Artificial Intelligence in Gastroenterology

Artif Intell Gastroenterol 2022 February 28; 3(1): 1-27

Published by Baishideng Publishing Group Inc
EDITORIAL

1 Gastrointestinal disorders in children with autism: Could artificial intelligence help?
Al-Biltagi M, Saeed NK, Qaraghuli S

MINIREVIEWS

13 Current advancements in application of artificial intelligence in clinical decision-making by gastroenterologists in gastrointestinal bleeding
Maulahela H, Annisa NG

21 Artificial intelligence and human liver allocation: Potential benefits and ethical implications
Mucenic M, de Mello Brandão AB, Marroni CA
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.

NAME OF JOURNAL
Artificial Intelligence in Gastroenterology

ISSN
ISSN 2644-3236 (online)

LAUNCH DATE
July 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Rajvinder Singh, Ferruccio Bonino

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2644-3236/editorialboard.htm

PUBLICATION DATE
February 28, 2022

COPYRIGHT
© 2022 Baishideng Publishing Group Inc

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com
Gastrointestinal disorders in children with autism: Could artificial intelligence help?

Mohammed Al-Biltagi, Nermin Kamal Saeed, Samara Qaraghuli

Abstract

Autism is one of the pervasive neurodevelopmental disorders usually associated with many medical comorbidities. Gastrointestinal (GI) disorders are pervasive in children, with a 46%-84% prevalence rate. Children with Autism have an increased frequency of diarrhea, nausea and/or vomiting, gastroesophageal reflux and/or disease, abdominal pain, chronic flatulence due to various factors as food allergies, gastrointestinal dysmotility, irritable bowel syndrome (IBS), and inflammatory bowel diseases (IBD). These GI disorders have a significant negative impact on both the child and his/her family. Artificial intelligence (AI) could help diagnose and manage Autism by improving children's communication, social, and emotional skills for a long time. AI is an effective method to enhance early detection of GI disorders, including GI bleeding, gastroesophageal reflux disease, Coeliac disease, food allergies, IBS, IBD, and rectal polyps. AI can also help personalize the diet for children with Autism by microbiome modification. It can help to provide modified gluten without initiating an immune response. However, AI has many obstacles in treating digestive diseases, especially in children with Autism. We need to do more studies and adopt specific algorithms for children with Autism. In this article, we will highlight the role of AI in helping children with gastrointestinal disorders, with particular emphasis on children.
with Autism.

**Key Words:** Autism; Gastrointestinal disorders; Artificial Intelligence; Children

©The Author(s) 2022. Published by Baishideng Publishing Group Inc. All rights reserved.

**Core Tip:** Autism is a prevalent childhood neurodevelopmental condition. Gastrointestinal (GI) disorders are pervasive in children, with a 46%-84% prevalence rate. The presence of GI can negatively impair children's management and education. Artificial intelligence (AI) could help diagnose and manage autism by improving children’s communication, social, and emotional skills for a long time. AI is an effective method to enhance early detection and management of GI disorders, including GI bleeding, gastroesophageal reflux disease, Coeliac disease, food allergies, irritable bowel syndrome, inflammatory bowel diseases, and rectal polyps. However, we still have some obstacles to increasing the benefit of AI in medicine, particularly in children with autism.

**Citation:** Al-Biltagi M, Saeed NK, Qaraghuli S. Gastrointestinal disorders in children with autism: Could artificial intelligence help? *Artif Intell Gastroenterol* 2022; 3(1): 1-12

**URL:** https://www.wjgnet.com/2644-3236/full/v3/i1/1.htm

**DOI:** https://dx.doi.org/10.35712/aig.v3.i1.1

**INTRODUCTION**

Since its description for the first time by Leo Kanner in 1943, the rate of autism has been on the rise and steadily increasing[1]. Autism is a neurodevelopmental condition. Autism, Asperger’s disorder, pervasive developmental disorder—not otherwise specified, form the autism spectrum disorders (ASD). At the same time, autism spectrum disorders, together with Rett’s disorder, childhood disintegrative disorder, and the overactive disorder accompanied with mental retardation and stereotyped movements, are a part of the pervasive developmental disorders[2]. The prevalence of autism varies from one country to another depending on the racial differences and the diagnostic facilities available, with an average of 1% worldwide. The autism incidence in the United States of America may reach up to 1/110, increasing to 1/64 in the United Kingdom[3]. In other parts of the world, the prevalence of autism may be underestimated. For example, in Bahrain, the prevalence of autism is 1/1000, with possible underestimation because of missed diagnosis and no official recording in some cases. Autism is also 4-5 times more common in boys than girls. Autism shows a wide range of prevalence according to the race, being more common in non-Hispanic white children, less in Hispanic and African American/black children, with wide variability in Asian/Pacific Residents[4].

The genesis of autism is still unclear. Nevertheless, we can assert that autism development is due to the complex interaction of several genetic, biological, advanced parental age, environmental, immunological, and psychosocial factors[5]. Recently, genetic studies discovered a wide variety of genetic mutations in most patients. These mutations do not necessarily follow the same pattern with a wide range of variability. However, these mutations can ultimately induce brain changes and inflammation[6]. This neuroinflammation can also occur in utero through defective placenta augmented by the immaturity of the blood-brain barrier of the fetus and the newly delivered baby. This neuroinflammation can be triggered either as a part of the maternal immune response to infection during pregnancy, premature delivery, as a part of postnatal encephalitis, or exposure to a toxic environment[1]. We still need to have more knowledge to understand the different causes and their effects on patients with autism.

The clinical presentation of autism is heterogeneous, formed mainly from a constellation of social, cognitive, motor, and perceptual symptoms, which usually appear before three years of age. Children with autism have a diverse range of behaviors, communications, interactions, and learning ways from most other children. The abnormal social communication and interaction skills are manifested by poor eye contact, a stern facial expression such as happiness or sadness, lack of interest with others, and lack of interest in playing or interacting with others. They also have restricted interests manifested by playing with the same toys the same way every time, getting upset with changing routine or minor changes, and focusing on certain parts of the toys or the body with obsessive interest. Additionally, they have repetitive or stereotyped behaviors such as constantly repeating words or phrases (i.e., echolalia), flapping hands, body rocking, or spinning self in circles). They also suffer from delayed language, movement, and cognitive or learning skills[7,8].
MEDICAL COMORBIDITIES

Besides the classic manifestations of autism, the affected patients may suffer the presence of many other medical comorbidities that are more common in people with ASD than in the general population. The presence of these comorbidities is one of the reasons for the significant increase of early mortality in patients with ASD, with death rates 3-10 times higher than the general population. These comorbidities may increase the risk of death in patients with autism and could affect their quality of life, impair proper diagnosis, interfere with their compromised learning capacity, and impair their ability to retain the acquired learning skills. Early recognition of these comorbidities helps improve the quality of life for both children and their families[9]. These comorbidities may include but are not limited to genetic, inborn errors of metabolism, congenital anomalies of the nervous system, neurologic disorders such as epilepsy and neuroinflammation, gastrointestinal (GI) disorders, and allergic disorders[10]. However, diagnosis of these comorbidities is not easily accessible due to communication impairments, occasional ambiguity of the symptoms, changes of the symptoms over time, and mimicking some of the classic symptoms of autism. A lack of available diagnostic instruments to screen these disorders further augments these difficulties[11].

