# Artificial Intelligence in

# *Gastroenterology*

## Contents

# Contents

## ABOUT COVER

Editorial Board Member of *Artificial Intelligence in Gastroenterology*, Natalia V Baryshnikova, MD, PhD, Associate Professor, Research Scientist, Senior Lecturer, Department of Internal Diseases, First Pavlov State Medical University, St-Petersburg 195257, Russia. baryshnikova_nv@mail.ru

## 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

There is currently no indexing.

## RESPONSIBLE EDITORS FOR THIS ISSUE

**Production Editor:** *Rui-Rui Wu*; **Production Department Director:** *Xiang Li*; **Editorial Office Director:** *Jin-Lei Wang*.

# Current advancements in application of artificial intelligence in clinical decision-making by gastroenterologists in gastrointestinal bleeding

Hasan Maulahela, Nagita Gianty Annisa

**Hasan Maulahela,** Department of Internal Medicine, Gastroenterology Division, Faculty of Medicine University of Indonesia - Cipto Mangunkusumo General Central National Hospital, Jakarta 10430, Indonesia

**Nagita Gianty Annisa,** Faculty of Medicine, University of Indonesia, Jakarta 10430, Indonesia

**Corresponding author:** Hasan Maulahela, MD, Assistant Professor, Department of Internal Medicine, Gastroenterology Division, Faculty of Medicine University of Indonesia - Cipto Mangunkusumo General Central National Hospital, Jl. Pangeran Diponegoro No. 71 RW.5, Kenari, Kec. Senen, Jakarta 10430, Indonesia. hasan.maulahela@yahoo.com

## Abstract

Artificial Intelligence (AI) is a type of intelligence that comes from machines or computer systems that mimics human cognitive function. Recently, AI has been utilized in medicine and helped clinicians make clinical decisions. In gastroenterology, AI has assisted colon polyp detection, optical biopsy, and diagnosis of *Helicobacter pylori* infection. AI also has a broad role in the clinical prediction and management of gastrointestinal bleeding. Machine learning can determine the clinical risk of upper and lower gastrointestinal bleeding. AI can assist the management of gastrointestinal bleeding by identifying high-risk patients who might need urgent endoscopic treatment or blood transfusion, determining bleeding stigmata during endoscopy, and predicting recurrence of gastrointestinal bleeding. The present review will discuss the role of AI in the clinical prediction and management of gastrointestinal bleeding, primarily on how it could assist gastroenterologists in their clinical decision-making compared to conventional methods. This review will also discuss challenges in implementing AI in routine practice.

**Key Words:** Gastrointestinal bleeding; Artificial intelligence; Machine learning; Artificial neural networks; Clinical decision making

**Core Tip:** Gastrointestinal bleeding is a common problem in the emergency department. Quick and appropriate clinical decision is needed in the management of gastrointestinal bleeding. Artificial intelligence, namely machine learning and deep learning, can utilize electronic health record data to provide insights which might help clinicians, especially gastroenterologists, in the management of gastrointestinal bleeding. The present review will discuss the roles of artificial intelligence in clinical prediction and management of gastrointestinal bleeding, and compare them to conventional methods. This review will also discuss challenges in the implementation of artificial intelligence in routine practice.

## INTRODUCTION

Artificial intelligence (AI) simulates human intelligence processes and cognitive function using machines or computer systems. Several terminologies need to be understood before talking about AI. Machine learning (ML) is a technique of AI in which a computer or a system can learn to improve its function using experience and data without explicit instruction. There are several machine learning methods, for example, CNN (convolutional neural network), that can perform image analysis. ANN (artificial neural network) consists of a hidden-layered connection between input and output. Meanwhile, deep learning is a class of machine learning which extracts higher-level information progressively using multiple layers of neural networks[1]. AI has transformed information technology by making it possible to analyse large-scale data within a short time[2].

Recently, AI has been utilized in medicine. AI has a broad role in medicine, from guiding treatment decisions using electronic health record data to assisting in performing surgeries and intelligent prostheses for people with disabilities[3]. In gastroenterology, AI has assisted in diagnosing and treating gastrointestinal (GI) diseases. AI also has roles in small intestinal endoscopy and endoscopic ultrasound, especially in evaluating and diagnosing lesions[4].

This review aims to discuss the roles of AI in GI bleeding, especially in clinical decision-making for gastroenterologists. More specifically, this review will discuss the advancements in the application of AI in clinical prediction and management of upper and lower GI bleeding and its limitations and future challenges.

## ARTIFICIAL INTELLIGENCE IN CLINICAL PREDICTION OF UPPER GASTROINTESTINAL BLEEDING

Several scoring systems or risk models have been developed to predict the clinical risk of GI bleeding. In patients using antithrombotic medications, these risk models include HAS-BLED (hypertension, abnormal kidney and liver function, stroke, bleeding, labile international normalized ratio, elder age, and drug or alcohol use), ATRIA (anticoagulation and risk factors in atrial fibrillation), ORBIT (Outcomes Registry for Better Informed Treatment of Atrial Fibrillation), and HEMORR2HAGES (hepatic or kidney disease, ethanol abuse, malignancy, older age, reduced platelet count or function, rebleeding, hypertension, anemia, genetic factors, excessive fall risk, and stroke)[5-7]. Among these models, HAS-BLED has the best performance to predict major bleeding events[8].

