ASIDE Gastroenterology

ASIDE GASTROENTEROLOGY

Review Article

Artificial Intelligence-Assisted EUS for Risk Stratification of Pancreatic Cystic Lesions: A Narrative Review of Current Evidence and Future Directions for Predicting High-Grade Dysplasia, Invasive Cancer, and the Need for Surgical Referral

Ahmed Salman[1,*]

1-Internal Medicine Department, Faculty of Medicine, Cairo University, Cairo, Egypt

ARTICLE INFO

Article history:

Received 7 May 2026

Received in revised form 20 Jun. 2026

Accepted 17 Jul. 2026

Published 28 Aug. 2026

Keywords:
Pancreatic cancer

High-grade dysplasia

Pancreatic cystic lesions

Endoscopic ultrasound

Artificial intelligence

Intraductal papillary mucinous neoplasm

ABSTRACT

Background: Pancreatic cystic lesions are increasingly detected on cross-sectional imaging, yet accurate risk stratification remains challenging. A clinically important minority, particularly intraductal papillary mucinous neoplasms and mucinous cystic neoplasms, may progress to high-grade dysplasia or invasive carcinoma. EUS plays a central role in cyst evaluation, but interpretation remains operator-dependent, and guideline-based risk categories have imperfect predictive accuracy. **Methods:** This narrative review was conducted using a structured literature search of PubMed, Scopus, Web of Science, and Embase with terms including artificial intelligence, deep learning, machine learning, endoscopic ultrasound, pancreatic cystic lesions, intraductal papillary mucinous neoplasm, high-grade dysplasia, and pancreatic cancer. Studies reporting AI applications in EUS image interpretation, multimodal risk prediction, radiomics, and cyst-fluid or molecular marker integration were included. A qualitative synthesis was performed given the heterogeneity of study designs and AI methodologies. **Results:** AI may enhance EUS-based assessment by extracting quantitative imaging features and integrating EUS data with CT, MRI, cyst-fluid biomarkers, cytology, molecular markers, and longitudinal cyst behavior. The most clinically meaningful goal — though not yet established — is individualized prediction of high-grade dysplasia and invasive cancer to support surgical referral decisions, particularly where guideline-based criteria are discordant or borderline. Current models remain preliminary and require prospective validation. Retrospective designs, small datasets, lack of external validation, and uncertainty regarding explainability and clinical integration limit evidence. **Conclusions:** Future prospective multicentre studies should determine whether AI-assisted EUS improves patient-centered outcomes by reducing unnecessary surgery while preventing delayed diagnosis of advanced neoplasia.

1. Introduction

Pancreatic cystic lesions are increasingly identified due to widespread use of cross-sectional abdominal imaging and aging populations. Although many pancreatic cysts are benign or indolent, a clinically important subset — particularly intraductal papillary mucinous neoplasms and mucinous cystic neoplasms — carries a risk of progression to high-grade dysplasia or invasive pancreatic cancer. The central challenge in clinical practice is to distinguish patients requiring surgery from those who can be safely managed with surveillance [1–3].

Current management relies on guideline-based risk stratification using clinical, radiological, and endoscopic features. The American Gastroenterological Association guideline recommends EUS-FNA in cysts with at least two high-risk features, including size ≥3 cm, a dilated main pancreatic duct, or an associated solid component [1]. European guidelines similarly emphasize absolute and relative indications for surgery, including jaundice, mural nodule enhancement, positive cytology, main pancreatic duct dilation, cyst size, growth rate, and symptoms [2]. The Kyoto guidelines refined IPMN management by emphasizing high-risk stigmata, worrisome features, surveillance intervals, and the need to balance cancer prevention with the risk of unnecessary pancreatic resection [3]. Despite these frameworks, risk prediction remains imperfect, and interobserver variability in cyst characterization continues to affect decision-making [1–3].

EUS plays a central role in evaluating pancreatic cystic lesions, providing high-resolution assessment of cyst morphology, mural nodules, septations, duct communication, solid components, and associated lymphadenopathy. EUS also enables cyst fluid aspiration for cytology, carcinoembryonic antigen (CEA), glucose, amylase, and molecular testing, thereby complementing CT and MRI findings [2–4]. However, EUS interpretation is operator-dependent, and features such as small mural nodules, subtle wall thickening, mucinous morphology, and early malignant transformation may be difficult to identify consistently. Cyst-fluid cytology has limited sensitivity, while molecular testing is not universally available and may not always provide definitive guidance [2–4].

* Corresponding author: Ahmed Salman, Internal Medicine Department, Faculty of Medicine, Cairo University, Cairo, Egypt. Email: awea844@gmail.com

Published and owned by PubPorta Publishing LLC. Academic oversight and scholarly guidance are provided by the American Society for Inclusion, Diversity, and Equity in Healthcare (ASIDE). ISSN (Print): 3066-4004, ISSN (Online): 3066-4012. Hosting by ASIDE Journals.

Citation: Salman A. Artificial Intelligence-Assisted EUS for Risk Stratification of Pancreatic Cystic Lesions: A Narrative Review of Current Evidence and Future Directions for Predicting High-Grade Dysplasia, Invasive Cancer, and the Need for Surgical Referral. ASIDE Gastroenterol. 2026;2(4):28-39, doi:10.71079/ASIDE.GI.082826830

AI has emerged as a potential tool to improve pancreatic cyst assessment by extracting quantitative imaging features, identifying subtle patterns beyond human visual perception, and integrating multimodal data into individualized risk prediction models. The most clinically meaningful AI application is the prediction of advanced neoplasia — specifically, high-grade dysplasia and invasive carcinoma — rather than simple cyst classification. Although high-grade dysplasia and invasive carcinoma are frequently grouped as advanced neoplasia, they differ in prognosis, operative urgency, and postoperative management; this review therefore distinguishes them wherever the clinical distinction is relevant and uses advanced neoplasia only as a deliberate composite term. This distinction is important because the ultimate purpose of risk stratification is to determine whether a patient should undergo surgery, undergo EUS-guided sampling, receive intensified surveillance, or undergo conservative follow-up [5, 6].

AI-assisted EUS may be particularly valuable, given that EUS is already a critical decision point in pancreatic cyst management. A clinically useful AI system could assist in detecting mural nodules, distinguishing mucinous from non-mucinous cysts, predicting malignant potential, and identifying patients whose predicted risk of advanced neoplasia may warrant surgical referral. Future systems may integrate real-time EUS images with cross-sectional imaging, cyst-fluid biomarkers, cytology, molecular alterations such as KRAS, GNAS, TP53, SMAD4, and CDKN2A, serum CA19-9, symptoms, cyst growth rate, age, comorbidity, and surgical fitness. Such multimodal models could shift pancreatic cyst management from categorical, guideline-based assessment to personalized, probability-based decision-making [4–6].

Existing reviews have addressed AI applications in pancreatic imaging broadly, radiomics-based risk prediction using CT and MRI, and machine-learning models for IPMN classification [5, 6]. However, no recent narrative review has focused specifically on AI-assisted EUS as a decision-support tool for predicting clinically meaningful outcomes — namely, high-grade dysplasia, invasive carcinoma, and the need for surgical referral — or distinguished between AI for EUS image interpretation and multimodal AI risk prediction as separate problems with different data requirements, validation standards, and readiness for clinical implementation. Given the rapid growth of this literature and the central role of EUS in pancreatic cyst management, a focused synthesis is now needed to clarify what has been achieved, what remains investigational, and what future studies must demonstrate before AI-assisted EUS can be responsibly integrated into clinical practice.

