Endoscopic Ultrasound-Based Artificial Intelligence Diagnosis of Pancreatic Cystic Neoplasms

Jin-Seok Park, Seok Jeong
Department of Internal Medicine, Inha University School of Medicine, Incheon, Korea

Pancreatic cystic neoplasms (PCNs) are precursors of pancreatic cancer, and the rate of their incidental detection has gradually increased recently with a reported prevalence from 2.4 to 13.5%. However, accurate diagnosis can be challenging because PCNs have morphologies ranging from benign to malignant disease, and as for other cancers, precise and timely management of premalignant PCN is essential to prevent malignant transformation. Endoscopic ultrasound (EUS) is a useful tool for the differential diagnosis PCN and treatment decision-making because its imaging features predict malignant transformation. However, its performance is suboptimal, and its accuracy for differentiating mucinous pancreatic cysts and other PCNs is only 65-75%, which has increased interest in the application of artificial intelligence (AI). AI has already provided tools that have improved diagnostic accuracies for many cancers, including colon, lung, and breast cancer, and recent studies have shown AI has the potential to differentiate mucinous and non-mucinous tumors and stratify the malignant potentials of PCNs. This article provides a review of the literature on EUS-based AI studies of PCNs.

INTRODUCTION

Endoscopic ultrasound (EUS) is used for diagnosing pancreatic cystic neoplasms (PCNs) such as intraductal papillary mucinous neoplasms (IPMNs), mucinous cystic neoplasms (MCNs), and serous cystic neoplasms (SCNs). Because of the improved quality and ubiquitous use of EUS abdominal imaging, the number of incidental PCNs has risen substantially recently, with estimated prevalence ranging from 2.4-13.5%. However, only a few subtypes of these cysts are considered precursor lesions of cancer. In clinical practice, accuracies for discriminating cystic subtypes of PCN using EUS range from 75 to 95%, and final diagnoses can only be
AI in Pancreatic Cysts

made after histopathological analysis. Importantly, PCN resection may dramatically reduce quality of life because it is associated with a 50% risk of complications and a 5% risk of death. Furthermore, it has been demonstrated that ~60% of surgeries are conducted unnecessarily on benign cysts. Given this challenging situation, artificial intelligence (AI) might improve the accuracy of PCN classification. Many studies have been performed on this topic over recent years, and the majority reported the use of AI improved the differential diagnosis of PCL. In this review, the efficacy of EUS-based AI for the diagnosis of PCL is discussed.

**ARTIFICIAL INTELLIGENCE**

AI enables the analysis of huge numbers of training images and the identification of patterns corresponding to specific clinical features, which can then be used to make diagnoses. Many different algorithms have been devised for this purpose and range in complexity from machine learning (ML) to deep learning (DL).  

**MACHINE LEARNING**

ML is a core discipline of AI and utilizes algorithms that detect patterns within existing data and then uses these patterns to ‘train’ itself to make predictions based on new data. Simple predictive models, such as logistic regression, have been used since the 1960s, but recent decades have witnessed the development of neural network algorithms that predict the risks for various cancers, including breast, lung, and colon cancer. ML uses mathematical algorithms to analyze input values and predict output values without being explicitly programmed to do so. These algorithms are designed to derive different predictive models using a training process during which parameters of an artificial neural network (ANN) model are estimated. The greatest advantage of ML is that it can integrate vast amounts of data and combine observed and predicted quantities in nonlinear and highly interactive ways. ML can be broadly classified based on the type of learning adopted, such as supervised, unsupervised, semi-supervised, or reinforcement learning. Examples of ML include linear discriminants, Bayesian networks, random forest, and support vector machines. Each ML method requires training with input and outcome datasets. In everyday life, ML is used for email spam filtering, online recommendation systems for movies or books, and weather prediction models.