Compared to the previous risk models, the prediction model using machine learning is hypothesized to have better performance since it can utilize more extensive and updated data sets. Herrin *et al*[9] tested three machine learning algorithms: Regularized Cox regression (RegCox), random survival forests, and extreme gradient boosting (XGBoost) on adult patients who were prescribed antithrombotic drugs (vitamin K antagonists, direct oral anticoagulants (DOACs), and/or thienopyridine antiplatelet agents) to predict the probability of GI bleeding at 6 and 12 mo. The data were obtained from medical and pharmacy claims data of 300000 patients. They also compared the performance of the machine learning algorithms to the HAS-BLED risk model.

In that study, all machine learning algorithms performed superiorly to HAS-BLED score in predicting GI bleeding at 6 and 12 mo. HAS-BLED score achieved an area under the curve (AUC) of 0.61 [95% confidence interval (CI): 0.59-0.62] for 6-mo GI bleeding risk and AUC of 0.60 (95%CI: 0.59-0.61) for 12-mo GI bleeding risk. Meanwhile, RegCox, the most superior algorithm from the three machine learning algorithms, had an AUC of 0.68 (95%CI: 0.66-0.70) for 6-mo GI bleeding risk and AUC of 0.67 (95%CI: 0.65-0.69) for 12-mo GI bleeding risk. HAS-BLED and the three machine learning algorithms obtained a

similar sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). However, all of them had an AUC less than 0.70, which is the conventional threshold for acceptable performance[9].

HAS-BLED score was derived to predict major bleeding events from patients treated with warfarin [10]. However, recently, antiplatelet agents and DOACs are more commonly used. Even though clinical extrapolation to calculate the risk of GI bleeding in patients taking antithrombotics is common, there are still concerns regarding the accuracy of HAS-BLED in predicting bleeding events in patients taking other types of anticoagulants or antiplatelets. Capodanno *et al*[11] found that HAS-BLED score could not predict major bleeding events in patients undergoing PCI (percutaneous coronary intervention) without artrial fibrillation who were discharged with dual antiplatelets. Although not specifically developed to predict GI bleeding events, several scoring systems have been developed for predicting bleeding events in patients taking dual antiplatelet therapy, such as CRUSADE, ACUITY, and PRECISE-DAPT. However, each scoring system has different accuracies in predicting short-term and long-term bleeding complications. For example, CRUSADE and ACUITY are better in predicting short-term complications, while PRECISE-DAPT is better in predicting long-term bleeding events[12].

Machine learning algorithms that utilize real-time data, such as RegCox, should better predict GI bleeding than the scoring systems mentioned above. Moreover, machine learning algorithms can provide time-to-event outcomes that can be used in the prediction of both short-term and long-term GI bleeding events. Herrin *et al*[9] used data sets from insurance claims and could not provide actual clinical values, which might contribute to low AUCs in their study. Data sets from electronic health record data that contain laboratory values and endoscopic reports might result in a better accuracy for clinical prediction of GI bleeding.

In patients presenting with upper GI bleeding, especially in the emergency department, it is important to stratify a patient's risk and predict mortality outcomes and the need for transfusion and other hemostatic interventions. Scoring systems such as the Glasgow-Blatchford score (GBS), Rockall score, and AIMS65 predict pre-endoscopic risk in patients with acute upper GI bleeding based on clinical, hemodynamic, and initial laboratory variables. Shung *et al*[13] conducted a systematic review that included 14 studies with 30 assessments of ML models. The median AUC for mortality, interventions, or rebleeding outcomes for ML models was 0.84. AUCs were higher in studies using ANNs than other models. They found that ML performed better than clinical risk scores for mortality in upper GI bleeding.

Recently, Shung *et al*[14] validated a machine learning model for upper GI bleeding that predicted composite outcomes of the need for hospital-based interventions (red blood cell transfusion, endoscopic hemostatic intervention, or surgery) and 30-d all-cause mortality. The chosen ML model was the XGBoost model. Different from previous studies, this study did not collect data from insurance records but through medical data that was directly entered by a nurse, physician, or medical student.

The ML model obtained an AUC of 0.91 (95%CI: 0.90-0.93) in the internal validation group, and an AUC of 0.90 (95%CI: 0.87-0.93) in the external validation group. The model performed better than GBS (AUC = 0.87, 95%CI: 0.84-0.91; *P* = 0.004), admission Rockall (AUC = 0.65, 95%CI: 0.60-0.71; *P* < 0.001), and AIMS65 (AUC = 0.64, 95%CI: 0.59-0.69; *P* < 0.001)[14].