GASTROINTESTINAL DISORDERS IN CHILDREN WITH AUTISM

Children with autism have a high prevalence of GI disorders occurring in 46%-84% of them. The interaction between autism and gastrointestinal disorders is shown in Figure 1. Unfortunately, many of these children cannot effectively communicate their symptoms or discomfort to their doctors. Chronic constipation occurs in about 50%. They have a restricted diet with low fibers, abnormal bowel training, increased intestinal transit time, and a high incidence of hypothyroidism, increasing the frequency of constipation. Diarrhea is three times more common in children with autism than the control due to increased prevalence of food sensitivities, gut dysbiosis, immune dysfunction, and the increased infection rate due to increased incidence of pica and abnormal child behavior[12,13]. They have also increased frequency of nausea and/or vomiting, gastroesophageal reflux and/or disease, and chronic flatulence due to various factors such as food allergies gastrointestinal dysmotility[14]. Abdominal pain is also frequent in children with autism which results from simple, functional disorders such as irritable bowel syndrome, or organic causes such as food allergies, food intolerance, parasitic infestations due to pica, colitis, ulcers, or inflammatory bowel diseases[15].

Food allergies occur in about one-quarter of children with autism compared to 5%-8% in the general Pediatric population[9,16]. The link between autism and Celiac diseases (CD) is debatable. However, some high-quality studies proved this link even in the absence of GI symptoms[17]. Given that children with autism are more prone to suffer from atopy & food allergies, possible non-coeliac gluten sensitivity (NCGS) or wheat sensitivity in those children needs to be considered, especially when irritable bowel syndrome symptoms are present[18]. Physicians should consider the possibility of NCGS in some patients with ASD, especially those presenting with atopic diseases, migraines, and mood and anxiety disorders. Therefore, investigating CD and non-coeliac gluten sensitivity even in the absence of typical GI symptoms could yield good results for children with autism[19].

Children with ASD are more likely to have various feeding disorders; behavioral, sensory-based, or medically related feeding problems. The behavioral feeding disorders may include aversive eating behaviors (such as food refusal, frequent choking or gagging, the expulsion of the food without a medical reason), and frequent Pica habits. The sensory-based feeding problems include restrictive or selective eating and textual refusal of specific foods, usually involving larger textures. The medically-related feeding disorder may affect oesophageal and swallowing disorders and motor delays[20]. Almost two-thirds of children with autism eat less than 20 types of foods and accept fewer foods from the primary food groups than typically developing children[21]. This high prevalence of feeding problems in ASD may be related to their propensity to concentrate on details, their fear of novelty, their way of perseveration, and impulsivity. The associated sensory impairments and the deficits in social compliance of children with autism augment their feeding disorders. These feeding behaviors could also be aggravated by specific biological food intolerance and parental anxiety, reinforcing negative feeding patterns[22]. These feeding disorders have tremendous effects on both children and their families. They may increase the risk of child abuse and the occurrence of specific nutritional deficiencies, but Weight and height are usually not affected. They also increase parental anxiety and stress, ending with child abuse[23].

IMPORTANCE AND DIFFICULTIES IN DIAGNOSING GI DISORDERS IN CHILDREN WITH AUTISM

It is essential to check for the presence of gastrointestinal disorders in children with autism, as they can
cause deterioration of autistic behaviors. For example, Abdominal pain related to especially reflux esophagitis and disaccharide malabsorption can cause irritation and discomfort to children with autism, which may contribute to the aggravation of their behavioral problems. It also could interfere with their learning abilities[24]. Meanwhile, gastrointestinal pain can cause behaviors that might be misdiagnosed as a behavior problem instead of a medical issue. For example, posturing, self-Injury, and/or outbursts without apparent cause could result from gastroesophageal reflux or esophagitis. The symptoms of GI dysfunction could induce sleep disturbances, which further aggravate the autistic manifestations[25].

Primary lactase deficiency that does not cause intestinal inflammation or injury is common in children with autism and may contribute to abdominal discomfort, pain, and observed aberrant behavior. Clinicians should screen for constipation and diarrhea or underwear soiling in children with autism who have prominent rigid–compulsive symptoms. If the GI disorder is recognized and medical treatment is effective, the behavioral problem may improve. When abdominal pain or discomfort is not alleviated, failure of psychotropic medications is more likely to occur. At the same time, these medications may even aggravate the problem if they have adverse gastrointestinal effects[9,26].

There is much evidence that modulation of the gut microbiota may be a manageable strategy for developing innovative therapies for complex CNS disorders, including autism[27]. The strong positive correlation of the gastrointestinal symptoms with the severity of autism indicates that children more severely affected by autism are likely to have severe gastrointestinal symptoms. Healthcare professionals should consider the possibility of gastrointestinal dysfunction in children with ASD, especially in those presenting with strange posturing or movements, sleep disorders, food intolerances, and aggressive or self-injurious behaviors[28]. The symptoms of GI dysfunction are associated with sleep disorders and food intolerance. Thus, it is essential to consider such association when evaluating and treating these comorbidities.

WHY IS IT DIFFICULT TO DIAGNOSE GI DISORDERS IN CHILDREN WITH AUTISM?

It is not always easy to detect GI manifestations in ASD. Children with autism have impaired communication skills, and many of them are nonverbal and cannot adequately express their pain, discomfort, or complaint through speech. Even those who can communicate verbally cannot adequately describe their symptoms. In addition, the symptoms of GI disorders may be missed as one of the classic symptoms and behavior commonly observed in children with autism. For example, toe-walking may be one of the typical stereotyped motor manifestations of autism to reduce feet overstimulation. It could also occur due to abdominal pain or loaded rectum or bladder. At the same time, GI disorders may present in atypical ways[29]. For example, suppose the child has abdominal pain or discomfort. In that case, he/she may touch his/her abdomen in a stereotyped way so that it can be easily missed with other stereotyped behaviors.

Moreover, GI disorders may present with non-GI manifestations. For example, sleep problems could be the manifestations of chronic GI disorders. It can be missed as being attributed to autism. The rate of sleep disorders increases from 30% of children with autism without GI disorders to reach 50% in the presence of GI disorders[30]. Children with autism may have hypersensitivity to various stimuli. On the other hand, they could occasionally have pain hyposensitivity with a high-pain threshold, affecting their
symptoms[31]. Unfortunately, no clinical practice guidelines exist to diagnose the presence of GI disorders in patients with ASD.

