli A**, Khademi-Zare H, Tavakkoli-Moghaddam R, Sadeghieh A, Sasanian M, Malekalipour Kordestanizadeh R. An integrated approach based on artificial intelligence and novel meta-heuristic algorithms to predict demand for dairy products: a case study. *Network* 2021; **32**: 1-35 [PMID: 33390063 DOI: 10.1080/0954898X.2020.1849841]

8   **Goli A**, Khademi Zare H, Tavakkoli-Moghaddam R, Sadeghieh A. (2019). Hybrid artificial intelligence and robust optimization for a multi-objective product portfolio problem. *Computers & Industrial Engineering* 2019; **137**: 106090 [DOI: 10.1016/j.cie.2019.106090]

9   **Goli A**, Malmir B. A Covering Tour Approach for Disaster Relief Locating and Routing with Fuzzy Demand. *International Journal of ITS Research* 2019; **18**: 140-152 [DOI: 10.1007/s13177-019-00185-2]

10 **Goli A**, Mohammadi H. Developing a sustainable operational management system using hybrid Shapley value and Multimoora method: case study petrochemical supply chain. *Environ Dev Sustain* 2021 [DOI: 10.1007/s10668-021-01844-9]

11 **Tirkolaee EB**, Goli A, Ghasemi P, Goodarzian F. Designing a sustainable closed-loop supply chain network of face masks during the COVID-19 pandemic: Pareto-based algorithms. *J Clean Prod* 2022; **333**: 130056 [PMID: 34924699 DOI: 10.1016/j.jclepro.2021.130056]

12 **Bhinder B**, Gilvary C, Madhukar NS, Elemento O. Artificial Intelligence in Cancer Research and Precision Medicine. *Cancer Discov* 2021; **11**: 900-915 [PMID: 33811123 DOI: 10.1158/2159-8290.CD-21-0090]

13 **Kann BH**, Hosny A, Aerts HJWL. Artificial intelligence for clinical oncology. *Cancer Cell* 2021; **39**: 916-927 [PMID: 33930310 DOI: 10.1016/j.ccell.2021.04.002]

14 **Huynh E**, Hosny A, Guthier C, Bitterman DS, Petit SF, Haas-Kogan DA, Kann B, Aerts HJWL, Mak RH. Artificial intelligence in radiation oncology. *Nat Rev Clin Oncol* 2020; **17**: 771-781 [PMID: 32843739 DOI: 10.1038/s41571-020-0417-8]

15 **Benzekry S**. Artificial Intelligence and Mechanistic Modeling for Clinical Decision Making in Oncology. *Clin Pharmacol Ther* 2020; **108**: 471-486 [PMID: 32557598 DOI: 10.1002/cpt.1951]

16 **Mármol I**, Sánchez-de-Diego C, Pradilla Dieste A, Cerrada E, Rodriguez Yoldi MJ. Colorectal Carcinoma: A General Overview and Future Perspectives in Colorectal Cancer. *Int J Mol Sci* 2017; **18** [PMID: 28106826 DOI: 10.3390/ijms18010197]

17 **Ou C**, Sun Z, Li X, Ren W, Qin Z, Zhang X, Yuan W, Wang J, Yu W, Zhang S, Peng Q, Yan Q, Xiong W, Li G, Ma J. MiR-590-5p, a density-sensitive microRNA, inhibits tumorigenesis by targeting YAP1 in colorectal cancer. *Cancer Lett* 2017; **399**: 53-63 [PMID: 28433598 DOI: 10.1016/j.canlet.2017.04.011]

18 **Mitsala A**, Tsalikidis C, Pitiakoudis M, Simopoulos C, Tsaroucha AK. Artificial Intelligence in Colorectal Cancer Screening, Diagnosis and Treatment. A New Era. *Curr Oncol* 2021; **28**: 1581-1607 [PMID: 33922402 DOI: 10.3390/curroncol28030149]

19 **Heald RJ**. The 'Holy Plane' of rectal surgery. *J R Soc Med* 1988; **81**: 503-508 [PMID: 3184105]

20 **Hohenberger W**, Weber K, Matzel K, Papadopoulos T, Merkel S. Standardized surgery for colonic cancer: complete mesocolic excision and central ligation--technical notes and outcome. *Colorectal Dis* 2009; **11**: 354-64; discussion 364 [PMID: 19016817 DOI: 10.1111/j.1463-1318.2008.01735.x]

21 **Moran BJ**. Stapling instruments for intestinal anastomosis in colorectal surgery. *Br J Surg* 1996; **83**: 902-909 [PMID: 8813772 DOI: 10.1002/bjs.1800830707]

22 **Jacobs M**, Verdeja JC, Goldstein HS. Minimally invasive colon resection (laparoscopic colectomy). *Surg Laparosc Endosc* 1991; **1**: 144-150 [PMID: 1688289]

23 **Weber PA**, Merola S, Wasielewski A, Ballantyne GH. Telerobotic-assisted laparoscopic right and sigmoid colectomies for benign disease. *Dis Colon Rectum* 2002; **45**: 1689-94; discussion 1695 [PMID: 12473897 DOI: 10.1007/s10350-004-7261-2]

24 **de Grey AD**. Artificial Intelligence and Medical Research: Time to Aim Higher? *Rejuvenation Res* 2016; **19**: 105-106 [PMID: 26993572 DOI: 10.1089/rej.2016.1827]

25 **Hamet P**, Tremblay J. Artificial intelligence in medicine. *Metabolism* 2017; **69S**: S36-S40 [PMID: 28126242 DOI: 10.1016/j.metabol.2017.01.011]

26 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

27 **Poplin R**, Varadarajan AV, Blumer K, Liu Y, McConnell MV, Corrado GS, Peng L, Webster DR. Prediction of cardiovascular risk factors from retinal fundus photographs *via* deep learning. *Nat Biomed Eng* 2018; **2**: 158-164 [PMID: 31015713 DOI: 10.1038/s41551-018-0195-0]

28 **Esteva A**, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer

with deep neural networks. *Nature* 2017; **542**: 115-118 [PMID: 28117445 DOI: 10.1038/nature21056]

29   **McKinney SM**, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, Back T, Chesus M, Corrado GS, Darzi A, Etemadi M, Garcia-Vicente F, Gilbert FJ, Halling-Brown M, Hassabis D, Jansen S, Karthikesalingam A, Kelly CJ, King D, Ledsam JR, Melnick D, Mostofi H, Peng L, Reicher JJ, Romera-Paredes B, Sidebottom R, Suleyman M, Tse D, Young KC, De Fauw J, Shetty S. International evaluation of an AI system for breast cancer screening. *Nature* 2020; **577**: 89-94 [PMID: 31894144 DOI: 10.1038/s41586-019-1799-6]

30   **Horie Y**, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, Hirasawa T, Tsuchida T, Ozawa T, Ishihara S, Kumagai Y, Fujishiro M, Maetani I, Fujisaki J, Tada T. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest Endosc* 2019; **89**: 25-32 [PMID: 30120958 DOI: 10.1016/j.gie.2018.07.037]

31   **Maier-Hein L**, Vedula SS, Speidel S, Navab N, Kikinis R, Park A, Eisenmann M, Feussner H, Forestier G, Giannarou S, Hashizume M, Katic D, Kenngott H, Kranzfelder M, Malpani A, März K, Neumuth T, Padoy N, Pugh C, Schoch N, Stoyanov D, Taylor R, Wagner M, Hager GD, Jannin P. Surgical data science for next-generation interventions. *Nat Biomed Eng* 2017; **1**: 691-696 [PMID: 31015666 DOI: 10.1038/s41551-017-0132-7]

32   **Marr B**. First FDA approval for clinical Cloud-Based Deep Learning in Healthcare. 2017. (accessed 1 Jun 2017). Available from: https://www.forbes.com/sites/bernardmarr/2017/01/20/first-fda-approval-for-clinical-cloud-based-deep-learning-in-healthcare/#7a0ed8dc161c

33   **Reismann J**, Romualdi A, Kiss N, Minderjahn MI, Kallarackal J, Schad M, Reismann M. Diagnosis and classification of pediatric acute appendicitis by artificial intelligence methods: An investigator-independent approach. *PLoS One* 2019; **14**: e0222030 [PMID: 31553729 DOI: 10.1371/journal.pone.0222030]

34   **Reichling C**, Taieb J, Derangere V, Klopfenstein Q, Le Malicot K, Gornet JM, Becheur H, Fein F, Cojocarasu O, Kaminsky MC, Lagasse JP, Luet D, Nguyen S, Etienne PL, Gasmi M, Vanoli A, Perrier H, Puig PL, Emile JF, Lepage C, Ghiringhelli F. Artificial intelligence-guided tissue analysis combined with immune infiltrate assessment predicts stage III colon cancer outcomes in PETACC08 study. *Gut* 2020; **69**: 681-690 [PMID: 31780575 DOI: 10.1136/gutjnl-2019-319292]

35   **Maeda Y**, Kudo SE, Mori Y, Misawa M, Ogata N, Sasanuma S, Wakamura K, Oda M, Mori K, Ohtsuka K. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc* 2019; **89**: 408-415 [PMID: 30268542 DOI: 10.1016/j.gie.2018.09.024]

36   **Ho TY**, Lin CW, Chang CC, Chen HT, Chen YJ, Lo YS, Hsiao PH, Chen PC, Lin CS, Tsou HK. Percutaneous endoscopic unilateral laminotomy and bilateral decompression under 3D real-time image-guided navigation for spinal stenosis in degenerative lumbar kyphoscoliosis patients: an innovative preliminary study. *BMC Musculoskelet Disord* 2020; **21**: 734 [PMID: 33172435 DOI: 10.1186/s12891-020-03745-w]

37   **Bhattacharya S**, Reddy Maddikunta PK, Pham QV, Gadekallu TR, Krishnan S SR, Chowdhary CL, Alazab M, Jalil Piran M. Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey. *Sustain Cities Soc* 2021; **65**: 102589 [PMID: 33169099 DOI: 10.1016/j.scs.2020.102589]

38   **Karako K**, Song P, Chen Y, Tang W. Realizing 5G- and AI-based doctor-to-doctor remote diagnosis: opportunities, challenges, and prospects. *Biosci Trends* 2020; **14**: 314-317 [PMID: 33100291 DOI: 10.5582/bst.2020.03364]

39   **Shiyam Sundar LK**, Muzik O, Buvat I, Bidaut L, Beyer T. Potentials and caveats of AI in hybrid imaging. *Methods* 2021; **188**: 4-19 [PMID: 33068741 DOI: 10.1016/j.ymeth.2020.10.004]

40   **Le Berre C**, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smaïl-Tabbone M, Danese S, Peyrin-Biroulet L. Application of Artificial Intelligence to Gastroenterology and Hepatology. *Gastroenterology* 2020; **158**: 76-94.e2 [PMID: 31593701 DOI: 10.1053/j.gastro.2019.08.058]

41   **Somashekhar SP**, Sepúlveda MJ, Norden AD, Rauthan A, Arun K, Patil P, Ethadka RY, Kumar RC. Early experience with IBM Watson for Oncology (WFO) cognitive computing system for lung and colorectal cancer treatment. [DOI: 10.1200/JCO.2017.35.15_suppl.8527]

42   **Jie Z**, Zhiying Z, Li L. A meta-analysis of Watson for Oncology in clinical application. *Sci Rep* 2021; **11**: 5792 [PMID: 33707577 DOI: 10.1038/s41598-021-84973-5]

43   **Yang SY**, Roh KH, Kim YN, Cho M, Lim SH, Son T, Hyung WJ, Kim HI. Surgical Outcomes After Open, Laparoscopic, and Robotic Gastrectomy for Gastric Cancer. *Ann Surg Oncol* 2017; **24**: 1770-1777 [PMID: 28357674 DOI: 10.1245/s10434-017-5851-1]

44   **Zhang R**, Zheng Y, Poon CCY, Shen D, Lau JYW. Polyp detection during colonoscopy using a regression-based convolutional neural network with a tracker. *Pattern Recognit* 2018; **83**: 209-219 [PMID: 31105338 DOI: 10.1016/j.patcog.2018.05.026]

45   **Su JR**, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, Li GC, Liu GQ, He YS, Zuo XL, Li YQ. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). *Gastrointest Endosc* 2020; **91**: 415-424.e4 [PMID: 31454493 DOI: 10.1016/j.gie.2019.08.026]

46   **Zhang S**, Han F, Liang Z, Tan J, Cao W, Gao Y, Pomeroy M, Ng K, Hou W. An investigation of CNN models for differentiating malignant from benign lesions using small pathologically proven datasets. *Comput Med Imaging Graph* 2019; **77**: 101645 [PMID: 31454710 DOI: 10.1016/j.compmedimag.2019.101645]

47   **Kim J**, Oh JE, Lee J, Kim MJ, Hur BY, Sohn DK, Lee B. Rectal cancer: Toward fully automatic discrimination of T2 and T3 rectal cancers using deep convolutional neural network. *Int J Imag Syst Tech* 2019; **29**: 247-259 [DOI: 10.1002/ima.22311]

48   **Wu QY**, Liu SL, Sun P, Li Y, Liu GW, Liu SS, Hu JL, Niu TY, Lu Y. Establishment and clinical application value of an automatic diagnosis platform for rectal cancer T-staging based on a deep neural network. *Chin Med J (Engl)* 2021; **134**: 821-828 [PMID: 33797468 DOI: 10.1097/CM9.0000000000001401]

49   **Ding L**, Liu GW, Zhao BC, Zhou YP, Li S, Zhang ZD, Guo YT, Li AQ, Lu Y, Yao HW, Yuan WT, Wang GY, Zhang DL, Wang L. Artificial intelligence system of faster region-based convolutional neural network surpassing senior radiologists in evaluation of metastatic lymph nodes of rectal cancer. *Chin Med J (Engl)* 2019; **132**: 379-387 [PMID: 30707177 DOI: 10.1097/CM9.0000000000000095]

50   **Schwyzer M**, Martini K, Benz DC, Burger IA, Ferraro DA, Kudura K, Treyer V, von Schulthess GK, Kaufmann PA, Huellner MW, Messerli M. Artificial intelligence for detecting small FDG-positive lung nodules in digital PET/CT: impact of image reconstructions on diagnostic performance. *Eur Radiol* 2020; **30**: 2031-2040 [PMID: 31822970 DOI: 10.1007/s00330-019-06498-w]

51   **Kim HJ**, Choi GS, Park JS, Park SY, Cho SH, Seo AN, Yoon GS. S122: impact of fluorescence and 3D images to completeness of lateral pelvic node dissection. *Surg Endosc* 2020; **34**: 469-476 [PMID: 31139999 DOI: 10.1007/s00464-019-06830-x]

52   **Kudo SE**, Ichimasa K, Villard B, Mori Y, Misawa M, Saito S, Hotta K, Saito Y, Matsuda T, Yamada K, Mitani T, Ohtsuka K, Chino A, Ide D, Imai K, Kishida Y, Nakamura K, Saiki Y, Tanaka M, Hoteya S, Yamashita S, Kinugasa Y, Fukuda M, Kudo T, Miyachi H, Ishida F, Itoh H, Oda M, Mori K. Artificial Intelligence System to Determine Risk of T1 Colorectal Cancer Metastasis to Lymph Node. *Gastroenterology* 2021; **160**: 1075-1084.e2 [PMID: 32979355 DOI: 10.1053/j.gastro.2020.09.027]

53   **Bedrikovetski S**, Dudi-Venkata NN, Kroon HM, Seow W, Vather R, Carneiro G, Moore JW, Sammour T. Artificial intelligence for pre-operative lymph node staging in colorectal cancer: a systematic review and meta-analysis. *BMC Cancer* 2021; **21**: 1058 [PMID: 34565338 DOI: 10.1186/s12885-021-08773-w]

54   **He J**, Wang Q, Zhang Y, Wu H, Zhou Y, Zhao S. Preoperative prediction of regional lymph node metastasis of colorectal cancer based on ^{18}F-FDG PET/CT and machine learning. *Ann Nucl Med* 2021; **35**: 617-627 [PMID: 33738763 DOI: 10.1007/s12149-021-01605-8]

55   **Spencer F**. Teaching and measuring surgical techniques: the technical evaluation of competence. *Bull Am Coll Surg* 1978; **63**: 9-12

56   **Suliburk JW**, Buck QM, Pirko CJ, Massarweh NN, Barshes NR, Singh H, Rosengart TK. Analysis of Human Performance Deficiencies Associated With Surgical Adverse Events. *JAMA Netw Open* 2019; **2**: e198067 [PMID: 31365107 DOI: 10.1001/jamanetworkopen.2019.8067]

57   **Pugh CM**, Santacaterina S, DaRosa DA, Clark RE. Intra-operative decision making: more than meets the eye. *J Biomed Inform* 2011; **44**: 486-496 [PMID: 20096376 DOI: 10.1016/j.jbi.2010.01.001]

58   **Hashimoto DA**, Axelsson CG, Jones CB, Phitayakorn R, Petrusa E, McKinley SK, Gee D, Pugh C. Surgical procedural map scoring for decision-making in laparoscopic cholecystectomy. *Am J Surg* 2019; **217**: 356-361 [PMID: 30470551 DOI: 10.1016/j.amjsurg.2018.11.011]

59   **Hashimoto DA**, Rosman G, Witkowski ER, Stafford C, Navarette-Welton AJ, Rattner DW, Lillemoe KD, Rus DL, Meireles OR. Computer Vision Analysis of Intraoperative Video: Automated Recognition of Operative Steps in Laparoscopic Sleeve Gastrectomy. *Ann Surg* 2019; **270**: 414-421 [PMID: 31274652 DOI: 10.1097/SLA.0000000000003460]

60   **Hashimoto DA**, Rosman G, Rus D, Meireles OR. Artificial Intelligence in Surgery: Promises and Perils. *Ann Surg* 2018; **268**: 70-76 [PMID: 29389679 DOI: 10.1097/SLA.0000000000002693]

61   **Harangi B**, Hajdu A, Lampe R, Torok P.   Recognizing ureter and uterine artery in endoscopic images using a convolutional neural network. In 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS) 2017 Jun 22 (pp. 726-727). IEEE

62   **Quellec G**, Lamard M, Cazuguel G, Droueche Z, Roux C, Cochener B.   Real-time retrieval of similar videos with application to computer-aided retinal surgery. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2011 Aug 30 (pp. 4465-4468). IEEE

63   **Kumazu Y**, Kobayashi N, Kitamura N, Rayan E, Neculoiu P, Misumi T, Hojo Y, Nakamura T, Kumamoto T, Kurahashi Y, Ishida Y, Masuda M, Shinohara H. Automated segmentation by deep learning of loose connective tissue fibers to define safe dissection planes in robot-assisted gastrectomy. *Sci Rep* 2021; **11**: 21198 [PMID: 34707141 DOI: 10.1038/s41598-021-00557-3]

64   **Gong J**, Holsinger FC, Noel JE, Mitani S, Jopling J, Bedi N, Koh YW, Orloff LA, Cernea CR, Yeung S. Using deep learning to identify the recurrent laryngeal nerve during thyroidectomy. *Sci Rep* 2021; **11**: 14306 [PMID: 34253767 DOI: 10.1038/s41598-021-93202-y]

65   **Madani A**, Namazi B, Altieri MS, Hashimoto DA, Rivera AM, Pucher PH, Navarrete-Welton A, Sankaranarayanan G, Brunt LM, Okrainec A, Alseidi A. Artificial Intelligence for Intraoperative Guidance: Using Semantic Segmentation to Identify Surgical Anatomy During Laparoscopic Cholecystectomy. *Ann Surg* 2020 [PMID: 33196488 DOI: 10.1097/SLA.0000000000004594]

66   **Mascagni P**, Vardazaryan A, Alapatt D, Urade T, Emre T, Fiorillo C, Pessaux P, Mutter D, Marescaux J, Costamagna G, Dallemagne B, Padoy N. Artificial Intelligence for Surgical Safety: Automatic Assessment of the Critical View of Safety in Laparoscopic Cholecystectomy Using Deep Learning. *Ann Surg* 2020 [PMID: 33201104 DOI: 10.1097/SLA.0000000000004351]

67   **Tokuyasu T**, Iwashita Y, Matsunobu Y, Kamiyama T, Ishikake M, Sakaguchi S, Ebe K, Tada K, Endo Y, Etoh T, Nakashima M, Inomata M. Development of an artificial intelligence system using deep learning to indicate anatomical landmarks during laparoscopic cholecystectomy. *Surg Endosc* 2021; **35**: 1651-1658 [PMID: 32306111 DOI: 10.1007/s00464-020-07548-x]

68   **Bernhardt S**, Nicolau SA, Soler L, Doignon C. The status of augmented reality in laparoscopic surgery as of 2016. *Med Image Anal* 2017; **37**: 66-90 [PMID: 28160692 DOI: 10.1016/j.media.2017.01.007]

69   **Ieiri S**, Uemura M, Konishi K, Souzaki R, Nagao Y, Tsutsumi N, Akahoshi T, Ohuchida K, Ohdaira T, Tomikawa M, Tanoue K, Hashizume M, Taguchi T. Augmented reality navigation system for laparoscopic splenectomy in children based on preoperative CT image using optical tracking device. *Pediatr Surg Int* 2012; **28**: 341-346 [PMID: 22130783 DOI: 10.1007/s00383-011-3034-x]

70   **Onda S**, Okamoto T, Kanehira M, Suzuki F, Ito R, Fujioka S, Suzuki N, Hattori A, Yanaga K. Identification of inferior pancreaticoduodenal artery during pancreaticoduodenectomy using augmented reality-based navigation system. *J Hepatobiliary Pancreat Sci* 2014; **21**: 281-287 [PMID: 23970384 DOI: 10.1002/jhbp.25]

71   **Galema HA**, Meijer RPJ, Lauwerends LJ, Verhoef C, Burggraaf J, Vahrmeijer AL, Hutteman M, Keereweer S, Hilling DE. Fluorescence-guided surgery in colorectal cancer; A review on clinical results and future perspectives. *Eur J Surg Oncol*

2021 [PMID: 34657780 DOI: 10.1016/j.ejso.2021.10.005]

72   **Jansen-Winkeln B**, Barberio M, Chalopin C, Schierle K, Diana M, Köhler H, Gockel I, Maktabi M. Feedforward Artificial Neural Network-Based Colorectal Cancer Detection Using Hyperspectral Imaging: A Step towards Automatic Optical Biopsy. *Cancers (Basel)* 2021; **13** [PMID: 33669082 DOI: 10.3390/cancers13050967]

73   **Köhler H**, Kulcke A, Maktabi M, Moulla Y, Jansen-Winkeln B, Barberio M, Diana M, Gockel I, Neumuth T, Chalopin C. Laparoscopic system for simultaneous high-resolution video and rapid hyperspectral imaging in the visible and near-infrared spectral range. *J Biomed Opt* 2020; **25** [PMID: 32860357 DOI: 10.1117/1.JBO.25.8.086004]

74   **Baltussen EJM**, Kok END, Brouwer de Koning SG, Sanders J, Aalbers AGJ, Kok NFM, Beets GL, Flohil CC, Bruin SC, Kuhlmann KFD, Sterenborg HJCM, Ruers TJM. Hyperspectral imaging for tissue classification, a way toward smart laparoscopic colorectal surgery. *J Biomed Opt* 2019; **24**: 1-9 [PMID: 30701726 DOI: 10.1117/1.JBO.24.1.016002]