**DEEP LEARNING**

DL differs from ML in that it can handle much larger datasets and has a neural-inspired architectural design. ANNs are AI algorithms that mimic human neurons. Each ANN has an input layer that receives a signal and an output layer that categorizes input signals. Just as neurons weigh action potentials to determine whether to propagate a signal, hidden layers between the input and output layers of ANNs weigh data characteristics and input from other neurons to determine outputs. DL systems are composed of multiple ANNs and can perform all the above-mentioned ML applications but are more complex, can provide decision-making capabilities, and handle extremely large datasets. DL algorithms have been applied to numerous specialties in the medical field. Its most prominent applications involve detecting abnormalities or diseases based on radiologic and pathologic data. DL can accomplish many tasks, such as identifying the location of a lesion of interest and generating differential diagnoses. Furthermore, convoluted neural networks (CNNs) were recently reported to outperform gastroenterologists at identifying pathologies in endoscopic images. CNNs differ from traditional neural networks because they incorporate many different types of layers, such as convolutional layers and pooling layers. CNNs use convolutional layers to extract features from input signals and identify important features. Pooling layers minimizes network size and maintain computer systems at computational thresholds. CNNs are much better at classifying large images than traditional neural networks.

**EUS-BASED ARTIFICIAL INTELLIGENCE AND THE MANAGEMENT OF CYSTIC PANCREATIC LESIONS**

Increased EUS availability has enabled pancreatic cyst surveillance. EUS provides more granularity and pancreatic images with better spatial resolutions than other imaging modalities, including computed tomography and magnetic resonance.
imaging, because of transducer proximity to pancreas. However, the learning curve of EUS is protracted, and thus to acquire competence in all aspects of EUS, it is suggested that clinicians perform 150 supervised cases. Thus, insufficient operator experience can result in misdiagnoses, and even experienced operators can be affected by fatigue due to the durations of some EUS procedures. In practice, the detection of small lesions by EUS is challenging, which makes tumor staging and differential diagnoses difficult because the sonographic appearances of many benign and malignant cystic lesions are similar. Moreover, the increased diagnostic accuracy of EUS-fine needle aspiration (FNA) cystic fluid analysis for differential diagnosis or prognosis prediction is also a concern. Given the amount of information provided by EUS procedures not perceived by the eye and beyond human interpretation, clinicians are developing specific ML and DL algorithms to aid the diagnosis and risk stratification of PCL. Unfortunately, most studies on the use of AI in PCN have used computer tomographic or MRI data, and few have investigated the use of EUS-based AI systems.

1. The application of artificial intelligence to pancreatic cyst differentiation

The first EUS-based AI evaluation of pancreatic lesions was performed in 2001. In this study, the program differentiated focal pancreatitis and malignant in 35 patients with a specificity, sensitivity, and accuracy of 50, 100, and 80%, respectively, versus histopathologic results, which compared with an accuracy of 89% achieved by two endosonographers. This study represented a monumental step in the use of AI, though, at the time, AI-based interpretations were no better than human interpretations. In the early days, ML techniques were mainly used in the AI algorithms used to analyze EUS images. However, thanks to advancements in AI, DL concepts such as CNN are now being applied to assist the differential diagnosis of PCL using EUS images. Nguon et al. developed DL-based algorithms for the differentiation of the MCN and SCN, and analyzed the EUS images of 109 patients (60 MCN, 49 SCN); MCN was defined based on histologic results obtained by EUS-FNA or surgery. An accuracy of up to 82.75% was achieved for differentiating MCN and SCN with an area under the receiver operating curve of 0.88. Vilas-Boas et al. developed a DL-based diagnostic system for differentiating mucinous and non-mucinous cysts using histologic or cystic fluid study results as the gold standard. Mucinous cysts were defined to have calcinoembryonic antigen fluid levels of >192 ng/mL, glucose levels <50 mg/dL, or mucinous epithelial cells by cytology. EUS images from 28 pancreatic cysts were used to train the model, and overall accuracy, sensitivity, and specificity were 98.5, 98.3, and 98.9%, respectively. However, given their various morphologies and heterogenous echogenicities, PCNs have been difficult to assess using AI until recently.