This review addresses two related but distinct roles of AI in pancreatic cyst management. The first is AI-assisted interpretation of EUS and nCLE images, where algorithms analyze cyst morphology, detect mural nodules, and classify cyst subtypes from imaging data alone. The second, and clinically more ambitious, is multimodal AI risk prediction, in which EUS-derived image features are integrated with cross-sectional imaging, cyst-fluid biomarkers, cytology, molecular markers, serum markers, and patient factors to generate individualized estimates of advanced neoplasia to inform surveillance or surgical referral decisions.

Cross-sectional imaging, cyst-fluid biomarkers, molecular markers, cytology, and clinical variables are discussed not as independent topics, but in the context of how they may complement and enhance EUS-based AI risk prediction. CT- or MRI-based radiomics and molecular testing are included only where they directly inform EUS-centered decision-making or contribute to integrated risk prediction models. The review will summarize current EUS-based risk stratification, emerging AI applications in pancreatic cyst assessment, multimodal prediction models, potential roles in surveillance and surgical decision-making, and the limitations that must be addressed before AI can be safely integrated into routine clinical practice.

2. Methods

This narrative review was conducted in accordance with the principles of transparent narrative synthesis. A structured narrative literature search was performed across four electronic databases — PubMed, Scopus, Web of Science, and Embase — from February 2000 to April 2025, which served as the final search date. The following search terms were used individually and in combination: artificial intelligence, machine learning, deep learning, neural network, endoscopic ultrasound, EUS, needle-based confocal laser endomicroscopy, nCLE, pancreatic cystic lesions, pancreatic cysts, intraductal papillary mucinous neoplasm, IPMN, mucinous cystic neoplasm, serous cystadenoma, high-grade dysplasia, invasive carcinoma, pancreatic cancer, radiomics, cyst-fluid biomarkers, molecular markers, risk stratification, and surgical decision-making. Reference lists of retrieved articles were also hand-searched to identify additional studies not captured by the database search. Within each database, these terms were grouped into two concept blocks — an artificial-intelligence block and a pancreatic-cyst/endoscopic-ultrasound block — combined using the Boolean operator OR within blocks and AND between blocks.

Studies were eligible if they reported original data or substantive discussion of AI applications in EUS image interpretation, nCLE analysis, multimodal risk prediction, radiomics, or cyst-fluid and molecular marker integration in pancreatic cystic lesion assessment. Studies were excluded if published in languages other than English, were conference abstracts without full-text availability, or did not specifically address pancreatic cystic lesions. No restrictions were placed on the study design, given the early and heterogeneous state of the literature. Records retrieved from all four databases were collated and de-duplicated; titles and abstracts were then screened against the eligibility criteria, and the full texts of potentially relevant articles were assessed for final inclusion. As this is a single-author narrative review, screening and selection were performed by the author.

Given the heterogeneity of study designs, AI methodologies, outcome definitions, and reported performance metrics, a formal meta-analysis was not performed. As this is a narrative review, formal risk-of-bias assessment using tools such as QUADAS-2 or PROBAST was not applied. However, methodological limitations — including retrospective design, small sample size, single-center recruitment, image selection bias, lack of external validation, and spectrum bias from surgically resected cohorts — are critically discussed where relevant throughout.

A key methodological challenge is the inconsistent use of outcome terminology across studies. The terms "malignancy," "advanced neoplasia," "high-grade dysplasia," "invasive carcinoma," and "support for surgical referral decision-making" are not interchangeable, yet are frequently used inconsistently across the reviewed studies. For this review, the following definitions apply. High-grade dysplasia refers to severely dysplastic epithelium without stromal invasion, confirmed on surgical pathology. Invasive carcinoma refers to pancreatic ductal adenocarcinoma arising in association with a cystic lesion, with histological evidence of stromal invasion. Advanced neoplasia is used as a composite term encompassing both high-grade dysplasia and invasive carcinoma, consistent with its use in the IPMN literature and major international guidelines [1–3]. Malignancy is used only when the reviewed study did not distinguish between high-grade dysplasia and invasive carcinoma, or when the term

reflects the language of the original publication. Support for surgical referral decision-making refers to the use of predicted biological risk, patient fitness, and the expected balance of surgical benefit against operative morbidity to inform whether a patient should be referred for consideration of resection, rather than to a histopathological endpoint per se. Because most reviewed models predict pathology or advanced neoplasia rather than a validated clinical decision to operate, we use the phrase "support for surgical referral decision-making" in preference to "surgical need" throughout this review, except where an individual study specifically evaluated surgical appropriateness. It should also be noted that the majority of AI evidence reviewed pertains specifically to IPMN, which carries the highest malignant potential among pancreatic cystic lesions. Mucinous cystic neoplasms, serous cystadenomas, pseudocysts, and cystic neuroendocrine tumors have distinct biology, natural history, and management thresholds; AI findings from IPMN cohorts should not be extrapolated to these entities without specific validation. Where individual studies used different or overlapping definitions, this is noted in the text, and original terminology is preserved to avoid misrepresentation.

Given the well-recognized methodological vulnerabilities of medical AI studies, the included studies were appraised narratively across the following domains. First, dataset characteristics were examined, including sample size, single- versus multi-center recruitment, consecutive versus selected patient inclusion, and proportion of high-risk or surgically resected cases. Second, model development and validation methodology was assessed, including whether internal, external, or prospective validation was performed, and whether appropriate train-test splitting or cross-validation procedures were described. Third, the risk of overfitting was assessed relative to dataset size and model complexity, particularly for deep learning models trained on small image datasets. Fourth, the risk of data leakage was noted where training and test sets were not clearly separated or where image-level rather than patient-level splitting was performed. Fifth, image selection methodology was examined — whether AI models were trained on curated still images or consecutive real-time examination frames. Sixth, class imbalance was considered, in which the proportion of high-risk or malignant cases substantially differed from the real-world surveillance prevalence. Seventh, performance metrics were evaluated critically, with preference given to studies reporting calibration, discrimination, and clinically relevant endpoints rather than accuracy alone. This appraisal was applied narratively rather than as a formal scoring instrument, and the findings are discussed in the relevant sections and summarised in the limitations section.

2.1. Current EUS-Based Risk Stratification of Pancreatic Cystic Lesions

Risk stratification of pancreatic cystic lesions is primarily aimed at identifying cysts that harbor, or are likely to progress to, high-grade dysplasia or invasive carcinoma. In clinical practice, this assessment relies on integrating patient factors, cross-sectional imaging, EUS morphology, cyst-fluid analysis, cytology, and, increasingly, molecular markers. EUS is usually reserved for cysts with concerning features on CT or MRI, indeterminate morphology, suspected mural nodules, or when cyst-fluid analysis may alter management [1–4].

The most important EUS-detectable high-risk features include enhancing mural nodules, solid components, main pancreatic duct dilation, abrupt change in duct caliber, suspicious lymphadenopathy, thickened cyst wall, and features suggestive of invasive disease. Mural nodules are particularly important as they may represent dysplastic epithelial proliferation and are strongly associated with advanced neoplasia in IPMN. EUS is often superior to CT or MRI in detecting small mural nodules and intracystic solid components, although interpretation may be affected by mucin, debris, blood clot, or artifacts [2–4, 7].