ROLE OF ARTIFICIAL INTELLIGENCE IN CHILDREN WITH AUTISM

Artificial intelligence (AI) enables a computer or computer-operated robot to perform tasks that humans usually do because they require human intelligence and judgment. The extensive application of artificial intelligence in various areas of life, including health, has begun to bear fruit. Whether we acknowledge it or not, artificial intelligence is inevitable and has a significant role in almost every aspect of our lives. The most important feature of AI is its ability to learn from its interaction, with the interaction-learning-interaction cycle. So, through pre-programmed flexible adaptation, AI can accurately interpret supplied external data, use these data to learn, and reuse the achieved learning to reach specific goals and duties. Machine learning is a part of AI and computer science that focuses on using data and algorithms to imitate how humans learn and gradually improve accuracy (Figure 2).

Deep learning, together with supervised and unsupervised learning, is a sub-class of machine learning that combines certain approaches that use specific algorithms to process and interpret data quicker, simpler, and more precisely[32]. In supervised learning, AI uses a computer algorithm to analyze predefined data to train and learn and then accurately names the new, hidden data. In unsupervised learning, the computer learns from massive, unlabelled led data and recognizes similarities and commonalities. Personalized medicine is an example of unsupervised medicine. The computers analyze the medical history, the result of the neck ultrasound or other radiology procedure, and the laboratory results for a patient with thyroid cancer to provide new perceptions for the treatment and the prognosis[33]. Medical sciences have greatly benefited from artificial intelligence, whether in diagnosing diseases, inventing appropriate medicines and treatments, or improving communication between doctors and patients[34]. There is much promise that AI will help improve healthcare services in many ways, including patient diagnosis, patient outcome, and drug invention, and assist the physician assistant and provide a better and more patient-tailored experience. This hope is driven by some of the emerging successful AI applications in healthcare[35].

It was a dream that AI could help diagnose and manage autism by improving children’s communication, social, and emotional skills for a long time. However, this dream starts to convert into reality despite not being the norm yet. Diagnosis of autism is subjective. Consequently, it becomes a real challenge in many situations. Parents and physicians may miss children with mild symptoms, while the more severe cases can simulate many other developmental disorders. Diagnosis of autism can be achieved using machine learning to provide a rapid, simple, and easy technique to provide for autism early diagnosis[36]. Machine learning can also help in improving the efficacy of behavioral health screening. The addition of machine learning techniques to complement the conventional methods in diagnosing autism helps fasten the diagnosis and reduce its cost[37]. An example of machine learning recognizes abnormal behavior using video monitor and artificial intelligence analysis of the body movement and behavior in children to detect early children with autism. Alcañiz Raya et al.[38] used machine-learning techniques to detect stereotyped and repetitive behaviors biomarkers, characteristics for autism. They used a depth sensor camera to track the body movements of the examined children. Consecutively, they exposed the children to different visual, auditory, and olfactory stimuli. They found that children with ASD had more significant body movements than typically developed children, especially in the head, trunk, and feet and for visual, followed by visual-auditive, and lastly for visual-auditive-olfactory stimuli.

An exciting study by Rahman et al.[39] aimed to study the ability of machine learning to predict the risk of autism during the neonatal period. They combined the machine learning techniques with electronic medical records using parental sociodemographic information and medical histories and the prescribed medications data to create features to train various machine learning algorithms. They succeeded in capturing early-life features that increase the risk of ASD. They were also able to uncover previously unknown features linked with increased ASD risk. An additional exciting study used fetal ultrasound features by a computer program to predict the child’s autism from the first day after birth. The fetal features included the baby’s head and stomach size, thighs length, and the time of acquiring a vertex presentation in preparation for delivery. The program also used the peri-labor data such as heart rate and body temperature and followed the children up to 6 years. Then the program can independently recognize the associations between different fetal characteristics and outcomes[40]. Artificial intelligence can obtain data on a large scale from all over the world, then re-study it and extract data used to increase the accuracy of autism diagnosis. Many applications are used in diagnosing autism. Artificial intelligence can use the data collected by these applications and process it, so we get accurate results that represent helpful diagnostic tools for their application in different parts of the world. It is also possible to determine the criteria of autism for each race according to its culture and customs. One of the widely used applications to diagnose autism worldwide was created by Dr. Fadi Fayez and Dr. Reza Shahamiri (Nelson Marlborough Institute of Technology, New Zealand). It uses ten questions for the four age groups, from toddlerhood, childhood, and adolescence to adulthood. They
AI is promising in treating children with autism, despite currently being costly. Robots can train and interact with the advantage of showing different facial expressions, proper social interaction, and response to different social cues with unlimited patience and the ability to repeat the cues in the same manner, unlimited times without variation. Some robots can show social expression by changing their eye color, raising their arms, or changing their voice tone[42]. Some children with autism have a better response to the robot than a human therapist. Some robots can incorporate data about individual children using video, audio, and measurements of vital signs such as heart rate and temperature and presence or absence of skin sweat to personalize their response to the child’s behaviors[43]. Despite being promising and effective, robotic intervention needs more wide-scale research, especially cost-effectiveness.

ROLE OF AI IN DIAGNOSIS OF GI DISORDERS

Diagnostic and therapeutic endoscopies have provided significant help in managing Pediatric gastrointestinal disorders for decades. Endoscopy with small bowel sampling is the gold standard to diagnose coeliac disease. Endoscopy also provides excellent assistance in diagnosing various GI diseases such as gastro-oesophageal reflux disease, eosinophilic oesophagitis, and inflammatory bowel disease. It is also used to stop GI bleeding, insert a gastrostomy tube, dilate a stricture, and remove a polyp. The recent marvelous endoscopic field achievements helped us reach previously non-reachable areas of the mid-small intestine using the wireless capsule video-endoscopy[44]. There was a broad jump in the endoscopic industry from the white light to the blue light endoscopy and recently endocytoscopy and endomicroscopy. These recent modalities helped visualize the mucosal structure at the cellular level with adequate histopathology determination. It helps gather a vast amount of data that needs many hours of interpretation by a highly experienced physician[45]. AI helps process these vast amounts of data and allows rapid and precious interpretation.

Gastroesophageal reflux disease (GERD) is a principal reason for abnormal behaviors in children with autism. Upper gastrointestinal endoscopy is one of the preferred modalities to diagnose and detect complications of GERD, including Barret esophagitis, by evaluating the oesophageal mucosa. It also can rule out other possible causes of the child’s symptoms, such as eosinophilic esophagitis[46]. AI helps improve the mucosal images’ quality and detect their exact anatomical location. Changes in the oesophageal mucosa such as Micro-erosions, Changes in intrapapillary capillary loops, and increased vascularity are landmarks for GERD detected by narrow-band imaging with the help of AI model using convolutional neural networks (CNNs)[47].