ML models could perform better than scoring systems in risk stratification in patients with upper GI bleeding because they could extract patterns from raw data and increase accuracy with additional data and experience. Moreover, ML models could analyze more complex and heterogeneous data.

## ARTIFICIAL INTELLIGENCE IN CLINICAL PREDICTION OF LOWER INTESTINAL BLE-EDING

AI also has roles in the clinical prediction of lower intestinal bleeding. In 2017, Loftus *et al*[15] conducted a study that compared ANN and a regression-based model to predict the severity of lower GI bleeding and the need for surgical intervention.

Loftus *et al*[15] performed the analysis retrospectively on 147 adult patients who underwent endoscopy, angiography, or surgery for acute lower intestinal bleeding. The regression-based model used was the Strate prediction rule. The ANN for prediction of severe bleeding incorporated six variables present on admission: Systolic blood pressure; hemoglobin; outpatient prescription of aspirin 325 mg daily; Charlson comorbidity index; base deficit ≥ 5 mEq/L; and international normalized ratio ≥ 1.5. Meanwhile, the ANN for prediction of the need for surgery combined three predictors from severe bleeding ANN with two additional variables, hemoglobin nadir and the occurrence of a 20% decrease in haematocrit[15].

The Strate risk factors in the study correlated significantly with severe bleeding (*r* = 0.29, *P* < 0.001). However, the Strate model was less accurate in predicting severe lower intestinal bleeding than the ANN [area under the receiver operating characteristic curve (AUROC) 0.66 (95%CI: 0.57-0.75) *vs* 0.98 (95%CI: 0.95-1.00)]. The ANN for predicting the need for surgical intervention also had good performance with an AUROC of 0.95 (95%CI: 0.90-1.00). ANN could perform better than the regression-based model because this program could incorporate intricate associations among variables into an

algorithm, similar to nonlinear statistical processing[15].

Ayaru *et al*[16] analyzed non-endoscopic variables from patients with acute lower GI bleeding in the emergency department for internal and external validation of the gradient boosting (GB) model. GB is a supervised machine learning algorithm used in regression and classification tasks with multiple simple learning algorithms used jointly to obtain better predictive performance. Their study compared GB model with BLEED classification, Strate prediction rule, and conventional multiple logistic regression in predicting severe bleeding, the need for therapeutic intervention, and recurrent bleeding in patients with acute lower GI bleeding.

Ayaru *et al*[16] found that the GB model performed better than other scoring systems with an accuracy of 88% for recurrent bleeding and therapeutic intervention and 78% for the need for therapeutic intervention. Meanwhile, conventional multiple logistic regression had an accuracy of 74% in predicting recurrent bleeding and the need for therapeutic intervention and an accuracy of 62% in predicting severe bleeding. BLEED classification and Strate prediction rule also performed more poorly than the GB model.

In their study, the GB model could provide variables contributing to the risk of severe acute lower GI bleeding and the contribution percentage. The variables and their contribution are platelet count (13.4%), activated partial thromboplastin time (13.0%), haematocrit (12.4%), urea (10.9%), creatinine (9.7%), prothrombin time (8.9%), diastolic blood pressure (6.8%), heart rate (4.1%), systolic blood pressure (3.9%), and alcohol abuse (3.9%)[16].

Both studies by Loftus *et al*[15] and Ayaru *et al*[16] found that AI performed better than scoring systems in predicting lower GI bleeding. Even though they used different algorithms, ANN and GB model both could perform better than other regression-based models and scoring systems. Moreover, the algorithms could provide variables contributing to the risk of bleeding and the need for therapeutic intervention. However, both studies were limited by their retrospective design. More prospective studies need to be conducted to determine the accuracy of ML models in lower GI bleeding prediction. More studies, including different AI algorithms, also need to be conducted to determine the better algorithm for predicting GI bleeding.

## ARTIFICIAL INTELLIGENCE IN MANAGEMENT OF UPPER AND LOWER GASTRO-INTESTINAL BLEEDING

AI has a broad role in the management of GI bleeding, starting from patient's admission, during endoscopy, to patient's care post-endoscopy or surgery. In patient admission and during pre-endoscopy, AI, especially machine learning, can be used in the risk stratification of patients with GI bleeding. Machine learning can also be used to determine whether the patient needs urgent endoscopy, blood transfusion, or surgical intervention, or if the patient can be safely observed and discharged from the emergency room[17].

Early identification of patients with high-risk GI bleeding is important and can reduce mortality and morbidity. To identify low-risk patients, a GBS score of 0 or 1 can be used to determine whether the patient can be safely discharged from the emergency room (sensitivity 98.6%, specificity 34.6%). However, GBS and other scoring systems such as Rockall and AIMS65 still perform poorly in predicting high-risk patients needing endoscopic treatment or surgical intervention[18].