Cyst size remains an important but imperfect marker of risk. Larger cysts, particularly those ≥3 cm or ≥4 cm depending on the guideline used, are associated with higher malignant potential, but size alone has limited specificity. Some large cysts remain indolent for many years, while smaller lesions may harbor advanced dysplasia if other high-risk features are present. Growth rate may provide useful longitudinal information, particularly when rapid enlargement occurs during surveillance. Growth thresholds differ across guidelines, and measurement variability between imaging modalities can complicate interpretation [1–3, 8].

Main pancreatic duct involvement is a key determinant of risk. Main-duct and mixed-type IPMNs carry a higher risk of high-grade dysplasia and invasive carcinoma than branch-duct IPMNs. EUS can assess duct communication, duct diameter, mural nodules within the duct, and associated parenchymal changes. Distinguishing branch-duct from mixed-type disease may be difficult when duct dilation is borderline or multifocal cysts are present. Accurate classification is clinically important as it directly influences the threshold for surgery [2, 3, 8].

EUS-guided cyst-fluid analysis adds biochemical, cytological, and molecular information to morphological assessment. CEA has traditionally been used to distinguish mucinous from non-mucinous cysts, although it does not reliably predict malignancy. Cyst-fluid glucose has emerged as a useful marker for mucinous differentiation and may perform similarly or better than CEA in some studies. Amylase may suggest pancreatic duct communication, but it is not specific. Cytology has high specificity when malignant or high-grade atypical cells are identified, but its sensitivity is limited in hypocellular samples [4, 9, 10].

Molecular testing can further refine risk assessment. KRAS and GNAS mutations support a diagnosis of mucinous cyst or IPMN, whereas alterations in TP53, SMAD4, CDKN2A, PIK3CA, PTEN, and aneuploidy are more concerning for advanced neoplasia. These markers are particularly valuable when imaging and cytology are equivocal. Molecular testing is limited by availability, cost, assay variability, and uncertainty about incorporation into routine surgical decision-making [3, 4, 11, 12].

Despite advances in guideline-based assessment, current EUS-based risk stratification remains limited by subjectivity and imperfect predictive accuracy. Interobserver variability affects recognition of mural nodules, duct communication, wall thickening, and solid components. Most guideline criteria are categorical rather than probabilistic, meaning patients may be placed into broad risk groups that do not fully account for the combined effect of imaging, biomarkers, molecular alterations, age, comorbidity, surgical fitness, and longitudinal cyst behavior. These limitations provide a strong rationale for AI-assisted EUS models that can integrate multiple variables and generate individualized predictions of high-grade dysplasia and invasive cancer to support surgical referral decisions [5, 6, 12].

A critical distinction exists between two roles AI may play relative to existing guidelines. In feature engineering, AI automates measurement of predefined guideline criteria — cyst size, duct diameter, and mural nodule detection — improving consistency without challenging the evidence base; explainability is less problematic as model inputs map directly onto established clinical variables.

Table 1: Key Variables Used in EUS-Based Risk Stratification of Pancreatic Cystic Lesions

Domain	Variable	Clinical relevance	Main limitation
Clinical features	Obstructive jaundice, pancreatitis, symptoms	May indicate high-risk disease or cyst-related complications [2, 3, 8]	May be nonspecific or unrelated to the cyst
Serum marker	CA19-9	Supports concern for invasive carcinoma when elevated [2, 3]	False elevation with cholestasis or inflammation
EUS morphology	Mural nodule or solid component	Strong predictors of advanced neoplasia and surgical consideration [2–4, 8]	May be confused with mucin, debris, or artifact
EUS morphology	Thickened/enhancing cyst wall	Suggests higher-risk morphology [2, 3]	Subjective and operator-dependent
Ductal features	Main pancreatic duct dilation or abrupt caliber change	Important risk features, especially in main-duct or mixed-type IPMN [2, 3, 8]	Thresholds vary between guidelines
Cyst behaviour	Size and growth rate	Larger or rapidly growing cysts may need closer surveillance or intervention [1–3, 8]	Size alone has limited specificity.
Cyst-fluid markers	CEA, glucose, amylase	Help distinguish mucinous from non-mucinous cysts and duct communication [4, 7, 9, 10]	Limited ability to predict high-grade dysplasia or cancer
Cytology	High-grade atypia or malignant cells	Highly specific for advanced neoplasia when present [2–4]	Low sensitivity due to hypocellular samples
Molecular markers	KRAS, GNAS, TP53, SMAD4, CDKN2A, PIK3CA, aneuploidy	Improve cyst classification and prediction of advanced neoplasia [3, 4, 11, 12]	Cost, availability, and assay variability
Patient context	Age, comorbidity, surgical fitness	Essential when balancing surveillance versus surgery [2, 3]	Not always included in cyst-risk algorithms

Abbreviations: CA19-9, carbohydrate antigen 19-9; CDKN2A, cyclin-dependent kinase inhibitor 2A; CEA, carcinoembryonic antigen; EUS, endoscopic ultrasound; GNAS, GNAS complex locus; IPMN, intraductal papillary mucinous neoplasm; KRAS, KRAS proto-oncogene, GTPase; PIK3CA, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha; SMAD4, SMAD family member 4; TP53, tumor protein p53.

In representation learning, deep learning models identify imaging patterns beyond current human classification, raising fundamental questions about transparency, biological plausibility, and the validity of acting on uninterpretable features. When a representation learning model conflicts with guideline-based assessment, the medicolegal implications are substantial — it remains unclear which should take precedence, who bears responsibility, and how conflicts should be communicated to patients. In current practice, where AI outputs conflict with established guideline criteria, the guideline-based assessment and multidisciplinary team judgment should take precedence, and AI-generated risk estimates should be treated as supportive rather than determinative until they have been prospectively validated. Future regulatory frameworks must explicitly address this tension before such models can be safely deployed.

The major EUS-based, cyst-fluid, cytological, molecular, and clinical variables currently used for pancreatic cyst risk stratification are summarised in **Table 1**.

2.2. AI for EUS Image Interpretation of Pancreatic Cystic Lesions

Artificial intelligence has the potential to enhance EUS-based assessment of pancreatic cystic lesions by reducing subjectivity and extracting image features that human observers may not consistently recognize. It should be noted that most published AI models have been developed and validated specifically in IPMN cohorts; evidence relating to mucinous cystic neoplasms, serous cystadenomas, pseudocysts, and cystic neuroendocrine tumors remains substantially more limited and should not be assumed to follow the same performance characteristics. Conventional EUS interpretation depends heavily on operator experience, image quality, lesion location, and the ability to distinguish true high-risk features from mimics such as mucus, debris, hemorrhage, or artifact. AI-based image analysis may support the endosonographer by improving recognition of cyst morphology, mural nodules, septations, solid components, duct communication, and features suggestive of mucinous or malignant pathology [13, 14].

AI for standard EUS image interpretation

Early AI applications in pancreatic cystic lesions have largely focused on classification tasks, including differentiating mucinous from non-mucinous cysts and identifying cyst subtypes such as IPMN, mucinous cystic neoplasm, serous cystadenoma, and cystic neuroendocrine tumor. Deep learning models can be trained to identify complex visual patterns in EUS images, potentially improving diagnostic consistency when cyst morphology is indeterminate. This may be particularly useful in branch-duct IPMN, where the presence or absence of mural nodules, solid components, and duct involvement can substantially alter management [14, 15].