Meanwhile, Takiyama et al[48] used CNNs to precisely recognize the anatomical location of esophagogastroduodenoscopy images. Pace et al[49] developed an artificial neural networks (ANN) model that can predict the presence of GERD without the need for invasive diagnostic techniques in
patients with GERD symptoms. We hope that AI will help classify patients with GERD provide personalized therapeutic approaches[30]. However, the ability of ANN to expect GERD diagnosis depending on the symptoms still requires more verifications in different clinical settings. Considering that patients with autism may have other ways of expressing GER- or GERD-related symptoms, they may need specific and additional protocols to be applied with AI.

Coeliac disease is a common but underdiagnosed autoimmune disorder affecting 1/100 people worldwide with a relatively higher incidence in children with autism. The presence of villous atrophic histology in duodenal mucosal biopsy samples obtained by endoscopy is the gold standard for diagnosis. Endoscopy can also detect unsuspected cases of coeliac disease by meticulous analysis of the small bowel mucosa and identification of subtle findings of villous atrophy. However, it needs multiple biopsies not to miss the lesion as it is patchy[31]. New endoscopic techniques such as the modified immersion technique under traditional white-light or narrow-band imaging significantly improve the visual confirmation of coeliac disease during endoscopy[52,53]. Video capsule endoscopy is reasonably well-tolerated and safe in children. It can identify mucosal lesions in the bowel, especially in the small intestine, with the risk of radiation or sedation[34]. Video capsule imaging could also help identify the coeliac disease[55]. Augmentation techniques using AI can help augment the obtained original mucosal images to avoid the effects of conditions that could affect the quality of images, such as the rotation of the endoscope or the effects of distant viewpoint from the mucosal wall changes. However, a patency capsule test should be done before video capsule endoscopy, especially in infants and young children[56]. Recent AI modalities using deep learning techniques such as convolutional neural network (CNN), Bayesian inference, or support vector machines are innovative computer technology that can aid computerized coeliac disease diagnosis[57]. Foers et al[58] used machine learning methods to classify intestinal T-cell receptor repertoires to detect patients with coeliac disease irrespective of their dietary gluten status.

Pediatric colonoscopy needs a high experience not to miss lesions and detect colonic lesions as early as possible. The significant progress in developing computerized vision during gastrointestinal endoscopy allowed the gathering and annotation of high-quality video information. The addition of AI to real-time endoscopy significantly improves the automated detection of colonic or rectal polyps, such as in juvenile polyp or familial adenomatous polyposis. As mentioned before, children with autism have an increased risk to develop inflammatory bowel disease (IBD). This increased risk is due to multifactorial pathogenesis, including an overactive immune system and disturbances of the brain-gut-microbiota axis[59]. A massive flow of data about IBD is currently available using electronic medical records, genetic analysis, and imaging modalities. Analysis, interpretation, and integration of these data with the help of AI can aid to build models that can predict the risk of IBD and increase its detection accuracy[60].

Ozawa et al[61] succeeded to develop a neural network trained on colonoscopy images from patients with ulcerative colitis. With the help of a computer-assisted system, this network was able to identify the normal mucosa, mucosal healing states, mucosa on remission, and mucosa in severe degrees of inflammation with high sensitivity and specificity. These findings will help the physicians personalize the treatment according to the patients’ conditions. In children, Mossotto et al[62] classified Crohn's and ulcerative colitis activity at diagnosis, using machine learning with integrating both endoscopic and histologic imaging. They were able to subtype the patients using this model with an accuracy reaching 80%, which significantly improves the diagnostic accuracy and permits a good option for targeted therapy. Dhaliwal et al[63] developed an algorithm using Random Forest Supervised and Unsupervised Machine learning in children to identify features that could help discriminate between ulcerative colitis and colonic Crohn's disease. They have a correct classification in 98% and 95% of children with ulcerative colitis and colonic Crohn disease, respectively.

Video capsule endoscopy is a safe, non-invasive procedure that can help diagnose IBD, especially for the patchy intestinal lesion of Crohn's disease. Nemeth et al[64] examined the accuracy and safety of video capsule endoscopy in 154 children and adolescents with suspected or established Crohn's disease. They found that video capsule endoscopy was safe and able to confirm the diagnosis of Crohn's disease with a significant impact on clinical management. However, interpretation of images obtained, and diagnosis based on Video capsule endoscopy is reader-dependant. As a result of the human concentration limitation, the lesion miss rate in capsule endoscopy ranges between 0.5% to 19% depending on the nature of the lesion[65]. AI can improve the accuracy of capsule endoscopy diagnosis by identifying distinct lesions and areas of interest with ease. However, there are many limitations to providing reliable classifications due to insufficient accuracy[66]. A convolutional neural network is used to analyze the large number of images obtained by capsule endoscopy to overcome these limitations. The convolutional neural network can differentiate normal intestinal mucosa, ulcers, erosion, polyps, and even worms with high accuracy reaching up to 96%[67,68].

Li et al[69] developed an AI system to automatically distinguish colorectal cancer early signs during colonoscopy with high sensitivity and specificity. Aguilar et al[70] found that transabdominal ultrasound augmented with a preceding AI model allows precise, fast, and non-invasive diagnosis of Buried bumper syndrome, complicating percutaneous endoscopic gastrostomy in children. Urban et al[71] successfully identified and removed rectal polyp using a deep neural network with a real-time accuracy rate of 96.4%. They used convolutional neural networks with an ordinary desktop machine
with a contemporary graphics processing unit. They concluded that their trained model could identify and locate polyps in real-time with high accuracy. However, caution should be taken when using convolutional neural networks as we cannot generalize the results of these studies to other situations, and we do not know the exact effect of using convolutional neural networks on the endoscopists' inspection behavior with overreliance on the technology. In addition, the direct and the indirect cost of these technologies and their acceptance to be a part of the diagnostic tools by the physician should also be considered[72].

ROLE OF AI IN MANAGING GI DISORDERS IN CHILDREN WITH AUTISM

Children with autism have a three-fold increase in the risk of gastrointestinal disorders than the typically developed children. However, the rate of parental reporting of these disorders is less in children with autism than in the typically developed ones. The wide varieties of GI manifestations in children with autism are related to the general heterogeneity of autism disorder and the underlying neurobiological mechanisms and disturbances of the neurotransmitters in both brain and gut[73]. Some GI symptoms are evident as diarrhea and constipation, while the others may vague and challenging to be recognized and can be missed as behavioral changes. Artificial intelligence can help detect and classify autism early, even in young infants. Meanwhile, AI can help detect and classify gastrointestinal disorders in children. Regrettably, according to the best of our knowledge, there are no currently available specific AI models to detect gastrointestinal disorders in children with autism. AI models designed to detect GI disorders in patients with autism should consider the differences in the symptomatology from the typically developed children and should design algorithms to detect these disorders.