75   **Ritschel K**, Pechlivanis I, Winter S. Brain tumor classification on intraoperative contrast-enhanced ultrasound. *Int J Comput Assist Radiol Surg* 2015; **10**: 531-540 [PMID: 24956998 DOI: 10.1007/s11548-014-1089-6]

76   **Hou F**, Yu Y, Liang Y. Automatic identification of parathyroid in optical coherence tomography images. *Lasers Surg Med* 2017; **49**: 305-311 [PMID: 28129441 DOI: 10.1002/lsm.22622]

77   **Ştefănescu D**, Streba C, Cârţână ET, Săftoiu A, Gruionu G, Gruionu LG. Computer Aided Diagnosis for Confocal Laser Endomicroscopy in Advanced Colorectal Adenocarcinoma. *PLoS One* 2016; **11**: e0154863 [PMID: 27144985 DOI: 10.1371/journal.pone.0154863]

78   **Krell RW**, Girotti ME, Fritze D, Campbell DA, Hendren S. Hospital readmissions after colectomy: a population-based study. *J Am Coll Surg* 2013; **217**: 1070-1079 [PMID: 24246621 DOI: 10.1016/j.jamcollsurg.2013.07.403]

79   **Lee SW**, Gregory D, Cool CL. Clinical and economic burden of colorectal and bariatric anastomotic leaks. *Surg Endosc* 2020; **34**: 4374-4381 [PMID: 31720809 DOI: 10.1007/s00464-019-07210-1]

80   **Karliczek A**, Harlaar NJ, Zeebregts CJ, Wiggers T, Baas PC, van Dam GM. Surgeons lack predictive accuracy for anastomotic leakage in gastrointestinal surgery. *Int J Colorectal Dis* 2009; **24**: 569-576 [PMID: 19221768 DOI: 10.1007/s00384-009-0658-6]

81   **Jafari MD**, Lee KH, Halabi WJ, Mills SD, Carmichael JC, Stamos MJ, Pigazzi A. The use of indocyanine green fluorescence to assess anastomotic perfusion during robotic assisted laparoscopic rectal surgery. *Surg Endosc* 2013; **27**: 3003-3008 [PMID: 23404152 DOI: 10.1007/s00464-013-2832-8]

82   **Mazaki J**, Katsumata K, Ohno Y, Udo R, Tago T, Kasahara K, Kuwabara H, Enomoto M, Ishizaki T, Nagakawa Y, Tsuchida A. A Novel Predictive Model for Anastomotic Leakage in Colorectal Cancer Using Auto-artificial Intelligence. *Anticancer Res* 2021; **41**: 5821-5825 [PMID: 34732457 DOI: 10.21873/anticanres.15400]

83   **Taha A**, Taha-Mehlitz S, Hendie A, Staudner T, Adamina M.   Development and external validation of an international, multicenter machine learning algorithm for prediction of anastomotic insufficiency after colonic or colorectal anastomosis The Prediction of Anastomotic Insufficiency risk after Colorectal surgery (PANIC) study. Available from: https://clinicaltrials.gov/ct2/show/NCT04985981

84   **Tian S**, Yin XC, Wang ZB, Zhou F, Hao HW. A VidEo-Based Intelligent Recognition and Decision System for the Phacoemulsification Cataract Surgery. *Comput Math Methods Med* 2015; **2015**: 202934 [PMID: 26693249 DOI: 10.1155/2015/202934]

85   **Fan B**, Li HX, Hu Y. An Intelligent Decision System for Intraoperative Somatosensory Evoked Potential Monitoring. *IEEE Trans Neural Syst Rehabil Eng* 2016; **24**: 300-307 [PMID: 26415181 DOI: 10.1109/TNSRE.2015.2477557]

Application of artificial intelligence in non-alcoholic fatty liver disease and viral hepatitis

Atchayaa Gunasekharan, Joanna Jiang, Ashley Nickerson, Sajid Jalil, Khalid Mumtaz

Atchayaa Gunasekharan, Department of Medicine, Jewish Hospital, Cincinatti, OH 43201, United States

Joanna Jiang, Ashley Nickerson, Sajid Jalil, Khalid Mumtaz, Department of Medicine, Ohio State University, Columbus, OH 43210, United States

Corresponding author: Khalid Mumtaz, MBBS, MSc, Associate Professor, Department of Medicine, Ohio State University, 395 W. 12th Avenue, Columbus, OH 43210, United States. khalid.mumtaz@osumc.edu

Abstract

Non-alcoholic fatty liver disease (NAFLD) and chronic viral hepatitis are among the most significant causes of liver-related mortality worldwide. It is critical to develop reliable methods of predicting progression to fibrosis, cirrhosis, and decompensated liver disease. Current screening methods such as biopsy and transient elastography are limited by invasiveness and observer variation in analysis of data. Artificial intelligence (AI) provides a unique opportunity to more accurately diagnose NAFLD and viral hepatitis, and to identify patients at high risk for disease progression. We conducted a literature review of existing evidence for AI in NAFLD and viral hepatitis. Thirteen articles on AI in NAFLD and 14 on viral hepatitis were included in our analysis. We found that machine learning algorithms were comparable in accuracy to current methods for diagnosis and fibrosis prediction (MELD-Na score, liver biopsy, FIB-4 score, and biomarkers). They also reliably predicted hepatitis C treatment failure and hepatic encephalopathy, for which there are currently no established prediction tools. These studies show that AI could be a helpful adjunct to existing techniques for diagnosing, monitoring, and treating both NAFLD and viral hepatitis.

Key Words: Non-alcoholic fatty liver disease; Non-alcoholic steatohepatitis; Fatty liver; Artificial intelligences; Steatosis; Fibrosis; Machine learning

Core Tip: Non-alcoholic fatty liver disease (NAFLD) exists on a spectrum from simple hepatocyte steatosis to non-alcoholic steatohepatitis (NASH) with ballooning and fibrosis. Given the lack of efficient screening methods and high rate of asymptomatic disease, it is challenging to identify patients with NAFLD in its various stages. Although liver biopsy remains the gold standard for diagnosing NASH, it is an invasive, costly, and painful procedure. Conventional imaging modalities including ultrasound, computed tomography, magnetic resonance imaging and transient elastography are limited by inter- and intra-observer variability depending on the stage of fibrosis. Similarly, despite recent progress in the prevention and treatment of viral hepatitis, predicting sustained virological response and disease progression remains challenging. Artificial intelligence (AI) is an exciting and increasingly pertinent field in medicine as clinicians incorporate augmenting technology into their daily practice. This review summarizes recent literature on the application of AI in NAFLD and viral hepatitis. Specifically, the review will assess the performance of AI as a non-invasive method for the diagnosis and staging of liver fibrosis and steatosis, as well as for the detection and treatment of chronic viral hepatitis. It will also aim to highlight the potential for AI based methods on their ability to develop therapeutic targets.

INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) exists on a spectrum from simple hepatocyte steatosis to inflammation, ballooning and fibrosis. Given the lack of efficient screening methods and high rate of asymptomatic disease, it is challenging to identify patients with various stages of NAFLD[1,2]. Non-alcoholic steatohepatitis (NASH) patients with significant fibrosis are at increased risk for cirrhosis and progressive liver failure, which has led NASH to become one of the leading causes of liver transplantation in the United States[3]. NASH affects approximately 3% to 6% of the US population, and this number continues to increase. It affects approximately 25% of the population worldwide[4].

Although liver biopsy remains the gold standard for diagnosing NASH, it is an invasive, costly, and painful procedure. Therefore, serial liver biopsies for surveillance are not always feasible. Conventional imaging modalities including ultrasound, computed tomography (CT), magnetic resonance imaging (MRI) and transient elastography are limited by inter- and intra-observer variability depending on the stage of fibrosis[1,2]. Similarly, despite recent progress in the prevention and treatment of viral hepatitis, predicting sustained virological response (SVR) and disease progression remains challenging.

Artificial intelligence (AI) is an exciting and increasingly pertinent field in medicine as clinicians incorporate augmenting technology into their daily practice. AI is the concept of teaching a computer to simulate the cognitive abilities of the human brain. Machine learning (ML) entails allowing the computer to simulate the human brain independently. It can either be supervised (through specific feedback from humans) or unsupervised, in which case there is no guidance provided and the computer is able to independently synthesize and analyze the output[1]. AI is increasingly applied to the diagnosis and prediction of various diseases. Researchers are developing machine learning (ML) algorithms to predict risk and outcomes using multiple demographic, clinical, biochemical, and imaging parameters for diagnosis and prognosis related to liver fibrosis and steatosis, including NAFLD and viral hepatitis[1].

Current methods of assessing liver fibrosis progression and mortality in both NAFLD and viral hepatitis have many limitations. These include the intra- and inter-observer variability in staging fibrosis, the inability to place fibrosis along a continuum, and the lack of identifiable markers for disease progression[1,2]. These limitations and the ability of ML models to overcome them will be discussed further in this review. This review will also highlight how ML models have the potential to present opportunities for drug discovery and prediction of therapeutic and toxic effects of drugs. Machine learning models based on AI provide promising features that could not only enhance screening for NAFLD, but also help with fibrosis staging in patients with NASH and viral hepatitis.

This review summarizes recent literature on the application of AI in NAFLD and viral hepatitis. The main objective is to assess the performance of AI as a non-invasive method for the diagnosis and staging of liver fibrosis and steatosis, as well as the detection and treatment of chronic viral hepatitis.

METHODS

A review of current literature in the areas of AI in NAFLD and viral hepatitis was conducted using two separate searches on PubMed. First, we used the search terms "non-alcoholic fatty liver disease", "NAFLD", and "deep learning" in combination with "artificial intelligence", "histology", "omics" and "radiology." The second search was conducted using the search terms "viral hepatitis" in combination with "hepatitis A", "hepatitis B", "hepatitis C", "hepatitis E", "machine learning", "artificial intelligence", "histology" and "radiology".

Most articles on NASH and NAFLD published between 2018 and 2021 were included in this review. Articles were excluded if they did not offer comparisons between AI modalities and existing methods for screening or prediction (MELD score, elastography, *etc*.). Twenty-seven articles were included in our review, 13 on NAFLD and 14 on chronic viral hepatitis. For studies on viral hepatitis, described machine learning algorithms fell into one of three categories: Predicting prevalence, screening for complications (including fibrosis, hepatocellular carcinoma, decompensated cirrhosis, and death), and predicting response to treatment.

USE OF AI FOR DIAGNOSING VIRAL HEPATITIS AND NAFLD/NASH

It is estimated that half of patients infected with hepatitis C worldwide are unaware of their diagnosis and only 17% have undergone liver fibrosis staging[5]. This rate is even lower for hepatitis B, for which only 10.5% of infected patients are aware of their status. In March 2020, the USPSTF recommended hepatitis C screening for all adults over 18; however, there are currently no population-based screening recommendations for hepatitis A and B. Primary care offices do not routinely test for hepatitis B. Machine learning has been used both to determine regional prevalence of chronic hepatitis and to identify undiagnosed cases.

Zheng *et al*[6] compared two algorithms (Elman neural network and autoregressive integrated moving average, or ARIMA) designed to predict incidence of hepatitis B in Guangxi, China. ARIMA is a type of model that can capture the randomness of data and is often used for infectious disease prediction. Predictions were compared to the reported cases of hepatitis B cases from the Health Commission of Guangxi, China. The neural network was the more predictive model, with a root-mean-square error (RMSE) of 0.89 and mean absolute error (MAE) of 0.70, while the ARIMA had an RSME of 0.94 and an MAE of 0.81.

A 2020 study by Doyle *et al*[7] aimed to predict chronic hepatitis C (HCV) positive status by using patient claims data to develop four algorithms, all with a predictive accuracy of over 95%. Algorithms included logistic regression, gradient boosted trees, a stacked ensemble, and random forests. The stacked ensemble performed the best, with a precision of 97% at recall levels > 50%. Key predictors of HCV infection included nonsteroidal anti-inflammatory drug use, opioids, healthcare utilization, patient age and osteoarthritis or glomerulonephritis treatment. We were unable to find any study to date using AI to screen for NAFLD/NASH.

USE OF AI TO ASSESS FIBROSIS IN VIRAL HEPATITIS AND NAFLD/NASH

Existing histologic models not only rely on scoring of fibrosis by a pathologist but are also unable to place fibrosis along a continuum. Artificial intelligence enables the placement of fibrosis along a continuum, identifies risk factors for progression of fibrosis, allows enhanced scoring of fibrosis stages, leading to better selection of patients for clinical trials This also allows for identification of therapeutic targets[2].

Lu *et al*[8] developed a light gradient-boosting machine model to predict liver fibrosis and cirrhosis in treatment-naive chronic hepatitis B patients at four centers in China. The model, named Fibro Box, outperformed transient elastography, APRI, and FIB-4, with area under the curve (AUC) 0.88 in external validation sets for significant fibrosis and 0.87 for cirrhosis. Input variables included fibroscan results, platelets, alanine aminotransferase (ALT), Prothrombin time (PT), and splenic vein diameter.

A 2013 study by Zheng *et al*[9], used an artificial neural network (ANN) to predict 3-month mortality of individuals with acute-on-chronic liver failure due to hepatitis B (HBV-ACLF). Patient characteristics included in this model were age, PT, serum sodium, total bilirubin, E antigen positivity status and hemoglobin. The ANN predicted mortality more accurately than MELD-based scoring systems, with area under the curve receiver operating characteristic (AUCROC) 0.765 in the validation cohort compared to 0.599 for MELD.

Similarly, Huo *et al*[10] developed ANNs to predict 28- and 90-d mortality in HBV-ACLF. Data were retrospectively reviewed from 684 patients admitted for ALF at 8 hospitals in various Chinese provinces with 423 cases in the training cohort and 261 in the validation cohort. In the training cohorts, the neural network had a significantly higher accuracy than MELD, MELD-Na, CLIF-ACLF, and Child-Pugh score, with AUC 0.948 and 0.913 for 28- and 90-d mortality, respectively. In the validation cohort, the model

performed significantly better than MELD and insignificantly better than other scoring systems, with AUC 0.748 and 0.754 for 28- and 90-d mortality. Significant mortality predictors included age, presence of HE, sodium, PT, gamma-glutamyl transpeptidase (GGT), e antigen, alkaline phosphatase, and bilirubin.

In another study, Wang *et al*[11] used deep learning radiomics of elastography (DLRE) to assess stages of liver fibrosis in patients with chronic hepatitis B. DLRE was compared to 2D shear wave elastography and biomarkers (AST: Platelet ratio, fibrosis index), with liver biopsy as the reference standard. 1990 images from 398 patients were used to develop the models. AUCROCs for DLRE were 0.97 for cirrhosis, 0.98 for advanced fibrosis, and 0.85 for significant fibrosis; this performed better than other methods except for elastography in severe fibrosis.

Like viral hepatitis, there are several studies establishing the role of AI in assessing fibrosis in NAFLD/NASH. In one study by Forlano *et al*[2], liver biopsy specimens were annotated by two expert pathologists using the clinical research network (CRN) score as a measurable scale of degree of steatosis, inflammation, ballooning and fibrosis. The machine learning model was built using 100 patients with NAFLD in the derivation group and 146 patients in the validation group. There was good concordance when the machine learning model was compared to the scoring of the expert histopathologist on the liver biopsy specimens; the interclass correlation coefficients were 0.97 (95%CI, 0.95-0.99; *P* value < 0.001) for steatosis, 0.96 (95%CI, 0.9-0.98; *P* value < 0.001) for inflammation, 0.94 (95%CI, 0.87-0.98; *P* value < 0.001 for ballooning, and 0.92 for fibrosis (95%CI, 0.88-0.96; *P* value < 0.001). A subgroup analysis showed that quantitative analysis performed better than the CRN score in differentiating between the various stages of NAFLD. Another CNN model developed by Qu *et al*[12], showed that a convolutional neural network (CNN) model had an area under the curve (AUC) of 63% for all four subsets of the NAFLD scoring, while the AUC's were 90.48% for steatosis, 81.06% for ballooning, 70.18% for inflammation and 83.85% for fibrosis. These studies underscore the utility of ML models in illustrating the heterogeneity of liver pathology in NAFLD[9,26].

In another study by Taylor-Weiner *et al*[13], a CNN model was developed that allowed for assessment of fibrosis along a continuum, which is not possible with pathologist scoring alone. The CRN and Ishak scores were applied to each pixel within a given image, allowing for evaluation of heterogeneity in fibrosis as well. In addition, the CNN served as a prediction model allowing for identification of features associated with disease progression. The model's predictions correlated significantly with the pathologist scoring in all three studies, the STELLAR-3, STELLAR-4, and ATLAS, whose participants were used to build and validate the ML model - steatosis, ρ = 0.60; *P* value < 0.001; lobular inflammation, ρ = 0.35; *P* value < 0.001; and HB, ρ = 0.41; *P* value < 0.001. The model's level of agreement with pathologist scoring was within the range of agreement between individual pathologists. The weighted Cohen's kappa was 0.801 for NASH CRN and 0.817 for the Ishak classifications.

Another study by Gawrieh *et al*[14] built a ML model using support vector machines (SVM) to better characterize architectural patterns in fibrosis. This ML model was built to differentiate between six different patterns of fibrosis and had a strong correlation with the pathologist's semi-quantitative scores for fibrosis, with a coefficient of determination of automated CPA ranging between 0.60 to 0.86 when compared with the pathologist score. The model was built using a trichrome-stained liver biopsy specimen which was marked with 987 annotations for different fibrosis types. As noted in the study, the model's AUROCs were 78.6% for detection of periportal fibrosis, 83.3% for pericellular fibrosis, 86.4% for portal fibrosis, and > 90% for detection of normal fibrosis, bridging fibrosis and presence of nodules/cirrhosis.

AI USING METABOLOMICS FOR NAFLD/NASH

There is an increasing number of studies focusing on metabolomics that allow for non-invasive identification of targets associated with development and progression of NAFLD. These biomarkers may differentiate between patients with and without cirrhosis, and between a healthy liver and NAFLD or NASH[3,15,16]. Several direct and indirect blood-based biomarkers currently exist to assess fibrosis. These have been incorporated to form scoring systems such as NAFLD fibrosis score (NFS), Fibrosis-4 (FIB-4), AST to platelet ratio index (APRI), BARD Score, FibroSURE and Enhanced liver fibrosis score [3]. ML allows for analysis of many multi-omics and clinical variables to screen for NASH and NAFLD and to build models for disease progression.

An eXtreme Gradient Boosting Model (XG Boost) was developed using the NIDDK database by Docherty *et al*[16], which contains a large real-world patient population. This model used confirmed NASH and non-NASH patients within this subset. The unique feature of this study is that it used several demographic variables and clinical biomarkers run through recursive feature elimination in combination with confirmed histologic cases to build an efficient model with a high specificity. When a greater number of markers were used in predicting patients with NASH, the AUROC was 0.82, sensitivity 81%, and precision 81%.

In a study of adults of European ancestry by Atabaski-Pasdar *et al*[15], patients with type 2 diabetes and others with high-risk features for the development of NASH were assessed for liver fat content

using MRI. Several multi-omics and clinical data, including laboratory markers, were entered into the least absolute shrinkage and selection operator to select the most relevant features, which then underwent random forest analysis for the development of the algorithm. The model developed using this method produced a cross-validated AUROC of 0.84 (95%CI 0.82, 0.86; *P* value < 0.001) and outperformed existing prediction tools for NAFLD. However, unlike other studies, the model was built in comparison to MRI fat content, which is not reflective of the continuum of NAFLD, and thus cannot be used to monitor disease progression.

Another study based in China by Ma *et al*[17] identified BMI, triglycerides, GGT, the serum ALT and uric acid as the most common features contributing to NAFLD when a Bayesian network model was used. The model had an accuracy of 83%, specificity of 0.878, sensitivity of 0.675, and F-measure score of 0.655. The F-measure score is an indicator of whether there can be a balance between precision and recall of these variables, and it was higher than for logistic regression models in machine learning.

AI IN IMAGE INTERPRETATION FOR NAFLD/NASH

Like markers discussed previously, many studies have combined machine learning with imaging modalities to more effectively assess liver fat content and to better define fibrosis scores. This would allow for more accurate monitoring of patients for disease progression and their selection for clinical trials.

Current modalities for estimation of liver fat content include conventional ultrasound (US), which is limited by variable accuracy, operator dependency, and its qualitative nature. The measurement of proton density fat fraction (PDFF) by MRI is proving to be an effective method for quantification of hepatic steatosis, but it is expensive and there is variability in results due to dependence on calibration. In a study by Han *et al*[18], one-dimensional CNN was applied to ultrasound radiofrequency signals for the diagnosis of NAFLD and quantitation of hepatic fat content with an AUC of 0.98 (95%CI: 0.94, 1.00). In diagnosing NAFLD, the model had an accuracy of 96%, sensitivity of 97%, and specificity of 94%, PPV of 97% and NPV of 94%. The ML model also correlated with MRI-PDFF with a Pearson correlation coefficient of 0.85 (*P* value < 0.001). The same method was applied to animal models in a study by Nguyen *et al*[19] and it showed that CNN outperformed quantitative ultrasound in differentiating between NAFLD and normal liver. Further support for ML comes from a recent study by Das *et al*[20] on pediatric patients which used an ensemble model comprising SVM, Neural Net and XG Boost that had an AUC of 0.92 (95%CI, 0.91-0.94) when tested in an external validation cohort.

Nonenhanced CT also remains superior to histopathologic quantification of liver fat content like MRI-PDFF, but it is also more commonly performed in clinical practice for other reasons when compared to MRI. It currently uses a manual region-of-interest (ROI) for estimation of liver fat content. A study by Graffy *et al*[21] developed a deep-learning based automated liver segmentation tool and applied it to estimate liver fat content using three-dimensional CNN, without having to depend on manual ROI. The pearson correlation coefficient was 0.93. This allows for large population level estimation of liver fat content to determine the prevalence of NAFLD. It would also determine normal liver fat content based on a large sample. Used in combination with other non-invasive modalities such as serum biomarkers, it could help identify patients who will need closer monitoring for NAFLD progression to cirrhosis. In a similar study by Hou *et al*[22], the automated liver attenuation ROI-based measurement model had a pearson coefficient of 0.94 when compared with manual ROI.

In addition to differentiating healthy liver from NAFLD, ML models have also been used to reduce variability in detecting fibrosis, specifically F2 fibrosis, which is a limiting feature of shear wave elastography. A study by Brattain *et al*[23] combined the use of shear wave elastography with CNN to better assess F2 fibrosis. This approach not only assessed image quality, but also selected ROI, unlike the previous studies. This ML model detected F2 fibrosis with AUC of 0.89 compared to AUC of 0.74 when image quality and ROI were not incorporated into a ML model. This demonstrates the importance of ML models once again in selecting patients for clinical trials, and in assessing response to treatment.

AI IN VIRAL HEPATITIS TREATMENT

The rate of SVR for hepatitis C with modern direct acting antiviral (DAA) regimens is estimated to be over 90%; however, variability remains in treatment length and efficacy. Patients with prior DAA exposure, cirrhosis, and other risk factors may require a longer treatment course[18,24]. Machine learning has been applied to predicting treatment response and duration based on patient-specific factors.