AI for needle-based confocal laser endomicroscopy

Needle-based confocal laser endomicroscopy has become an important platform for AI development in assessing pancreatic cysts. Because nCLE provides real-time microscopic imaging of the intracystic epithelium during EUS, it produces highly detailed images suitable for computer-assisted interpretation. CNN-based analysis of nCLE images has shown promising performance in differentiating pancreatic cystic lesions and stratifying IPMN risk, suggesting that AI may help convert subjective optical interpretation into more reproducible image-based diagnosis [16, 17].

Beyond cyst subtype classification, a more clinically meaningful role for AI is predicting high-grade dysplasia and invasive carcinoma. In IPMN, the key question is not simply whether a cyst is mucinous, but whether it harbors high-grade dysplasia or invasive carcinoma. AI models may support this task by integrating EUS image features with clinical variables, cyst-fluid findings, cytology, and molecular data to generate individualized risk estimates, rather than relying on categorical features such as cyst size or duct diameter alone [18, 19].

However, current evidence remains preliminary. Many AI studies in pancreatic cystic lesions are retrospective, single-center, and based on small image datasets. Model performance may be affected by image selection bias, inconsistent annotation, differences in EUS platforms, and limited external validation. A recent systematic review found that most studies used small datasets and focused on differential diagnosis or risk stratification rather than real-time clinical implementation. AI-assisted EUS should therefore be viewed as investigational and supportive rather than a replacement for expert endosonographic assessment [20].

A fundamental limitation of current AI studies in EUS and nCLE is the near-exclusive reliance on curated still-image datasets rather than the full video stream generated during a real examination [16, 17, 20, 21]. EUS and nCLE are inherently dynamic procedures: diagnostic information emerges from continuous probe movement, real-time acoustic adjustment, and sequential image acquisition across the lesion. Selecting representative frames introduces sampling bias, may miss intermittently visible features such as small mural nodules or subtle wall thickening, and does not reflect the conditions under which AI would need to perform in clinical practice [20, 21]. Video-based AI analysis, using recurrent neural networks, temporal convolutional models, or transformer-based architectures, would more accurately represent the examination as it occurs. nCLE in particular generates continuous microscopic video that is inherently temporal; analyzing this as isolated static frames discards the tissue dynamics, probe movement patterns, and evolving image quality that characterize real-time interpretation [16, 17]. Future AI systems should therefore be developed and validated using full procedural video, with performance assessed in real-time or near-real-time settings.

An ideal AI platform would not only characterize cyst morphology but also quantify the probability of high-grade dysplasia or invasive cancer, highlight the image regions driving the prediction, identify uncertainty, and recommend whether further sampling, short-interval surveillance, multidisciplinary review, or surgical referral should be considered. In this way, AI-assisted EUS could evolve from a diagnostic image classifier into a clinically useful decision-support tool for pancreatic cyst risk stratification [20].

2.3. Predicting High-Grade Dysplasia and Invasive Cancer: Implications for Surgical Referral

Several predictive models have aimed to quantify risk in the assessment of pancreatic cysts. Nomogram approaches combining clinical and imaging variables have been used to predict high-grade dysplasia or invasive carcinoma in IPMN, supporting the principle that risk should be evaluated using integrated models rather than single features [22–24].

More recent machine-learning studies have examined whether AI can enhance preoperative prediction of malignant IPMN. Kang et al. studied a multinational cohort of 828 patients with resected IPMN, comparing logistic regression against random forest, support vector machine, and artificial neural network models [25]. All demonstrated comparable discrimination, with AUC values ranging from 0.79 to 0.81; no model achieved statistically superior performance over logistic regression, and confidence intervals overlapped across methods. External validation was not performed, and the cohort was restricted to surgically resected cases, limiting generalisability to the broader surveillance population. These findings suggest that AI is not automatically superior to conventional statistical modeling, but may provide a useful framework when larger, richer datasets incorporating variables beyond standard preoperative imaging become available [25].

AI may be particularly valuable when the clinical question is not whether malignancy exists, but whether surgery is appropriate. Pancreatic surgery carries substantial morbidity, and a significant proportion of resected cysts harbor low-grade dysplasia or benign pathology. Conversely, delayed surgery in patients with high-grade dysplasia or early invasive carcinoma may forfeit the opportunity for curative resection. A well-designed AI model should therefore assess whether the expected benefit of surgery outweighs the risks of continued surveillance, integrating predicted histological risk with patient age, frailty, comorbidity, operative fitness, life expectancy, and patient preference [26, 27].

While low-grade dysplasia, high-grade dysplasia, minimally invasive carcinoma, and advanced invasive carcinoma are clinically distinct, many models group high-grade dysplasia and invasive cancer as a composite "advanced neoplasia" endpoint. Although appropriate for general surgical decision support, future models should ideally distinguish high-grade dysplasia from invasive carcinoma, and early from advanced invasive disease. This distinction has implications for counseling, operative planning, lymph node assessment, and postoperative surveillance [27].

Explainability is particularly important in this context, given that AI predictions may influence decisions about major pancreatic surgery. Models that identify the most relevant contributing variables — such as mural nodule size, duct diameter, serum CA19-9, cyst location, and molecular alterations — are more likely to gain clinical acceptance. Interpretability also assists clinicians in judging when a prediction is biologically plausible, when it conflicts with clinical judgment, and when additional imaging, EUS-guided sampling, or expert review is warranted [27, 28].

AI-based prediction of advanced neoplasia to guide surgical referral remains an evolving area rather than an established standard of care. Most models are developed using surgically resected cohorts that over-represent high-risk cysts and may not generalize to the broader surveillance population, creating spectrum bias and potentially inflating apparent performance. Models trained on retrospective datasets may also perform less reliably in patients with small incidental cysts, multifocal disease, equivocal imaging, or long-term stability. Prospective validation in real-world cohorts against clinically relevant endpoints — including unnecessary surgery avoided, cancers missed, progression during surveillance, and survival benefit — is required before AI can meaningfully guide surgical referral decisions [22, 28].

2.4. Multimodal AI and Personalized Risk Stratification

The future of AI-assisted EUS for pancreatic cystic lesions is unlikely to depend solely on EUS images. Pancreatic cyst risk is biological, radiological, biochemical, molecular, and patient-specific; the most useful AI systems will therefore likely be multimodal, integrating EUS findings with CT or MRI features, cyst-fluid biomarkers, cytology, molecular alterations, serum markers, symptoms, cyst growth rate, and patient-level factors to generate a personalised estimate of high-grade dysplasia and invasive carcinoma to inform surgical referral consideration [29].