Shung *et al*[19] developed multiple natural language processing (NLP)-based approaches to identify patients with acute GI bleeding in the emergency room. They used electronic health record-based phenotyping algorithms and compared the performance with the Systematized Nomenclature of Medicine, a standard method to identify patients' conditions. They found that the NLP-based approach performed better than the Systematized Nomenclature of Medicine [PPV 85% (95%CI: 83%-87%) *vs* 69% (95%CI: 66%-72%); *P* < 0.001] in identifying patients with acute GI bleeding.

Seo *et al*[20] developed four machine learning algorithms to predict adverse events and hemodynamic instability in patients with initially stable non-variceal upper GI bleeding. The four machine learning algorithms were logistic regression with regularization, random forest classifier (RF), GB classifier, and voting classifier (VC). The adverse events analyzed included hypotension, mortality, and rebleeding within 7 d. The algorithms were compared with the standard scoring system GBS and Rockall scores. Among the machine learning algorithms, the RF model showed the best performance in predicting mortality (AUC: RF 0.917 *vs* GBS 0.710), while the VC model had the highest accuracies in predicting hypotension (AUC: VC 0.757 *vs* GBS 0.668) and rebleeding within 7 d (AUC: VC 0.733 *vs* GBS 0.694).

In the intensive care unit (ICU), Deshmukh *et al*[21] developed a machine learning model to calculate mortality risk in patients admitted with GI bleeding. They compared the model with the APACHE IVa risk score and found that the model performed better in classifying low-risk patients [AUC: 0.85 (95%CI: 0.80-0.90) *vs* 0.80 (95%CI: 0.73-0.86)]. The model achieved a sensitivity of 100% and specificity of 27%, compared with APACHE IVa risk score with a sensitivity of 100% and specificity of 4%.

Levi *et al*[22] also developed a machine learning algorithm to predict the need for blood transfusion in ICU patients with GI bleeding. Existing scoring systems such as GBS and Rockall score focus on predicting mortality and the need for intervention. They do not assist in determining the level of

monitoring needed for hospitalized patients. Moreover, these scoring systems were validated only for upper GI bleeding. Levi *et al*[22] trained the algorithm on different data sets: MIMIC-III (Medical Information Mart for Intensive Care-III); eICU-CRD (eICU Collaborative Research Database v.2.0); or both. All models performed well with an AUROC > 0.80. A similar study by Shung *et al*[23] also found that a long short-term memory model, a type of Recurrent Neural Network, performed better than a regression-based model (AUROC: 0.65 *vs* 0.56; *P* < 0.001) in determining high-risk GI bleeding patients requiring red blood cell transfusion in the ICU.

In patients with acute lower GI bleeding, Das *et al*[24] constructed ANN and multiple logistic regression models to predict the outcomes of intervention for control of hemorrhage, recurrent bleeding, and death. The models classify patients with lower GI bleeding as high-risk and low-risk patients. The study found that ANN was significantly better than BLEED (accuracy for predicting death 87% *vs* 21%; for recurrent bleeding 89% *vs* 41%; and for intervention 96% *vs* 46%) in internal validation. ANN was also better than multiple logistic regression models in predicting the three outcomes in the external validation (for death 97% *vs* 70%; for recurrent bleeding 93% *vs* 73%; and for intervention 94% *vs* 70%).

Shung *et al*[23], Seo *et al*[20], Deshmukh *et al*[21], and Das *et al*[24] showed that machine learning models could be used in risk stratification for patients with acute upper and lower GI bleeding. More advanced interventions, such as endoscopic or surgical intervention, could be considered in high-risk patients. Therefore, AI could help emergency physicians and gastroenterologists decide patients who might need urgent endoscopic or surgical intervention and help prepare the necessary interventions earlier. Meanwhile, Levi *et al*[22] showed that AI could help determine which patients need tighter monitoring. Many patients with GI bleeding admitted to ICU stop bleeding and do not require further intervention. In hospitals with limited ICU capacities, AI might help determine patients with GI bleeding who may or may not require ICU-level care.

All studies mentioned above used electronic health record data to train the models, making the results readily applicable for the hospital setting. These studies used different machine learning models. Interestingly, Seo *et al*[20] found that different models had different accuracies in determining the risk of different outcomes. Choosing the appropriate machine learning algorithm or model is essential to achieve the highest accuracy. However, there are still not many studies that compare the accuracies between different machine learning models.

During endoscopy, AI might help identify endoscopic characteristics of hemorrhage, such as determining the Forrest classification of peptic ulcer, which will help determine the management needed for the patient. Yen *et al*[25] compared the performance of deep learning with expert and novice endoscopists. They retrieved endoscopic still images of 1694 patients with peptic ulcer bleeding. Four deep learning models were pre-trained with ImageNet. In the end, the Mobile Net V2 model was chosen with the most optimum performance and compared with expert and novice endoscopists. For the 3-class categories, the sensitivity and specificity were 94.83% and 92.36%, respectively. Meanwhile, for the 4-class categories, the sensitivity and specificity were 95.40% and 92.70%, respectively. The deep learning model also had a higher interobserver agreement with expert endoscopists compared to novice endoscopists.