Haga *et al*[24] applied nine machine learning algorithms to identify the optimized combination of HCV genotypic variants that predict SVR after DAA therapy. HCV genomes were sequenced from the serum of 173 patients (including 64 without SVR). The support vector machine algorithm was found to be the most predictive, with a validation accuracy of 0.95. Feldman *et al*[25] used data from 60 million beneficiaries of a managed care plan (including 3943 cases of hepatitis C who received sofos-
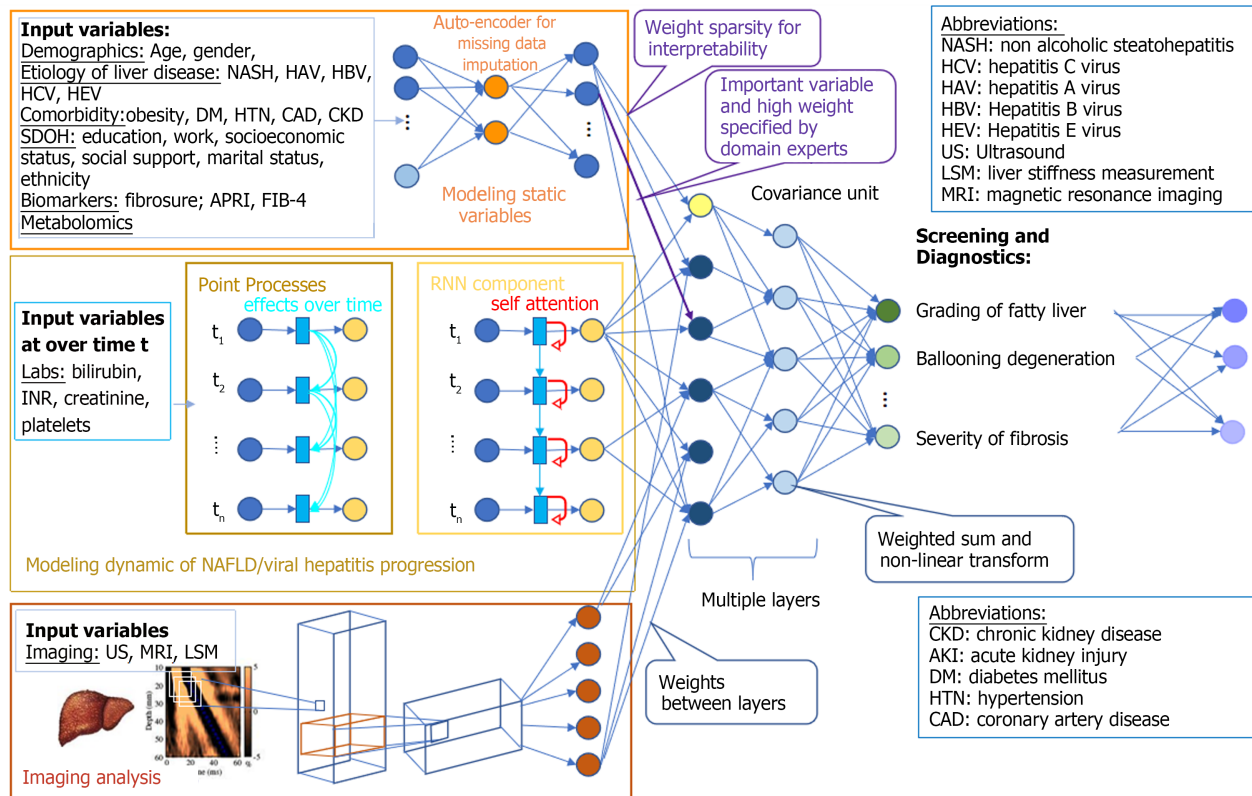
Figure 1 Framework of artificial intelligence based dynamic of non-alcoholic fatty liver disease/viral hepatitis diagnosis, progression and outcomes.

buvir/ledipasvir), to identify demographic and medical factors that may predict a prolonged course of DAA. Machine learning algorithms included extreme gradient boosting (XG Boost), random forest and support vector machine, with XG Boost being the optimal predictive model at an AUC of 0.745. Patient age, comorbidity burden, and type 2 diabetes status were significant predictors. Wei *et al*[26] developed an ANN and logistic regression model to predict fibrosis reversal after 78 wk of hepatitis B treatment. Significant predictors included AST and ALT, platelets, WBC, gender, and Fibroscan results. The ANN outperformed the logistic regression model, with an AUC of 0.81 *vs* 0.75.

The only approved treatment for NAFLD is weight reduction. We were unable to find AI based algorithms and predictive models for NAFLD due to lack of pharmacologic management options.

DISCUSSION

Among the algorithms described, more complex models performed better, with machine learning consistently outperforming more basic logistic regression models. The highest-performing models incorporated both demographic and radiologic/serologic variables. AI models also predicted complications more accurately than biomarkers and scoring systems like MELD and FIB-4. These models could be used to predict the incidence and prevalence of viral hepatitis in regions without robust, widespread screening programs. Additionally, they could be helpful in the initiation of treatment and predicting response to antivirals for individual patients, for which no gold standard currently exists.

Limitations of the current AI models are notably due to the lack of large scale, randomized controlled trials. Further research is necessary to demonstrate the utility of AI. With further advancements, ML models could potentially be incorporated into all aspects of a patient's care, from screening the general population for NAFLD or NASH, to monitoring disease progression and treatment response in clinical trials by enhancing classification of steatosis, ballooning, inflammation, and fibrosis. In this regard, more population-based studies are needed to study the applications of ML models in screening. Additionally, large scale, randomized controlled trials are needed to study serologic and histologic markers for disease progression. Further studies are also warranted to explore the potential of ML algorithms to provide target-specific medications, yielding efficacious pharmacotherapy in a disease such as NASH where good treatment options are lacking at this time. Though AI is promising in terms of its potential to develop therapeutic targets, we were unable to find any studies to date describing the use of AI in drug discovery.

Future directions also include using AI to actively improve outcomes with viral hepatitis by increasing adherence to DAAs or identifying individuals at risk for contracting viral hepatitis. Machine learning models could also help identify barriers to accessing treatment.

CONCLUSION

Machine learning models focus on various aspects of liver disease, including demographics, biochemical labs, histologic assessment and patterns, identification of non-invasive biomarkers, and liver imaging techniques (Figure 1). Overall, the studies outlined above are promising in their reliance on non-invasive methods as opposed to conventional liver biopsy to study the stages of fibrosis, as well as their ability to place fibrosis along a continuum and identify markers for disease progression. This could reduce healthcare costs by allowing better selection of patients in whom a liver biopsy is performed. It would also benefit patients by decreasing the number of them who undergo this invasive procedure. AI can also improve efficiency of pathologist and sonographer scoring of samples when added to existing methods. This will allow for a better understanding of the pathophysiology of diseases like NAFLD, which would not only allow for appropriate screening for disease progression, but also improve the ability to develop therapeutic targets.

FOOTNOTES

ORCID number: Atchayaa Gunasekharan 0000-0001-7868-6513; Joanna Jiang 0000-0002-7995-6499; Ashley Nickerson 0000-0002-4494-9985; Sajid Jalil 0000-0001-6123-153X; Khalid Mumtaz 0000-0001-7868-6514.

REFERENCES

1   **Aggarwal P**, Alkhouri N. Artificial Intelligence in Nonalcoholic Fatty Liver Disease: A New Frontier in Diagnosis and Treatment. *Clin Liver Dis (Hoboken)* 2021; **17**: 392-397 [PMID: 34386201 DOI: 10.1002/cld.1071]

2   **Forlano R**, Mullish BH, Giannakeas N, Maurice JB, Angkathunyakul N, Lloyd J, Tzallas AT, Tsipouras M, Yee M, Thursz MR, Goldin RD, Manousou P. High-Throughput, Machine Learning-Based Quantification of Steatosis, Inflammation, Ballooning, and Fibrosis in Biopsies From Patients With Nonalcoholic Fatty Liver Disease. *Clin Gastroenterol Hepatol* 2020; **18**: 2081-2090.e9 [PMID: 31887451 DOI: 10.1016/j.cgh.2019.12.025]

3   **Perakakis N**, Polyzos SA, Yazdani A, Sala-Vila A, Kountouras J, Anastasilakis AD, Mantzoros CS. Non-invasive diagnosis of non-alcoholic steatohepatitis and fibrosis with the use of omics and supervised learning: A proof of concept study. *Metabolism* 2019; **101**: 154005 [PMID: 31711876 DOI: 10.1016/j.metabol.2019.154005]

4   **Leow WQ**, Bedossa P, Liu F, Wei L, Lim KH, Wan WK, Ren Y, Chang JP, Tan CK, Wee A, Goh GB. An Improved qFibrosis Algorithm for Precise Screening and Enrollment into Non-Alcoholic Steatohepatitis (NASH) Clinical Trials. *Diagnostics (Basel)* 2020; **10** [PMID: 32872090 DOI: 10.3390/diagnostics10090643]

5   **Yehia BR**, Schranz AJ, Umscheid CA, Lo Re V 3rd. The treatment cascade for chronic hepatitis C virus infection in the United States: a systematic review and meta-analysis. *PLoS One* 2014; **9**: e101554 [PMID: 24988388 DOI: 10.1371/journal.pone.0101554]

6   **Zheng Y**, Zhang L, Zhu X, Guo G. A comparative study of two methods to predict the incidence of hepatitis B in Guangxi, China. *PLoS One* 2020; **15**: e0234660 [PMID: 32579598 DOI: 10.1371/journal.pone.0234660]

7   **Doyle OM**, Leavitt N, Rigg JA. Finding undiagnosed patients with hepatitis C infection: an application of artificial intelligence to patient claims data. *Sci Rep* 2020; **10**: 10521 [PMID: 32601354 DOI: 10.1038/s41598-020-67013-6]

8   **Lu XJ**, Yang XJ, Sun JY, Zhang X, Yuan ZX, Li XH. FibroBox: a novel noninvasive tool for predicting significant liver fibrosis and cirrhosis in HBV infected patients. *Biomark Res* 2020; **8**: 48 [PMID: 33005419 DOI: 10.1186/s40364-020-00215-2]

9   **Zheng MH**, Shi KQ, Lin XF, Xiao DD, Chen LL, Liu WY, Fan YC, Chen YP. A model to predict 3-month mortality risk of acute-on-chronic hepatitis B liver failure using artificial neural network. *J Viral Hepat* 2013; **20**: 248-255 [PMID: 23490369 DOI: 10.1111/j.1365-2893.2012.01647.x]

10   **Huo Y**, Terry JG, Wang J, Nair S, Lasko TA, Freedman BI, Carr JJ, Landman BA. Fully automatic liver attenuation estimation combing CNN segmentation and morphological operations. *Med Phys* 2019; **46**: 3508-3519 [PMID: 31228267 DOI: 10.1002/mp.13675]

11   **Wang K**, Lu X, Zhou H, Gao Y, Zheng J, Tong M, Wu C, Liu C, Huang L, Jiang T, Meng F, Lu Y, Ai H, Xie XY, Yin LP, Liang P, Tian J, Zheng R. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. *Gut* 2019; **68**: 729-741 [PMID: 29730602 DOI: 10.1136/gutjnl-2018-316204]

12   **Qu H**, Minacapelli CD, Tait C, Gupta K, Bhurwal A, Catalano C, Dafalla R, Metaxas D, Rustgi VK. Training of computational algorithms to predict NAFLD activity score and fibrosis stage from liver histopathology slides. *Comput Methods Programs Biomed* 2021; **207**: 106153 [PMID: 34020377 DOI: 10.1016/j.cmpb.2021.106153]

13   **Taylor-Weiner A**, Pokkalla H, Han L, Jia C, Huss R, Chung C, Elliott H, Glass B, Pethia K, Carrasco-Zevallos O, Shukla C, Khettry U, Najarian R, Taliano R, Subramanian GM, Myers RP, Wapinski I, Khosla A, Resnick M, Montalto MC, Anstee QM, Wong VW, Trauner M, Lawitz EJ, Harrison SA, Okanoue T, Romero-Gomez M, Goodman Z, Loomba R, Beck AH, Younossi ZM. A Machine Learning Approach Enables Quantitative Measurement of Liver Histology and Disease Monitoring in NASH. *Hepatology* 2021; **74**: 133-147 [PMID: 33570776 DOI: 10.1002/hep.31750]

14   **Gawrieh S**, Sethunath D, Cummings OW, Kleiner DE, Vuppalanchi R, Chalasani N, Tuceryan M. Automated quantification and architectural pattern detection of hepatic fibrosis in NAFLD. *Ann Diagn Pathol* 2020; **47**: 151518 [PMID: 32531442 DOI: 10.1016/j.anndiagpath.2020.151518]

15   **Atabaki-Pasdar N**, Ohlsson M, Viñuela A, Frau F, Pomares-Millan H, Haid M, Jones AG, Thomas EL, Koivula RW, Kurbasic A, Mutie PM, Fitipaldi H, Fernandez J, Dawed AY, Giordano GN, Forgie IM, McDonald TJ, Rutters F, Cederberg H, Chabanova E, Dale M, Masi F, Thomas CE, Allin KH, Hansen TH, Heggie A, Hong MG, Elders PJM, Kennedy G, Kokkola T, Pedersen HK, Mahajan A, McEvoy D, Pattou F, Raverdy V, Häussler RS, Sharma S, Thomsen HS, Vangipurapu J, Vestergaard H, 't Hart LM, Adamski J, Musholt PB, Brage S, Brunak S, Dermitzakis E, Frost G, Hansen T, Laakso M, Pedersen O, Ridderstråle M, Ruetten H, Hattersley AT, Walker M, Beulens JWJ, Mari A, Schwenk JM, Gupta R, McCarthy MI, Pearson ER, Bell JD, Pavo I, Franks PW. Predicting and elucidating the etiology of fatty liver disease: A machine learning modeling and validation study in the IMI DIRECT cohorts. *PLoS Med* 2020; **17**: e1003149 [PMID: 32559194 DOI: 10.1371/journal.pmed.1003149]

16   **Docherty M**, Regnier SA, Capkun G, Balp MM, Ye Q, Janssens N, Tietz A, Löffler J, Cai J, Pedrosa MC, Schattenberg JM. Development of a novel machine learning model to predict presence of nonalcoholic steatohepatitis. *J Am Med Inform Assoc* 2021; **28**: 1235-1241 [PMID: 33684933 DOI: 10.1093/jamia/ocab003]

17   **Ma H**, Xu CF, Shen Z, Yu CH, Li YM. Application of Machine Learning Techniques for Clinical Predictive Modeling: A Cross-Sectional Study on Nonalcoholic Fatty Liver Disease in China. *Biomed Res Int* 2018; **2018**: 4304376 [PMID: 30402478 DOI: 10.1155/2018/4304376]

18   **Han A**, Byra M, Heba E, Andre MP, Erdman JW Jr, Loomba R, Sirlin CB, O'Brien WD Jr. Noninvasive Diagnosis of Nonalcoholic Fatty Liver Disease and Quantification of Liver Fat with Radiofrequency Ultrasound Data Using One-dimensional Convolutional Neural Networks. *Radiology* 2020; **295**: 342-350 [PMID: 32096706 DOI: 10.1148/radiol.2020191160]

19   **Nguyen TN**, Podkowa AS, Park TH, Miller RJ, Do MN, Oelze ML. Use of a convolutional neural network and quantitative ultrasound for diagnosis of fatty liver. *Ultrasound Med Biol* 2021; **47**: 556-568 [PMID: 33358553 DOI: 10.1016/j.ultrasmedbio.2020.10.025]

20   **Das A**, Connell M, Khetarpal S. Digital image analysis of ultrasound images using machine learning to diagnose pediatric nonalcoholic fatty liver disease. *Clin Imaging* 2021; **77**: 62-68 [PMID: 33647632 DOI: 10.1016/j.clinimag.2021.02.038]

21   **Graffy PM**, Sandfort V, Summers RM, Pickhardt PJ. Automated Liver Fat Quantification at Nonenhanced Abdominal CT for Population-based Steatosis Assessment. *Radiology* 2019; **293**: 334-342 [PMID: 31526254 DOI: 10.1148/radiol.2019190512]

22   **Hou Y**, Zhang Q, Gao F, Mao D, Li J, Gong Z, Luo X, Chen G, Li Y, Yang Z, Sun K, Wang X. Artificial neural network-based models used for predicting 28- and 90-day mortality of patients with hepatitis B-associated acute-on-chronic liver failure. *BMC Gastroenterol* 2020; **20**: 75 [PMID: 32188419 DOI: 10.1186/s12876-020-01191-5]

23   **Brattain LJ**, Telfer BA, Dhyani M, Grajo JR, Samir AE. Objective Liver Fibrosis Estimation from Shear Wave Elastography. *Annu Int Conf IEEE Eng Med Biol Soc* 2018; **2018**: 1-5 [PMID: 30440285 DOI: 10.1109/EMBC.2018.8513011]

24   **Haga H**, Sato H, Koseki A, Saito T, Okumoto K, Hoshikawa K, Katsumi T, Mizuno K, Nishina T, Ueno Y. A machine learning-based treatment prediction model using whole genome variants of hepatitis C virus. *PLoS One* 2020; **15**: e0242028 [PMID: 33152046 DOI: 10.1371/journal.pone.0242028]

25   **Feldman TC**, Dienstag JL, Mandl KD, Tseng YJ. Machine-learning-based predictions of direct-acting antiviral therapy duration for patients with hepatitis C. *Int J Med Inform* 2021; **154**: 104562 [PMID: 34482150 DOI: 10.1016/j.ijmedinf.2021.104562]

26   **Wei W**, Wu X, Zhou J, Sun Y, Kong Y, Yang X. Noninvasive Evaluation of Liver Fibrosis Reverse Using Artificial Neural Network Model for Chronic Hepatitis B Patients. *Comput Math Methods Med* 2019; **2019**: 7239780 [PMID: 31428186 DOI: 10.1155/2019/7239780]

Machine learning in endoscopic ultrasonography and the pancreas: The new frontier?

Cem Simsek, Linda S Lee

Cem Simsek, Department of Gastroenterology, Hepatology and Endoscopy, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02215, United States

Linda S Lee, Division of Gastroenterology, Hepatology and Endoscopy, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, United States

Corresponding author: Linda S Lee, MD, Associate Professor, Division of Gastroenterology, Hepatology and Endoscopy, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115, United States. lslee@partners.org

Abstract

Pancreatic diseases have a substantial burden on society which is predicted to increase further over the next decades. Endoscopic ultrasonography (EUS) remains the best available diagnostic method to assess the pancreas, however, there remains room for improvement. Artificial intelligence (AI) approaches have been adopted to assess pancreatic diseases for over a decade, but this methodology has recently reached a new era with the innovative machine learning algorithms which can process, recognize, and label endosonographic images. Our review provides a targeted summary of AI in EUS for pancreatic diseases. Included studies cover a wide spectrum of pancreatic diseases from pancreatic cystic lesions to pancreatic masses and diagnosis of pancreatic cancer, chronic pancreatitis, and autoimmune pancreatitis. For these, AI models seemed highly successful, although the results should be evaluated carefully as the tasks, datasets and models were greatly heterogenous. In addition to use in diagnostics, AI was also tested as a procedural real-time assistant for EUS-guided biopsy as well as recognition of standard pancreatic stations and labeling anatomical landmarks during routine examination. Studies thus far have suggested that the adoption of AI in pancreatic EUS is highly promising and further opportunities should be explored in the field.

Key Words: Artificial intelligence; Pancreas; Endoscopic ultrasonography; Pancreatic cancer; Autoimmune pancreatitis; Pancreatic cystic lesions

Core Tip: Several reviews in the literature have discussed the use of artificial intelligence in pancreatic disease. However, this is the first review that focuses on the application of artificial intelligence (AI) specifically to endoscopic ultrasonography (EUS) of the pancreas, including pancreatic cystic lesions, pancreatic cancer, chronic pancreatitis, and autoimmune pancreatitis, where it appears to enhance EUS diagnosis. AI may also offer real-time assistance during procedures to direct biopsy towards the highest yield areas as well augment EUS training.

INTRODUCTION

Pancreatic diseases create a substantial burden on society. Pancreatic cancer is the third leading cause of cancer-related death in the United States, and its death count is expected to rise to 460000 by 2040, becoming the second leading cause of cancer related death in 2040[1-3]. Chronic pancreatitis is another cause of the burden with significant morbidity from chronic pain, diabetes mellitus, and even pancreatic cancer[4,5]. Additionally, pancreatic cystic lesions are reported to be detected up to 20% of abdominal imaging studies[6]. Endoscopic ultrasonography (EUS) has surpassed magnetic resonance imaging (MRI), computed tomography (CT) and transabdominal ultrasonography in the diagnosis of pancreatic diseases; however, there remains room for improvement in the diagnostic sensitivity of EUS[7]. In this regard, utilization of artificial intelligence (AI) with EUS has emerged as a promising strategy (Figure 1). Although EUS has better performance than the alternative radiology imaging methods, it is also more operator dependent. The endosonographer's experience and skills can significantly alter the diagnostic or therapeutic outcomes of an EUS procedure. AI may decrease this operator dependency as it can assist the endosonographer in several tasks that include, but are not limited, to identifying anatomical landmarks, detecting lesions, interpreting sonographic findings, and guiding obtaining optimal tissue biopsy with higher diagnostic yield. Because AI algorithms use higher resolution EUS imaging data, they might distinguish patterns and identify details from the images which may not be recognizable with human detection alone currently. Finally, AI research with EUS is more convenient because imaging data used to train the AI models often have readily available definitive histologic diagnoses.

Targeted summary of AI and research

AI is an umbrella term for the computerized performance of complex tasks that normally require human intelligence, such as visual perception, learning, pattern recognition and decision-making[8] (Figure 2). Current medical applications using AI have made significant progress due to advancements in computer technology, data science, and the digitalization of health care. From the development of more complex machine learning algorithms, AI has progressed rapidly to its current front-line role in image-based diagnosis, speech recognition, robotic surgery, drug discovery and patient monitoring[9]. However, the progress of AI in medicine has just begun and has yet to realize its full potential.

Machine learning (ML) is a field of artificial intelligence in which algorithms learn and improve from interactions with the data, obviating the need for explicit programming. Deep learning (DL) is a subfield of ML inspired by the organization and working principle of the human brain and is made up of individual neurons which form multilayered artificial neural networks (ANN). These networks are comprised of input and output layers each of which can execute simple tasks and sequentially interact with one another to produce a conclusion. Among ANNs, Multi-Layered Perceptron are earlier models that are simpler with fewer layers and can only use linear functions[10]. Convolutional neural networks (CNN) include more layers that can also operate in a non-linear fashion allowing more complex tasks such as image classification and have been the most popular DL algorithm. CNNs were inspired by the human visual cortex and designed to process grid pattern data such as images. They have serial neural network layers to recognize and extract features from the input data, learn the patterns of features, and perform hierarchical organization through the layers to search for the intended output (Figure 3)[11]. Most commonly used CNN algorithms are AlexNet, ResNet, U-Net, which all work using the same principle, and the technical details are beyond the scope of this review[12]. Another type of ANN is recurrent neural network (RNN), which also contains a multi-layered structure. In addition, each neuron in this network has its own internal memory, which taken altogether constitutes a collective memory of the network. This neural network can remember previous input data and use it to process subsequent inputs. Therefore, these algorithms are beneficial in processing sequential data such as before and after an intervention or time series data. An example of RNN is the long short-term memory model[13].