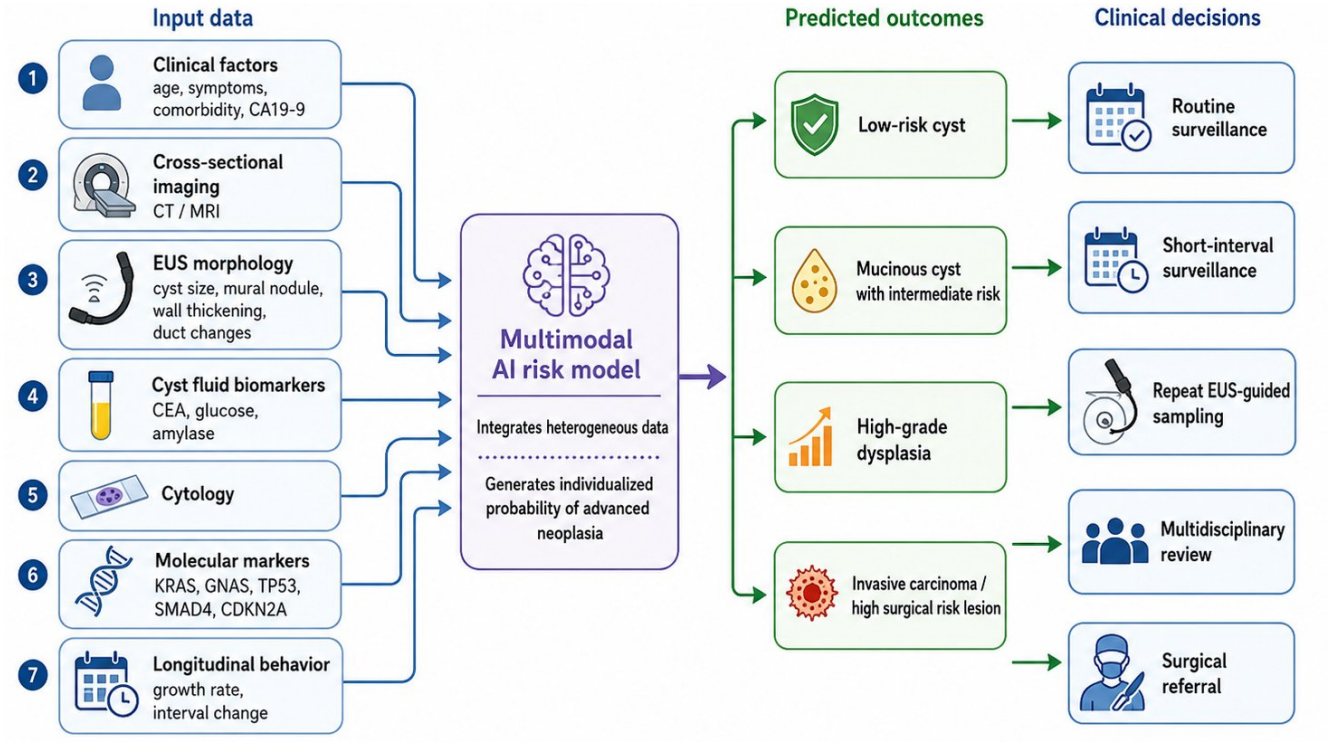

Figure 1: Conceptual multimodal AI-assisted pathway for EUS-based pancreatic cyst risk stratification This figure depicts a conceptual framework illustrating how multimodal AI may integrate heterogeneous data sources to generate individualized risk estimates for pancreatic cystic lesions. The pathway has not been prospectively validated and does not represent a clinically approved or guideline-endorsed workflow. All outputs should be interpreted in conjunction with clinical judgment, institutional expertise, and multidisciplinary team review. Surgical referral and surveillance decisions must remain physician-led and should not be determined by AI prediction alone pending prospective validation in representative clinical populations. Abbreviations: AI, Artificial intelligence; CA19-9, Carbohydrate antigen 19-9; CEA, Carcinoembryonic antigen; CT, Computed tomography; EUS, Endoscopic ultrasound; MRI, Magnetic resonance imaging.

2.5. Radiomics and cross-sectional imaging

Radiomics provides a key link between traditional imaging and AI-based risk prediction. CT- and MRI-based radiomic models can extract quantitative features of cyst shape, wall irregularity, internal texture, ductal anatomy, and surrounding pancreatic parenchyma that may not be captured by standard visual assessment. Early radiomic studies in IPMN demonstrated that quantitative imaging features can predict malignant potential and may complement standard clinical and radiological criteria [29–31]. In branch-duct IPMN, MRI-based radiomic nomograms that integrate imaging characteristics with the main pancreatic duct diameter and serum CA19-9 have been developed [32, 33]. Since MRI is commonly used for long-term cyst surveillance and EUS is typically reserved for indeterminate or concerning findings, a combined MRI–EUS AI pathway could enable MRI to detect evolving risk and trigger targeted EUS when the predicted probability of high-grade dysplasia or invasive carcinoma increases [32, 33].

2.6. Integrated multimodal AI models

A multimodal AI model may help reconcile discordant findings. A patient may present with a cyst>3 cm but no definite mural nodule, mild duct dilation, and negative cytology, or with a cyst that is rapidly growing and has elevated CA19-9. Current guidelines categorize patients into broad risk groups; AI could instead weight these characteristics simultaneously and generate a continuous risk score [32–34]. Molecular and cyst-fluid data may further enhance multimodal predictions — mutations supporting mucinous differentiation help characterize cyst type, while markers of advanced neoplasia increase predicted risk. Molecular results should not be interpreted in isolation but in the context of imaging, cytology, cyst behavior, and patient surgical fitness. AI may aggregate these diverse data streams into a unified decision-support framework rather than requiring clinicians to interpret each variable independently [11–13, 34].

It is important to acknowledge that many pancreatic cysts encountered in routine surveillance are straightforward to manage using existing guideline-based criteria and do not require complex multimodal AI integration [1–3]. A fully integrated model combining EUS morphology, nCLE findings, cross-sectional imaging, cyst-fluid biomarkers, molecular markers, and longitudinal behavior represents an aspirational framework rather than an immediately achievable clinical tool. Practical challenges include the absence of standardized data acquisition protocols, limited availability of molecular testing and nCLE in routine practice, and the fundamental dependency on surgical pathology as the only reliable ground truth for model training and validation [22, 23]. AI-assisted EUS is most likely to add value in cases where guidelines provide discordant, borderline, or insufficient guidance — including cysts meeting size thresholds without mural nodules, mild main pancreatic duct dilation of 5–9 mm, equivocal cytology or indeterminate cyst-fluid CEA, discordant imaging findings, and patients with competing operative risk or frailty. In these settings, a calibrated AI probability estimate of high-grade dysplasia or invasive carcinoma could meaningfully supplement guideline-based assessment and support multidisciplinary decision-making [1–3, 26, 27].

A practical AI-supported pathway would begin with standard CT or MRI surveillance, followed by EUS when imaging or clinical features suggest increased risk. During EUS, AI could assist in morphological evaluation, highlight suspicious intracystic regions, and support targeted cyst-fluid sampling. Post-procedure, a multimodal model could integrate EUS findings, imaging features, cyst-fluid results, cytology, molecular markers, and patient characteristics to stratify patients into actionable categories: routine surveillance, shortened surveillance interval, repeat EUS-guided sampling, multidisciplinary review, or surgical referral. The proposed framework is presented in **Figure 1** [34].

2.7. Limitations, Validation, and Future Directions

Available studies are largely retrospective, single-center, and based on selected image datasets rather than consecutive real-world EUS examinations. A critical consequence is reference-standard bias: models trained on surgically resected cohorts — where the prevalence of high-grade dysplasia or invasive carcinoma may reach 30–50% — learn to discriminate among already high-risk cysts rather than identify the rare dangerous lesion within a predominantly low-risk surveillance population. This spectrum bias inflates the apparent sensitivity and positive predictive value, and current models are likely to perform considerably worse when applied to unselected surveillance patients, in whom most cysts are small, incidental, and biologically indolent [21, 35].

Obtaining definitive ground truth for AI training and validation inherently requires surgical resection, since histopathological examination of the resected specimen remains the only reliable standard for confirming high-grade dysplasia or invasive carcinoma [22, 23]. Future studies should incorporate centralized, blinded pathology review with predefined grading criteria to minimize interobserver variability in dysplasia classification, alongside rigorous lesion-level matching between preoperative imaging and histological specimens to ensure AI training labels accurately reflect the biological state of the imaged lesion. This creates an unavoidable tension: resected cohorts are, by definition, enriched for high-risk disease and do not represent the broader surveillance population [21, 35]. Addressing this will require studies that combine surgical pathology-confirmed labels with prospective follow-up data from non-resected patients, enabling AI models to be trained and validated across the full spectrum of pancreatic cyst biology.

