

75135_Auto_Edited.docx

Name of Journal: *Artificial Intelligence in Gastrointestinal Endoscopy*

Manuscript NO: 75135

Manuscript Type: MINIREVIEWS

Artificial intelligence in the endoscopic approach of biliary tract diseases: A current review

Fábio Pereira Correia, Luís Carvalho Lourenço

Abstract

In recent years there have been major developments in the field of artificial intelligence. The different areas of medicine have taken advantage of this tool to make various diagnostic and therapeutic methods more effective, safe, and user-friendly. In this way, artificial intelligence has been an increasingly present reality in medicine. In the field of Gastroenterology, the main application has been in the detection and characterization of colonic polyps, but an increasing number of studies have been published on the application of deep learning systems in other pathologies of the gastrointestinal tract. Evidence of the application of artificial intelligence in the assessment of biliary tract is still scarce. Some studies support the usefulness of these systems in the investigation and treatment of choledocholithiasis, demonstrating that they have the potential to be integrated into clinical practice and endoscopic procedures, such as endoscopic retrograde cholangiopancreatography. Its application in cholangioscopy for the investigation of undetermined biliary strictures also seems to be promising. Assessing the bile duct through endoscopic ultrasound can be challenging, especially for less experienced operators, thus becoming an area of potential interest for artificial intelligence. In this review, we summarize the state of the art of artificial intelligence in the endoscopic diagnosis and treatment of biliary diseases.

INTRODUCTION

The concept of artificial intelligence (AI) began to be explored in 1950, by Alan Turing, when he proposed to think about the question: 'Can machines think?' [1]. This concept was defined as the ability of a computer to achieve human performance in cognitive tasks [2]. AI systems have evolved over the years, with increasingly complex algorithms and increasingly similar performance to the human brain. From this evolution came machine learning and, later, deep learning, two subfields of artificial intelligence. Machine learning identifies and learns patterns and applies them to information in similar future scenarios. Deep learning, currently one of the most used systems, is based on an artificial neural network capable of learning and making decisions by itself, like a human brain [3].

In recent years, several AI systems have been developed for application in several areas of medicine. They are used in the most diverse functions, such as to help assess medical scans, pathology slides, skin lesions, retinal images, electrocardiograms, endoscopy, faces, and vital signs [4]. Gastroenterology, a versatile medical specialty with a wide area of knowledge and an important intervention component, has been one of the areas where AI has been applied more frequently. Some of the application in Gastroenterology are the investigation of dysplasia in areas of Barrett's esophagus, the diagnosis of gastroesophageal reflux disease, the differentiation of acute and chronic pancreatitis, the detection and classification of colorectal polyps, the characterization of colic inflammatory activity in patients with inflammatory bowel disease, among others [3,5].

Despite the many studies on the application of AI in Gastroenterology, the evidence of the use of this technology in the diagnosis and treatment of biliary tract diseases is still scarce. In this review, we conduct research, across multiple platforms and with no time limit, on the application of AI in the diagnosis and treatment of biliary pathology.

ARTIFICIAL INTELLIGENCE AND COMMON BILE DUCT STONES

Gallstones are a very prevalent pathology in the Western population, often asymptomatic, however in some cases complicating with choledocholithiasis,

cholangitis or acute pancreatitis. The diagnosis of choledocholithiasis is not always immediate and linear and may involve several diagnostic methods, from less invasive tests such as ⁷ abdominal ultrasound, computed tomography (CT) and magnetic resonance cholangiopancreatography to more invasive methods such as endoscopic ultrasound (EUS). Until a few years ago, endoscopic retrograde choangiopancreatography (ERCP) was the first-line method in the diagnosis and treatment of common bile duct (CBD) stones. Since this is an invasive procedure with associated risks of complications (for example, post-ERCP pancreatitis, bleeding, and perforation), ERCP is no longer used for an exclusively diagnosis purposes, maintaining an important therapeutic role in patients with a high likelihood or confirmed choledocholithiasis [6].

The importance of correctly selecting patients with an indication for ERCP has led to the development of several models to predict the presence of stones in the CBD. Currently, it is known that models based on Artificial Neural Networks (ANNs) are more suitable than logistic regression models (used in predictive models for dichotomous outcomes) in the evaluation of biological systems. As such, these models have also been proven to be the most effective at predicting the likelihood of CBD stones and thus discriminating patients who will benefit from ERCP [7]. In addition to ANN models based on clinical data, artificial intelligence systems are currently being developed to facilitate the detection of gallstones in imaging exams (CT and abdominal ultrasound) [8,9] and, in this way, contribute to a more careful selection of patients for ERCP.

One of the most important steps for the success of ERCP is the cannulation of the major papilla. Several studies report failure in selective biliary cannulation in up to 20% of cases, even when performed by experienced endoscopists ³ [10]. The European Society of Gastrointestinal Endoscopy has defined criteria for difficult biliary cannulation (at least one of the following): (1) more than five contacts with the papilla whilst attempting to cannulate; (2) time to cannulation greater than five minutes; or (3) more than one unintentional pancreatic cannulation. In these cases, longer manipulation of the papilla and multiple attempts at cannulation increase the risk of post-ERCP pancreatitis [11].