Advantages of AI in EUS
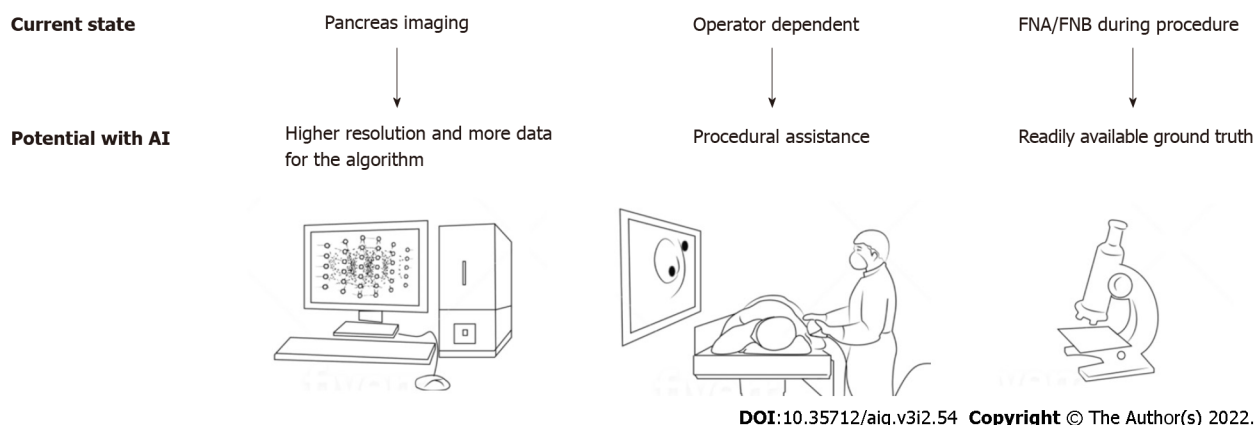
Figure 1 Potential benefit of artificial intelligence in pancreatic endoscopic ultrasonography. Current state of the pancreatic endoscopic ultrasonography (EUS) demonstrates that the procedure yields high resolution pancreatic imaging data, it is operator dependent, and allows acquisition of fine-needle aspiration (FNA) and fine-needle biopsy (FNB). Potentials with artificial intelligence (AI) implementation are utilizing of this higher resolution imaging data for training the algorithms with the readily available histologic ground truth from the FNA and FNB, as well as providing procedural assistance to address operator dependency.
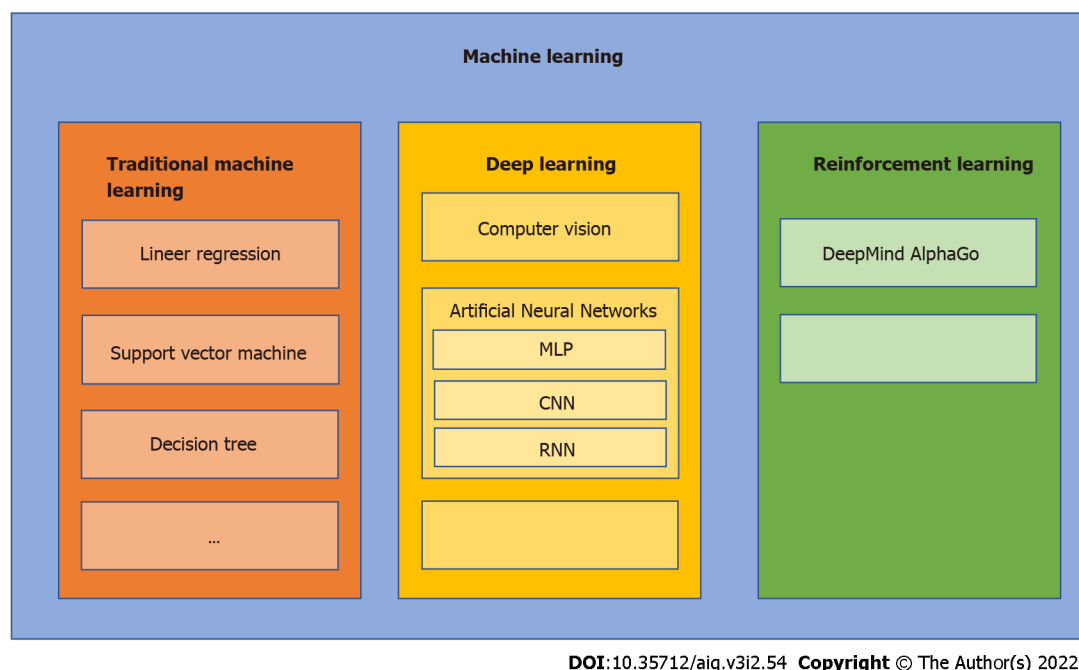
Figure 2 Overview of machine learning domains. Traditional machine learning algorithms rely on being trained by annotated and processed datasets to perform simpler tasks such as classification and regression. Deep learning algorithms are more autonomous, generally do not require annotation and processing of data for training and can perform more complicated tasks such as image detection and speech recognition. Reinforcement learning algorithms are self-teaching systems that can perform actions and learn by trial and error to achieve the best outcome; they perform most complex tasks such as game playing and learning to walk. CNN: Convolutional neural networks; MLP: Multi-layered perceptron; RNN: Recurrent neural networks.

Machine learning can perform two different types of tasks: Supervised and unsupervised. *Supervised* algorithms aim to reach a previously defined targeted outcome and are used for classification and prediction tasks. Labeled input data is presented to the algorithm and the model is trained with direct feedbacks to predict corresponding outputs. The spectrum of supervised approaches includes statistical methods such as logistic regression, linear regression, decision trees as well as support vector machines and random forest. *Unsupervised algorithms* do not have a predefined target and are used for clustering and dimensionality reduction. Unsupervised models are currently used for disease subtype and biomarker discovery studies[14,15]. Supervised learning has been more commonly used in EUS research; therefore, several important nuances will be summarized to better understand the presented literature. To train supervised learning algorithms, the dataset should be pre-annotated for the targeted

Deep neural network

Input layer   Muliple hidden layers   Output layer

DOI:10.35712/aig.v3i2.54 **Copyright** © The Author(s) 2022.
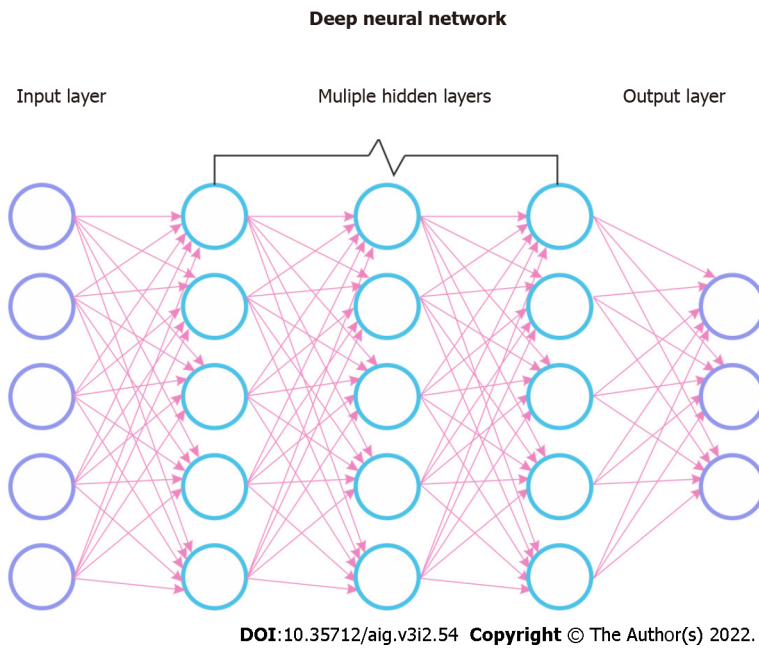
Figure 3 Design overview of deep neural network model including input, output, and multiple hidden feature-detecting layers.

outcome, which may be a diagnosis, class, or feature. The algorithm aims to optimize its feature detection ability to match the presented inputs to this annotated targeted output, which is defined as "ground truth". This optimization, or training, task requires large datasets, therefore, learning algorithms are data hungry. However, such datasets are not commonly available, which necessitates data augmentation techniques be used to expand the dataset by inserting slightly changed copies of previously collected data or by creating new synthetic data with computerized approaches.

During training of the algorithm, available data is split into three sets: Training, validation, and test. Training and validation sets are used to develop and fine-tune the model, whereas the test set is used to assess the performance of the final model product. Of note, this validation is different from its conventional use in medicine and seeks to optimize parameters of the model during the training phase. Two of the most common validation approaches in medical AI research are cross-validation and hold-out validation. *Cross-validation* occurs when the dataset is randomly resampled and split repetitively – the number of repetitions is designated with *k*- into training and test sets. Each training and test set is then used to develop a new model, and *k* repetitions yield new *k* models. In contrast, *hold-out validation* is a constant single split of a training set and an independent test set to develop one final model which is simpler to perform but brings an increased risk of sampling error. Another important concept in machine learning is *overfitting*, which is defined as a falsely superior performance of the model caused by learning irrelevant features of the dataset or 'noise' as well as the intended signals. Therefore, a separate test set is important to accurately assess the model's performance.

There are several nuances in the performance assessment of a machine learning model. Sensitivity (*recall*), specificity, positive predictive value (*precision*), negative predictive value and area under the rule operator characteristic (AUC) curve are commonly used for assessing the performance of classification. The *area under the precision-recall curve* (*AUPRC*) is used instead of the AUC when observations are not equally distributed for two groups. The *Dice coefficient* (*F1 score*) is the harmonic mean of precision and recall. It is commonly used to assess the labeling performance of an image recognition model. In a model where a ground truth area X is labeled by an image recognition model as area Y, Dice coefficient equals the overlap of X and Y areas divided by the total of X and Y areas, multiplied by two. Another similar metric is the Jaccard index, or intersection over union (IoU), defined as the ratio of overlap and union of two areas: the algorithm labeled area and the ground truth area. Both Jaccard index and Dice coefficient's values range from 0 to 1 signifying 0% to 100% accuracy of labeling with 1 being the highest level of accuracy for both.

While AI has been utilized to investigate numerous gastrointestinal diseases, the study of pancreatic diseases using AI and EUS is limited[5]. In this review, we provide a targeted overview of AI with a summary of the current literature on the use of AI in EUS for the diagnosis of pancreatic diseases.

METHODS

A nonsystematic search of the current literature was performed for 2015 and 2021 in the MEDLINE,

PubMed, Google Scholar, Scopus, Web of Science and Embase databases with the following terms: Machine learning, deep learning, artificial intelligence, EUS, endosonography, endoscopic ultrasound, pancreas, pancreatic disease, pancreatitis, and pancreatic cancer. Review articles were manually screened for any additional studies of interest. Congress abstracts, reviews, correspondences, editorials, and book chapters were excluded. Two authors reviewed all the studies after the initial search and confirmed the appropriateness of each study for inclusion. Our literature search yielded fifteen studies with modern machine learning algorithms (Table 1). Of note, five of the fifteen studies were published in 2021 with only two prospective clinical trials from the same group.

APPLICATIONS OF AI IN PANCREATIC EUS

The application of AI was divided into sonographic image recognition, procedural assistance, and training. Endosonographic images contain cues that may not be recognizable by human visual perception. In this context, deep learning algorithms are promising tools to recognize the patterns from these cues. As such, several important diagnostic challenges in pancreatic diseases with EUS have been addressed, including the classification and risk stratification of pancreatic cysts and the diagnosis of autoimmune pancreatitis (AIP) and pancreatic ductal adenocarcinoma (PDAC).

Pancreatic cystic neoplasms

Pancreatic cysts are increasingly detected in patients undergoing abdominal cross-sectional imaging with up to 20% detection rate on MRI[6,16,17]. Since pancreatic cysts carry a risk of malignancy, this risk should be stratified to guide clinical management. However, in most cases, imaging results are not sufficient for the classification of pancreatic cysts, especially for small lesions[18]. Additionally, assessing the risk of malignant progression remains challenging with current imaging modalities, clinical criteria, cyst fluid analysis or their combinations[18,19]. In this context, ML may help classify pancreatic cysts.

Several studies have investigated the utility of EUS ML models in pancreatic cysts, focusing on malignancy risk assessment and classification. Two studies by Kuwahara *et al*[23] and Nguon *et al*[21] used still images of EUS examinations with data augmentation, while Springer *et al*[20] and Kurita *et al* [22] applied multimodality approaches that included cyst fluid analyses and clinical data[20-23].

The 2019 study by Kuwahara *et al*[23] assessed the accuracy of ML to predict malignant intraductal papillary mucinous neoplasms (IPMN). This single-center study included 50 IPMN patients who underwent surgical resection. Therefore, all diagnoses were made from histopathological examination of surgical specimens. A total of 3970 still images were collected from 50 EUS examinations, and the CNN was fed over 500000 images using data augmentation. Ten-fold cross-validation was performed for training. For each case, the output of the CNN model was given as a predictive continuous value ranging from 0 to 1 for benign and malignant assigned probabilities, respectively. When the final model's predictive values were compared with the surgical diagnoses, predictive values for the benign cases were significantly lower than values for the malignant cases (0.104 *vs* 0.808, respectively). The optimal cutoff for the predictive value was determined using the Youden Index. This cutoff value (0.49) generated an AUC of 98% for the diagnosis of malignancy. The accuracy of the final model (94%) was significantly higher than that of human preoperative diagnosis which incorporated contrast enhanced EUS examination findings of mural nodule size, diameter main pancreatic duct, cyst size, and growth rate (56%). Multivariate analysis showed that the AI predictive value was the only significant factor for diagnosing malignant IPMN. ML outperformed currently used criteria, including serum CA 19-9, presence of mural nodule, and type of IPMN. This study demonstrated the promise of EUS ML algorithms in predicting malignant IPMNs. However, further prospective studies with larger sample sizes that do not rely solely on internal validation are necessary.

Kurita *et al*[22] used a multimodality approach to differentiate benign from malignant cysts. This single center study used 85 patients with pancreatic cystic lesions and final diagnosis from surgical pathology or combination of cyst fluid analysis, radiology imaging, and clinical follow-up. The input data consisted of sex, cyst fluid protein markers, cytologic diagnosis and EUS imaging features of the cyst. A Multi-layered Perceptron was used as the ML model. The final model achieved 95.7% sensitivity, 91.9% specificity, and 0.97 AUC for classifying lesions as benign or malignant, which was the primary endpoint. The model showed 92.9% accuracy which was significantly higher than carcinoembryonic antigen (CEA) (71.8%) and cytology (85.9%) alone[22]. An external data set was not available to test the algorithm. In addition, it is unclear why the algorithm did not mention inclusion of known high-risk features including enhancing nodule, solid mass, and dilated main pancreatic duct.

Another large multicenter study used a ML based approach called CompCyst to guide the management of pancreatic cystic lesions and relied heavily on molecular analysis of cyst fluid in addition to clinical and radiologic imaging features. The study population consisted of 862 patients recruited from 16 centers who underwent surgical resection with final diagnosis based on histologic analysis. DNA from cyst fluid were extracted and evaluated for four types of molecular abnormalities including mutations, loss of heterozygosity, aneuploidy as well as protein markers CEA and vascular

Table 1 Summary of included machine learning studies on endoscopic ultrasonography in pancreatic disease

Field	Ref.	Study population used for training (*n*)	Task	Machine learning method	Performance (in test population if available)
Pancreatic Cysts	Kuwahara *et al*[23], 2019	Benign IPMN (27); Malignant IPMN (23)	Differentiate benign from malignant IPMN	Convolutional neural network	AUC = 0.98
	Springer *et al*[20], 2019	Mucinous cystic neoplasms (153); Serous Cystic Neoplasms (148); IPMN (447); Malignant cysts (114)	Guide clinical management by classify into three risk groups: No risk of malignancyLow risk of progression. High-risk of progression or malignant	Not available	First group: 100% specificity, 46% sensitivity. Second group: 54% specificity, 91% sensitivity. Third group: 30% specificity, 99% sensitivity.
	Kurita *et al*[22], 2019	Mucinous cystic neoplasms (23); Serous Cystic Neoplasms (15); IPMN (30); Other cyst types (17)	Differentiate benign from malignant cyst	Multi-layered perceptron	AUC = 0.96, sensitivity: 95%, specificity: 91.9%
	Nguon *et al*[21], 2021	Mucinous cystic neoplasms (59); Serous Cystic Neoplasms (49)	Differentiate mucinous cystic neoplasm and serous cystadenoma	Convolutional neural network	AUC = 0.88
Pancreatic Cancer	Saftouiu *et al*[27], 2008	PDAC (32); Normal pancreas (22); Chronic pancreatitis (11); Pancreatic neuroendocrine tumor (3)	Differentiate benign from malignant masses	Multi-layered perceptron	AUC = 0.96
	Saftoiu *et al*[28], 2012	PDAC (211); Chronic pancreatitis (47)	Differentiate cancer from benign masses	Multi-layered perceptron	AUC = 0.94
	Ozkan *et al*[30], 2016	PDAC (202); Normal pancreas (130)	Differentiate cancer from normal pancreas	Multi-layered perceptron	Accuracy: 87.5%, sensitivity: 83.3%, and specificity: 93.3%
	Udristou *et al*[31], 2021	PDAC (30); Chronic pancre-atitis (20); Pancreatic neuroendocrine tumor (15)	Diagnose focal pancreatic mass	Convolutional neural network and long short-term memory	Mean AUC = 0.98 (Includes PDAC, CP and PNET)
	Tonozuka *et al*[32], 2021	PDAC (76); Chronic pancre-atitis (34); Control (29)	Differentiate pancreatic cancer from chronic pancreatitis and normal pancreas	Convolutional neural network and pseudo-colored heatmap	AUC = 0.94
Autoimmune pancreatitis	Zhu *et al*[34], 2015	AIP (81); Chronic pancre-atitis (100)	Differentiate AIP from chronic pancreatitis	Support Vector Machine	Accuracy: 89.3%, sensitivity: 84.1%, and specificity: 92.5%
	Marya *et al*[36], 2021	AIP (146); PDAC (292); Chronic pancreatitis (72); Normal pancreas (73)	Differentiate of AIP from PDAC	Convolutional neural network and pseudo-colored heatmap	AUC for AIP from all other = 0.92
Procedural assistance	Iwasa *et al*[38], 2021	Pancreatic mass (100)	Segmentation of pancreatic masses	Convolutional neural network	Intersection over unit = 0.77
	Zhang *et al*[40], 2020	EUS videos (339)	Recognition of stations, and segmentation of anatomical landmarks	Convolutional neural network	Accuracy for classification of stations (average) = 0.824, Dice coefficient for segmentation of pancreas (average) = 0.715

AUC: Area under the rule operator characteristic; AIP: Autoimmune pancreatitis; CP: Chronic pancreatitis; EUS: Endoscopic ultrasonography; IPMN: Intraductal papillary mucinous neoplasms; PDAC: Pancreatic ductal adenocarcinoma; PNET: Primitive neuroectodermal tumors.

endothelial growth factor-A (VEGF-A). Then the CompCyst test was used to classify cysts into one of the three following groups using a combination of molecular and imaging features. The first group was defined as cysts without any malignant potential which would not need surveillance. VHL and GNAS were used in this step and achieved 100% specificity and 46% sensitivity. The second group was cysts with small risk of malignant progression which would require surveillance. Multiple gene mutations and solid component in imaging was used in this step yielding 91% sensitivity and 54% specificity in the test cohort. The third group included cysts with high likelihood of malignant progression or malignancy which should be resected. VEGF-A protein expression was used in this step with 99% sensitivity and 30% specificity. The system was compared to standard of care and demonstrated significantly higher accuracy (69% *vs* 56%, respectively)[20]. This study used a separate validation set and a comprehensive model that incorporated clinical and radiologic findings, however, the wide-ranging molecular analysis is not readily available for routine clinical use.

A recent 2021 study focused on differentiating mucinous cystic neoplasms from serous cystadenomas using a total of 109 cases from two centers[21]. Final diagnoses were determined by endosonographers with over 5 years of experience. Additional cyst fluid or histopathologic examinations were available for only 44% of patients. A total of 221 still images were obtained followed by data augmentation, but the final number of input images was not provided in the study. The ResNet framework was used as the CNN model. Three hold-out validations were performed with 10 cases for testing, and the remaining cases used for training. The result of the study showed 82.75% accuracy and 0.88 AUC to correctly classify mucinous cystic neoplasms and serous cystadenomas from the still EUS images. A pseudo-colored decision map [gradient weighted class activation mapping (GradCAM)] was used to visualize the decision-making process. Presentation of the pseudo-colored decision map is an important asset because it highlights and color codes (red for higher impact and blue for lower impact) the areas in the image which affected the algorithm's final decision; therefore, this allows clinicians to better comprehend the decision-making process by the model. However, this study has several limitations. First, the most commonly encountered cyst, IPMN, was not included in the dataset that decreases the generalizability of the model. Second, ground truth was endosonographers' expert opinion and only 44% of patients had cyst fluid or histologic confirmation of diagnosis. Despite various limitations, the studies presented demonstrate the feasibility of image recognition ML models to perform classification tasks for pancreatic cysts and guide clinical management.

Pancreatic cancer

PDAC is currently the fourth leading cause of cancer-related mortality in Western countries and is predicted to become the second by 2030[24]. Most cases are diagnosed at later stages with 5-year survival rates less than 10%. A promising strategy is earlier diagnosis to combat this disease[25]. For this, EUS with FNA has superseded the cross-sectional imaging modalities such as CT and MRI, especially in the earlier diagnosis of PDAC[26]. However, EUS is operator dependent, and EUS diagnosis of PDAC is more challenging in patients with baseline abnormal pancreatic imaging (*e.g.*, chronic pancreatitis) who also carry a higher risk. Within this context, ML has been used to improve the diagnostic performance of EUS for pancreatic masses. Four studies used histologically confirmed PDAC cases with normal pancreas as control. Additional control groups were used in different studies to reflect clinical scenarios including chronic pancreatitis and neuroendocrine tumors. EUS images served as inputs for the algorithms. Additional EUS diagnostic technology, such as elastography, digital characteristics, contrast-enhancement, and Doppler imaging were also used. Regarding ML methods, Support-Vector-Machines were used in earlier studies to select the best combination of digital imaging features. In later studies the preferred methods were neural networks with different complexity levels depending on the year of the study. Although the models and populations varied, all studies achieved over 80% specificity and 0.94 AUC, demonstrating the feasibility of ML in this area.