2. The application of artificial intelligence to pancreatic cystic neoplasms malignant risk stratification

AI helps clinicians stratify PCN malignant risk by identifying deeper characteristics in EUS images. Similarly, EUS-delineated cyst morphology has been utilized to develop AI models that differentiate dysplasia in PCNs. In 2019 a retrospective study was conducted using a DL algorithm to analyze 3790 EUS images of IPMN patients with the aim of evaluating the ability of the algorithm to differentiate malignant potential in patients with a pathological diagnosis that underwent resection. The DL program returned malignant potentials of 95.7, 96.2, and 94% for the sensitivity, specificity, and accuracy, respectively. Schultz et al. developed CNNs to differentiate IPMNs with low- or high-grade dysplasia. EUS images were collected from 43 patients with IPMN, as determined by histopathology of surgical specimens. CNN models were used to analyze EUS images, and the algorithm competently classified advanced neoplasia with a surprising accuracy of 99.6%, which was considerably higher than international guidelines (American Gastroenterology Association, American College of Gastroenterology, and Fukuoka guidelines); international guidelines had accuracies ranging from 51.8 to 70.3% for this cohort. The results of AI studies on the differentiation of pancreas cyst and risk stratification are summarized in Table 1.
AI in Pancreatic Cysts

3. EUS-guided needle-based confocal laser endomicroscopy artificial intelligence applications

Confocal laser endomicroscopy (CLE) is a new technique that couples cystic epithelial tissue illumination with real-time, high-resolution microscopic imaging and allows endoscopists to obtain real-time optical biopsies of PCNs in multiple locations. Major trials have established reference standards and safety profiles for the EUS-guided needle-based CLE (EUS-nCLE) of PCNs. This technique provides significant benefits for the differentiation of IPMNs and other PCNs. In addition, certain characteristics of EUS-nCLE enable the prediction of advanced neoplasia, such as high-grade dysplasia or carcinoma, in IPMNs. However, EUS-nCLE video files are often difficult to interpret because they are typically large with multiple frames, which results in inter-observer variations. AI provides the ability to interpret such large amounts of data and detect advanced neoplasia. Furthermore, AI algorithms have been used to stratify EUS-nCLE images. For example, Machicado et al. developed a CNN algorithm to risk-stratify IPMNs. EUS-nCLE video frames from 35 patients with histologically proven IPMNs were used to design a CNN model that measured papillary epithelial thickness and extracted nCLE features for risk stratification. The algorithm had higher accuracy (85.7%), sensitivity (83.3%), and specificity (88.2%) than those required by international guidelines.

4. The applications of artificial intelligence for cystic fluid analysis

AI algorithms have been used indirectly to identify biomarkers in cystic fluid associated with high-grade dysplasia in IPMN. Current clinical guidelines cannot accurately identify patients with IPMN of the pancreas at high risk of pancreatic cancer; thus, a bio-signature that can accurately predict IPMN with high malignant potential is required. Maker et al. used Lasso-penalized logistic regression to analyze the cystic fluid levels of various mRNAs and miRNAs and of KRAS and GNAS in 59 IPMN patients to identify an optimal bio-signature for high-grade dysplasia in IPMN. They found that II.1 β, MUC4, and prostaglandin E synthase 2 mRNA levels most accurately differentiated high- and low-risk IPMN with an area under the curve of up to 0.86 ($p=0.02$). However, few studies have addressed the diagnostic or risk stratification potentials of AI-based genomic studies.

5. Application of artificial intelligence to pancreatic cystic neoplasms treatment decision-making

PCNs are common and often pose a management dilemma because some cysts are precancerous, whereas others have little risk of developing into invasive cancers. Thus, the accurate differentiation of low-risk PCNs and high-risk PCLs is essential to prevent unnecessary pancreatic surgery and for the appropriate management of high-risk patients. AI has demonstrated its ability to assist clinician decision-making by integrating clinical characteristics, imaging features, and genomic analysis results. Springer et al. applied ML to develop the CompCyst algorithm to differentiate patients requiring surgery, routine monitoring, or not requiring further surveillance. CompCyst was trained using the clinical features, imaging characteristics, and the cystic fluid genetic and biochemical markers of 436 patients with pancreatic cyst and