Recently, Kim T *et al* ^[12] developed an artificial intelligence system that predict the location of the ampulla of Vater (AOV) and its difficulty to cannulate. In this model, the identification of the papilla is not based on a bounding box, but on a pixel-wise soft mask, which is a density map where each pixel has a probability of belonging to an AOV (figure 1). In a fivefold cross-validation study, the ¹ model detected the ampulla with mean intersection-over-union 64.1%, precision 76.2%, recall 78.4%, and centroid distance 0.021. These results demonstrate a comparable performance with the human expert in recognizing the range of AOV and to pinpoint the location of AOV, although expert achieve a better deletion of unnecessary parts (precision 91.7% *vs* 78.9%).

¹ Regarding the prediction of cannulation difficulty, the results were not as consistent: high performance for estimating easy cases for selective cannulation with the average precision and recall of 0.802 and 0.719, respectively, but low recall of 0.611 in the selection of difficult cases. The study showed, however, a good performance in predicting the need for additional cannulation techniques during the performance of ERCP.

After cannulation, there are several factors associated to more complex procedure and a lower probability of complete clearance of gallstones, including a ⁶ more acute distal CBD angulation and a shorter length of the distal CBD arm ^[13]. With the aim to predict the technical difficulty of retrieving CBD stones and help the endoscopist to select the best therapeutic approach and accessories during the ERCP, Huang Li ^[14] developed a system based in deep convolutional neural networks, named intelligent difficulty scoring and assistance system (DSAS). This system was evaluated in a retrospective study where 1954 cholangiograms were used - 1381 images for training and 573 images for validation (internal and external). The system showed good accuracy, sensitivity, and specificity (91.45%, 94.57% and 81.13%, respectively) in detecting common bile duct stones, in addition to good results in image segmentation of the stone, common bile duct and duodenoscope - mean Intersection over Union was 68.35%, 86.42% and 95.85%, respectively. In the assessment of technical difficulty scoring of CBD stone extraction during ERCP, the DSAS was consistent with expert's endoscopists. This

system provides a score value, with scores ≥ 2 being associated with greater difficulty in achieving complete CBD clearance (stone clearance rate - score < 2 : 86%; score ≥ 2 : 36%) and more frequently associated with the use of endoscopic papillary-balloon dilation.

ARTIFICIAL INTELLIGENCE AND INDETERMINATE BILIARY STRICTURES

Indeterminate biliary strictures still represent a diagnostic challenge nowadays. Despite the wide differential diagnosis, including benign and malignant causes, the main concern remains the exclusion of a potential malignant cause [15,16]. The methods initially used in the investigation of these strictures, which include imaging, laboratory evaluation and ERCP, although having a high specificity, they have a low sensitivity. Thus, it is difficult to definitively rule out a malignant pathology, which compromises the subsequent approach to the patient [16]. A meta-analysis confirmed the low sensitivity of both cytology (45%) and intraductal biopsies (48.1%) guided by ERCP, in the diagnosis of biliary strictures. Even combining both techniques, the sensitivity is suboptimal [17].

Cholangioscopy has emerged in recent years as a valuable tool in the characterization of these lesions, allowing direct visualization of the stricture and guided biopsies. A recent meta-analysis confirmed the high sensitivity (94%), specificity (95%) and accuracy (94%) of the cholangioscopy in the visual interpretation of biliary malignancies [18]. There are some features suggesting a malignant pathology, namely irregular and tortuous vessels, masses, papillary projections, or infiltrative lesions. Currently, there is no widely accepted system for the visual diagnosis of the stricture, which leads to some non-negligible degree of interobserver variability [19,20].

To overcome that problem, Saraiva M *et al* [21] developed a convolutional neural network-based algorithm with the aim of automatically detecting and differentiating between benign and malignant strictures during cholangioscopy. To train and validate this system, they used 11855 images - 9695 for malignant strictures and 2160 for benign findings (benign biliary strictures or normal segments of the biliary tract). In a 5-fold cross validation study, the sensitivity, specificity, accuracy, and AUC in differentiating

malignant from benign lesions was 94.7%, 92.1%, 94.9%, and 0.988 respectively, with a processing speed of 7 ms per frame. Due to its potential for use in real-time, this system may be useful in choosing the area to be biopsied, to obtain a better histological sample. Ghandour B *et al* [22] also developed an artificial intelligence system that detects features suggestive of malignancy in cholangioscopy images with a sensitivity of 81%, specificity of 91%, positive predictive value of 93%, negative predictive value of 77% and AUC of 0.86. Ribeiro T *et al* [23] created a system for automatic detection of papillary projections in cholangioscopic images, which, like the previous ones, showed very promising results. Although these studies show very promising results regarding the application of artificial intelligence in cholangioscopy, only isolated images were used, and they need to be validated using full videos in real time and in clinical practice.