Fascinating use of AI in children with autism is the use of 'SMART TOILET' to monitor bowel health and to help to detect irritable bowel syndrome and inflammatory bowel disease. An artificial intelligence tool with a camera and microcomputer are attached to the traditional toilet to help evaluate patients' stool, including form, defecation time, urination, and the presence or absence of blood[74]. This 'SMART TOILET' will significantly help manage toilet disorders, common in children with autism, as they cannot correctly report their bowel habits, dysfunction, or defecatory disorders. The use of microbiota transfer therapy (MTT) showed significant potential in alleviating the symptoms associated with GI complications and reducing the severity of behavioral symptoms in children with autism[75]. Children with autism who had MTT also showed changes in their plasma metabolite profile to be nearly similar to the typically developing peers[76]. Qureshi et al[77] examined the differences in gut microbial metabolites between children with autism and GI disorders vs the typically developing children without GI disorders and determined the effects of gut MTT on the fecal metabolites of the group with autism. They used machine learning to create 5-metabolite fecal models for classification, which showed significant changes before and after MTT. The developed multivariate metabolite models showed the potential of fecal metabolite panels to effectively categorize children with autism from the typically developed. Similar machine learning models can diagnose children with autism using their gut microbiome data compared to subjects with and without autism.

About 10% of children with autism are on a special diet. Despite no diet specific for autism, children with autism are frequently put on a gluten-free, casein-free diet. However, children should not start a special diet except when it is evidence-based. Dietary management can help alleviate many of the functional gastrointestinal symptoms in patients with irritable bowel syndrome, which is relatively common in children with autism. One of these dietary managements is restoring the imbalance in gut microbiota. Karakan et al[78] studied the efficacy of artificial intelligence-based personalized modification of the dietary microbiome in patients with IBS. They designed an AI-based diet to optimize a personalized nutritional strategy using an algorithm about the individual gut microbiome features. According to the IBS index score, they developed the algorithm to design the diet. The algorithm assessing an IBS index score used the microbiome composition to design the optimized diets based on the microbiome modification to match the observed with the healthy scores.

Gluten sensitivity is common in children with autism. AI can help produce allergen-free gluten in plants with high gluten content, such as wheat and corn. This new gluten retains its unique beneficial quality regarding texture, taste, and nutritional value without the ability to stimulate the autoimmune response and cascade of gluten sensitivity or coeliac disease. Another way to overcome coeliac disease and gluten sensitivity is to create an oral enzyme able to degrade the ingested gluten. The proposed enzyme should be stable and active in both stomach and duodenum, rapidly neutralize the gluten-peptides that can activate T-cell, and be safe to be ingested by humans. Many enzymes, including cysteine proteases, prolyl endopeptidases, and subtilisin’s, could split the non-digestible gluten peptides in vitro and vitro. AI can help develop new techniques like enteric coating to protect the enzyme or genetic modification, increasing its production and enhancing its stability in the GI tract[79].
LIMITATION FOR THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence has many obstacles in treating digestive diseases, especially in children with autism. Among these obstacles are the ethical aspects and the confidence of the medical staff in the mechanisms of artificial intelligence. The development of artificial intelligence also requires a robust infrastructure with enhanced patient confidentiality controls. Also, committees must be established to control the work of artificial intelligence to avoid the inappropriate and unethical use of artificial intelligence. Another significant limitation is the difference in symptoms of GI disorders in children with autism than the typically developed children. Children with autism may need a minimum level of communication abilities and cognitive function to use AI-directed models. When building an algorithm, it should be tailored to children with autism.

CONCLUSION

Autism is a neurodevelopmental condition with multiple comorbidities. Besides the classic manifestations of autism, the affected patients may suffer the presence of many other medical comorbidities that are more common in people with ASD than in the general population. Children with autism have a high prevalence of GI disorders occurring in 46%-84% of them with a bilateral mutual pathway between autism and GI disorders. Children with autism have an increased frequency of diarrhea, nausea and/or vomiting, gastroesophageal reflux and/or disease, abdominal pain, chronic flatulence due to various factors as food allergies, gastrointestinal dysmotility, IBS, and IBD. AI could help diagnose and manage autism by improving children's communication, social, and emotional skills for a long time. AI is an effective method to enhance early detection of GI disorders, including GI bleeding, gastroesophageal reflux disease, Coeliac disease, food allergies, IBS, IBD, and rectal polyps. AI can also help personalize the diet for children with autism by microbiome modification. AI can help to provide modified gluten without the ability to initiate an immune response. However, AI has many obstacles in treating digestive diseases, especially in children with autism. There is a need to do more studies and adapt specific algorithms for children with autism.

ACKNOWLEDGEMENTS

We thank the anonymous referees for their valuable suggestions.

FOOTNOTES

Author contributions: Al-Biltagi M, Saeed NK, and Qaraghuli S wrote and revised the manuscript.

Conflict-of-interest statement: The authors declare that they have no conflict of interest.

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: Egypt

ORCID number: Mohammed Al-Biltagi 0000-0002-7761-9536; Nermin Kamal Saeed 0000-0001-7875-8207; Samara Qaraghuli 0000-0002-6909-2658.

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

REFERENCES


AIG | https://www.wjgnet.com

February 28, 2022 | Volume 3 | Issue 1

21410934 DOI: 10.1186/1471-230X-11-22


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

Hasan Maulahela, Nagita Gianty Annisa

Specialty type: Gastroenterology and hepatology
Provenance and peer review: Invited article; externally peer reviewed.
Peer-review model: Single blind
Peer-review report’s scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): B
Grade C (Good): 0
Grade D (Fair): D, D
Grade E (Poor): 0

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

©The Author(s) 2022. Published by Baishideng Publishing Group Inc. All rights reserved.
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.

Citation: Maulahela H, Annisa NG. Current advancements in application of artificial intelligence in clinical decision-making by gastroenterologists in gastrointestinal bleeding. Artif Intell Gastroenterol 2022; 3(1): 13-20
DOI: https://dx.doi.org/10.35712/aig.v3.i1.13

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, elder 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 arterial 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 ANN’s 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 BLEEDING

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
Maulabela H et al. AI in gastrointestinal 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 GB 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 startups 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

Author contributions: Maulahela H proposed the idea of the manuscript, and wrote and edited the manuscript; Annisa NG performed data accusation and manuscript writing.

