74192_Auto_Edited.docx

Role of artificial intelligence in early detection and screening for pancreatic adenocarcinoma

Kenneth Weicong Lin, Tiing Leong Ang, James Weiquan Li

Abstract

Pancreatic adenocarcinoma remains to be one of the deadliest malignancies in the world despite treatment advancement over the past few decades. Its low survival rates and poor prognosis can be attributed to ambiguity in recommendations for screening and late symptom onset, contributing to its late presentation. In the recent years, Artificial Intelligence (AI) as emerged as a field to aid in the process of clinical decision making. Considerable efforts have been made in the realm of AI to screen for and predict future development of PDAC. This review discusses the use of AI in early detection and screening for pancreatic adenocarcinoma, and factors which may limit its use in a clinical setting.

INTRODUCTION

The global incidence of pancreatic cancer is increasing, and it remains one of the leading causes of cancer-related death, with 495,773 new cases of pancreatic cancer diagnosed and accounting for 466,003 deaths in 2020 [1]. Although the 5-year survival rates for pancreatic ductal adenocarcinoma (PDAC) have improved, it remains low at approximately 9% [2,3], and the overall prognosis of PDAC is poor. This is partly due to the late stage of presentation of PDAC, which is largely dependent on patient symptoms for suspicion of the disease [4,5]. Early cases are asymptomatic and there is a

lack of a simple screening tool for clinical use unlike the case of colorectal cancer screening where screening can be performed in the primary care setting with the use of fecal immunohistochemical test. In the case of PDAC, cross-sectional imaging tests such as computed tomography (CT) or magnetic resonance imaging (MRI) are needed for detection, making widespread population screening unfeasible. Germline mutations and a family history of PDAC have been identified as the strongest risk factors for the disease [6, 7]. As such, efforts in screening programmes have focused their attention on this group of patients [8]. Pancreatic cysts, increased age, and smoking are also known risk factors for PDAC [5, 9, 10], although it may not be practical to conduct routine surveillance for patients with these risk factors. There is an interest in selecting higher risk patients for screening, as the appropriate use biomarkers and imaging may result in detection of early-stage PDAC amenable to curative resection [2, 3, 11-15].

Artificial intelligence (AI) is a branch in computer science where computer systems are designed to perform tasks which would require human intelligence. It is recognised as a potential tool as part of the screening efforts and building predictive models [16]. Most progress for AI in endoscopy has been made in the field of colonoscopy, where polyp detection and characterisation has been studied [17]. Computer-aided diagnosis has also been extended to detection and screening of PDAC [18] in endoscopic ultrasound (EUS) [19, 20], magnetic resonance imaging (MRI) [21] and cytology from fine needle sampling [22]. In recent years, various groups have harnessed the potential of AI in creating prediction models. These include The Felix Project [23], the Pancreatic-Cancer Collective [24], and the Early Detection Research Network [25] effort.

This mini-review aims to study the role of AI in the early detection and screening for pancreatic cancer, as well as factors which may limit its use.

METHODS

A comprehensive literature search was performed in the PubMed, MEDLINE and EMBASE electronic databases from the inception of the databases up to and including 30 November 2021. The key words used were "artificial intelligence", "pancreatic

cancer", "pancreatic adenocarcinoma", "pancreatic ductal adenocarcinoma", "pancreatic carcinoma", "screening", and "early detection". These were supplemented with manual searches of references from retrieved articles. Publications in English were considered for this mini-review.

AI BASIC PRINCIPLES AND TERMINOLOGIES

AI is a term that refers to the ability of a computer programme to imitate the human mind to perform tasks such as problem solving and learning [26, 27].

Machine learning (ML) is the commonest branch of AI used in medicine and refers to a mathematical model that aims to generate a prediction based on a set of data provided [28, 29]. In supervised learning, the data points are labelled and the ML model "learns" from these labels and identifies new data points. In contrast, labels are not provided in unsupervised learning, and the model recognises the patterns of the data by learning its unknown properties and identifying crucial data checkpoints. This is especially important when the gold standard is not available [29].

Deep learning (DL) is subset of ML that employs the use of Artificial Neural Networks (ANN). Like the human brain, ANN consists of layers of artificial neurons that are interlinked. Each layer receives a weighted signal from the previous layer(s) and these signals will be propagated to the next layer when a specific threshold is exceeded [29]. In the setting of a pancreatic lesion or cancer, DL first identifies the basics of the lesion (e.g., location) in its initial layers before moving on to next layer for further characterisation (e.g., size, shape, colour). A final prediction of the pancreatic lesion is made after a systematic assessment *via* multiple layers of neural network [29].