Gastric ulcer is a common medical condition, with a yearly incidence of more than 5 in 1000 adults. However, gastric ulcer also has a risk to develop into gastric cancer. The malignancy rate in endoscopically diagnosed gastric ulcers ranges from 2.4% to 21%. Therefore, early detection of malignant ulcers is important for further treatment and a better prognosis. Several studies have developed AI algorithms to differentiate between malignant and benign gastric ulcers. For example, Klang *et al*[26] developed a CNN model with an AUC of 0.91 (95%CI: 0.85-0.96) with a sensitivity of 92% and specificity of 75%. Similar studies were also conducted by Namikawa *et al*[27], Yoon *et al*[28], and Wu *et al*[29] using the CNN model to differentiate gastric ulcers and early gastric cancers with satisfying performances.

AI also aids in the diagnosis of *Helicobacter pylori* (*H. pylori*) infection. Itoh *et al*[30] developed a CNN model to diagnose *H. pylori* infection, using 149 training images and 30 test images from upper GI endoscopy images. The sensitivity of CNN for detection of *H. pylori* infection was 86.7%, while the specificity was 86.7%, with an AUC of 0.956. Mohan *et al*[31] conducted a systematic review consisting of five studies using CNN for detection of *H. pylori* infection. Images used for the diagnosis were from a combination of white-light, blue laser imaging, and linked color imaging. The pooled accuracy of AI for detecting *H. pylori* infection was 87.1% (95%CI: 81.8-91.1) with a sensitivity of 86.3% and specificity of 87.1%. Meanwhile, endoscopists achieved an accuracy of 82.9% (95%CI: 76.7-87.7), with a sensitivity of 79.6% and specificity of 83.8%.

AI also aids the detection of small bowel bleeding using wireless capsule endoscopy. Le Berre *et al*[32] reviewed 12 studies using various AI classifiers such as color spectrum transformation, MLP (multilayer perceptron network), SVM (support vector machine, a type of machine learning model), joint diagonalization, PCA (principal component analysis), and CNN. The sensitivity from various studies ranged from 87.8% to 100%, while the specificity ranged from 85.8% to 99.9%. The highest accuracy of 99.6% was obtained in a study by Xiao *et al*[33] using deep CNN and 10000 images (8200 training and 1800 test images).

After management in the hospital, AI can be used in identifying the risk of recurrent bleeding in patients with GI bleeding. Wong *et al*[34] developed a machine learning model to predict recurrent bleeding. The model was built based on six parameters (age, baseline haemoglobin, presence of gastric

ulcer, GI diseases, malignancies, and infections). The model identified patients with recurrent ulcer bleeding within 1 year with an AUROC of 0.775 and overall accuracy of 84.3%.

## CONCLUSIONS AND FUTURE CHALLENGES

As discussed above, AI, especially machine learning and deep learning, has broad roles in clinical prediction and management of GI bleeding by utilizing data that could help clinicians in their decision-making. Even though AI can utilize a large set of electronic health record data, they might not be able to utilize several important data such as patient's behavior or endoscopic images, which might not be stated in electronic health records or stored in different servers[35].

Since machine learning outcomes depend on the data set, the outcome might not be replicable in other centers. For example, factors that influence the risk of GI bleeding might be different in different centers with different data sets using the same AI algorithm. The data set used for the algorithm training could influence the algorithm's performance. Hence, it is crucial to have a high-quality data set that is well-integrated with the AI system before establishing an AI system[35]. Once established, the integrated electronic health record and AI algorithm system could be copied to be used by different centers.

Adopting AI also has several barriers, especially in developing countries, such as insufficient technological infrastructure and difficulty integrating AI in the routine workflow. Adequate data warehouses, secure analytic platforms, and informatics and machine learning experts must be employed. Some clinicians might be reluctant to substitute clinical judgment with computational analysis. It is important to ensure the healthcare providers' trust before implementing the tool. A contingency plan concerning patients' safety should be established if the algorithm makes an error. Legal framework regarding clinical decision-making by AI and its responsibility is currently unavailable[35,36].

An issue related to the safety of AI is the "black-box" algorithms. Black box AI is any AI system whose inputs and operations are not visible to its users. Many machine learning models are considered a black box, and it is difficult to understand how the algorithm arrived at its conclusion, even for those who trained it. Clinicians who use the algorithm might not realize whether a clinical decision suggested by an AI model is wrong because they do not know how the model arrived at the conclusion. Moreover, AI is still prone to biases. A diagnosis or prognostic algorithm trained with data from mostly Caucasian patients, for example, might not be as accurate for Black or Asian patients. An algorithm developed in high resource settings might not recommend accurate or fair treatment in settings with more limited resources[37].