External validation is a critical unmet requirement. AI systems must be tested across different institutions, EUS platforms, image acquisition protocols, endosonographers, patient populations, and cyst subtypes. Future studies should prioritize prospective multicentre validation using consecutive patients and clinically relevant endpoints, including high-grade dysplasia, invasive carcinoma, progression during surveillance, and avoidance of unnecessary surgery [35, 36].

The implementation of AI in EUS faces procedural barriers distinct from those in radiology- or pathology-based AI. Image quality, frame selection, probe angulation, acoustic shadowing, and lesion accessibility vary substantially between endosonographers, institutions, and equipment manufacturers. Yet, most AI models have been developed on curated still images from a limited number of EUS platforms. Annotation standards for key EUS features remain poorly defined, introducing labeling noise into training datasets. Integration into real-time workflow requires sufficiently low latency to influence decision-making during the examination without disrupting technique or prolonging procedure time. Addressing these barriers will require standardized image acquisition protocols, multi-platform validation datasets, and regulatory frameworks that account for the operator-dependent and dynamic nature of EUS [35, 36].

A practical limitation of multimodal AI frameworks is the assumption that all input variables are available. In real-world practice, molecular testing, nCLE, cyst-fluid cytology, and serial imaging may be unavailable or deferred, particularly in lower-resource settings or when cysts are small. Most reviewed studies do not report model performance under incomplete data conditions. Future studies should evaluate robustness to missing variables, incorporate appropriate missing-data handling strategies, and compare complex multimodal models with simpler, parsimonious alternatives that use only routinely available inputs [35, 36].

A high AUC alone does not establish clinical utility. Most reviewed studies report discrimination without calibration data, decision-curve analysis, or prospectively defined thresholds. Calibration — the agreement between predicted probabilities and observed outcomes — is essential for individualized risk communication but is rarely reported. Decision-curve analysis quantifies net clinical benefit across decision thresholds, directly accounting for the asymmetric consequences of missed cancers versus unnecessary pancreatic resections. Future studies should report calibration curves, Brier scores, and decision-curve analyses alongside AUC. They should define clinically meaningful thresholds prospectively rather than optimizing cutpoints on the same dataset used for model development [35, 36].

The potential harms of AI-driven care escalation deserve explicit acknowledgment. Poorly calibrated or false-positive AI outputs could trigger unnecessary repeat EUS-FNA, nCLE, molecular testing, short-interval surveillance, and unwarranted surgical referral, exposing patients to procedural risk, anxiety, and financial burden. Pancreatic surgery carries substantial morbidity and mortality, and an AI system that systematically overestimates malignant risk could cause net harm even if discrimination metrics appear acceptable. Future studies should therefore measure harm alongside benefit, reporting rates of unnecessary procedures, avoidable operations, patient-reported outcomes, and cost-effectiveness [35, 36].

Most AI models have been trained on retrospectively selected still images rather than full procedural video, and it remains unknown how these systems will perform in real-time EUS or nCLE, where image quality, motion, probe angulation, and temporal dynamics differ substantially from those in curated image sets [20, 21, 35, 36]. Development of video-native AI architectures capable of processing the temporal dimension of EUS and nCLE procedures represents a critical priority for future research.

Explainability remains a fundamental challenge. AI predictions may influence whether a patient undergoes major pancreatic surgery or continues surveillance; clinicians, therefore, need to understand which variables contributed most to the risk estimate and whether the prediction is biologically plausible. Explainable models that highlight relevant EUS regions, identify influential variables, and provide calibrated risk probabilities are likely to be more acceptable than opaque black-box outputs [36, 37].

Table 2: Summary of key AI studies in EUS-based pancreatic cyst assessment

Study [Ref]	Modality	N	Cyst type	Algorithm	Target outcome	Reference standard	Validation	Performance	Main limitations
Kuwahara et al. [5]	EUS images	84	IPMN	CNN (deep learning)	Malignancy prediction	Surgical pathology	Internal only	AUC 0.89	Single-center; small dataset; still images only
Machicado et al. [6]	nCLE frames	33	IPMN	CNN	IPMN risk stratification	Surgical pathology	Internal only	Sens 94%, Spec 100%	Very small sample; single-center; still-frame analysis of video
Kurita et al. [14]	EUS cyst fluid	64	Mixed PCL	CNN (deep learning)	Malignant vs benign	Surgical pathology/cytology	Internal only	AUC 0.99	Very small sample; single-center; selected dataset
Lee et al. [15]	nCLE images	38	Mixed PCL	Deep learning CAD	Cyst subtype classification	Surgical pathology	Internal only	Accuracy 90.5%	Very small sample; retrospective; still-frame analysis
Springer et al. [13]	Molecular + clinical	130	Mixed PCL	Logistic regression	Cyst classification	Surgical pathology	Internal only	AUC 0.99	Single-center; selected surgical cohort
Correa-Gallego et al. [22]	Clinical + imaging	289	IPMN	Nomogram	HGD or invasive carcinoma	Surgical pathology	Internal only	AUC 0.79	Retrospective; single-center; surgical cohort only
Li et al. [23]	Clinical + imaging	261	IPMN	Nomogram	HGD or invasive carcinoma	Surgical pathology	Internal only	AUC 0.854	Retrospective; single-center; surgical cohort only
Kang et al. [25]	Clinical + imaging	828	IPMN	ML vs logistic regression	Malignant IPMN	Surgical pathology	Internal only	AUC 0.79–0.81	Retrospective; no external validation; surgical cohort
He et al. [27]	Clinical + imaging	218	IPMN	Prediction model	Degree of malignancy	Surgical pathology	Internal only	AUC 0.841	Retrospective; single-centre; surgical cohort
Lavista Ferres et al. [28]	Multimodal	359	Mixed PCL	Explainable AI	Cyst management guidance	Surgical pathology/follow-up	Internal only	AUC 0.82	Retrospective; single-center; limited external validity

continued on the next page

Table 2 (continued)

Study [Ref]	Modality	N	Cyst type	Algorithm	Target outcome	Reference standard	Validation	Performance	Main limitations
Hanania et al. [29]	CT radiomics	54	IPMN	Quantitative imaging	Malignant potential	Surgical pathology	Internal only	AUC 0.86	Very small sample; single-center; CT only
Permuth et al. [30]	CT radiomics + miRNA	45	IPMN	Radiomic-molecular model	Malignant pathology	Surgical pathology	Internal only	AUC 0.91	Very small sample; single-center; no external validation
Chakraborty et al. [31]	CT radiomics	94	IPMN	Radiomic model	High-risk IPMN	Surgical pathology	Internal only	AUC 0.85	Single-center; CT only; surgical cohort
Cui et al. [32]	MRI radiomics + clinical	226	BD-IPMN	Radiomic nomogram	IPMN grade	Surgical pathology	Internal + external	AUC 0.91	Branch-duct IPMN only; MRI only
Flammia et al. [33]	MRI radiomics	99	BD-IPMN	Radiomic model	Malignant degeneration	Surgical pathology	Internal only	AUC 0.83	Small sample; single-center; MRI only
Lou et al. [34]	CT radiomics + clinical	398	IPMN	Radiomic-clinical model	Benign vs malignant IPMN	Surgical pathology	Internal only	AUC 0.88	Retrospective; single-centre; CT only

Abbreviations: AUC, area under the receiver operating characteristic curve; BD-IPMN, branch-duct intraductal papillary mucinous neoplasm; CAD, computer-aided diagnosis; CNN, convolutional neural network; HGD, high-grade dysplasia; IPMN, intraductal papillary mucinous neoplasm; ML, machine learning; nCLE, needle-based confocal laser endomicroscopy; PCL, pancreatic cystic lesion; Sens, sensitivity; Spec, specificity.