ANNs are first trained using the training data set, where the model learns to identify specific patterns to obtain a relationship between the input and the output. Hyperparameters refer to all settings that are pre-determined by the investigator and are used to construct the model for optimal execution of a particular task or on a specific dataset. The validation data set involves a different data set that is used to fine-tune the hyperparameters of the model. Finally, the test data set refers to a data set

whose purpose is to evaluate the performance of the model against unseen data and determine its generalizability [29]. This set needs to be unseen by the model during training and validation. However in certain studies, the test set is sometimes a subset of the training or validation data set, which many result in overfitting of the model. This may lead to a discrepancy in the performance of the model when tested in the same centre and a decline in performance when validated externally.

MODEL FOR SCREENING FOR AND EARLY IDENFICATION OF DEVELOPING PDAC

Early detection of pancreatic cancer requires a step wise approach in order to systematically screen for risk factors and identify high-risk groups. Figure 1 is a schematic diagram showing the workflow and neural network to be designed for an early detection protocol. It represents the complex interplay between each of the input(s) to be processed for the next neural layer(s) until a final output is obtained. We will be discussing the role of AI in of early detection of pancreatic cancer based on this model.

AI IN CLINICAL DECISION MAKING USING HEALTH RECORDS

The identification of risk factors for pancreatic cancer is essential in identifying the specific population which would benefit from screening [18, 30, 31]. Factors such as diabetes, hemoglobin A1C (HbA1c) value, weight, body mass index (BMI), blood type, smoking status, alcohol use and family history of pancreatic cancer influences the age of onset of screening for an individual [13, 32]. These factors are easily available in the primary care setting and could potentially predict the development of pancreatic cancer within 5 years, even before any changes to the pancreas can be detected on imaging [30]. However, most of the data is stored in health records, which are often proprietary or internet-separated to protect patient data. The retrieval and subsequent integration of data from different platforms remains a manual and laborious process for physicians [30]. Even after retrieval, there are no validated scoring systems to assess these risk

factors and stratify patients. On the other hand, AI, with the aid of Natural Language Processing (NLP), can facilitate this process [33-38]. In a case-control study, Malhotra *et al* created an algorithm based on electronic health records (EHR) obtained from primary care to identify 41.3% of patients (≤ 60 years old) who had significant risk of developing pancreatic cancer up to 20 mo prior to diagnosis with a sensitivity, specificity, area under the receiver operating characteristic (AUROC) curve of 72.5%, 59.0% and 0.66, respectively [33]. Similarly, Appelbaum *et al* was able to train an ANN using 101,381 EHRs to predict the development of PDAC one year before the diagnosis in a population of high-risk patients (AUROC 0.68, confidence interval (CI): 0.65 - 0.71) [35].

Despite its potential benefits, research in AI for the above purpose is still preliminary as they are mostly based on retrospective data from single institutions or registries, and hence not ready for use in a wider clinical setting [33-38]. One of the major limitations would be the lack validation in the real-world setting or at least in populations derived from different centres to overcome the risk of bias and overfitting.

The use of AI in EHR faces other challenges. Various institutions' medical records are built on different healthcare systems and encoding systems, making the task of harmonising them difficult [30]. There is also a lack of standardised clinical research data collection models. To overcome this, efforts are made to build a model of processing and integrating data across institutions. The i2b2 was created to review medical records, retrieve specific data of interest and repurpose it for research [39]. The Observational Health Data Sciences and Informatics was developed from the Observational Medical Outcomes Partnership (OMOP), an initiative that develops the Common Data Model aiming to gather information from different data sets or medical repositories and systemically analyse them in a common platform [40]. Similarly, the National Patient-centered Clinical research network (pcorNet) is another example which was developed in United States to access millions of HER and create a common data set for research purposes [41]. A common dataset with a standardised format for input of data relevant to PDAC would enable AI systems to leverage on big data to identify changing risk profiles in PDAC, enabling the clinician to channel resources for screening to the

appropriate cohorts of patient depending on the population from which this data has been derived.

While these are upcoming and promising initiatives, concerns surrounding restrictions in data sharing, privacy issues, and maintenance costs could hinder data collection efforts [18]. EHRs are also stored in different languages in different regions of the world, making the integration of data difficult. Besides, once data sets are gathered, obtaining IRB approval from the various sites for research may be difficult.

AI AND THE USE OF NON-INVASIVE BIOMARKERS

CA19.9 and CEA are the most widely used markers for screening of PDAC, but have also been proven to lack the specificity when applied individually and without clinical context [42, 43]. On the other hand, a combined measurement can potentially increase its sensitivity and specificity up to 1 year before the diagnosis of PDAC [44-46]. Capitalising on this concept, Yang *et al*, developed an algorithm (with 658 subjects in its training set) to diagnose pancreatic cancer by using ANN to combine CA 19.9, CA125 and CEA values. This model was subsequently evaluated against the test set and was able to yield an AUROC of 0.905 (95%CI 0.868-0.942) and a high diagnostic accuracy of 83.5% for pancreatic cancer [47].