In an early 2008 study by Saftoiu *et al*[27], ML for EUS elastography images was evaluated to discriminate pancreatic tumors from 'pseudotumoral' chronic pancreatitis. The prospective study enrolled 68 patients including PDAC, pancreatic neuroendocrine tumor, chronic pancreatitis, and normal pancreas. Final diagnoses were confirmed with additional pathology, imaging findings, and 6-mo follow-up of patients. From each patient, EUS elastography images were converted to vector data. As the sample size was small, 10-fold cross-validation was performed. The vector data was then analyzed with simple three and four layered ANNs. This ML algorithm yielded an AUC of 0.93 to classify malignant tumors from normal and pseudotumoral pancreatitis. This study was followed by a larger prospective blinded study in 2012 with 258 patients enrolled from 13 European centers. The population consisted of 211 PDAC confirmed by pathology diagnosis and 47 chronic pancreatitis patients diagnosed by clinical, imaging and EUS criteria (at least four of the following: hyperechoic foci, hyperechoic strands, lobularity, calcifications, hyperechoic duct wall, dilated main pancreatic duct, irregular main pancreatic duct, dilated side branches, and cysts). EUS elastography images of the regions of interests were converted to vector data and then analyzed with similar ANNs. One hundred training iterations were performed with the model to increase the statistical power of the results. The mean performance of one hundred models to correctly classify PDAC from chronic pancreatitis showed 0.94 (0.91-0.97) AUC with 85.6% sensitivity and 82.9% specificity compared with 0.85 AUC for hue histogram analysis[28]. These two studies present an excellent example for the roadmap of ML research with an initial proof-of-concept study followed by a larger prospective study. Of note, less complex neural networks were used with fewer layers. Multi-layered Perceptron only accepts numeric data as the input unlike newer CNN algorithms that can directly process the image itself. Therefore, the performance of ML in these studies can be improved.

An early study in 2013 used analysis of digital image characteristics as input to the ML model[29]. The study population consisted of 262 PDAC patients diagnosed by cytology with 126 chronic pancreatitis controls diagnosed by standard EUS criteria and over 2-year follow up. Regions of interests were manually selected by blinded endosonographers. Then 105 digital imaging characteristics of these images were extracted with dedicated software. The final combination of 16 characteristics yielded a strong discriminative performance with 94.2% accuracy, 96% sensitivity and 93% specificity.

Another older study evaluated the use of ML to classify PDAC from normal pancreas[30]. This retrospective study in 2016 included 202 PDAC patients and 130 patients with normal pancreas as controls. The regions of interests from EUS images were annotated by endosonographers. Then digital characteristics of the images (wavelet decomposition energy, boundary fractal, gray level cooccurrence matrix, standard statistical) were extracted. Among 112 digital characteristics, 20 were identified as more effective for classification, and therefore served as the input for the ML algorithm. A three-layered Multi-layered Perceptron model was used as the neural network, which is a less-complex approach accepting numerical data such as the digital characteristics of EUS images and does not require extra image processing. As such, because the images themselves are not being used, important information may not be included in the model. The final model yielded 83% sensitivity, 93% specificity and 87% accuracy for differentiating PDAC from normal pancreas. This model also only compared PDAC images to normal pancreatic tissue and not to other commonly encountered differential diagnoses such as chronic pancreatitis, which limits its adoptability to clinical use.

A recent study in 2021 evaluated the performance of ML to classify focal solid lesions. The study population consisted of 30 patients with PDAC, 20 patients with pseudo tumors in chronic pancreatitis, and 15 patients with pancreatic neuroendocrine tumors[31]. The final diagnoses were confirmed with histologic evaluations of fine-needle specimens and clinical follow-ups. From each EUS examination, 5 sets of images were extracted including grayscale images, color Doppler, contrast-enhanced imaging, and elastography. A total of 1300 collected images was increased to 3360 with data augmentation. Regarding the ML method, a CNN algorithm was combined with a Long Short-Term Memory model. Long Short-Term Memory model is a supervised ML model that has additional feedback learning functions and allows the use of sequential pre- and post-contrast appearance from the same EUS images. Cross-validation was performed for each dataset with 80% of images used as training and 20% as test sets. The final combined model's overall specificity was 96.4%, and sensitivity was 98.6% for classifying the pancreatic masses. For PDAC cases, the algorithm yielded 96.7% specificity, 98.1% sensitivity, 97.6% accuracy, and 0.97 AUC. When compared to previous studies, Udristoiu *et al*[31] used a more complex, combined ML approach with CNN and Long Short-Term Memory allowing inclusion of temporal data with contrast-enhanced imaging.

Tonozuka *et al*[32] also evaluated their own ML algorithm for its performance in classifying pancreatic masses. The 139 total patients included 76 with PDAC, 34 with chronic pancreatitis and 29 normal controls. PDAC was diagnosed using histology from EUS-fine needle biopsy or surgery, and chronic pancreatitis was diagnosed using the Rosemont criteria. All patients were followed for over 6 mo. Ten still images of lesions were chosen from each EUS examination, and the input dataset was increased to over 80000 after data augmentation. From 1390 still images, 920 were used for training and cross-validation, while the remaining 470 images were used for testing. A CNN algorithm with seven layers was used. In addition to the CNN model, a pseudo-colored feature mapping was used to highlight the areas in the image with greater impact on the final model, which makes the decision-making process more comprehensible to the endosonographer. In the test dataset, the model yielded 84.1% specificity, 92.4% sensitivity and 0.94 AUC.

Autoimmune pancreatitis

AIP is an increasingly recognized entity that may be challenging to diagnose. Accurate diagnosis is particularly important as the differential often includes PDAC with its different prognostic and management implications. Many diagnostic algorithms have been developed that include clinical, serologic, imaging, and histopathologic criteria, but their performance remains limited. While EUS with biopsy is the most effective diagnostic tool, its diagnostic yield also is suboptimal[33]. Image processing may enhance our ability to diagnose AIP by extracting data and learning from the cues in sonographic images. Two studies have studied the utility of ML in differentiating AIP from other diagnoses, including chronic pancreatitis and PDAC. The studies by Zhu *et al*[34] and Marya *et al*[35] used different ML approaches, but both achieved over 80% sensitivity and specificity for diagnosing AIP only from EUS images[34,35].

The earlier 2015 retrospective study by Zhu *et al*[34] studied a ML algorithm to differentiate AIP from chronic pancreatitis using an EUS image dataset of 81 AIP and 100 chronic pancreatitis cases. AIP diagnoses were based on HISORt criteria. Chronic pancreatitis was diagnosed by standard EUS criteria. Experienced endosonographers selected regions of interest in EUS images, and 115 digital parameters were extracted from each image. Then, a supervised Support Vector Machine algorithm was used to select the best combination of these digital parameters for discriminating AIP from chronic pancreatitis. The final combination of digital parameters yielded 90.6% accuracy, 84.1% sensitivity and 94.0% specificity.

A recent study examined the additive performance of ML with EUS to distinguish AIP from PDAC as well as chronic pancreatitis and normal pancreas. The study included 583 patients (146 AIP, 292 PDAC, 72 chronic pancreatitis, and 74 normal) with all available videos and still images of the pancreatic and peripancreatic regions included in the analysis regardless of whether they included regions of interest [36]. A total of 1174461 still images were extracted from the images and videos. Since all portions of EUS videos were included, there was a risk of oversimplification of diagnosis from certain aspects of the examination, such as presence of metastasis, which were removed from the dataset. The classification

was performed with two datasets: the first one included still images obtained from both EUS videos and captured images, while the second dataset only included EUS videos. The CNN algorithm was trained for both datasets. Pseudo-colored feature mapping was also used to visualize decision making. For comparison, seven independent EUS experts evaluated each case using videos. In the final analysis, ML showed 87% specificity, 90% sensitivity and 0.9 AUC for distinguishing AIP from PDAC in the image-only dataset. In the video-only dataset, the metrics were 90%, 93% and 0.96 for specificity, sensitivity, and AUC, respectively. The ML model was superior to expert endosonographers, who had 82.4% specificity and 53.8% sensitivity in differentiating AIP from PDAC. ML also had high sensitivity (99%) and specificity (98%) for distinguishing AIP from normal pancreas. It had inferior performance in separating AIP from chronic pancreatitis (94% sensitivity, 71% specificity, 0.89 AUC). The heatmap analysis yielded interesting results, which may help guide endosonographers, showing that visualizing a hyperechoic plane between the parenchyma and duct or vessel was highly predictive of AIP while post acoustic enhancement deep to a dilated pancreatic duct or vessel was consistent with PDAC. Regarding AI technology, these two studies differ with respect to their approach of utilizing ML with EUS data. Zhu *et al*[34] used an older ML algorithm, support vector machine, which is a supervised algorithm that classifies two numeric data points. As such, EUS images are converted into numerical data by extracting digital parametric features, and then the ML model is trained with these features. On the other hand, Marya *et al*[35] used a CNN algorithm, ResNet, with 50 layers that can work directly on the EUS images itself.

Procedural assistance and training

EUS is the leading modality for assessing and obtaining tissue from the pancreas with approximately 90% specificity and sensitivity for solid masses[36]. However, interobserver reliability remains an issue in EUS as accuracy relies on the endosonographers' skills and experience and carries the risk of false-negative results. Pancreatic EUS also has a steep learning curve. ML approaches have been developed to potentially augment the diagnostic performance of EUS and biopsy as well as aid in training.

Iwasa *et al*[38] tested ML to augment contrast enhanced EUS by dividing the sonographic image into regions with similar appearance and then differentiating regions of interest, also called automatic segmentation. For this study, videos from 100 contrast enhanced EUS examinations of solid pancreatic masses with histologic diagnosis were used. Each video was transformed into 900 still images as input for a U-Net CNN algorithm. The borders of the lesions were manually annotated by two endosono-graphers and served as the ground truth. IoU was used as the performance output of the algorithm with median IoU for all cases being 0.77, which is greater than the acceptable 0.5 threshold value[37]. The EUS videos were also classified into different categories to understand the effect of respiratory movements and visibility of boundaries of the lesions by the endosonographers. IoU significantly improved to 0.91 in cases with the most visible boundaries and decreased to 0.13 for cases with the least visible boundaries[39]. On the other hand, respiratory movements did not change the performance of the algorithm. This proof-of-concept study suggests that ML can provide real-time assistance in the detection of pancreatic lesions. The classification of exams with respect to the ease of detecting the border of lesions is an important aspect of this study because it demonstrated that ML can also be affected by the quality of the EUS examination and the sonographic characteristics of the lesion, reflected in this case by how well the border was visible.

A case report suggested that a ML model may help target areas to biopsy within pancreatic masses that have the highest diagnostic yield by avoiding areas of necrosis. A CNN algorithm was used to label and highlight the more cellular region in a 6.5 cm solid pancreatic mass, which was predicted to have the highest probability of yielding a diagnosis by discriminating it from neighboring necrotic or inflam-matory regions. EUS-fine needle aspiration was performed and yielded a positive diagnosis for PDAC. The technical details, training dataset and methods, validation and model characteristics were not presented in the report[39]. This is a novel idea that may provide valuable intra-procedural assistance, however, needs further evaluation.

ML may aid EUS training by guiding the steps of routine diagnostic EUS evaluation of the pancreas. A novel AI system aimed to assist recognition of fundamental stations and identification of pancreatic and vascular anatomical landmarks. This was performed in four steps: Identifying images, filtering suitable images, recognizing pancreas stations, and segmenting anatomical landmarks and monitoring for loss of visualization of the pancreas. Two expert endosonographers decided on the criteria for suitable images and annotated video clips that served as ground truth. A ResNet model was used as the CNN algorithm. A separate set of prospective EUS examinations were used as a test set. Three different endosonographers classified each image for comparison with the AI model. The final model was tested using an external test set and demonstrated an accuracy of 82.4% to identify six anatomical stations (abdominal aorta, pancreatic body, pancreatic tail, confluence, pancreatic head from stomach, or pancreatic head from descending duodenum), and a Dice of 0.715 to label pancreas and vessels. Comparison of the AI model with the three expert endosonographers yielded strong interobserver agreement with kappa values of 0.846, 0.853 and 0.826[40]. The results of this study demonstrated that a ML model may aid in recognizing stations and anatomic landmarks in sonographic images. This has the potential to assist procedural navigation during EUS examination and improve cognitive aspects of EUS skills. However, the impact of such real-time procedural assistance on the endosonographer's

performance was not assessed in this study and warrants further evaluation.

CONCLUSION

In this review, we summarize the current literature regarding the use of ML in EUS for diagnosing pancreatic diseases. Our review defined two main areas for AI in the field: visual recognition-classification and procedural assistance and training. AI has been more utilized in transabdominal ultrasonography for detecting liver fibrosis and in CT scans for lesion classification, which have been extensively reviewed elsewhere[41-45]. ML appears to have great potential in assisting EUS examination of the pancreas as sonographic imaging contains vital visual information that the human eye cannot distinguish. The diagnostic accuracy of EUS imaging is highly operator dependent and requires both technical and cognitive skills. Acquisition of these skills currently requires dedicated training with proctorship and procedural experience, which remains limited, apart from dedicated advanced endoscopy fellowship programs. These issues in training limit the widespread adoption of EUS, which is the leading tool for diagnosing pancreatic disorders, including PDAC. AI may assist in the development of cognitive skills and augmentation of procedural efficiency in relatively less experienced endosonographers.

Further opportunities should be explored with AI and pancreatic EUS. However, several limitations exist in the field. First, the number of EUS procedures and the prevalence of pancreatic diseases are lower, which makes it more difficult to train data-hungry machine learning algorithms. Second, annotation of EUS data is more challenging compared to other imaging modalities as the number of experts endosonographers is relatively limited. Third, EUS examinations with histopathologic or cytologic diagnosis is harder to obtain for certain pancreatic diseases and have issues with sensitivity, which further limits the number of studies for AI training. However, these limitations may be overcome with multi-center collaborations and prospective data collection, which will hopefully lead to improved image recognition, procedural assistance, and training for pancreatic EUS.

FOOTNOTES

Author contributions: Simsek C collected data and wrote the paper; Lee L carried out data collection; both authors read, edited, and approved the final manuscript.

Conflict-of-interest statement: Cem Simsek is co-founder of Algomedicus Inc.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: https://creativecommons.org/Licenses/by-nc/4.0/

Country/Territory of origin: United States

ORCID number: Cem Simsek 0000-0002-7037-5233; Linda S Lee 0000-0001-9921-8920.

S-Editor: Liu JH
L-Editor: A
P-Editor: Liu JH

REFERENCES

1   **Miller KD**, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, Jemal A, Kramer JL, Siegel RL. Cancer treatment and survivorship statistics, 2019. *CA Cancer J Clin* 2019; **69**: 363-385 [PMID: 31184787 DOI: 10.3322/caac.21565]

2   **Bray F**, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**: 394-424 [PMID: 30207593 DOI: 10.3322/caac.21492]

3   **Rahib L**, Wehner MR, Matrisian LM, Nead KT. Estimated Projection of US Cancer Incidence and Death to 2040. *JAMA Netw Open* 2021; **4**: e214708 [PMID: 33825840 DOI: 10.1001/jamanetworkopen.2021.4708]

4   **Xiao AY**, Tan ML, Wu LM, Asrani VM, Windsor JA, Yadav D, Petrov MS. Global incidence and mortality of pancreatic diseases: a systematic review, meta-analysis, and meta-regression of population-based cohort studies. *Lancet Gastroenterol Hepatol* 2016; **1**: 45-55 [PMID: 28404111 DOI: 10.1016/S2468-1253(16)30004-8]

5   **Kichler A**, Jang S. Chronic Pancreatitis: Epidemiology, Diagnosis, and Management Updates. *Drugs* 2020; **80**: 1155-1168

[PMID: 32647920 DOI: 10.1007/s40265-020-01360-6]

6   **Megibow AJ**, Baker ME, Gore RM, Taylor A. The incidental pancreatic cyst. *Radiol Clin North Am* 2011; **49**: 349-359 [PMID: 21333781 DOI: 10.1016/j.rcl.2010.10.008]

7   **Zhang B**, Zhu F, Li P, Yu S, Zhao Y, Li M. Endoscopic ultrasound elastography in the diagnosis of pancreatic masses: A meta-analysis. *Pancreatology* 2018; **18**: 833-840 [PMID: 30093353 DOI: 10.1016/j.pan.2018.07.008]

8   **Russell S**, Norvig P.   Artificial intelligence: a modern approach. 2002

9   **Yu KH**, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng* 2018; **2**: 719-731 [PMID: 31015651 DOI: 10.1038/s41551-018-0305-z]

10   **Penny W**, Frost D. Neural networks in clinical medicine. *Med Decis Making* 1996; **16**: 386-398 [PMID: 8912300 DOI: 10.1177/0272989X9601600409]

11   **Yamashita R**, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 2018; **9**: 611-629 [PMID: 29934920 DOI: 10.1007/s13244-018-0639-9]

12   **Zhang**, HM. and B. Dong, A Review on Deep Learning in Medical Image Reconstruction. *Journal of the Operations Research Society of China* 2020; **8**: 311-340 [DOI: 10.1007/s40305-019-00287-4]

13   **Van Houdt G**, Mosquera C, Nápoles G. A review on the long short-term memory model. *Artif Intell Rev* 2020; **53**: 5929-5955 [DOI: 10.1007/s10462-020-09838-1]

14   **Rintala TJ**, Federico A, Latonen L, Greco D, Fortino V. A systematic comparison of data- and knowledge-driven approaches to disease subtype discovery. *Brief Bioinform* 2021; **22** [PMID: 34396389 DOI: 10.1093/bib/bbab314]

15   **Sinkala M**, Mulder N, Martin D. Machine Learning and Network Analyses Reveal Disease Subtypes of Pancreatic Cancer and their Molecular Characteristics. *Sci Rep* 2020; **10**: 1212 [PMID: 31988390 DOI: 10.1038/s41598-020-58290-2]

16   **Zerboni G**, Signoretti M, Crippa S, Falconi M, Arcidiacono PG, Capurso G. Systematic review and meta-analysis: Prevalence of incidentally detected pancreatic cystic lesions in asymptomatic individuals. *Pancreatology* 2019; **19**: 2-9 [PMID: 30503370 DOI: 10.1016/j.pan.2018.11.014]

17   **Zaheer A**, Pokharel SS, Wolfgang C, Fishman EK, Horton KM. Incidentally detected cystic lesions of the pancreas on CT: review of literature and management suggestions. *Abdom Imaging* 2013; **38**: 331-341 [PMID: 22534872 DOI: 10.1007/s00261-012-9898-y]

18   **Pinho DF**, Rofsky NM, Pedrosa I. Incidental pancreatic cysts: role of magnetic resonance imaging. *Top Magn Reson Imaging* 2014; **23**: 117-128 [PMID: 24690615 DOI: 10.1097/RMR.0000000000000018]

19   **Wu J**, Wang Y, Li Z, Miao H. Accuracy of Fukuoka and American Gastroenterological Association Guidelines for Predicting Advanced Neoplasia in Pancreatic Cyst Neoplasm: A Meta-Analysis. *Ann Surg Oncol* 2019; **26**: 4522-4536 [PMID: 31617119 DOI: 10.1245/s10434-019-07921-8]

20   **Springer S**, Masica DL, Dal Molin M, Douville C, Thoburn CJ, Afsari B, Li L, Cohen JD, Thompson E, Allen PJ, Klimstra DS, Schattner MA, Schmidt CM, Yip-Schneider M, Simpson RE, Fernandez-Del Castillo C, Mino-Kenudson M, Brugge W, Brand RE, Singhi AD, Scarpa A, Lawlor R, Salvia R, Zamboni G, Hong SM, Hwang DW, Jang JY, Kwon W, Swan N, Geoghegan J, Falconi M, Crippa S, Doglioni C, Paulino J, Schulick RD, Edil BH, Park W, Yachida S, Hijioka S, van Hooft J, He J, Weiss MJ, Burkhart R, Makary M, Canto MI, Goggins MG, Ptak J, Dobbyn L, Schaefer J, Sillman N, Popoli M, Klein AP, Tomasetti C, Karchin R, Papadopoulos N, Kinzler KW, Vogelstein B, Wolfgang CL, Hruban RH, Lennon AM. A multimodality test to guide the management of patients with a pancreatic cyst. *Sci Transl Med* 2019; **11** [PMID: 31316009 DOI: 10.1126/scitranslmed.aav4772]

21   **Nguon LS**, Seo K, Lim JH, Song TJ, Cho SH, Park JS, Park S. Deep Learning-Based Differentiation between Mucinous Cystic Neoplasm and Serous Cystic Neoplasm in the Pancreas Using Endoscopic Ultrasonography. *Diagnostics (Basel)* 2021; **11** [PMID: 34201066 DOI: 10.3390/diagnostics11061052]

22   **Kurita Y**, Kuwahara T, Hara K, Mizuno N, Okuno N, Matsumoto S, Obata M, Koda H, Tajika M, Shimizu Y, Nakajima A, Kubota K, Niwa Y. Diagnostic ability of artificial intelligence using deep learning analysis of cyst fluid in differentiating malignant from benign pancreatic cystic lesions. *Sci Rep* 2019; **9**: 6893 [PMID: 31053726 DOI: 10.1038/s41598-019-43314-3]

23   **Kuwahara T**, Hara K, Mizuno N, Okuno N, Matsumoto S, Obata M, Kurita Y, Koda H, Toriyama K, Onishi S, Ishihara M, Tanaka T, Tajika M, Niwa Y. Usefulness of Deep Learning Analysis for the Diagnosis of Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas. *Clin Transl Gastroenterol* 2019; **10**: 1-8 [PMID: 31117111 DOI: 10.14309/ctg.0000000000000045]

24   **Rahib L**, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* 2014; **74**: 2913-2921 [PMID: 24840647 DOI: 10.1158/0008-5472.CAN-14-0155]

25   **Howlader N**, Noone AM, Krapcho M, Garshell J, Miller D, Altekruse SF, Kosary Cl, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA.   SEER cancer statistics review, 1975–2012. National Cancer Institute, 2015. Available from: https://seer.cancer.gov/archive/csr/1975_2012/

26   **Kitano M**, Yoshida T, Itonaga M, Tamura T, Hatamaru K, Yamashita Y. Impact of endoscopic ultrasonography on diagnosis of pancreatic cancer. *J Gastroenterol* 2019; **54**: 19-32 [PMID: 30406288 DOI: 10.1007/s00535-018-1519-2]

27   **Săftoiu A**, Vilmann P, Gorunescu F, Gheonea DI, Gorunescu M, Ciurea T, Popescu GL, Iordache A, Hassan H, Iordache S. Neural network analysis of dynamic sequences of EUS elastography used for the differential diagnosis of chronic pancreatitis and pancreatic cancer. *Gastrointest Endosc* 2008; **68**: 1086-1094 [PMID: 18656186 DOI: 10.1016/j.gie.2008.04.031]

28   **Săftoiu A**, Vilmann P, Gorunescu F, Janssen J, Hocke M, Larsen M, Iglesias-Garcia J, Arcidiacono P, Will U, Giovannini M, Dietrich CF, Havre R, Gheorghe C, McKay C, Gheonea DI, Ciurea T; European EUS Elastography Multicentric Study Group. Efficacy of an artificial neural network-based approach to endoscopic ultrasound elastography in diagnosis of focal pancreatic masses. *Clin Gastroenterol Hepatol* 2012; **10**: 84-90.e1 [PMID: 21963957 DOI: 10.1016/j.cgh.2011.09.014]