Table 1. EUS-based AI models used for pancreatic cystic neoplasm evaluation

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>Model</th>
<th>Task</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nguon et al.</td>
<td>109</td>
<td>CNN</td>
<td>Differentiate MCN and SCN</td>
<td>83%</td>
</tr>
<tr>
<td>Schults et al.</td>
<td>43</td>
<td>CNN</td>
<td>Differentiate low and high grade IPMN</td>
<td>96.6%</td>
</tr>
<tr>
<td>Kuwahara et al.</td>
<td>50</td>
<td>Deep learning</td>
<td>Determine the malignant potential of IPMN</td>
<td>94%</td>
</tr>
<tr>
<td>Machicado et al.</td>
<td>35</td>
<td>CNN</td>
<td>Differentiate low and high grade IPMN</td>
<td>82%</td>
</tr>
<tr>
<td>Vilas-Boas et al.</td>
<td>28</td>
<td>Deep learning</td>
<td>Differentiate mucinous and non-mucinous cysts</td>
<td>98.5%</td>
</tr>
</tbody>
</table>

AI, artificial intelligence; CNN, convoluted neural network; EUS, endoscopic ultrasound; IPMN, intraductal papillary mucinous neoplasm; MCN, mucinous cystic neoplasm; SCN, serous cystic neoplasm.
then tested it in an independent cohort of 426 patients using histopathology as the gold standard. Clinical management informed by CompCyst was more accurate than management determined using conventional clinical and imaging criteria. In particular, in a subset of patients without advanced neoplasia (n=140), CompCyst recommended surveillance in 68%, whereas standard of care recommended surveillance in 34% (p=0.02). Thus, compliance with CompCyst results would have spared more than half of the patients (60%) that underwent cyst resection.

However, standard management and CompCyst similarly identified lesions requiring surgery with high accuracy (89 and 91%, respectively).

CONCLUSIONS

The use of AI is in its infancy in the medical field, and much effort will be required from clinicians to develop algorithms suitable for clinical practice. New modalities based on the use of AI tools will undoubtedly allow patient risk stratification for malignancy, improve the management of PCNs, and enable clinicians to diagnose pancreas cysts more accurately.

REFERENCES


Conflicts of Interest

The authors have no conflicts to disclose.

ORCID

Jin-Seok Park https://orcid.org/0000-0001-9911-8823
Seok Jeong https://orcid.org/0000-0001-6178-8338

요 약

췌장낭성병변은 최근 영상기술의 발전으로 우연히 발견되는 비율이 점차 증가하고 있으며, 유방암은 복부 컴퓨터단층촬영을 시행한 사람에서 많게는 13.5%까지 보고되었다. 그나지 췌장낭성질환의 정확한 진단은 영상에서 악성 증상까지 다양한 형태로 보일 수 있어 영상학적 진단만으로는 매우 어렵다. 처음파내시경은 췌장낭성병변을 비교적 정확하게 진단하고 치료를 결정하는 데 매우 중요한 도구로 사용된다. 그러나 내시경 초음파는 악성으로 진행 가능한 악성 췌장낭종과 다른 췌장낭성질환을 구분하는 데 정확도가 65-75%에 불과하다. 인공지능은 대장암, 폐암, 유방암과 같은 여러 종류의 암 진단의 정확도를 향상시키는데 효과적인 도구로 사용되고 있으며, 최근 연구에서는 췌장낭성병변에서도 악성성 종양과 비양성성 종양을 구분하고 악성으로 진행 위험도를 평가하는 데 도움이 되는 것으로 보고되고 있다. 인공지능의 적용은 영상분석에도 국한되지 않고 최근에는 췌장낭중의 액체 분석, 유전자 분석, confocal laser endomicroscopy 등 다양한 분야에서 활용되고 있으며 기대되는 연구 결과를 발표하고 있다. 인공지능은 의료 분야에서 아직 시작단계에 있어, 임상에 적용하기 위해 적절한 알고리즘을 개발하는 데에는 개발자들의 큰 노력이 필요하다. 그러나 이러한 기술은 앞으로 췌장낭중 병변을 보다 정확하게 진단하고 효과적이고 효율적으로 관리하는 데 도움을 줄 수 있는 잠재력을 가지고 있다고 하겠다.