New biomarkers for PDAC such as MicroRNAs and gene expressions have generated much interest in the recent years [45, 48-52]. MircoRNAs are non-coding RNAs that are involved in the regulation of biological pathways, and when altered, could lead to the development of PDAC [53]. MicroRNA can potentially predict future PDAC [54] or detect early stage pancreatic cancer. However, they have the same limitations in sensitivity and specificity when applied without clinical context and as independent test [55, 56]. A combination of the commonly used biomarkers and newer biomarkers may address the problem of low sensitivity and specificity [56], and in particular can be combined with clinical and demographic information as described earlier to increase its usefulness.

While AI is able to make use of plasma microRNA panels and specific gene expressions to diagnose pancreatic cancer [57, 58], studies on their use on predicting future pancreatic

cancer are not available [55]. By integrating Particle Swarm Optimization (PSO), Artificial Neural Network (ANN) and Neighborhood Component Analysis (NCA) iterations on a list of microRNAs that are most commonly expressed by pancreatic cancer, Alizadeh *et al* created a model consisting of 5 MicroRNAs (miR-663a, miR-1469, miR-92a-2-5p, miR-125b-1-3p and miR-532–5p) to diagnose pancreatic cancer (Accuracy: 0.93, Sensitivity: 93%, and Specificity: 92%) [59]. Similarly in a multicentre study by Cao *et al*, a machine learning approach was able to identify 2 panels of microRNAs to differentiate pancreatic cancer from chronic pancreatitis with an accuracy of above 80% [57].

Gene expressions have gained popularity in diagnosing pancreatic cancer [13, 60]. Using a machine learning approach, Khatri *et al* analysed the results from transcriptomics-based meta-analysis to create a nine-gene panel to diagnose pancreatic cancer. This panel was able to differentiate PDAC from chronic pancreatitis with a specificity of 89%, sensitivity of 78%, and accuracy of 83% and an AUROC of 0.95. As compared to a normal pancreas, it was also used to identify stage I and II PDACs with a sensitivity of 74%, specificity of 75%, and an AUROC of 0.82 [61]. In another study, a machine learning algorithm was formulated based on the biochemical differences in the serum of 2 groups of subjects (PDAC group and High risk group) detected *via* the use of Probe Electrospray Ionization Mass Spectrometry (PESI-MS) to identify early stages of pancreatic cancer [62]. It was able to differentiate healthy controls from subjects with earlier stage of PDAC with sensitivity of 81.2% and specificity of 96.8% respectively and an accuracy of 92.9%.

At present, these studies have shown that AI can offer the advantage of identifying specific microRNA and genetic combinations to identifying pancreatic cancer at a faster speed, making this process less laborious. However, these studies lack external validation, limiting their application in modern practice. Besides, studies utilising AI to formulate specific sequences to accurately predict future pancreatic cancer development are still lacking. More studies are required to analyse its ability in predicting future pancreatic cancer for high risk groups especially during the latency period.

CURRENT EVIDENCE IN PREDICTING THE DEVELOPMENT OF PANCREATIC LESIONS INTO PDAC IN THE FUTURE[KL1]

Various studies have been conducted using AI to diagnose pancreatic cancer and has yield promising results. Table 1 summarises these studies so far. In a retrospective study, Liu et. al was able to train a CNN to identify pancreatic cancer on contrast-enhanced CT and achieve an AUROC of 0.9, with more than 90% for its sensitivity and specificity for its test set. It maintained good sensitivity of 91.3%, specificity of 84.5%, an accuracy of 85.6% and AUROC of 0.955 (95%CI 0.955–0.956) with the validation set [63]. Further analysis revealed that with CNN, radiologists missed 7% of the pancreatic cancers, of which majority were accurately diagnosed by CNN [63]. By enhancing the CNN, Liu et. al was able to process the CT images and obtain the diagnosis faster than the radiologists (3 s for CNN VS 8 mins for a radiologist) with an AUROC of 0.9632 [64], proving that AI is comparable to radiologists.

Besides CT, EUS has been frequently utilised to diagnosed pancreatic cancer. Table 2 summaries these studies so far. The EUS-CAD based CNN was developed in a retrospective study by Tonozuka *et al* to identify lesions harbouring pancreatic cancer in patients with chronic pancreatitis with a sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of 90.2%, 74.9%, 80.1%, and 88.7%, respectively, and an AUROC of 0.924 [65]. Similar findings were also echoed in Zhu et. al who utilised SVM to obtain a sensitivity, specificity, PPV and NPV of over 90% for diagnosis of pancreatic cancer in chronic pancreatitis [66].