29   **Zhu M**, Xu C, Yu J, Wu Y, Li C, Zhang M, Jin Z, Li Z. Differentiation of pancreatic cancer and chronic pancreatitis using computer-aided diagnosis of endoscopic ultrasound (EUS) images: a diagnostic test. *PLoS One* 2013; **8**: e63820 [PMID: 23704940 DOI: 10.1371/journal.pone.0063820]

30   **Ozkan M**, Cakiroglu M, Kocaman O, Kurt M, Yilmaz B, Can G, Korkmaz U, Dandil E, Eksi Z. Age-based computer-aided diagnosis approach for pancreatic cancer on endoscopic ultrasound images. *Endosc Ultrasound* 2016; **5**: 101-107 [PMID: 27080608 DOI: 10.4103/2303-9027.180473]

31   **Udriştoiu AL**, Cazacu IM, Gruionu LG, Gruionu G, Iacob AV, Burtea DE, Ungureanu BS, Costache MI, Constantin A, Popescu CF, Udriştoiu Ș, Săftoiu A. Real-time computer-aided diagnosis of focal pancreatic masses from endoscopic ultrasound imaging based on a hybrid convolutional and long short-term memory neural network model. *PLoS One* 2021; **16**: e0251701 [PMID: 34181680 DOI: 10.1371/journal.pone.0251701]

32   **Tonozuka R**, Itoi T, Nagata N, Kojima H, Sofuni A, Tsuchiya T, Ishii K, Tanaka R, Nagakawa Y, Mukai S. Deep learning analysis for the detection of pancreatic cancer on endosonographic images: a pilot study. *J Hepatobiliary Pancreat Sci* 2021; **28**: 95-104 [PMID: 32910528 DOI: 10.1002/jhbp.825]

33   **Facciorusso A**, Barresi L, Cannizzaro R, Antonini F, Triantafyllou K, Tziatzios G, Muscatiello N, Hart PA, Wani S. Diagnostic yield of endoscopic ultrasound-guided tissue acquisition in autoimmune pancreatitis: a systematic review and meta-analysis. *Endosc Int Open* 2021; **9**: E66-E75 [PMID: 33403238 DOI: 10.1055/a-1293-7279]

34   **Zhu J**, Wang L, Chu Y, Hou X, Xing L, Kong F, Zhou Y, Wang Y, Jin Z, Li Z. A new descriptor for computer-aided diagnosis of EUS imaging to distinguish autoimmune pancreatitis from chronic pancreatitis. *Gastrointest Endosc* 2015; **82**: 831-836.e1 [PMID: 25952089 DOI: 10.1016/j.gie.2015.02.043]

35   **Marya NB**, Powers PD, Chari ST, Gleeson FC, Leggett CL, Abu Dayyeh BK, Chandrasekhara V, Iyer PG, Majumder S, Pearson RK, Petersen BT, Rajan E, Sawas T, Storm AC, Vege SS, Chen S, Long Z, Hough DM, Mara K, Levy MJ. Utilisation of artificial intelligence for the development of an EUS-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut* 2021; **70**: 1335-1344 [PMID: 33028668 DOI: 10.1136/gutjnl-2020-322821]

36   **Facciorusso A**, Wani S, Triantafyllou K, Tziatzios G, Cannizzaro R, Muscatiello N, Singh S. Comparative accuracy of needle sizes and designs for EUS tissue sampling of solid pancreatic masses: a network meta-analysis. *Gastrointest Endosc* 2019; **90**: 893-903.e7 [PMID: 31310744 DOI: 10.1016/j.gie.2019.07.009]

37   **Rezatofighi H**, Tsoi N, Gwak JY, Sadeghian A, Reid I, Savarese S.   Generalized intersection over union: A metric and a loss for bounding box regression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019

38   **Iwasa Y**, Iwashita T, Takeuchi Y, Ichikawa H, Mita N, Uemura S, Shimizu M, Kuo YT, Wang HP, Hara T. Automatic Segmentation of Pancreatic Tumors Using Deep Learning on a Video Image of Contrast-Enhanced Endoscopic Ultrasound. *J Clin Med* 2021; **10** [PMID: 34441883 DOI: 10.3390/jcm10163589]

39   **Yu X**, Zhang Z, Zhu N, Tang A, Hu S, Wang X, Tian L. Application of a novel artificial intelligence system in guiding the targeted puncture of a pancreatic mass. *Endoscopy* 2021 [PMID: 34624923 DOI: 10.1055/a-1625-3396]

40   **Zhang J**, Zhu L, Yao L, Ding X, Chen D, Wu H, Lu Z, Zhou W, Zhang L, An P, Xu B, Tan W, Hu S, Cheng F, Yu H. Deep learning-based pancreas segmentation and station recognition system in EUS: development and validation of a useful training tool (with video). *Gastrointest Endosc* 2020; **92**: 874-885.e3 [PMID: 32387499 DOI: 10.1016/j.gie.2020.04.071]

41   **Chen Y**, Luo Y, Huang W, Hu D, Zheng RQ, Cong SZ, Meng FK, Yang H, Lin HJ, Sun Y, Wang XY, Wu T, Ren J, Pei SF, Zheng Y, He Y, Hu Y, Yang N, Yan H. Machine-learning-based classification of real-time tissue elastography for hepatic fibrosis in patients with chronic hepatitis B. *Comput Biol Med* 2017; **89**: 18-23 [PMID: 28779596 DOI: 10.1016/j.compbiomed.2017.07.012]

42   **Yang CM**, Shu J. Cholangiocarcinoma Evaluation *via* Imaging and Artificial Intelligence. *Oncology* 2021; **99**: 72-83 [PMID: 33147583 DOI: 10.1159/000507449]

43   **Gatos I**, Tsantis S, Spiliopoulos S, Karnabatidis D, Theotokas I, Zoumpoulis P, Loupas T, Hazle JD, Kagadis GC. A Machine-Learning Algorithm Toward Color Analysis for Chronic Liver Disease Classification, Employing Ultrasound Shear Wave Elastography. *Ultrasound Med Biol* 2017; **43**: 1797-1810 [PMID: 28634041 DOI: 10.1016/j.ultrasmedbio.2017.05.002]

44   **Lee JH**, Joo I, Kang TW, Paik YH, Sinn DH, Ha SY, Kim K, Choi C, Lee G, Yi J, Bang WC. Deep learning with ultrasonography: automated classification of liver fibrosis using a deep convolutional neural network. *Eur Radiol* 2020; **30**: 1264-1273 [PMID: 31478087 DOI: 10.1007/s00330-019-06407-1]

45   **Kuwahara T**, Hara K, Mizuno N, Haba S, Okuno N, Koda H, Miyano A, Fumihara D. Current status of artificial intelligence analysis for endoscopic ultrasonography. *Dig Endosc* 2021; **33**: 298-305 [PMID: 33098123 DOI: 10.1111/den.13880]

Artificial intelligence in critically ill diabetic patients: current status and future prospects

Deven Juneja, Anish Gupta, Omender Singh

Deven Juneja, Anish Gupta, Omender Singh, Institute of Critical Care Medicine, Max Super Speciality Hospital, Saket, New Delhi 110092, India

Corresponding author: Deven Juneja, FCCP, MBBS, Director, Institute of Critical Care Medicine, Max Super Speciality Hospital, Saket, 1, Press Enclave Road, Saket, New Delhi 110092, India. devenjuneja@gmail.com

Abstract

Recent years have witnessed increasing numbers of artificial intelligence (AI) based applications and devices being tested and approved for medical care. Diabetes is arguably the most common chronic disorder worldwide and AI is now being used for making an early diagnosis, to predict and diagnose early complications, increase adherence to therapy, and even motivate patients to manage diabetes and maintain glycemic control. However, these AI applications have largely been tested in non-critically ill patients and aid in managing chronic problems. Intensive care units (ICUs) have a dynamic environment generating huge data, which AI can extract and organize simultaneously, thus analysing many variables for diagnostic and/or therapeutic purposes in order to predict outcomes of interest. Even non-diabetic ICU patients are at risk of developing hypo or hyperglycemia, complicating their ICU course and affecting outcomes. In addition, to maintain glycemic control frequent blood sampling and insulin dose adjustments are required, increasing nursing workload and chances of error. AI has the potential to improve glycemic control while reducing the nursing workload and errors. Continuous glucose monitoring (CGM) devices, which are Food and Drug Administration (FDA) approved for use in non-critically ill patients, are now being recommended for use in specific ICU populations with increased accuracy. AI based devices including artificial pancreas and CGM regulated insulin infusion system have shown promise as comprehensive glycemic control solutions in critically ill patients. Even though many of these AI applications have shown potential, these devices need to be tested in larger number of ICU patients, have wider availability, show favorable cost-benefit ratio and be amenable for easy integration into the existing healthcare systems, before they become acceptable to ICU physicians for routine use.

Key Words: Artificial intelligence; Blood glucose; Critical care; Diabetes mellitus; Intensive care unit; Machine learning

Core Tip: Increasing number of applications and devices based on artificial intelligence are being tested and approved for medical care. These devices have the potential to change the way we presently manage chronic diseases like diabetes. Moreover, their application in data rich and dynamic intensive care unit environment may have great implications in detecting hypo or hyperglycemia and reducing glycemic variability, while improving safety and accuracy and reducing nursing workload. Devices like artificial pancreas and continuous glucose monitoring regulated insulin infusion systems have shown promise as comprehensive glucose control solutions and may change the future of care for critically ill diabetic patients.

INTRODUCTION

As per the International Diabetes Federation 2021 estimates, about 537 million people are living with diabetes signifying a 10% prevalence rate worldwide with an estimated 6.7 million deaths in 2021. This number will rise exponentially in the coming years which will place a heavy burden on the already stressed healthcare system[1]. These patients are at increased risk of developing complications like sepsis, diabetes keto-acidosis and other complications necessitating intensive care unit (ICU) admission. In addition, critically ill diabetic patients are at an increased risk of developing nosocomial infections, having a longer ICU stay and increased ICU mortality[2-4].

All components of diabetes care including prevention and management of hyperglycemia and hypoglycemia, are essential to improve outcomes. In critically ill patients, these complications may be multifactorial and may also occur in non-diabetic patients, complicating their disease course. In addition to hyper- and hypoglycemia, glycemic variability (GV) and time in target range (TITR) are recently recognized components of dysglycemia which may affect patient outcomes[5-7]. However, the exact target for blood glucose (BG) control in ICU is not well established. Moreover, targeting tight glucose control necessitates frequent blood sampling and adjustment of insulin dose, increasing the work-load on ICU staff. In addition, targeting tight glucose control has not shown to have any mortality benefit but is associated with five-fold increased risk of hypoglycemia[8].

It has been difficult to establish a safe blood sugar level but as per American Diabetes Association (ADA) a BG level below 180 mg/dL is acceptable[9]. The surviving sepsis guidelines further recommend a target BG levels between 140-180 mg/dL in patients with sepsis[10].

Artificial intelligence (AI) is a rapidly evolving science which is gradually changing the landscape of many industries including healthcare. As ICUs have a dynamic environment which generates a huge amount of data, AI has a tremendous scope and now is increasingly being used in advanced mechanical ventilation, weaning from ventilation, predicting development of sepsis, antibiotic dosing and radiological assessment and monitoring[11-15]. In this review, we will be discussing the current applications and potential role AI may have in managing critically ill diabetic patients.

ARTIFICIAL INTELLIGENCE

There is no standard definition of AI but as per the Encyclopaedia Britannica, AI refers to "a system endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience"[16]. Basically, AI based systems should be able to perform tasks comparable to human intelligence.

AI has great potential and has been used in the field of medicine for discovery of new drug molecules, diagnostics, radiology and imaging, molecular biology, bioinformatics and therapeutics. AI has the ability to analyze and scrutinize massive amounts of data and help understand disease patterns. The human brain can store a limited amount of information at any one time and may be unable to analyze and visualize patterns embedded in vast quantities of data[17]. In contrast computers have a large storage capacity and can discern even small associations within the data. However, computer programming has limitations as they are able to follow only certain specific patterns, as per the programming instructions. AI in contrast differs from traditional computer programming as it learns from exposure to various experiences and inputs, assimilates the data and can improve on its own

intelligence and modify the output behavior.

AI consists of a wide spectrum of complex algorithms and is broadly divided into machine learning (ML), deep learning, and cognitive computing. In ML, AI systems are trained with large repository of data and algorithms to enable them to follow a format to examine relationships and learn from them. Deep learning based systems develop insights by conducting complex interventions on the available data while cognitive AI systems are the most complex and try and match the human intelligence by understanding, reasoning, interacting, and learning from the data. Such systems are able to process and interpret exponential amounts of data (both structured and unstructured) and thus help in proposing any valid connections or hypothesis[18].

The AI functioning can be broken down in a systematic way and the processes involved can be divided into 3 main functions which occur in succession, which are knowledge discovery followed by learning and finally reasoning.

Knowledge discovery/ retrieval

The discovery of knowledge is the essence of AI. It works by creating algorithms for acquiring relevant and potential information from databases and is referred to as knowledge discovery in databases (KDD). For KDD to be effective it should have an in-depth knowledge of the area of interest as it will evaluate and interpret patterns and models to decide what data constitutes knowledge and what does not. KDD, hence plays a pivotal role in identifying information which is useful and valid.

Learning

Once the KDD process is complete the next step is learning from the knowledge or information acquired. Systems are allowed to automatically learn without human intervention or assistance. It usually consists of an inductive component which could be a simple process or could consist of a convolutional neural network (CNN). The various techniques used are artificial neural networks (ANNs), support vector machines (SVMs), random forest (RF), evolutionary algorithms, deep learning, Naive Bayes (NB), decision trees, and regression algorithms.

Certain types of AI algorithms are more commonly employed in healthcare settings than others. SVMs are used to predict clearly defined outcomes and adherence to medications. ANNs are algorithms which have been inspired by neuronal organization of animal brains, and have been employed to analyze data from computed tomography images, mammograms *etc.*, to predict complications and outcomes. Logistic regression, is a ML algorithm which has been used to predict and classify probability of an event using predictor variables. Using data from electronic records or patient's medical history, RF algorithms have been used to predict risk of disease, and NB are the most advanced ML algorithms which have been used recently to predict development of disease in specific patient populations[19].

Reasoning

Reasoning is the final step in the AI process and involves the use of logical techniques to come to a conclusion from the available data. The primary objective of reasoning is to perform tasks at the level of a human intelligence and in a specialized manner with the final objective to generate inferences in the most precise manner.

AI algorithms

AI is a rapidly evolving technology with increasing number of subsets being introduced regularly, each having their own advantages and limitations. For prediction and management of diabetes, commonly used AI algorithms include linear regression (LR), classification/decision trees (DTs), RF, SVMs, ANNs, and NB.

LR is a regression model which analyses the data and predicts a continuous output, finding solution following a linear curve. DTs are predictive models which predict outcome from the given data, but can find solution using both linear and non-linear curves. DTs also fare better than LR models for categorical independent variables. RF is a variation of DT, supporting both linear and non-linear solutions, but is better at handling of missing values and outliers. It is more favorable than DTs as it is more robust, accurate and provides a more generalized solution.

SVMs are supervised learning algorithms which are recently gaining popularity for their applications in healthcare settings. Even though they are mostly used for classification problems in ML, they can also be applied for regression problems. They also support linear and non-linear solutions and are better than LR in handling outliers and analyzing data with large number of features.

ANN is an advanced technology based on the brain and the nerves and programmed to mimic the biological neural system. ANNs can also find non-linear solutions and are sub-classified as convolutional (feedforward networks) and recurrent (feedback loop) neural networks. ANNs have better accuracy but require larger training data as compared to LR.

As compared to LR, DT and RF, which are discriminative models, NB is a generative model which works well even with small data sets. This supervised learning algorithm is based on Bayes theorem and can provide solutions to classification problems. It is easy, fast and performs well in case of categorial data. However, it is a bad estimator and its probability outputs are not reliable.

ROLE OF AI IN MANAGEMENT OF DIABETES MELLITUS

Medical management forms only a small part of the entire spectrum of diabetes care, as diabetes mellitus (DM) is mainly a life-style disorder. Apart from medications, education on self-management (meal schedules, calorie counting, exercising, routine BG monitoring) and continuous medical care is paramount not only to prevent acute complications but also to minimize the risk of long-term complications like nephropathy, retinopathy, diabetic foot, cardiovascular disease, or stroke. As a result, diabetes care is complex and various medical and life-style related factors need to be taken into account to optimize management.

The use of AI in DM is not new and a number of studies have shown the role of AI applications in the care of diabetic patients[20-24]. A number of complex AI systems, and their clinical applications have been described (Table 1). Deep-learning based AI algorithms may help in early diagnosis of diabetic retinopathy using retinal photographs with a reported sensitivity and specificity of more than 90%[25]. IDx-DR is the first such AI-based device approved by US-FDA for screening of diabetic patients for retinopathy[26]. As it does not require a clinician to interpret the results, this automated system can help the non-eye specialists to recognize early signs of retinopathy and send the patients to eye-specialists only if indicated, thereby simplifying the process and achieving higher patient satisfaction[27].

Dreamed Advisor pro assimilates data regarding the glucose levels, insulin dose and carbohydrate intake and using AI-based MD-Logic algorithms it then makes recommendations for insulin dose adjustments. These recommendations have been shown to be similar to those given by experienced physicians in the real-world settings validating the use of such devices in day-to-day clinical practice[23, 28]. Several real-time Continuous Glucose Monitoring (CGM) devices like Medtronic Guardian Connect and Dexcom G6 CGM systems, are commercially available which can act as self-monitoring tools for diabetic patients (Table 1). These devices can provide real-time glucose values which can be displayed on the patient's mobile phones and can raise an alarm if the BG levels go beyond the predefined range. These devices can further be connected to insulin pumps and hence aid in insulin dose adjustments. However, these devices require repeated calibrations with the capillary blood glucose levels, to be measured by finger pricks. Use of these glucose sensors for more than 70% of the time, has shown to improve the HbA1c by 0.4 to 0.6% and reduce the incidence of hypoglycemic episodes[29]. Presently, these devices and applications have not been validated in ICU patients but can be further modified and tested to be applied in the management of critically ill patients.

AI IN DIABETES MANAGEMENT IN ICU

Hyperglycemia is a common phenomenon in the ICU irrespective of the reason for admission and may occur even in the absence of pre-existing DM. The pathophysiology of hyperglycemia in ICU is multifactorial and can occur secondary to release of stress hormones (corticosteroids and catecholamines), proinflammatory mediators, administration of exogenous drugs (corticosteroids, vasopressors, ascorbic acid), parenteral solutions containing dextrose, stress hyperglycemia and use of commercial dietary feeds or supplements[30]. Irrespective of cause, hyperglycemia is associated with an increase in ICU stay, hospitalization costs, morbidity, and mortality[4,31].

Apart from hyperglycemia, hypoglycemia and GV have also been shown to be associated with increase in mortality in critically ill patients[5,6]. Use of variable insulin protocols which are not clinically validated and inaccurate blood sugar measurements are responsible for this GV seen in the ICUs. In addition, insulin sensitivity in critically ill patients follows a very erratic course and is plagued with frequent changes which could be secondary to the underlying illness, dietary changes or medications.

TITR has been recognized as another domain of dysglycemia in critically ill patients[7]. It may be defined as the total time spent in the target range and is expressed as the percentage of time. Data suggests that critically ill patients having more than 70% TITR, have significantly higher survival rates [32]. However, the exact cut-offs for TITR remain unclear with different studies suggesting TITR ranging from 50-80% for improving outcomes[33,34].

In spite of several widely accepted applications for out-patient and long-term management of DM, AI applications in management of critically ill patients are limited. The possible applications of AI in critically ill diabetes patients are given in Table 2[35].

Blood glucose monitoring and prediction

Blood glucose management requires frequent sampling and insulin dose adjustments. Capillary BG monitoring still remains the most commonly employed method, even in critically ill patients. However, its accuracy may be affected in patients with subcutaneous oedema, shock, and hypoxemia, which commonly affect ICU patients. Hence, using arterial blood is preferred but it requires repeated arterial punctures or presence of an invasive arterial line. The characteristics of an ideal method to monitor BG is given in the Table 3.

Table 1 Clinical uses of artificial intelligence in management of diabetes

AI applications	Examples of AI devices	Clinical uses
Retinal screening	IDx-DR device	Screening and diagnosis of diabetic retinopathy
Clinical diagnosis	Advisor Pro	Detection and monitoring of diabetes and its associated complications. Fine-tuning insulin dose
Patient self-management tools	Medtronic Guardian Connect System, Dexcom G6 CGM systems; Mobile applications	Improve blood glucose control, activity and dietary tracking
Risk stratification	AI using random forest and; gradient boosting techniques	Prediction of new-onset diabetes; Prediction of subpopulations at risk for complications, non-compliance to therapy and hospitalization

AI: Artificial intelligence.

Table 2 Possible critical care applications of artificial intelligence in diabetes management

Blood glucose monitoring and prediction

Detection of adverse glycemic events

Blood glucose control strategies

Insulin bolus calculators and advisory systems

Risk and patient stratification

Table 3 Characteristics of an ideal tool to monitor blood glucose in intensive care unit

Ease to use

Minimal burden on staff

Automated data entry

High rate of adherence

Allow for minimal sampling

Comfortable to use for the patient

Use of a proven algorithm to calculate insulin dosage

Quickly correct hyperglycemia

Consistently maintain glucose within the predetermined optimal range

Ensure minimal glycemic variability

Prevent episodes of hypoglycemia

Provide easy interface with other patient measurements and data

Easy to integrate into existing hospital systems

Avoid the need for repeated data entry

Maintain results in a comprehensive, standardized database to facilitate multi-center comparison

Continuous glucose monitoring

Continuous Glucose Monitoring has been employed in the management of DM for more than a decade. Several CGM devices have been developed and are presently commercially available and approved for in-hospital use (Table 4). They can be broadly classified as transdermal (non-invasive), subcutaneous (minimally invasive) and intra-vascular (invasive) devices. Subcutaneous and transdermal devices are not considered ideal in critically ill patients because the presence of subcutaneous oedema, hypoxemia, and shock may affect their accuracy. Hence, intravascular devices may be preferable in these patients. However, the continuous subcutaneous flash glucose monitoring (FGM) system (FreeStyle Libre) has been recently tried in critically ill patients and has shown to have high test-retest reliability and acceptable accuracy[36-38].

Table 4 Continuous glucose monitoring devices

Type of device	Name of device	Comments
Intravenous	GlucoClear by Edwards Lifesciences; (Irvine, CA)	Approved in Europe
Intravenous	Glysure System by Glysure (Abingdon, UK)	Approved in Europe
Intravenous	Eirus by Maquet Getinge Group (Rastatt, Germany)	Approved in Europe
Intravenous	OptiScanner 5000 by OptiScan; (Hayward, CA)	Approved in EuropeFDA-approved for use in US hospitals
Intravenous	GlucoScout (International Biomedical, Austin, TX)	FDA-approved for use in US hospitals
Intravenous	Dexcom G	FDA-approved and CEA approved
Intravenous	Guardian™ Connect system by Medtronic (San Diego, CA)	FDA-approved for use in US hospitals
Subcutaneous	Freestyle Libre by Abbott Diabetes Care	US FDA approved

FDA: Food and Drug Administration; CEA: Carcinoembryonic antigen.