The black box algorithms also raise legal concerns. It is still unclear if it could be considered medical malpractice when a clinician gives a wrong treatment recommended by a black-box algorithm because they could not review the basis of recommendation. Lawsuits might also be brought to the hospitals that implement the AI algorithm or even to the technology companies that develop the algorithm[37]. Currently, it is recommended to use AI to support a clinical decision that has been already made instead of using AI to create a new clinical decision.

Another ethical concern regarding the use of AI in medicine is patients' privacy. Personal health condition is one of the most legally protected forms of data. Meanwhile, AI is usually provided by start-ups or private technological companies. Previous cases of data breaches or technological companies monetizing their customers' personal information are concerns that need to be addressed. Companies need to provide technical safeguards to maintain data privacy to prevent breaches. Patients should be informed of data uses, and patients should give their consent before their data is used[38].

To prevent misuse of patients' medical information, legal frameworks need to be updated to suit the rapid improvement of AI. Health Insurance Portability and Accountability Act (HIPAA) privacy rule is the United States national standard for protecting individual medical records and other individual health information. An example of a loophole in the regulation is if a genetic company sells their data to pharmaceutical or insurance firms, the HIPAA privacy rule could not apply because DNA information is not legally counted as healthcare[39]. Therefore, regulations concerning patients' privacy and safety need to be revisited and updated to catch up with the improvement of technology. Strict legal penalties should be implemented for those who break the regulations.

## FOOTNOTES

# REFERENCES

1    **Yang YJ**, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019; **25**: 1666-1683 [PMID: 31011253 DOI: 10.3748/wjg.v25.i14.1666]

2    **Parasher G**, Wong M, Rawat M. Evolving role of artificial intelligence in gastrointestinal endoscopy. *World J Gastroenterol* 2020; **26**: 7287-7298 [PMID: 33362384 DOI: 10.3748/wjg.v26.i46.7287]

3    **Amisha**, Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. *J Family Med Prim Care* 2019; **8**: 2328-2331 [PMID: 31463251 DOI: 10.4103/jfmpc.jfmpc_440_19]

4    **Alagappan M**, Brown JRG, Mori Y, Berzin TM. Artificial intelligence in gastrointestinal endoscopy: The future is almost here. *World J Gastrointest Endosc* 2018; **10**: 239-249 [PMID: 30364792 DOI: 10.4253/wjge.v10.i10.239]

5    **Pisters R**, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. *Chest* 2010; **138**: 1093-1100 [PMID: 20299623 DOI: 10.1378/chest.10-0134]

6    **Apostolakis S**, Lane DA, Guo Y, Buller H, Lip GY. Performance of the HEMORR(2)HAGES, ATRIA, and HAS-BLED bleeding risk-prediction scores in patients with atrial fibrillation undergoing anticoagulation: the AMADEUS (evaluating the use of SR34006 compared to warfarin or acenocoumarol in patients with atrial fibrillation) study. *J Am Coll Cardiol* 2012; **60**: 861-867 [PMID: 22858389 DOI: 10.1016/j.jacc.2012.06.019]

7    **Senoo K**, Proietti M, Lane DA, Lip GY. Evaluation of the HAS-BLED, ATRIA, and ORBIT Bleeding Risk Scores in Patients with Atrial Fibrillation Taking Warfarin. *Am J Med* 2016; **129**: 600-607 [PMID: 26482233 DOI: 10.1016/j.amjmed.2015.10.001]

8    **Roldán V**, Marín F, Fernández H, Manzano-Fernandez S, Gallego P, Valdés M, Vicente V, Lip GYH. Predictive value of the HAS-BLED and ATRIA bleeding scores for the risk of serious bleeding in a "real-world" population with atrial fibrillation receiving anticoagulant therapy. *Chest* 2013; **143**: 179-184 [PMID: 22722228 DOI: 10.1378/chest.12-0608]

9    **Herrin J**, Abraham NS, Yao X, Noseworthy PA, Inselman J, Shah ND, Ngufor C. Comparative Effectiveness of Machine Learning Approaches for Predicting Gastrointestinal Bleeds in Patients Receiving Antithrombotic Treatment. *JAMA Netw Open* 2021; **4**: e2110703 [PMID: 34019087 DOI: 10.1001/jamanetworkopen.2021.10703]

10   **Faye A**, Hung K, Cheng K, Parikh N, Knotts R, Freedberg D, Lebwohl B. 626 HAS-BLED Scores Underestimate Gastrointestinal Bleeding Risk Among Those With *H. pylori*. *Am J Gastroenterol* 2019; **114**: S364-S364. [DOI: 10.14309/01.ajg.0000592040.17612.c1]

11   **Capodanno D**, Rossini R, Musumeci G, Lettieri C, Senni M, Valsecchi O, Angiolillo DJ, Lip GY. Predictive accuracy of CHA2DS2-VASc and HAS-BLED scores in patients without atrial fibrillation undergoing percutaneous coronary intervention and discharged on dual antiplatelet therapy. *Int J Cardiol* 2015; **199**: 319-325 [PMID: 26241637 DOI: 10.1016/j.ijcard.2015.07.064]