ARTIFICIAL INTELLIGENCE AND ENDOSCOPIC ULTRASOUND - BILIARY DUCT

The intimate location of the distal stomach, proximal duodenum and biliary tract makes EUS a great diagnostic method for biliary tract conditions. The relevance of EUS has increased in recent decades, including its application in investigation of hepatobiliary diseases. EUS has shown an excellent performance in the diagnosis of several biliary pathologies, namely choledocholithiasis, microlithiasis, biliary strictures, biliary obstruction, or cholangiocarcinoma [24,25]. EUS-guided interventions have also grown in the last years, being an option, in experienced centers, for drainage of biliary obstruction when ERCP fails, as well as in acute cholecystitis, biliary leaks and bilomas [26,27].

EUS is a challenging advanced endoscopic technique with a long learning curve [28]. As such, the development of systems that facilitate the interpretation of ultrasound endoscopy findings appears to be essential for the wide adoption of EUS. Yao Liwen *et al* [29] developed a deep learning-based system, BP MASTER, which, in real-time, recognizes the stations (the fundus of stomach; body of stomach and antrum; duodenal bulb; and descending duodenum) where the transducer is located and provide the

corresponding operation instructions, delineates the bile duct, and gives an estimate of its diameter (figure 2). To train the model, the authors used 10681 images in the bile duct station recognition and 2529 images in the bile duct annotation. For model validation, 2425 images and 515 video clips were used for internal validation and 799 images for external validation. This system showed an accuracy of 93.3% for station recognition in image validation set and 90.1% in video validation set and a Dice of 0.77 in the bile duct segmentation. The results obtained with this system were comparable to those of expert endoscopists. Furthermore, in a crossover study, this system showed an improvement in trainees' accuracy from 60.8% to 76.3%.

Another application of ultrasound endoscopy is the evaluation of polypoid lesions of the gallbladder. Recently, an artificial intelligence system applied to EUS was developed [30] that allows the distinction between gallstones and polypoid lesions with an accuracy of 95.7% and the differentiation of neoplastic and non-neoplastic polyps with an accuracy of 89.8%. At this last point, ⁵ the accuracy of the EUS-AI was between mid-level and expert EUS endoscopists.

Despite being promising systems, with the potential to reduce endoscopic procedures with greater risks and even surgeries, further studies are needed to validate the results obtained.

CONCLUSION

The diagnostic and therapeutic complexity associated with bile tract diseases makes this an attractive area for the development of AI systems.

In choledocholithiasis, AI systems have proved to be useful both in diagnosis, allowing a more careful selection of patients with indication for ERCP; as well as treatment, assisting the endoscopist in the critical steps of the procedure (eg, cannulation). The application of AI in cholangioscopy showed interest in the possibility of a more objective characterization of indeterminate biliary strictures and of directing biopsies to areas where the findings are more suspicious. Endoscopic ultrasound, an intervention

area with a long learning curve, could benefit from the introduction of this technology, especially for less experienced endoscopists.

Despite this, there are still few studies focused on biliary condition, and most of them are retrospective, with small samples and high risk of bias. In the future, it is essential to continue to invest in the development of systems that optimize the diagnosis and facilitate the treatment of biliary pathologies.

8%

SIMILARITY INDEX

PRIMARY SOURCES

- | | | |
|----------|--|---------------|
| 1 | www.ncbi.nlm.nih.gov
<small>Internet</small> | 58 words — 2% |
| <hr/> | | |
| 2 | www.wjgnet.com
<small>Internet</small> | 50 words — 2% |
| <hr/> | | |
| 3 | Brian M. Fung, Teodor C. Pitea, James H. Tabibian.
"Difficult Biliary Cannulation in Endoscopic
Retrograde Cholangiopancreatography: Definitions, Risk
Factors, and Implications", EMJ Hepatology, 2021
<small>Crossref</small> | 22 words — 1% |
| <hr/> | | |
| 4 | Sommer Hammoud, Asheesh Bedi, Erin Magennis,
William C. Meyers, Bryan T. Kelly. "High Incidence of
Athletic Pubalgia Symptoms in Professional Athletes With
Symptomatic Femoroacetabular Impingement", Arthroscopy:
The Journal of Arthroscopic & Related Surgery, 2012
<small>Crossref</small> | 17 words — 1% |
| <hr/> | | |
| 5 | Sung Ill Jang, Young Jae Kim, Eui Joo Kim, Huapyeong
Kang, Seung Jin Shon, Yu Jin Seol, Dong Ki Lee,
Kwang Gi Kim, Jae Hee Cho. "Diagnostic performance of
endoscopic ultrasound - artificial intelligence using deep
learning analysis of gallbladder polypoid lesions", Journal of
Gastroenterology and Hepatology, 2021
<small>Crossref</small> | 14 words — 1% |
| <hr/> | | |
| 6 | www.ejohg.com
<small>Internet</small> | |

14 words — 1%

7 [ia803208.us.archive.org](#)
Internet

13 words — 1%

EXCLUDE QUOTES ON
EXCLUDE BIBLIOGRAPHY ON

EXCLUDE MATCHES OFF