Conflict-of-interest statement: The authors declare that they have no conflict of interest to disclose.

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, and build upon this work non-commercially, provided the original work is properly cited. For commercial re-use, please contact journals@preventdisease.com.

Legal entity responsible for publication: World Journal of Gastrointestinal Pathology.
REFERENCES


Maulahela H et al. AI in gastrointestinal bleeding

DOI: 10.1136/bmj.j6432


Artificial intelligence and human liver allocation: Potential benefits and ethical implications

Marcos Mucenic, Ajacio Bandeira de Mello Brandão, Claudio Augusto Marroni

Abstract
Since its implementation almost two decades ago, the urgency allocation policy has improved the survival of patients on the waiting list for liver transplantation worldwide. The Model for End-Stage Liver Disease score is widely used to predict waiting list mortality. Due to some limitations related to its use, there is an active investigation to develop other prognostic scores. Liver allocation (LA) entails complex decision-making, and grafts are occasionally not directed to the recipients who are more likely to survive. Prognostic scores have, thus far, failed to predict post-operative survival. Furthermore, the increasing use of marginal donors is associated with worse outcomes. Adequate donor-recipient pairing could help avoid retransplantation or futile procedures and reduce postoperative complications, mortality, hospitalization time, and costs. Artificial intelligence has applications in several medical fields. Machine learning algorithms (MLAs) use large amounts of data to detect unforeseen patterns and complex interactions between variables. Artificial neural networks and decision trees were the most common forms of MLA tested on LA. Some researchers have shown them to be superior for predicting waiting list mortality and graft failure than conventional statistical methods. These promising techniques are increasingly being considered for implementation.

Key Words: Liver transplantation; Liver cirrhosis; Artificial intelligence; Prognosis; Survival; Machine learning

©The Author(s) 2022. Published by Baishideng Publishing Group Inc. All rights reserved.
Core Tip: This review discusses the ethical aspects and current advancements in liver allocation (LA). It summarizes the concept of artificial intelligence and focuses on the latest developments of machine learning algorithms as applied to predicting waiting list mortality and LA. To date, only a few research groups have published works on this field; they also wrote reviews on the subject. Our minireview offers a thorough and impartial view of the topic, and we hope this will alert other potential researchers to this promising field.

INTRODUCTION

Liver transplantation (LT) is the treatment of choice for patients with terminal liver disease[1]. LT is increasingly performed worldwide; however, organ scarcity remains a significant challenge for transplant teams[2], placing greater weight on the need for efficient liver allocation (LA). Therefore, correct organ allocation is of paramount importance. An optimal allocation system for LT should balance considerations of equity (equal opportunity to receive the graft), need (to reduce waiting list mortality), utility (maximizing the overall life-years gained), and benefit (optimizing outcomes from each organ transplanted)[3].

Urgency criteria are based on need, prioritizing grafts to the most critically ill patients. Survival without LT is estimated through prognostic models, such as the model for end-stage liver disease (MELD). MELD is a validated score derived solely from laboratory test results (total bilirubin, serum creatinine, and prothrombin time). MELD is simple and accurate, and it predicts the 3-mo mortality of candidates with an area under the receiver operating characteristic curve (AUROC) of 0.83[4]. Since 2002, the MELD score was adopted by the United Network for Organ Sharing (UNOS) to rank waitlisted patients in order of urgency in the United States (US). Several countries followed this organ allocation system (sickest first)[5]. While post-transplant survival for the sickest is lesser than that of patients with better physiological reserves, they are the ones who benefit the most from LT. Patients with a MELD score of 31-34 had a relative life expectancy 43 times higher than those who remained on the list, and patients with a MELD score of 35-40 were 128 times more likely to survive[6]. MELD was further refined after studies showed adding serum sodium concentration to the formula (MELD-Na) improves risk stratification. This system replaced MELD for LA in some countries, such as the USA, Canada, and Brazil[5,7-9]. Godfrey et al[10] advised of a possible loss of predictive accuracy of the MELD score over time, reaching an AUROC of only 0.70 in 2015. This may be due to changes in the epidemiology and treatment of liver diseases and increasing age and comorbidities. Despite several valid concerns about the model, it remains the most widely used.

Urgency allocation models have no value in predicting survival after LT[11,12]. Additionally, the donor pool has been expanded in the last two decades. Although the use of marginal donors (e.g., older donors, steatotic livers, and donation after cardiac death) has been necessary in this regard, it increases the rate of graft failure and postoperative complications, adding further complexity to the matter of allocation[13,14]. Living donor liver transplantation is another strategy to expand the donor pool; however, it poses an inherent risk to healthy donors. Its proportion to the total number of LT is small[3].

The MELD score does not reflect mortality risk in compensated patients with hepatocellular carcinoma (HCC). Exception points are granted to candidates with HCC, one of the leading LT indications worldwide. Currently, the prioritization of HCC candidates varies from one country to another, and there is no international consensus on the matter[15]. Due to the excessive advantage conferred by these exception points, there have been some changes in global allocation policies[16-18]. Notwithstanding these revisions, HCC candidates still have increased transplant rates, decreased risk of delisting, and worse post-transplant prognosis[15].

Outcomes after LT depend on both the preoperative condition of the recipient and donor “quality”. Utility criteria have been sought to offer grafts to recipients with greater chances of survival, estimating the outcome based on donor and recipient characteristics[3,16]. While this would decrease the odds for older and sicker patients to receive an organ, the overall post-transplant survival could improve. Better selection avoids retransplantation or futile procedures and reduces postoperative morbidity, hospitalization time, and costs. The survival benefit is quantifiable by estimating waiting list survival and post-transplant outcomes. A benefit-based system could balance urgency and utility in allocation decisions. For a benefit-based allocation to be successful, transplant teams would need an accurate model to predict post-transplant survival.
Although the concept of applying donor-related variables to an algorithm had been used before, Feng et al.[19] devised the term “donor risk index” (DRI), wherein they identified seven donor characteristics that predicted graft failure. Other researchers further investigated this interesting concept, adding cold ischemia time and organ origin (national or regional). However, DRI has not been widely adopted owing to the following main reasons cited by surgical teams: the inaccuracy to predict survival, exclusion of other relevant risk factors, and difficulty of explaining its concept to the recipients[20].