12   **Choi SY**, Kim MH, Cho YR, Sung Park J, Min Lee K, Park TH, Yun SC. Performance of PRECISE-DAPT Score for Predicting Bleeding Complication During Dual Antiplatelet Therapy. *Circ Cardiovasc Interv* 2018; **11**: e006837 [PMID: 30545256 DOI: 10.1161/CIRCINTERVENTIONS.118.006837]

13   **Shung D**, Simonov M, Gentry M, Au B, Laine L. Machine Learning to Predict Outcomes in Patients with Acute Gastrointestinal Bleeding: A Systematic Review. *Dig Dis Sci* 2019; **64**: 2078-2087 [PMID: 31055722 DOI: 10.1007/s10620-019-05645-z]

14   **Shung DL**, Au B, Taylor RA, Tay JK, Laursen SB, Stanley AJ, Dalton HR, Ngu J, Schultz M, Laine L. Validation of a Machine Learning Model That Outperforms Clinical Risk Scoring Systems for Upper Gastrointestinal Bleeding. *Gastroenterology* 2020; **158**: 160-167 [PMID: 31562847 DOI: 10.1053/j.gastro.2019.09.009]

15   **Loftus TJ**, Brakenridge SC, Croft CA, Smith RS, Efron PA, Moore FA, Mohr AM, Jordan JR. Neural network prediction of severe lower intestinal bleeding and the need for surgical intervention. *J Surg Res* 2017; **212**: 42-47 [PMID: 28550920 DOI: 10.1016/j.jss.2016.12.032]

16   **Ayaru L**, Ypsilantis PP, Nanapragasam A, Choi RC, Thillanathan A, Min-Ho L, Montana G. Prediction of Outcome in Acute Lower Gastrointestinal Bleeding Using Gradient Boosting. *PLoS One* 2015; **10**: e0132485 [PMID: 26172121 DOI: 10.1371/journal.pone.0132485]

17   **Yen HH**, Wu PY, Chen MF, Lin WC, Tsai CL, Lin KP. Current Status and Future Perspective of Artificial Intelligence in the Management of Peptic Ulcer Bleeding: A Review of Recent Literature. *J Clin Med* 2021; **10** [PMID: 34441823 DOI: 10.3390/jcm10163527]

18   **Stanley AJ**, Laine L, Dalton HR, Ngu JH, Schultz M, Abazi R, Zakko L, Thornton S, Wilkinson K, Khor CJ, Murray IA, Laursen SB; International Gastrointestinal Bleeding Consortium. Comparison of risk scoring systems for patients presenting with upper gastrointestinal bleeding: international multicentre prospective study. *BMJ* 2017; **356**: i6432 [PMID:

28053181 DOI: 10.1136/bmj.i6432]

19    **Shung D**, Tsay C, Laine L, Chang D, Li F, Thomas P, Partridge C, Simonov M, Hsiao A, Tay JK, Taylor A. Early identification of patients with acute gastrointestinal bleeding using natural language processing and decision rules. *J Gastroenterol Hepatol* 2021; **36**: 1590-1597 [PMID: 33105045 DOI: 10.1111/jgh.15313]

20    **Seo DW**, Yi H, Park B, Kim YJ, Jung DH, Woo I, Sohn CH, Ko BS, Kim N, Kim WY. Prediction of Adverse Events in Stable Non-Variceal Gastrointestinal Bleeding Using Machine Learning. *J Clin Med* 2020; **9** [PMID: 32796647 DOI: 10.3390/jcm9082603]

21    **Deshmukh F**, Merchant SS. Explainable Machine Learning Model for Predicting GI Bleed Mortality in the Intensive Care Unit. *Am J Gastroenterol* 2020; **115**: 1657-1668 [PMID: 32341266 DOI: 10.14309/ajg.0000000000000632]

22    **Levi R**, Carli F, Arévalo AR, Altinel Y, Stein DJ, Naldini MM, Grassi F, Zanoni A, Finkelstein S, Vieira SM, Sousa J, Barbieri R, Celi LA. Artificial intelligence-based prediction of transfusion in the intensive care unit in patients with gastrointestinal bleeding. *BMJ Health Care Inform* 2021; **28** [PMID: 33455913 DOI: 10.1136/bmjhci-2020-100245]

23    **Shung D**, Huang J, Castro E, Tay JK, Simonov M, Laine L, Batra R, Krishnaswamy S. Neural network predicts need for red blood cell transfusion for patients with acute gastrointestinal bleeding admitted to the intensive care unit. *Sci Rep* 2021; **11**: 8827 [PMID: 33893364 DOI: 10.1038/s41598-021-88226-3]