A recently published meta-analysis reported that the use of CGM was associated with significantly reduced HbA1c values and reduced risk of severe hypoglycaemia[39]. In addition, use of FGM was associated with significant reduction in episodes of mild hypoglycemia and was associated with increased treatment satisfaction in patients with type-I diabetes. Hence, it is suggested that real time monitoring with CGM or FGM has the potential to achieve better control in short-time fluctuations in BG levels, improve glycemic control and may also reduce healthcare costs[40]. Although several studies have been conducted testing these devices in critically ill patients, their impact on reducing length of stay in ICU or overall patient outcomes remains unknown[41].

While these devices may not benefit all ICU patients, they may be particularly useful in specific patient populations like those on intravenous insulin or corticosteroids, patients with end stage renal or liver disease, neurosurgery or traumatic brain injury patients and post-transplant patients[42-44]. However, these devices need to be further tested in larger patient cohorts before they find mainstream application.

Detection of adverse glycemic events

Detection of adverse events in the form of both hypoglycemia and hyperglycemia using AI technologies have been studied by various research groups mainly in type 1 and type 2 diabetes patients[35]. The studies used either CGM devices or self-monitoring of blood glucose monitors to detect the individual events. The results were based on the sensitivity and specificity of the modalities used. For example the DCBPN algorithm used by Zhang *et al*[45] provided an accuracy of 88.5% in predicting the BG levels. In the study by Otto *et al*[46], identification of episodes of hypoglycemia, hyperglycemia, severe hypoglycemia, and severe hyperglycemia were 120%, 46%, 123%, and 76% more likely after pattern identification as compared to periods when no pattern was identified. Another study by Nguyen *et al* [47] used electrocardiographic (ECG) parameters to detect episodes of hyperglycemia with a reported sensitivity and specificity of 70.59% and 65.38%, respectively. The results suggested that ECG signal and ANN patterns could be used to detect adverse hyperglycemic events in diabetic patients. Overall, AI has a potential role to predict adverse events and thus help modify treatment protocols so as to rectify them.

Blood glucose control strategies

There are various AI methodologies, fuzzy logic (FL), ANN, RF, which have been used for sugar control. Out of these FL is the most commonly used methodology as it mimics the management strategies by actual diabetes caregivers. Various studies have been performed using the FL methodology for BG control, mainly in type 1 diabetic patients[48,49]. The results have shown better control of nocturnal glucose levels with a low risk of hypoglycaemia as compared to standard insulin pump treatment.

Now, more complex methodologies are being proposed for BG control such as complimentary AI algorithms to support traditional AI controllers. The latest technology is the development of neural networks for regulation of BG[50,51].

From the above data it is evident that AI may potentially help to control BG but similar research in critically ill patients is limited. The LOGIC-1 trial was a single centre randomized control trial (RCT) which compared LOGIC-Insulin computerized algorithm to expert nurses in BG control for critically ill patients[52]. LOGIC-Insulin improved the efficacy of tight glucose control without increasing the risk of hypoglycemia. Encouraged by the results, a larger multi-center RCT, the LOGIC-2 trial, was conducted comparing software guided glucose control to nurse directed orders. This trial also showed better control of BG without an increase in hypoglycemia[53].

Hence, research shows that algorithmic based approach may be beneficial to control BG levels. Even the ability to anticipate excursions in sugar levels could provide early warnings regarding ineffective treatments. Newer CGM could lead to prediction of future glucose levels but reliability may be affected due various physiological and technical factors. Pappada *et al*[54] studied a neural network model for predicting glucose levels in a surgical critical care setting and found CGM to be useful in this patient population. However, further research and studies may be required in real time to test their validity in other critically ill patients.

Artificial pancreas

For BG control one of the most extensively researched modality is the artificial pancreas (AP) which consists of a glucose sensor, a closed-loop control algorithm, and an insulin infusion device. The glucose sensor estimates the BG level which in turn is fed to the control unit with the closed loop algorithm. This is turn directs the infusion device to inject the programmed amount of insulin. Thus, it has been developed to mimic the Islet cells of the pancreas which secrete insulin based on the BG levels. The majority of algorithms used by AP have been derived from control engineering theory and include proportional-integral-derivative (PID), model-predictive control, adaptive control, and FL control[55, 56]. However, the major limiting factor is a reliable glucose sensor and hence, now AI is being used to develop better models of AP.

At present, AP are of two types viz a viz single hormone (insulin only) and dual hormone (insulin and glucagon) systems. Overall, AP has been shown to be safe and effective in controlling BG, reducing episodes of hypoglycemia and hyperglycemia, and increase the proportion of TITR. Weisman *et al*[57] conducted a meta-analysis which showed that AP improves the TITR by 12.59% (equivalent to 172 minutes per day) compared to conventional treatment. Furthermore, this analysis showed that dual-hormone AP systems were associated with greater improvements, especially with respect to hypoglycemic events as compared to single hormone systems. The average time spent in hypoglycemia was reduced by 35 minutes/day. These benefits were more pronounced at night time.

In critically ill patients, use of AP to control BG has shown to reduce the frequency for sampling, reduce the nursing workload, achieve stable glycemic control with reduced episodes of hypo or hyperglycemia, and cause less GV[58-62]. In addition, its use has been associated with significant reduction in postoperative infectious complications in patients undergoing major surgeries[62]. However, use of AP was unable to achieve any significant improvement in mean glucose concentration, improve clinical outcome or show a favorable cost-benefit ratio.

Insulin bolus calculators and advisory systems

Insulin dependent patients routinely require calculation of insulin dosages based on their consumption of carbohydrates. The bolus doses are based on multiple factors like previous insulin dose, BG measurements, approximate calorie count *etc.* This may be a challenging task and could lead to errors in judgement and calculation, eventually leading to adverse glycemic events. Various applications are being developed to simplify this daunting task. Various research groups have used the case-based reasoning methodology for these calculations which has proved to be a safe decision tool. Some studies have also shown that complimenting this system to an AP leads to an improvement in glycemic control [62,63]. Since the cause of hyperglycemia in ICU is multifactorial, probably a combination of an AP with case-based methodology may be of help as glucose excursions could be treated in a more standardized way with better control.

MD-Logic controller, developed on the FL systems, have shown to provide superior glycemic control with fewer nocturnal hypoglycemic episodes as compared to insulin pump treatment[49]. However, it still needs to be validated in ICU patients.

Software based algorithms for insulin dosing

Software based algorithms have been developed to determine insulin dosage depending on the BG levels. These programs, although more complicated than the paper-based protocols, can reduce errors and improve adherence. The simplest of these are based on PID models. Devices based on this model titrate insulin administration based on the previous BG values and predicting the changes in glucose value for a given insulin dose using a dynamic multiplier response to insulin sensitivity. The advantages of this model include the need for minimal patient related information for initiation and its ability to provide real-time dose adjustments. However, this model necessitates multiple blood sampling, which may be up to 18 times per day for BG measurements[64,65].

A more complex modification of software is Glucose Regulation for Intensive Care Patients which not only takes into account the BG values and insulin infusion rates but also includes the change in these values over time. This may increase its effectiveness and may potentially reduce overtreatment and hence, hypoglycemic episodes[66,67].

The most recent algorithms are classified as model predictive controls, which not only include insulin sensitivity and dextrose administration but also include several patient-specific parameters like their age and diabetes status. Based on these factors, these algorithms try to predict the patient's response to hyperglycemia and insulin therapy and adjust the insulin dose accordingly. As the number of

parameters required to be entered at the time of initiation are more, the devices based on these algorithms are more complicated and time consuming but they have advantages of increased accuracy, significantly reduced need for repeated blood sampling and may offer a more individualized insulin therapy[68-70].

CGM regulated insulin infusion system

Newer technologies like CGM which have been validated in non-critically ill patients are now increasingly been used with increased accuracy in ICU patients. Integration of these CGM devices with automated insulin suspension with AI algorithms (Basal-IQ™ technology) have been approved by US-FDA. Use of these predictive low-glucose suspend (PLGS) algorithms offer clinical advantage over the more conventional threshold suspend systems which stop insulin only when the predefined threshold of glucose is breached. Glucose values are obtained by the integrated CGM device (Dexcom G6™) and the Basal-IQ™ has the ability to predict when the glucose value is going to drop below the predefined level and it stops the insulin infusion[71]. Control-IQ is a more advanced hybrid closed-loop system which also uses activity and sleep settings to adjust the insulin requirements. Basal-IQ™ and Control-IQ™ algorithms can predict hypoglycemic events up to 30 minutes in advance and hence, can titrate the insulin dose accordingly.

Integration of CGM with an automated insulin suspension has shown to reduce the frequency and duration of hypoglycaemia with a reported relative risk reduction of 45%[72]. This effect has been shown to exist across different age groups, and is persistent over multiple weeks with real-world use. A large randomized crossover trial comparing the PLGS with sensor-augmented insulin pump showed 31% reduction in time spent in hypoglycemia (< 70 mg/dL) with no increase in incidence of rebound hyperglycemia[73]. It may be suggested that, use of this technology may be feasible and effective for patients with difficult to control DM and those at higher risk for developing hypoglycemia[72].

Risk and patient stratification

Diabetes is a chronic disease associated with many complications. Even though most of the complications develop over a period of time, diabetic patients are also prone to develop acute life-threatening complications like nosocomial infections, acute kidney injury and even cardiovascular complications. AI using deep-learning techniques have been able to produce algorithms which are able to predict long-term micro-angiopathic complications like diabetic retinopathy, diabetic foot, diabetic neuropathy and diabetic nephropathy, with reasonable accuracy[74-77]. Role of AI in predicting the development of macro-angiopathic complications like acute myocardial infarction has also been assessed but there is a dearth of data regarding its role in predicting other acute complications, especially in critically ill patients[78].

AI has been used effectively to determine patients at risk for developing sepsis and life-threatening nosocomial infections like catheter related blood stream infections and *Clostridium difficile* infections and also to predict which ward patients may deteriorate and require ICU admission. However, such models currently do not exist specifically for diabetes patients[13,79-81].

A few studies have also used AI in predicting mortality in critically ill diabetes patients. In their study, Ye *et al*[82] using the MIMIC-III database, reported that AI using CNN was highly accurate in predicting mortality in critically ill diabetes patients with an area under the curve (AUC) of 0.97. Using the same MIMIC-III database, Anand *et al*[83] developed simple predictive tools with AI, to predict mortality in critically ill diabetics. Their models could achieve AUCs of 0.787 and 0.785 to predict mortality. However, these models need to be compared to more widely used and validated models for mortality prediction in ICU patients like acute physiology and chronic health evaluation and sequential organ failure and assessment scores.

Coronavirus disease critical care

The recent pandemic of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has put an unprecedented strain on the healthcare with enhanced need for infection control and patient isolation. Separate coronavirus disease 2019 (COVID-19) ICUs had to be developed with negative pressure chambers with treating staff wearing personal protection equipment at all times. Diabetes is one of the most common comorbidities among COVID-19 patients. Diabetic patients developing COVID-19 are at higher risk for requiring ICU admission and have poorer outcomes. The need for personal protection and risk of transmission of infection has put immense pressure on already limited clinical workforce. In such a scenario, labour intensive work like frequent BG monitoring and insulin dose adjustments may get seriously hampered. AI may be especially helpful by reducing the burden on the healthcare workers (HCWs) and reducing their risk of exposure.

Computerized algorithms, automated closed loop systems and remote monitoring may all be used effectively to manage critically ill COVID-19 patients. CGM devices are capable of continuous BG tracking enabling real-time monitoring of BG levels while reducing the need for bedside monitoring, thereby reducing the risk of exposure for the HCWs. The efficacy and safety of CGM in managing critically ill COVID-19 patients has been tested and verified and it has been reported to reduce the need for bedside BG testing by up to 71%. In addition, the efficacy of CGM devices was not significantly

affected by presence of fever, hypoxemia, need for vasopressors, acidosis or with use of corticosteroid or parenteral nutrition[84-86]. Based on this, US-FDA has allowed the use of CGM in COVID-19 ICUs to reduce the exposure of HCWs[87].

AI based devices have the potential to improve patient care and outcomes by providing a better glucose control without increasing the nursing workload and avoiding risk of transmission of infection. Hence, it is recommended to prefer CGM to reduce the need for frequent nurse contact for patients with active COVID-19 infection[88]. Moreover, AI has also been instrumental in achieving glycemic control in COVID-19 patient on extracorporeal membrane oxygenation support by using AP[89].

STRENGTHS OF AI

AI-based devices have the potential to improve glycemic control, reduce GV, increase the TITR, and reduce episodes of hyper and hypoglycemia, thus providing comprehensive diabetes care. AI may allow us to achieve a better and more individualized glycemic control taking into account specific patient requirements as per their calorie intake, exercise and underlying comorbidities. In addition, AI may be better suited to care for patients at risk for adverse effects and those with changing needs, like those in critical care areas. It may enable HCWs to monitor their patients remotely with reduced need for close contact thereby, reducing their workload and exposure to infective patients. By reducing the need for frequent blood sampling and providing close glucose monitoring and insulin dose titration, AI-based algorithms may increase patient safety and satisfaction.

LIMITATIONS OF AI

Healthcare applications of AI are rapidly increasing. However, it still has several limitations affecting its widespread applicability (Table 5). Even though many AI applications have found acceptability in out-patients and ward patients with diabetes, data regarding its safety and accuracy in critically ill patients remains limited. As AI application is largely data-driven, involving collection of sensitive personal data, it may have privacy issues leading to medico-legal problems. Lack of regulations, recommendations and guidelines pertaining to use of AI further limit its applicability. These safety, liability and reliability issues prevent widespread use of AI in critical care practice. In addition, challenges of integrating AI into existing healthcare infrastructure and user acceptance also persist.

FUTURE DIRECTIONS

The future of healthcare development is in AI. Its large-scale applicability requires widespread availability, low cost and ease of use. In addition, AI needs to be adapted gradually in the existing healthcare system and HCWs need to be trained not only to better utilize AI but also to be aware of how to avoid any medico-legal issues arising from its application. Changes in the laws and regulations are also required to safeguard patient's interest and avoid any violation of patient's privacy. With techno-logical improvements in AI, the dosing algorithms for insulin delivery may become individualized for closed-loop control of glycemia. Larger studies, evaluating their efficacy and safety, especially in critically ill patients, along with standardization of AI algorithms and techniques need to be done to improve the acceptability of AI.

CONCLUSION

Many currently available devices and techniques which have proven their role in management of non-critically ill patients, may soon be available for ICU patients, with improved accuracy. CGM is already being recommended for use in critically ill COVID-19 patients and soon may be available for use in all critically ill patients. Its integration with automated insulin suspension holds greater promise. Use of AP may also provide a comprehensive glycemic control option. AI has the potential of reducing the workload of HCWs, provide better glycemic control and prevent related complications, however, larger RCTs may be required before we implement these techniques in our day-to-day critical care. Even though presently AI might not be in its prime for managing critically ill diabetic patients, it is the future of healthcare.

Table 5 Limitations of artificial intelligence

Factors	
Human factors	Inhibition, lack of experience
Technical factors	Cost, availability and implementation
Data limitation	Lack of data in ICU patients, lack of large scale randomized trials
Design limitation	Devices tried in certain patient populations may not be applicable in ICU patients
Ethical	Lack of guidelines

ICU: Intensive care unit.

FOOTNOTES

REFERENCES

1   IDF Diabetes Atlas 10th edition Internet. 2217. Cited 26 January 2022. Available from: https://idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html

2   **Kompoti M**, Michalia M, Salma V, Diogou E, Lakoumenta A, Clouva-Molyvdas PM. Glycated hemoglobin at admission in the intensive care unit: clinical implications and prognostic relevance. *J Crit Care* 2015; **30**: 150-155 [PMID: 25239822 DOI: 10.1016/j.jcrc.2014.08.014]

3   **Abu-Ashour W**, Twells L, Valcour J, Randell A, Donnan J, Howse P, Gamble JM. The association between diabetes mellitus and incident infections: a systematic review and meta-analysis of observational studies. *BMJ Open Diabetes Res Care* 2017; **5**: e000336 [PMID: 28761647 DOI: 10.1136/bmjdrc-2016-000336]

4   **Egi M**, Bellomo R, Stachowski E, French CJ, Hart GK, Hegarty C, Bailey M. Blood glucose concentration and outcome of critical illness: the impact of diabetes. *Crit Care Med* 2008; **36**: 2249-2255 [PMID: 18664780 DOI: 10.1097/CCM.0b013e318181039a]

5   **NICE-SUGAR Study Investigators**, Finfer S, Liu B, Chittock DR, Norton R, Myburgh JA, McArthur C, Mitchell I, Foster D, Dhingra V, Henderson WR, Ronco JJ, Bellomo R, Cook D, McDonald E, Dodek P, Hébert PC, Heyland DK, Robinson BG. Hypoglycemia and risk of death in critically ill patients. *N Engl J Med* 2012; **367**: 1108-1118 [PMID: 22992074 DOI: 10.1056/NEJMoa1204942]

6   **Krinsley JS**. Understanding glycemic control in the critically ill: three domains are better than one. *Intensive Care Med* 2011; **37**: 382-384 [PMID: 21210079 DOI: 10.1007/s00134-010-2110-3]

7   **Krinsley JS**. Glycemic control in the critically ill: What have we learned since NICE-SUGAR? *Hosp Pract (1995)* 2015; **43**: 191-197 [PMID: 26224425 DOI: 10.1080/21548331.2015.1066227]

8   **Yamada T**, Shojima N, Noma H, Yamauchi T, Kadowaki T. Glycemic control, mortality, and hypoglycemia in critically ill patients: a systematic review and network meta-analysis of randomized controlled trials. *Intensive Care Med* 2017; **43**: 1-15 [PMID: 27637719 DOI: 10.1007/s00134-016-4523-0]

9   **American Diabetes Association**. 15. Diabetes Care in the Hospital: *Standards of Medical Care in Diabetes-2021*. *Diabetes Care* 2021; **44**: S211-S220 [PMID: 33298426 DOI: 10.2337/dc21-S015]

10   **Evans L**, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, Machado FR, Mcintyre L, Ostermann M, Prescott HC, Schorr C, Simpson S, Joost Wiersinga W, Alshamsi F, Angus DC, Arabi Y, Azevedo L, Beale R, Beilman G, Belley-Cote E, Burry L, Cecconi M, Centofanti J, Yataco AC, De Waele J, Dellinger RP, Doi K, Du B, Estenssoro E,

Ferrer R, Gomersall C, Hodgson C, Møller MH, Iwashyna T, Jacob S, Kleinpell R, Klompas M, Koh Y, Kumar A, Kwizera A, Lobo S, Masur H, McGloughlin S, Mehta S, Mehta Y, Mer M, Nunnally M, Oczkowski S, Osborn T, Papathanassoglou E, Perner A, Puskarich M, Roberts J, Schweickert W, Seckel M, Sevransky J, Sprung CL, Welte T, Zimmerman J, Levy M. Executive Summary: Surviving Sepsis Campaign: International Guidelines for the Management of Sepsis and Septic Shock 2021. *Crit Care Med* 2021; **49**: 1974-1982 [PMID: 34643578 DOI: 10.1097/CCM.0000000000005357]

11   **Sottile PD**, Albers D, Higgins C, Mckeehan J, Moss MM. The Association Between Ventilator Dyssynchrony, Delivered Tidal Volume, and Sedation Using a Novel Automated Ventilator Dyssynchrony Detection Algorithm. *Crit Care Med* 2018; **46**: e151-e157 [PMID: 29337804 DOI: 10.1097/CCM.0000000000002849]

12   **Kuo HJ**, Chiu HW, Lee CN, Chen TT, Chang CC, Bien MY. Improvement in the Prediction of Ventilator Weaning Outcomes by an Artificial Neural Network in a Medical ICU. *Respir Care* 2015; **60**: 1560-1569 [PMID: 26329358 DOI: 10.4187/respcare.03648]

13   **Nemati S**, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU. *Crit Care Med* 2018; **46**: 547-553 [PMID: 29286945 DOI: 10.1097/CCM.0000000000002936]

14   **Turnidge J**. Pharmacodynamics and dosing of aminoglycosides. *Infect Dis Clin North Am* 2003; **17**: 503-528, v [PMID: 14711074 DOI: 10.1016/s0891-5520(03)00057-6]

15   **Rachh P**, Levey AO, Lemmon A, Marinescu A, Auffermann WF, Haycook D, Berkowitz EA. Reducing STAT Portable Chest Radiograph Turnaround Times: A Pilot Study. *Curr Probl Diagn Radiol* 2018; **47**: 156-160 [PMID: 28705527 DOI: 10.1067/j.cpradiol.2017.05.012]

16   **Copeland BJ**.   "Artificial Intelligence". Encyclopedia Britannica. Cited 22 January 2022. Available from: https://www.britannica.com/technology/artificial-intelligence

17   **Gobet F**, Clarkson G. Chunks in expert memory: evidence for the magical number four ... or is it two? *Memory* 2004; **12**: 732-747 [PMID: 15724362 DOI: 10.1080/09658210344000530]

18   **Raghupathi V**, Raghupathi W. Healthcare Expenditure and Economic Performance: Insights From the United States Data. *Front Public Health* 2020; **8**: 156 [PMID: 32478027 DOI: 10.3389/fpubh.2020.00156]

19   **Shillan D**, Sterne JAC, Champneys A, Gibbison B. Use of machine learning to analyse routinely collected intensive care unit data: a systematic review. *Crit Care* 2019; **23**: 284 [PMID: 31439010 DOI: 10.1186/s13054-019-2564-9]

20   **Nomura A**, Noguchi M, Kometani M, Furukawa K, Yoneda T. Artificial Intelligence in Current Diabetes Management and Prediction. *Curr Diab Rep* 2021; **21**: 61 [PMID: 34902070 DOI: 10.1007/s11892-021-01423-2]

21   **Ellahham S**. Artificial Intelligence: The Future for Diabetes Care. *Am J Med* 2020; **133**: 895-900 [PMID: 32325045 DOI: 10.1016/j.amjmed.2020.03.033]

22   **Abràmoff MD**, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018; **1**: 39 [PMID: 31304320 DOI: 10.1038/s41746-018-0040-6]

23   **Nimri R**, Battelino T, Laffel LM, Slover RH, Schatz D, Weinzimer SA, Dovc K, Danne T, Phillip M; NextDREAM Consortium. Insulin dose optimization using an automated artificial intelligence-based decision support system in youths with type 1 diabetes. *Nat Med* 2020; **26**: 1380-1384 [PMID: 32908282 DOI: 10.1038/s41591-020-1045-7]

24   **Dankwa-Mullan I**, Rivo M, Sepulveda M, Park Y, Snowdon J, Rhee K. Transforming Diabetes Care Through Artificial Intelligence: The Future Is Here. *Popul Health Manag* 2019; **22**: 229-242 [PMID: 30256722 DOI: 10.1089/pop.2018.0129]

25   **Lam C**, Yu C, Huang L, Rubin D. Retinal Lesion Detection With Deep Learning Using Image Patches. *Invest Ophthalmol Vis Sci* 2018; **59**: 590-596 [PMID: 29372258 DOI: 10.1167/iovs.17-22721]