Clinical integration requires careful workflow design. AI should support, not replace, expert endosonographic assessment and multidisciplinary decision-making. An effective system should provide real-time assistance, identify uncertainty, and recommend appropriate next steps - repeat imaging, targeted EUS-guided sampling, molecular testing, short-interval surveillance, or surgical review - while remaining contextual to patient age, comorbidity, frailty, operative risk, life expectancy, and patient preference [37, 38]. Furthermore, AI platforms, nCLE, molecular testing, and high-quality EUS expertise are currently concentrated in tertiary academic centers; without deliberate attention to scalability, reimbursement pathways, and equitable access, widespread adoption of AI-assisted EUS risks widening rather than reducing existing disparities in pancreatic cyst care. Clear governance is also required to define medicolegal responsibility when AI-generated risk estimates diverge from guideline-based recommendations or multidisciplinary judgment: accountability for the final management decision must remain with the treating clinicians and the multidisciplinary team rather than with the algorithm, and institutions should establish documented pathways for reconciling discordant AI outputs so that such conflicts do not disrupt the clinical workflow or create ambiguous lines of liability.

Future research should move beyond diagnostic accuracy alone. Studies should evaluate whether AI-assisted EUS changes management, improves risk calibration, reduces unnecessary pancreatic resections, detects high-grade dysplasia or early invasive cancer earlier, and remains safe during longitudinal follow-up. A minimum evidence threshold for clinical implementation should include prospective multicentre validation in consecutive patients, calibration testing, decision-curve analysis confirming net benefit over existing guidelines, workflow integration testing in real-time procedural settings, demonstrated impact on clinical management rather than diagnostic labeling alone, and a post-deployment monitoring plan to detect performance degradation over time or across subpopulations. Regulatory approval, medicolegal responsibility, data governance, cybersecurity, and equity across diverse populations will also be essential before these tools can be adopted routinely. Ultimately, AI-assisted EUS will be clinically valuable only if it improves patient-centered outcomes rather than merely increasing model performance metrics [38, 39]. The methodological characteristics and performance metrics of key AI studies included in this review are summarised in **Table 2**.

3. Conclusion

AI-assisted EUS represents a promising but early-stage approach to pancreatic cyst risk stratification. Current evidence suggests that AI may enhance EUS image interpretation, improve recognition of high-risk morphological features, and support individualized estimation of the risk of high-grade dysplasia and invasive carcinoma when integrated with cyst-fluid biomarkers, molecular markers, cross-sectional imaging, and clinical variables. However, available studies are predominantly retrospective, single-center, and based on surgically resected cohorts that do not represent the broader surveillance population. No AI model has been prospectively validated, regulatory approved, or demonstrated to improve patient-centered outcomes in clinical practice. The distinction between statistical discrimination and genuine clinical utility remains largely unaddressed, and implementation barriers specific to EUS — including device variability, annotation standards, real-time integration, and operator dependence — have not been systematically resolved. AI-assisted EUS should therefore be regarded as investigational at this stage. These tools should not replace guideline-based risk assessment, expert endosonographic interpretation, or multidisciplinary team decision-making. Future prospective multicentre studies with consecutive patient enrolment, external validation, and clinically meaningful endpoints are needed before AI-assisted EUS can be responsibly integrated into routine pancreatic cyst management.

Conflicts of Interest

The author declares no conflict of interest.

Funding Source

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgments

None.

Institutional Review Board (IRB)

Not applicable.

Large Language Model

Generative artificial intelligence tools were used in a limited capacity to assist with language refinement and grammar checking during the preparation of this manuscript. The author reviewed and verified all content and takes full responsibility for the integrity, accuracy, and scientific content of the manuscript.

Authors Contribution

The author conceived and designed the review, performed the literature search, drafted and critically revised the manuscript, and approved the final version for submission.

Data Availability

No new datasets were generated or analysed for this review article. All information discussed is derived from previously published studies and publicly available literature.

References

1. Vege SS, Ziring B, Jain R, et al. American Gastroenterological Association Institute guideline on the diagnosis and management of asymptomatic neoplastic pancreatic cysts. Gastroenterology. 2015. [PMID: 25805375, doi:10.1053/j.gastro.2015.01.015].
2. European Study Group on Cystic Tumors of the Pancreas. European evidence-based guidelines on pancreatic cystic neoplasms. Gut. 2018;67(5):789-804. [PMID: 29574408, PMCID: PMC5890653, doi:10.1136/gutjnl-2018-316027].
3. Ohtsuka T, Fernandez-Del Castillo C, et al. International evidence-based Kyoto guidelines for the management of intraductal papillary mucinous neoplasm of the pancreas. Pancreatology. 2024;24(2):255-70. [PMID: 38182527, doi:10.1016/j.pan.2023.12.009].
4. Guilabert L, Nikolìc S, de Madaria E, et al. Endoscopic ultrasound for pancreatic cystic lesions: a narrative review. BMJ Open Gastroenterol. 2025. [PMID: 40998473, PMCID: PMC12481305, doi:10.1136/bmjgast-2025-001893].
5. Kuwahara T, Hara K, Mizuno N, et al. Usefulness of Deep Learning Analysis for the Diagnosis of Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas. Clin Transl