24    **Das A**, Ben-Menachem T, Cooper GS, Chak A, Sivak MV Jr, Gonet JA, Wong RC. Prediction of outcome in acute lower-gastrointestinal haemorrhage based on an artificial neural network: internal and external validation of a predictive model. *Lancet* 2003; **362**: 1261-1266 [PMID: 14575969 DOI: 10.1016/S0140-6736(03)14568-0]

25    **Yen HH**, Wu PY, Chen MF, Lin WC, Tsai CL, Lin KP. Current Status and Future Perspective of Artificial Intelligence in the Management of Peptic Ulcer Bleeding: A Review of Recent Literature. *J Clin Med* 2021; **10** [PMID: 34441823 DOI: 10.3390/jcm10163527]

26    **Klang E**, Barash Y, Levartovsky A, Barkin Lederer N, Lahat A. Differentiation Between Malignant and Benign Endoscopic Images of Gastric Ulcers Using Deep Learning. *Clin Exp Gastroenterol* 2021; **14**: 155-162 [PMID: 33981151 DOI: 10.2147/CEG.S292857]

27    **Namikawa K**, Hirasawa T, Nakano K, Ikenoyama Y, Ishioka M, Shiroma S, Tokai Y, Yoshimizu S, Horiuchi Y, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Artificial intelligence-based diagnostic system classifying gastric cancers and ulcers: comparison between the original and newly developed systems. *Endoscopy* 2020; **52**: 1077-1083 [PMID: 32503056 DOI: 10.1055/a-1194-8771]

28    **Yoon HJ**, Kim S, Kim JH, Keum JS, Oh SI, Jo J, Chun J, Youn YH, Park H, Kwon IG, Choi SH, Noh SH. A Lesion-Based Convolutional Neural Network Improves Endoscopic Detection and Depth Prediction of Early Gastric Cancer. *J Clin Med* 2019; **8** [PMID: 31454949 DOI: 10.3390/jcm8091310]

29    **Wu L**, Wang J, He X, Zhu Y, Jiang X, Chen Y, Wang Y, Huang L, Shang R, Dong Z, Chen B, Tao X, Wu Q, Yu H. Deep learning system compared with expert endoscopists in predicting early gastric cancer and its invasion depth and differentiation status (with videos). *Gastrointest Endosc* 2022; **95**: 92-104.e3 [PMID: 34245752 DOI: 10.1016/j.gie.2021.06.033]

30    **Itoh T**, Kawahira H, Nakashima H, Yata N. Deep learning analyzes Helicobacter pylori infection by upper gastrointestinal endoscopy images. *Endosc Int Open* 2018; **6**: E139-E144 [PMID: 29399610 DOI: 10.1055/s-0043-120830]

31    **Mohan BP**, Khan SR, Kassab LL, Ponnada S, Mohy-Ud-Din N, Chandan S, Dulai PS, Kochhar GS. Convolutional neural networks in the computer-aided diagnosis of *Helicobacter pylori* infection and non-causal comparison to physician endoscopists: a systematic review with meta-analysis. *Ann Gastroenterol* 2021; **34**: 20-25 [PMID: 33414617 DOI: 10.20524/aog.2020.0542]

32    **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]

33    **Xiao Jia**, Meng MQ. A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images. *Annu Int Conf IEEE Eng Med Biol Soc* 2016; **2016**: 639-642 [PMID: 28268409 DOI: 10.1109/EMBC.2016.7590783]

34    **Wong GL**, Ma AJ, Deng H, Ching JY, Wong VW, Tse YK, Yip TC, Lau LH, Liu HH, Leung CM, Tsang SW, Chan CW, Lau JY, Yuen PC, Chan FK. Machine learning model to predict recurrent ulcer bleeding in patients with history of idiopathic gastroduodenal ulcer bleeding. *Aliment Pharmacol Ther* 2019; **49**: 912-918 [PMID: 30761584 DOI: 10.1111/apt.15145]

35    **Sheikhtaheri A**, Sadoughi F, Hashemi Dehaghi Z. Developing and using expert systems and neural networks in medicine: a review on benefits and challenges. *J Med Syst* 2014; **38**: 110 [PMID: 25027017 DOI: 10.1007/s10916-014-0110-5]

36    **Shung DL**. Advancing care for acute gastrointestinal bleeding using artificial intelligence. *J Gastroenterol Hepatol* 2021; **36**: 273-278 [PMID: 33624892 DOI: 10.1111/jgh.15372]

37    **Gerke S**, Minssen T, Cohen G.  Ethical and legal challenges of artificial intelligence-driven healthcare. In: Artificial Intelligence in Healthcare [Internet]. Elsevier; 2020 [cited 2021 Dec 11]. 295–336. Available from: https://Linkinghub.elsevier.com/retrieve/pii/B9780128184387000125

38    **Murdoch B**. Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC Med Ethics* 2021; **22**: 122 [PMID: 34525993 DOI: 10.1186/s12910-021-00687-3]

39    **Ahuja AS**. The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ* 2019; **7**: e7702 [PMID: 31592346 DOI: 10.7717/peerj.7702]

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