Despite numerous studies looking at using AI to diagnose pancreatic cancer (as shown in Table 1 and Table 2), only a few attempted to predict the development to pancreatic cancer. On average, CT changes for early pancreatic cancer starts approximately 12 to 18 mo before diagnosis [67]. Yet, pancreatic cancer can advance from being undetectable to metastatic in a short period of time even before the next surveillance imaging [68, 69]. AI-based imaging itself cannot be used to predict pancreatic cancer and should be combined with other markers.

An ideal AI model for predicting pancreatic cancer is one that integrates multiple biochemical and radiological and clinical data [70]. In a retrospective proof-of-concept study, Springer et. al. developed a supervised machine learning-based approach (CompCyst) based on a combination of patient-reported symptoms, imaging results (including CT, MRI and EUS images), cyst fluid and molecular characteristics to calculate its malignant potential and subsequently determine the management of pancreatic cyst(s). When tested against the validation set, CompCyst outperformed the current standard of care (accuracy 56%) in its ability to identify patients who required surgery, close monitoring or can be discharged (accuracy 69%) [71]. CompCyst correctly identified 60% of the surgeries that were not warranted and could have been avoided, while not compromising on its ability to identifying those who truly require surgery. With CompCyst, 71% of the pancreatic lesions were correctly identified as PDAC as compared to 58% based on clinical suspicion [71].

While this study has proven that AI has the potential to incorporate various clinical characteristics, biomarkers, and imaging characteristics to assess for the malignant potential of a pancreatic lesion, it has a number of limitations. Firstly, the imaging characteristics and molecular biomarkers that were identified as high risk features were obtained at the time of surgery and not during screening. These features may not be present early enough to be identified by routine screening. Secondly, important risk factors (including age and diabetes) that were crucial in the early detection of PDAC (as shown in Figure 1) were not included in its learning process, representing a missed step in the screening process. Finally, CompCyst is yet to be externally validated and cannot be applied to the clinical setting currently.

While CompCyst is a potential tool to aid in clinical decision making, future studies aiming at early detection of PDAC face a myriad of challenges. Firstly, the pancreas is a complex organ. Unlike the other organs, the pancreas can be highly variable in its anatomy and location. Moreover, the training data set is highly dependent on the quality of the images provided. Hence, automated segmentation of the pancreas *via* a deep learning approach remains challenging [72]. Secondly, the lack of databases limits

the ability to develop new training sets. There are currently only a few open-access databases [73], and there are issues regarding sharing of images across various institutions as pointed out by the Alliance of Pancreatic Cancer Consortia (APaCC) imaging working group [70]. Finally, the algorithm for early detection of PDAC will have to evaluate images of pancreatic lesion(s) across different time points of surveillance and from different 3 imaging modalities (namely CT, MRI, and EUS). Unlike CompCyst which looks at images at one time point (i.e. at surgery), combining multiple images obtained from periodical surveillance *via* these 3 imaging modalities will require a very large data base and multiple layers.

There is a major gap that needs to be bridged before AI systems for early detection of pancreatic cancer can be developed. Given sufficient training data and cooperation, AI-based image analyzers could match or even outperform physicians in image classification and lesion detection [70].

[KL1]Revision of subtitle

CONCLUSION

Despite the recent advances to predict future PDAC, the use of AI in screening for pancreatic cancer remains limited in the clinical setting. Much of the efforts are made in the research setting and lack external validation and generalisability. However, this field remains promising as we recognise the challenges ahead to bridge the necessary gaps. The hope to develop an integrated AI model to screen for PDAC remains a reality, and it will play a complementary role in assisting physicians in their clinical decision making process but not replace it.

74192_Auto_Edited.docx

5%

SIMILARITY INDEX

PRIMARY SOURCES

1   www.ncbi.nlm.nih.gov
    Internet   43 words — 1%

2   www.wjgnet.com
    Internet   38 words — 1%

3   Matthew R. Young, Natalie Abrams, Sharmistha   32 words — 1%
    Ghosh, Jo Ann S. Rinaudo, Guillermo Marquez, Sudhir
    Srivastava. "Prediagnostic Image Data, Artificial Intelligence,
    and Pancreatic Cancer", Pancreas, 2020
    Crossref

4   Behrouz Alizadeh Savareh, Hamid Asadzadeh   19 words — 1%
    Aghdaie, Ali Behmanesh, Azadeh Bashiri et al. "A
    machine learning approach identified a diagnostic model for
    pancreatic cancer through using circulating microRNA
    signatures", Pancreatology, 2020
    Crossref

5   www.pubfacts.com
    Internet   17 words — < 1%

6   Ananya Malhotra, Bernard Rachet, Audrey   14 words — < 1%
    Bonaventure, Stephen P. Pereira, Laura M.
    Woods. "Can we screen for pancreatic cancer? Identifying a
    sub-population of patients at high risk of subsequent diagnosis
    using machine learning techniques applied to primary care
    data", PLOS ONE, 2021

7 **cwww.intechopen.com**
Internet

14 words — < 1%