Since neither candidate nor donor factors are predictive of survival following transplantation, scores that include both variables have been described over the previous years. Halldorson et al.[21] proposed adding donor age to the MELD score, creating the D-MELD, demonstrating a survival disadvantage when combining higher donor age with higher MELD recipients. Schaubel et al.[22] proposed the balance of risk (BAR) score, which included donor, recipient, and procurement surgery variables. For an estimated 5-year survival, the c-statistic reached an AUROC of 0.63. Rana et al.[23] developed a complex score named survival outcomes following liver transplantation (SOFT), containing > 100 variables. It reached an AUROC of 0.70 to predict 3-mo survival after LT. In 2018, the United Kingdom introduced allocation rules based on benefit. Each graft is offered nationally to the recipient predicted to have the greatest survival benefit from that specific graft. LA is based on the transplant benefit score (TBS), calculated by 21 and 7 receptor and donor criteria, respectively. TBS reflects the difference in days between expected 5-year post-operative survival and expected 5-year waiting list survival. The model reduced deaths on the waiting list and maximized post-operative life-years[24].

The external validity of these scores is limited by several factors, such as ethnicity, regional differences in the allocation and transplantation procedures, and changes in practice over time. Since the scores are based on logistic regression (LR) models, they depend upon the assumption of independence of each variable and are limited when facing nonlinear variable interactions. Complex donor-related models are computationally difficult to implement, and their accuracy can be limited by the large number of variables that impact survival and possible undetected confounding factors. They have not been validated by other researchers or found wide acceptance to date. Moreover, they were not designed for an ideal donor-recipient matching[23,25]. Therefore, transplant teams are faced with a complex decision-making process when having to choose recipients for LA. Objective criteria would exempt the medical staff from difficult decisions and assess whether patients are excessively sick to be transplanted[3,7,26]. “Artificial intelligence” (AI) or machine learning algorithms (MLAs) are under increasingly active investigation for this use[27].

**AI**

AI is a general term used to describe any application wherein computer systems perform tasks normally associated with human intelligence. It can be a substitute for human subjectivity and limitations[28]. AI encompasses simple automated tasks and increasingly complex fields, such as machine learning, deep learning, and artificial neural networks (ANNs).

In classical programming, the computer is supplied with an algorithm and a dataset to provide an output. Machine learning, in contrast, supplies the computer with data and associated outputs, which it uses to create an algorithm that describes the relationship between the two. These MLAs can detect patterns and improve their analysis over time with further data[29]. MLAs can analyze any number of variables and are not driven (or limited) by hypothesis. This method detects nonlinear patterns within large datasets wherein multiple interactions between variables can occur. MLA can accommodate numerous interdependent variables and improve as more cases are increasingly analyzed[30,31].

Typically, MLA applied to healthcare fall into the category of supervised learning techniques. These algorithms learn the associations between input and labeled outcome data. The following are the basic steps of supervised ML: (1) Acquire a dataset and split it into separate training, validation, and test datasets; (2) use training and validation datasets to create a model that analyzes the association between data and outcomes; and (3) evaluate the model via the test dataset to determine how well it predicts outcomes. There are other techniques used, such as unsupervised learning, wherein data are not labeled to find out previously unknown patterns. Semi-supervised learning is particularly useful for datasets that contain both labeled and unlabeled data. Reinforcement learning uses the consequences of their actions to learn to determine the optimal behavior for a given context[29,32].

The decision tree (DT) is a supervised learning technique primarily used for classification tasks (categorical variables). It consists of a hierarchically organized structure of nodes that makes predictions by splitting (branching) the data. Each split can connect to a new root node or attach to a terminal or “leaf” node. A random forest (RF) is an ensemble method that produces multiple DTs[29].

One of the advantages of DTs for healthcare applications is their interpretability. However, each node is determined in isolation without considering the possible impact of future splits. This can fail to capture the dataset’s underlying characteristics. This disadvantage stimulated the development of optimal classification trees (OCTs). This type of DT is formed entirely in a single step, allowing each split to be determined with full knowledge of all other splits[33].
ANN is an MLA inspired by biological neural networks. Each ANN contains nodes (analogous to cell bodies) that communicate with other nodes via connections (analogous to axons and dendrites), with multiple layers (an input layer, an output layer, and a hidden layer between them) of connected mathematical functions. ANNs can capture complex nonlinear relationships in data, allowing for sophisticated supervised and unsupervised learning tasks[28,29,32].

Support vector machine (SVM) is another type of MLA. This method organizes data by variable classes in a nonlinear modality, subsequently separating by a hyperplane and forming multidimensional planes in space using these data. It can be used for classification or regression problems[34].

### AI APPLIED FOR THE PREDICTION OF MORTALITY IN THE WAITING LIST

MLA has shown promising results for predicting 3-mo mortality on the waiting list. The simulation model was based on OCTs. OCTs were fed with > 1.6 million observations and trained, validated, and tested. The result showed a slightly superior AUROC than MELD-Na for predicting death or unsuitability for LT (0.859 vs 0.841). The authors argue that this system would save at least 418 more lives annually in the US[35]. An interesting point in this simulation model compared with MELD was the increased allocation of livers to non-HCC patients and a decreased number of waitlist deaths and removals for both HCC and non-HCC patients. Their results await further validation.

Cucchetti et al[36] applied an ANN model to predict the 3-mo mortality of patients awaiting LT in the pre-MELD era. The analysis included only laboratory values (liver biochemical and function tests, creatinine, and hemogram). The participants were randomly divided into training and testing groups in a proportion of 75%-25%. After each of the 10 training sessions, ANN was tested on the remaining individuals who were not selected for training. The most accurate ANN system was tested in a retrospective cohort in another LT center. The performance of ANN in predicting the 3-mo mortality was superior to that of MELD (AUROC, 0.98 vs 0.86). Results were similar for the external validation cohort (0.96 and 0.86, respectively).

### AI APPLIED FOR LA

Hundreds of variables contribute to multiple decisions made in an organ transplant. For each record, the UNOS database collects > 400 parameters. AI can theoretically improve the outcomes of allocation strategies[30,31]. An optimal outcome would be a decreased number of retransplant procedures, excellent graft and overall survival, and decreasing rate of waiting list mortality.

In a large multicenter Spanish study, Briceño et al[37] applied 57 variables (26, 19, 6, and 6 from the recipient, donor, retrieval procedure, and transplant procedure, respectively) for each donor-recipient pair to predict 3-mo graft survival. A total of 1003 liver transplants were analyzed. This sample had been previously described in a pilot study by the same group of researchers[38]. The following were the two models of ANN used: a positive-survival (PS) model to predict the 3-mo graft survival rate after LT and a negative-survival (NS) model to predict the 3-mo graft failure rate. These ANN models are MLAs that simulate a biological neural system. In this study, 90% of the data was used for training and 10% for testing, which was repeated ten times to allow all patterns to participate in both phases. Subsequently, the model that correctly classified the most D-R pairs was chosen. PS methodology was slightly superior to common statistical methods (multiple regression, MR) to predict graft survival (90.8% vs 87.7%). NS methodology performed worse for predicting graft loss; however, it was far superior to MR (71.4% vs 3.4%). Finally, the AUROC curves were compared with previously reported scores (MELD, D-MELD, DRI, P-SOFT, SOFT, and BAR). In the PS model, ANN had an AUROC of 0.81, significantly higher than that of other conventional statistical methods (which varied from 0.42 to 0.68). In the NS model, NN had an AUROC of 0.82, which was also significantly higher than that of the other scores (which varied from 0.42 to 0.61). Of the previously reported scores, BAR showed the best AUROC results.