26   **US Food and Drug Administration**.   FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. Accessed April 18, 2022. Available from: https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm604357.htm

27   **Keel S**, Lee PY, Scheetz J, Li Z, Kotowicz MA, MacIsaac RJ, He M. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. *Sci Rep* 2018; **8**: 4330 [PMID: 29531299 DOI: 10.1038/s41598-018-22612-2]

28   **Nimri R**, Dassau E, Segall T, Muller I, Bratina N, Kordonouri O, Bello R, Biester T, Dovc K, Tenenbaum A, Brener A, Šimunović M, Sakka SD, Nevo Shenker M, Passone CG, Rutigliano I, Tinti D, Bonura C, Caiulo S, Ruszala A, Piccini B, Giri D, Stein R, Rabbone I, Bruzzi P, Omladič JŠ, Steele C, Beccuti G, Yackobovitch-Gavan M, Battelino T, Danne T, Atlas E, Phillip M. Adjusting insulin doses in patients with type 1 diabetes who use insulin pump and continuous glucose monitoring: Variations among countries and physicians. *Diabetes Obes Metab* 2018; **20**: 2458-2466 [PMID: 29885025 DOI: 10.1111/dom.13408]

29   **Gómez AM**, Henao Carrillo DC, Muñoz Velandia OM. Devices for continuous monitoring of glucose: update in technology. *Med Devices (Auckl)* 2017; **10**: 215-224 [PMID: 28979168 DOI: 10.2147/MDER.S110121]

30   **Dungan KM**, Braithwaite SS, Preiser JC. Stress hyperglycaemia. *Lancet* 2009; **373**: 1798-1807 [PMID: 19465235 DOI: 10.1016/S0140-6736(09)60553-5]

31   **Siegelaar SE**, Hermanides J, Oudemans-van Straaten HM, van der Voort PH, Bosman RJ, Zandstra DF, DeVries JH. Mean glucose during ICU admission is related to mortality by a U-shaped curve in surgical and medical patients: a retrospective cohort study. *Crit Care* 2010; **14**: R224 [PMID: 21143980 DOI: 10.1186/cc9369]

32   **Signal M**, Le Compte A, Shaw GM, Chase JG. Glycemic levels in critically ill patients: are normoglycemia and low variability associated with improved outcomes? *J Diabetes Sci Technol* 2012; **6**: 1030-1037 [PMID: 23063028 DOI: 10.1177/193229681200600506]

33   **Penning S**, Chase JG, Preiser JC, Pretty CG, Signal M, Mélot C, Desaive T. Does the achievement of an intermediate glycemic target reduce organ failure and mortality? *J Crit Care* 2014; **29**: 374-379 [PMID: 24679489 DOI: 10.1016/j.jcrc.2014.01.013]

34   **Omar AS**, Salama A, Allam M, Elgohary Y, Mohammed S, Tuli AK, Singh R. Association of time in blood glucose range with outcomes following cardiac surgery. *BMC Anesthesiol* 2015; **15**: 14 [PMID: 25670921 DOI: 10.1186/1471-2253-15-14]

35   **Contreras I**, Vehi J. Artificial Intelligence for Diabetes Management and Decision Support: Literature Review. *J Med Internet Res* 2018; **20**: e10775 [PMID: 29848472 DOI: 10.2196/10775]

36   **Zhang Y**, Liu X, Zhang J, Fu J, Li S, Chen S, Chen Y, Lu B. Evaluation for the feasibility and accuracy of Freestyle Libre Flash Glucose Monitoring System used by COVID-19 patients in intensive care unit. *J Diabetes* 2021; **13**: 603-605 [PMID: 33787006 DOI: 10.1111/1753-0407.13181]

37   **Kotzapanagiotou E**, Tsotridou E, Volakli E, Dimitriadou M, Chochliourou E, Kalamitsou S, Kotzapanagiotou F, Sdougka M, Christoforidis A. Evaluation of continuous flash glucose monitoring in a pediatric ICU setting. *J Clin Monit Comput* 2020; **34**: 843-852 [PMID: 31482363 DOI: 10.1007/s10877-019-00384-y]

38   **Ancona P**, Eastwood GM, Lucchetta L, Ekinci EI, Bellomo R, Mårtensson J. The performance of flash glucose monitoring in critically ill patients with diabetes. *Crit Care Resusc* 2017; **19**: 167-174 [PMID: 28651513]

39   **Dicembrini I**, Cosentino C, Monami M, Mannucci E, Pala L. Effects of real-time continuous glucose monitoring in type 1 diabetes: a meta-analysis of randomized controlled trials. *Acta Diabetol* 2021; **58**: 401-410 [PMID: 32789691 DOI: 10.1007/s00592-020-01589-3]

40   **Wood A**, O'Neal D, Furler J, Ekinci EI. Continuous glucose monitoring: a review of the evidence, opportunities for future use and ongoing challenges. *Intern Med J* 2018; **48**: 499-508 [PMID: 29464891 DOI: 10.1111/imj.13770]

41   **Umpierrez GE**, Klonoff DC. Diabetes Technology Update: Use of Insulin Pumps and Continuous Glucose Monitoring in the Hospital. *Diabetes Care* 2018; **41**: 1579-1589 [PMID: 29936424 DOI: 10.2337/dci18-0002]

42   **Wallia A**, Umpierrez GE, Nasraway SA, Klonoff DC; PRIDE Investigators. Round Table Discussion on Inpatient Use of Continuous Glucose Monitoring at the International Hospital Diabetes Meeting. *J Diabetes Sci Technol* 2016; **10**: 1174-1181 [PMID: 27286715 DOI: 10.1177/1932296816656380]

43   **Wallia A**, Umpierrez GE, Rushakoff RJ, Klonoff DC, Rubin DJ, Hill Golden S, Cook CB, Thompson B; DTS Continuous Glucose Monitoring in the Hospital Panel. Consensus Statement on Inpatient Use of Continuous Glucose Monitoring. *J Diabetes Sci Technol* 2017; **11**: 1036-1044 [PMID: 28429611 DOI: 10.1177/1932296817706151]

44   **Krinsley JS**, Chase JG, Gunst J, Martensson J, Schultz MJ, Taccone FS, Wernerman J, Bohe J, De Block C, Desaive T, Kalfon P, Preiser JC. Continuous glucose monitoring in the ICU: clinical considerations and consensus. *Crit Care* 2017; **21**: 197 [PMID: 28756769 DOI: 10.1186/s13054-017-1784-0]

45   **Zhang Y**, Zhu JM, Liang YB, Chen HB, Yin SM, Chen ZC. Non-invasive blood glucose detection system based on conservation of energy method. *Physiol Meas* 2017; **38**: 325-342 [PMID: 28107204 DOI: 10.1088/1361-6579/aa50cf]

46   **Otto EA**, Tannan V. Evaluation of the utility of a glycemic pattern identification system. *J Diabetes Sci Technol* 2014; **8**: 830-838 [PMID: 24876425 DOI: 10.1177/1932296814532210]

47   **Nguyen LL**, Su S, Nguyen HT. Neural network approach for non-invasive detection of hyperglycemia using electrocardiographic signals. *Annu Int Conf IEEE Eng Med Biol Soc* 2014; **2014**: 4475-4478 [PMID: 25570985 DOI: 10.1109/EMBC.2014.6944617]

48   **Nimri R**, Atlas E, Ajzensztejn M, Miller S, Oron T, Phillip M. Feasibility study of automated overnight closed-loop glucose control under MD-logic artificial pancreas in patients with type 1 diabetes: the DREAM Project. *Diabetes Technol Ther* 2012; **14**: 728-735 [PMID: 22853723 DOI: 10.1089/dia.2012.0004]

49   **Phillip M**, Battelino T, Atlas E, Kordonouri O, Bratina N, Miller S, Biester T, Stefanija MA, Muller I, Nimri R, Danne T. Nocturnal glucose control with an artificial pancreas at a diabetes camp. *N Engl J Med* 2013; **368**: 824-833 [PMID: 23445093 DOI: 10.1056/NEJMoa1206881]

50   **Fernandez de Canete J**, Gonzalez-Perez S, Ramos-Diaz JC. Artificial neural networks for closed loop control of in silico and ad hoc type 1 diabetes. *Comput Methods Programs Biomed* 2012; **106**: 55-66 [PMID: 22178070 DOI: 10.1016/j.cmpb.2011.11.006]

51   **Catalogna M**, Cohen E, Fishman S, Halpern Z, Nevo U, Ben-Jacob E. Artificial neural networks based controller for glucose monitoring during clamp test. *PLoS One* 2012; **7**: e44587 [PMID: 22952998 DOI: 10.1371/journal.pone.0044587]

52   **Van Herpe T**, Mesotten D, Wouters PJ, Herbots J, Voets E, Buyens J, De Moor B, Van den Berghe G. LOGIC-insulin algorithm-guided versus nurse-directed blood glucose control during critical illness: the LOGIC-1 single-center, randomized, controlled clinical trial. *Diabetes Care* 2013; **36**: 188-194 [PMID: 22961576 DOI: 10.2337/dc12-0584]

53   **Dubois J**, Van Herpe T, van Hooijdonk RT, Wouters R, Coart D, Wouters P, Van Assche A, Veraghtert G, De Moor B, Wauters J, Wilmer A, Schultz MJ, Van den Berghe G, Mesotten D. Software-guided versus nurse-directed blood glucose control in critically ill patients: the LOGIC-2 multicenter randomized controlled clinical trial. *Crit Care* 2017; **21**: 212 [PMID: 28806982 DOI: 10.1186/s13054-017-1799-6]

54   **Pappada SM**, Cameron BD, Rosman PM, Bourey RE, Papadimos TJ, Olorunto W, Borst MJ. Neural network-based real-time prediction of glucose in patients with insulin-dependent diabetes. *Diabetes Technol Ther* 2011; **13**: 135-141 [PMID: 21284480 DOI: 10.1089/dia.2010.0104]

55   **Bothe MK**, Dickens L, Reichel K, Tellmann A, Ellger B, Westphal M, Faisal AA. The use of reinforcement learning algorithms to meet the challenges of an artificial pancreas. *Expert Rev Med Devices* 2013; **10**: 661-673 [PMID: 23972072 DOI: 10.1586/17434440.2013.827515]

56   **Bertachi A**, Ramkissoon CM, Bondia J, Vehí J. Automated blood glucose control in type 1 diabetes: A review of progress and challenges. *Endocrinol Diabetes Nutr (Engl Ed)* 2018; **65**: 172-181 [PMID: 29279252 DOI: 10.1016/j.endinu.2017.10.011]

57   **Weisman A**, Bai JW, Cardinez M, Kramer CK, Perkins BA. Effect of artificial pancreas systems on glycaemic control in patients with type 1 diabetes: a systematic review and meta-analysis of outpatient randomised controlled trials. *Lancet Diabetes Endocrinol* 2017; **5**: 501-512 [PMID: 28533136 DOI: 10.1016/S2213-8587(17)30167-5]

58   **Mibu K**, Yatabe T, Hanazaki K. Blood glucose control using an artificial pancreas reduces the workload of ICU nurses. *J Artif Organs* 2012; **15**: 71-76 [PMID: 21947674 DOI: 10.1007/s10047-011-0611-7]

59   **DeJournett L**, DeJournett J. In Silico Testing of an Artificial-Intelligence-Based Artificial Pancreas Designed for Use in the Intensive Care Unit Setting. *J Diabetes Sci Technol* 2016; **10**: 1360-1371 [PMID: 27301982 DOI: 10.1177/1932296816653967]

60   **Hanazaki K**, Tanioka N, Munekage M, Uemura S, Maeda H. Closed-loop artificial endocrine pancreas from Japan. *Artif*

Organs 2021; **45**: 958-967 [PMID: 34105784 DOI: 10.1111/aor.14008]

61 **Yang Z**, Tao G, Guo M, Sun B, Gong L, Ding Y, Ye S, Liu W, Yang X. [Efficacy and safety of simulated artificial pancreas in modulating stress hyperglycemia in critically ill patients: a prospective randomized controlled study]. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue* 2018; **30**: 165-169 [PMID: 29402368 DOI: 10.3760/cma.j.issn.2095-4352.2018.02.014]

62 **Herrero P**, Pesl P, Reddy M, Oliver N, Georgiou P, Toumazou C. Advanced Insulin Bolus Advisor Based on Run-To-Run Control and Case-Based Reasoning. *IEEE J Biomed Health Inform* 2015; **19**: 1087-1096 [PMID: 24956470 DOI: 10.1109/JBHI.2014.2331896]

63 **Herrero P**, Bondia J, Adewuyi O, Pesl P, El-Sharkawy M, Reddy M, Toumazou C, Oliver N, Georgiou P. Enhancing automatic closed-loop glucose control in type 1 diabetes with an adaptive meal bolus calculator - in silico evaluation under intra-day variability. *Comput Methods Programs Biomed* 2017; **146**: 125-131 [PMID: 28688482 DOI: 10.1016/j.cmpb.2017.05.010]

64 **Juneja R**, Roudebush C, Kumar N, Macy A, Golas A, Wall D, Wolverton C, Nelson D, Carroll J, Flanders SJ. Utilization of a computerized intravenous insulin infusion program to control blood glucose in the intensive care unit. *Diabetes Technol Ther* 2007; **9**: 232-240 [PMID: 17561793 DOI: 10.1089/dia.2006.0015]

65 **Newton CA**, Smiley D, Bode BW, Kitabchi AE, Davidson PC, Jacobs S, Steed RD, Stentz F, Peng L, Mulligan P, Freire AX, Temponi A, Umpierrez GE. A comparison study of continuous insulin infusion protocols in the medical intensive care unit: computer-guided vs. standard column-based algorithms. *J Hosp Med* 2010; **5**: 432-437 [PMID: 20945468 DOI: 10.1002/jhm.816]

66 **Vogelzang M**, Loef BG, Regtien JG, van der Horst IC, van Assen H, Zijlstra F, Nijsten MW. Computer-assisted glucose control in critically ill patients. *Intensive Care Med* 2008; **34**: 1421-1427 [PMID: 18389221 DOI: 10.1007/s00134-008-1091-y]

67 **Hoekstra M**, Schoorl MA, van der Horst IC, Vogelzang M, Wietasch JK, Zijlstra F, Nijsten MW. Computer-assisted glucose regulation during rapid step-wise increases of parenteral nutrition in critically ill patients: a proof of concept study. *JPEN J Parenter Enteral Nutr* 2010; **34**: 549-553 [PMID: 20852185 DOI: 10.1177/0148607110372390]

68 **Pachler C**, Plank J, Weinhandl H, Chassin LJ, Wilinska ME, Kulnik R, Kaufmann P, Smolle KH, Pilger E, Pieber TR, Ellmerer M, Hovorka R. Tight glycaemic control by an automated algorithm with time-variant sampling in medical ICU patients. *Intensive Care Med* 2008; **34**: 1224-1230 [PMID: 18297268 DOI: 10.1007/s00134-008-1033-8]

69 **Hoekstra M**, Vogelzang M, Verbitskiy E, Nijsten MW. Health technology assessment review: Computerized glucose regulation in the intensive care unit--how to create artificial control. *Crit Care* 2009; **13**: 223 [PMID: 19849827 DOI: 10.1186/cc8023]

70 **Amrein K**, Ellmerer M, Hovorka R, Kachel N, Fries H, von Lewinski D, Smolle K, Pieber TR, Plank J. Efficacy and safety of glucose control with Space GlucoseControl in the medical intensive care unit--an open clinical investigation. *Diabetes Technol Ther* 2012; **14**: 690-695 [PMID: 22694176 DOI: 10.1089/dia.2012.0021]

71 Tandem Diabetes Care Announces Health Canada Approval of t:slim X2 Insulin Pump with Basal-IQ Predictive Low-Glucose Suspend Technology. Cited 6 January 2022. Available from: http://investor.tandemdiabetes.com/news-releases/news-release-details/tandem-diabetes-care-announces-health-canadaapproval-tslim-x2/

72 **Müller L**, Habif S, Leas S, Aronoff-Spencer E. Reducing Hypoglycemia in the Real World: A Retrospective Analysis of Predictive Low-Glucose Suspend Technology in an Ambulatory Insulin-Dependent Cohort. *Diabetes Technol Ther* 2019; **21**: 478-484 [PMID: 31329468 DOI: 10.1089/dia.2019.0190]

73 **Forlenza GP**, Li Z, Buckingham BA, Pinsker JE, Cengiz E, Wadwa RP, Ekhlaspour L, Church MM, Weinzimer SA, Jost E, Marcal T, Andre C, Carria L, Swanson V, Lum JW, Kollman C, Woodall W, Beck RW. Predictive Low-Glucose Suspend Reduces Hypoglycemia in Adults, Adolescents, and Children With Type 1 Diabetes in an At-Home Randomized Crossover Study: Results of the PROLOG Trial. *Diabetes Care* 2018; **41**: 2155-2161 [PMID: 30089663 DOI: 10.2337/dc18-0771]

74 **Ting DSW**, Cheung CY, Nguyen Q, Sabanayagam C, Lim G, Lim ZW, Tan GSW, Soh YQ, Schmetterer L, Wang YX, Jonas JB, Varma R, Lee ML, Hsu W, Lamoureux E, Cheng CY, Wong TY. Deep learning in estimating prevalence and systemic risk factors for diabetic retinopathy: a multi-ethnic study. *NPJ Digit Med* 2019; **2**: 24 [PMID: 31304371 DOI: 10.1038/s41746-019-0097-x]

75 **Cruz-Vega I**, Hernandez-Contreras D, Peregrina-Barreto H, Rangel-Magdaleno JJ, Ramirez-Cortes JM. Deep Learning Classification for Diabetic Foot Thermograms. *Sensors (Basel)* 2020; **20** [PMID: 32235780 DOI: 10.3390/s20061762]

76 **Williams BM**, Borroni D, Liu R, Zhao Y, Zhang J, Lim J, Ma B, Romano V, Qi H, Ferdousi M, Petropoulos IN, Ponirakis G, Kaye S, Malik RA, Alam U, Zheng Y. An artificial intelligence-based deep learning algorithm for the diagnosis of diabetic neuropathy using corneal confocal microscopy: a development and validation study. *Diabetologia* 2020; **63**: 419-430 [PMID: 31720728 DOI: 10.1007/s00125-019-05023-4]

77 **Norouzi J**, Yadollahpour A, Mirbagheri SA, Mazdeh MM, Hosseini SA. Predicting Renal Failure Progression in Chronic Kidney Disease Using Integrated Intelligent Fuzzy Expert System. *Comput Math Methods Med* 2016; **2016**: 6080814 [PMID: 27022406 DOI: 10.1155/2016/6080814]

78 **Yamada T**, Iwasaki K, Maedera S, Ito K, Takeshima T, Noma H, Shojima N. Myocardial infarction in type 2 diabetes using sodium-glucose co-transporter-2 inhibitors, dipeptidyl peptidase-4 inhibitors or glucagon-like peptide-1 receptor agonists: proportional hazards analysis by deep neural network based machine learning. *Curr Med Res Opin* 2020; **36**: 403-409 [PMID: 31855074 DOI: 10.1080/03007995.2019.1706043]

79 **Churpek MM**, Yuen TC, Winslow C, Hall J, Edelson DP. Differences in vital signs between elderly and nonelderly patients prior to ward cardiac arrest. *Crit Care Med* 2015; **43**: 816-822 [PMID: 25559439 DOI: 10.1097/CCM.0000000000000818]

80 **Beeler C**, Dbeibo L, Kelley K, Thatcher L, Webb D, Bah A, Monahan P, Fowler NR, Nicol S, Judy-Malcolm A, Azar J. Assessing patient risk of central line-associated bacteremia *via* machine learning. *Am J Infect Control* 2018; **46**: 986-991 [PMID: 29661634 DOI: 10.1016/j.ajic.2018.02.021]

81 **Li BY**, Oh J, Young VB, Rao K, Wiens J. Using Machine Learning and the Electronic Health Record to Predict Complicated *Clostridium difficile* Infection. *Open Forum Infect Dis* 2019; **6**: ofz186 [PMID: 31139672 DOI:

10.1093/ofid/ofz186]

82   **Ye J**, Yao L, Shen J, Janarthanam R, Luo Y. Predicting mortality in critically ill patients with diabetes using machine learning and clinical notes. *BMC Med Inform Decis Mak* 2020; **20**: 295 [PMID: 33380338 DOI: 10.1186/s12911-020-01318-4]

83   **Anand RS**, Stey P, Jain S, Biron DR, Bhatt H, Monteiro K, Feller E, Ranney ML, Sarkar IN, Chen ES. Predicting Mortality in Diabetic ICU Patients Using Machine Learning and Severity Indices. *AMIA Jt Summits Transl Sci Proc* 2018; **2017**: 310-319 [PMID: 29888089]

84   **Sadhu AR**, Serrano IA, Xu J, Nisar T, Lucier J, Pandya AR, Patham B. Continuous Glucose Monitoring in Critically Ill Patients With COVID-19: Results of an Emergent Pilot Study. *J Diabetes Sci Technol* 2020; **14**: 1065-1073 [PMID: 33063556 DOI: 10.1177/1932296820964264]

85   **Faulds ER**, Boutsicaris A, Sumner L, Jones L, McNett M, Smetana KS, May CC, Buschur E, Exline MC, Ringel MD, Dungan K. Use of Continuous Glucose Monitor in Critically Ill COVID-19 Patients Requiring Insulin Infusion: An Observational Study. *J Clin Endocrinol Metab* 2021; **106**: e4007-e4016 [PMID: 34100545 DOI: 10.1210/clinem/dgab409]

86   **Agarwal S**, Mathew J, Davis GM, Shephardson A, Levine A, Louard R, Urrutia A, Perez-Guzman C, Umpierrez GE, Peng L, Pasquel FJ. Continuous Glucose Monitoring in the Intensive Care Unit During the COVID-19 Pandemic. *Diabetes Care* 2021; **44**: 847-849 [PMID: 33361145 DOI: 10.2337/dc20-2219]

87   **Dexcom**.   Fact sheet for healthcare providers: use of Dexcom continuous glucose monitoring systems during the COVID-19 pandemic. Cited 1 February 2022. Available from: https://www.dexcom.com/hospitalfacts

88   **Galindo RJ**, Umpierrez GE, Rushakoff RJ, Basu A, Lohnes S, Nichols JH, Spanakis EK, Espinoza J, Palermo NE, Awadjie DG, Bak L, Buckingham B, Cook CB, Freckmann G, Heinemann L, Hovorka R, Mathioudakis N, Newman T, O'Neal DN, Rickert M, Sacks DB, Seley JJ, Wallia A, Shang T, Zhang JY, Han J, Klonoff DC. Continuous Glucose Monitors and Automated Insulin Dosing Systems in the Hospital Consensus Guideline. *J Diabetes Sci Technol* 2020; **14**: 1035-1064 [PMID: 32985262 DOI: 10.1177/1932296820954163]

89   **Hinoue T**, Yatabe T, Fujiwara H, Nishida O. Glucose control using an artificial pancreas in a severe COVID-19 patient on extracorporeal membrane oxygenation: a case report. *J Anesth* 2021; **35**: 586-590 [PMID: 34169361 DOI: 10.1007/s00540-021-02965-1]

Baishideng®

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com