Aylón et al[30] applied D-R pairing with ANN on 858 cases in a large-volume LT center (King’s College Hospital). They used the same PS and NS models described by Briceño et al[37], with some differences in the included variables. AUROCs for PS (0.94) and NS (0.94) 3 mo after LT were significantly more accurate than that of BAR, which is the second-best score (AUROC, 0.84). Furthermore, the researchers performed a 12-mo analysis, and when ANN was used to predict graft survival and loss (0.78 and 0.82, respectively), their results were better than that of the best prediction achieved by other scores (BAR, 0.71).

Lau et al[39] applied ML techniques in an Australian single-center study with 180 LTs. A bootstrap sample containing approximately 63% of the cases was used for the training set, and the remaining data were used for testing. This process was repeated 1000 times. RF classifiers and ANN were used on the overall top 15 ranked characteristics to determine the performance as measured by AUROC values. Graft failure (NS) within 30 d was the primary outcome, and NS at 3 mo postoperatively was the secondary outcome. The results were subsequently compared with those of DRI and SOFT scores. The AUROC for the 30-d NS was 0.818 with RF and 0.835 with NN compared with 0.64 and 0.68 with SOFT.
Table 1 Overview of original works on artificial intelligence applied to liver allocation

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Sample size and location</th>
<th>AI model(s)</th>
<th>Outcomes analyzed</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Briceño et al [37], 2014</td>
<td>1003 LT recipients (multicenter in Spain)</td>
<td>ANN with PS and NS model with D-R pairing</td>
<td>3-mo Graft survival (PS); 3-mo Graft failure (NS)</td>
<td>AUROC 0.81 (PS); AUROC 0.82 (NS)</td>
<td>Superior to BAR score (0.68 for PS, 0.61 for NS). Other conventional statistics fared worse</td>
</tr>
<tr>
<td>Ayllón et al [30], 2018</td>
<td>858 LT recipients (single-center in England)</td>
<td>ANN (PS and NS) with D-R pairing</td>
<td>3-mo Graft survival (PS); 3-mo Graft failure (NS)</td>
<td>AUROC 0.90 (PS); AUROC 0.90 (NS)</td>
<td>Superior to BAR score (AUROC 0.71). Same model above on a different population (external validation)</td>
</tr>
<tr>
<td>Lau et al [39], 2017</td>
<td>180 LT recipients (single-center in Australia)</td>
<td>ANN and RF</td>
<td>30-d and 3-mo Graft failure (NS)</td>
<td>30-d prediction: AUROC 0.82 (RF) AUROC 0.855 (ANN)</td>
<td>Superior to SOFT and DRI scores</td>
</tr>
<tr>
<td>Guigo-Rubio et al [40], 2021</td>
<td>20456 LT recipients (5-yr survival) to 37646 LT recipients (3-mo survival) UNOS database</td>
<td>ANN, RF, DT, SVM, MLP</td>
<td>3-mo, 1 yr, 2 yr, 5 yr survival</td>
<td>AUROC up to 0.618 (3-mo), 0.614 (1-yr), 0.611 (2-yr), 0.644 (5-yr)</td>
<td>No superiority compared to conventional statistics (LR was slightly superior)</td>
</tr>
</tbody>
</table>

AI: Artificial intelligence; ANN: Artificial neural network; AUROC: Area under the receiver operating characteristic curve; BAR: Balance of risk; DRI: Donor risk index; DT: Decision tree; D-R: Donor-receptor; MLP: Multilayer perceptron; NS: Negative-survival; LT: Liver transplantation; LR: Logistic regression; PS: Positive-survival; SOFT: Survival outcomes following liver transplantation; SVM: Support vector machines; RF: Random forest; UNOS: United network of organ sharing.

and DRI scores, respectively. The AUROC decreased to 0.715 with RF and 0.56 with NN to predict the 3-mo NS (including 90 cases in the analysis).

Contrastingly, MLA did not prove to be superior to LR for predicting survival after adult LT using donor-recipient matching in a large database [40]. Four different survival endpoints were analyzed using the UNOS database, including 3-mo and 1-, 2-, and 5-year survivals, varying from 37646 transplants in the 3-mo analysis to 20456 transplants in the 5-year analysis. A total of 28 variables were considered, including recipient, donor, and matching variables. Several types of MLA were used, including ANN, RF, DT, and SVM. The researchers suggested that this lack of accuracy of MLA could be ascribed to database limitations. The highest AUROCs were obtained with LR, followed by RF.

Table 1 summarizes the original works on AI applied to LA that were discussed above.

A systematic review of AI for predicting post-transplant survival was performed by Wingfield et al [34]. Nine publications were included, and articles were considered of good quality overall. ANN and LR were the most common types of MLA and conventional statistical methods, respectively. MLAs were similar or superior to conventional statistics.

CONCLUSION

Although prioritization criteria have successfully reduced mortality in the waiting list, there is room for refinement in mortality prediction and a growing need for improving LA guidelines. Conventional statistical methods have, thus far, failed to provide a useful and widely applicable allocation score. AI can bring meaningful insights to this field. Paradoxically, MLA could help improve the ethics of LA, increasing waitlist and post-transplant survival, preferably with quality-adjusted life-years gained. The results obtained, thus far, are promising; however, we must consider the limitations of AI in medicine. First, its accuracy depends upon the availability of accurate, organized, and thorough datasets. In this regard, the algorithms also depend upon the data used to feed them, and regional particularities can limit their validation. Further, the clinical relevance of the results must be properly evaluated by experts in the field. Moreover, it can be challenging for the lay population to understand and accept LA decisions based on AI analysis. Finally, the health providers must make the final decision, at least while the concepts of ethics and justice rest upon the human mind.

FOOTNOTES

Author contributions: de Mello Brandão AB wrote about liver allocation; Marroni CA wrote about artificial intelligence; Mucenic M wrote about the applications of artificial intelligence on liver allocation and on the prediction of waiting list mortality, wrote abstract and conclusions, and revised the writing style; all three writers revised the paper.

Conflict-of-interest statement: There is no conflict of interest associated with any of the authors of this manuscript.
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: Brazil

ORCID number: Marcos Mucenic 0000-0001-9389-2236; Ajacio Bandeira de Mello Brandão 0000-0001-8411-5654; Claudio Augusto Marroni 0000-0002-1718-6548.

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

REFERENCES