Gastroenterol. 2019. [PMID: 31117111, PMCID: PMC6602761, doi:10.14309/ctg.0000000000000045].
6. Machicado JD, Chao WL, Carlyn DE, et al. High performance in risk stratification of intraductal papillary mucinous neoplasms by confocal laser endomicroscopy image analysis with convolutional neural networks (with video). Gastrointest Endosc. 2021;94(1):78-87.e2. [PMID: 33465354, doi:10.1016/j.gie.2020.12.054].
7. Brugge WR, Lewandrowski K, Lee-Lewandrowski E, et al. Diagnosis of pancreatic cystic neoplasms: a report of the cooperative pancreatic cyst study. Gastroenterology. 2004. [PMID: 15131794, doi:10.1053/j.gastro.2004.02.013].
8. Tanaka M, Fernández-Del Castillo C, Kamisawa T, et al. Revisions of international consensus Fukuoka guidelines for the management of IPMN of the pancreas. Pancreatology. 2017;17(5):738-53. [PMID: 28735806, doi:10.1016/j.pan.2017.07.007].
9. McCarty TR, Garg R, Rustagi T. Pancreatic cyst fluid glucose in differentiating mucinous from nonmucinous pancreatic cysts: a systematic review and meta-analysis. Gastrointest Endosc. 2021;94(4):698-712.e6. [PMID: 33964311, doi:10.1016/j.gie.2021.04.025].
10. Carr RA, Yip-Schneider MT, Simpson RE, et al. Pancreatic cyst fluid glucose: rapid, inexpensive, and accurate diagnosis of mucinous pancreatic cysts. Surgery. 2018;163(3):600-5. [PMID: 29241991, doi:10.1016/j.surg.2017.09.051].
11. Singhi AD, McGrath K, Brand RE, et al. Preoperative next-generation sequencing of pancreatic cyst fluid is highly accurate in cyst classification and detection of advanced neoplasia. Gut. 2018;67(12):2131-41. [PMID: 28970292, PMCID: PMC6241612, doi:10.1136/gutjnl-2016-313586].
12. Pflüger MJ, Jamouss KT, Afghani E, et al. Predictive ability of pancreatic cyst fluid biomarkers: A systematic review and meta-analysis. Pancreatology. 2023;23(7):868-77. [PMID: 37230894, doi:10.1016/j.pan.2023.05.005].
13. Springer S, Wang Y, Dal Molin M, et al. A combination of molecular markers and clinical features improves the classification of pancreatic cysts. Gastroenterology. 2015;149(6):1501-10. [PMID: 26253305, PMCID: PMC4782782, doi:10.1053/j.gastro.2015.07.041].
14. Kurita Y, Kuwahara T, Hara K, et al. Diagnostic ability of artificial intelligence using deep learning analysis of cyst fluid in differentiating malignant from benign pancreatic cystic lesions. Sci Rep. 2019. [PMID: 31053726, PMCID: PMC6499768, doi:10.1038/s41598-019-43314-3].
15. Lee TC, Angelina CL, Kongkam P, et al. Deep-Learning-Enabled Computer-Aided Diagnosis in the Classification of Pancreatic Cystic Lesions on Confocal Laser Endomicroscopy. Diagnostics (Basel). 2023;13(7):1289. [PMID: 37046507, PMCID: PMC10093377, doi:10.3390/diagnostics13071289].
16. Nakai Y, Isayama H, Shinoura S, et al. Confocal laser endomicroscopy in gastrointestinal and pancreatobiliary diseases. Dig Endosc. 2014;26 Suppl 1:86-94. [PMID: 24033351, doi:10.1111/den.12152].
17. Wu CCH, Lim SJM, Tan DMY. Endoscopic ultrasound-guided needle-based confocal laser endomicroscopy for pancreatic cystic lesions: current status and prospects. Clin Endosc. 2024;57(4):434-45. [PMID: 38978396, PMCID: PMC11294861, doi:10.5946/ce.2023.157].
18. Rangwani S, Ardeshna DR, Rodgers B, et al. Application of Artificial Intelligence in the Management of Pancreatic Cystic Lesions. Biomimetics (Basel). 2022;7(2):79. [PMID: 35735595, PMCID: PMC9221027, doi:10.3390/biomimetics7020079].
19. Balduzzi A, Janssen BV, De Pastena M, et al. Artificial intelligence-based models to assess the risk of malignancy on radiological imaging in patients with intraductal papillary mucinous neoplasm of the pancreas: scoping review. Br J Surg. 2023. [PMID: 37402951, PMCID: PMC10638536, doi:10.1093/bjs/znad201].
20. Qadir MI, Baril JA, Yip-Schneider MT, et al. Artificial intelligence in pancreatic intraductal papillary mucinous neoplasm imaging: a systematic review. PLOS Digit Health. 2025. [PMID: 40700385, PMCID: PMC12286379, doi:10.1371/journal.pdig.0000920].
21. Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ. 2020. [PMID: 32213531, PMCID: PMC7190037, doi:10.1136/bmj.m689].
22. Correa-Gallego C, Do R, Lafemina J, et al. Predicting dysplasia and invasive carcinoma in intraductal papillary mucinous neoplasms of the pancreas: development of a preoperative nomogram. Ann Surg Oncol. 2013;20(13):4348-55. [PMID: 24046103, doi:10.1245/s10434-013-3207-z].
23. Li B, Shi X, Gao S, et al. Nomogram for the Prediction of High-Grade Dysplasia and Invasive Carcinoma in Patients With Intraductal Papillary Mucinous Neoplasms of the Pancreas Based on Variables of Noninvasive Examination. Front Oncol. 2021. [PMID: 33767983, PMCID: PMC7985057, doi:10.3389/fonc.2021.609187].
24. Pergolini I, Sahora K, Ferrone CR, et al. Long-term risk of pancreatic malignancy in patients with branch-duct intraductal papillary mucinous neoplasm in a referral center. Gastroenterology. 2017;153(5):1284-94. [PMID: 28739282, doi:10.1053/j.gastro.2017.07.019].
25. Kang JS, Lee C, Song W, et al. Risk prediction for malignant intraductal papillary mucinous neoplasm of the pancreas: logistic regression versus machine learning. Sci Rep. 2020. [PMID: 33208887, PMCID: PMC7676251, doi:10.1038/s41598-020-76974-7].
26. Lennon AM, Wolfgang CL, Canto MI, et al. The early detection of pancreatic cancer: what will it take to diagnose and treat curable pancreatic neoplasia? Cancer Res. 2014;74(13):3381-9. [PMID: 24924775, PMCID: PMC4085574, doi:10.1158/0008-5472.CAN-14-0734].
27. He X, Fan R, Sun J, et al. A model for predicting degree of malignancy in patients with intraductal papillary mucinous neoplasm. Front Oncol. 2023. [PMID: 36761937, PMCID: PMC9902908, doi:10.3389/fonc.2023.1087852].
28. Lavista Ferres JM, Oviedo F, Robinson C, et al. Performance of explainable artificial intelligence in guiding the management of patients with a pancreatic cyst. Pancreatology. 2024;24(7):1182-91. [PMID: 39261223, doi:10.1016/j.pan.2024.09.001].
29. Hanania AN, Bantis LE, Feng Z, et al. Quantitative imaging to evaluate malignant potential of IPMNs. Oncotarget. 2016. [PMID: 27588410, PMCID: PMC5349873, doi:10.18632/oncotarget.11769].
30. Permuth JB, Choi J, Balarunathan Y, et al. Combining radiomic features with a miRNA classifier may improve prediction of malignant pathology for pancreatic intraductal papillary mucinous neoplasms. Oncotarget. 2016. [PMID: 27589689, PMCID: PMC5349874, doi:10.18632/oncotarget.11768].
31. Chakraborty J, Midya A, Gazit L, et al. CT radiomics to predict high-risk intraductal papillary mucinous neoplasms of the pancreas. Med Phys. 2018;45(11):5019-29. [PMID: 30176047, PMCID: PMC8050835, doi:10.1002/mp.13159].
32. Cui S, Tang T, Su Q, et al. Radiomic nomogram based on MRI to predict grade of branching type intraductal papillary mucinous neoplasms of the pancreas: a multicenter study. Cancer Imaging. 2021. [PMID: 33750453, PMCID: PMC7942000, doi:10.1186/s40644-021-00395-6].
33. Flammia F, Innocenti T, Galluzzo A, et al. Branch duct-intraductal papillary mucinous neoplasms (BD-IPMNs): an MRI-based radiomic model to determine the malignant degeneration potential. Radiol Med. 2023;128(4):383-92. [PMID: 36826452, doi:10.1007/s11547-023-01609-6].
34. Lou F, Li M, Chu T, et al. Comprehensive analysis of clinical data and radiomic features from contrast-enhanced CT for differentiating benign and malignant pancreatic intraductal papillary mucinous neoplasms. Sci Rep. 2024. [PMID: 39060387, PMCID: PMC11282090, doi:10.1038/s41598-024-68067-6].
35. Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health. 2019;1(6):e271-97. [PMID: 33323251, doi:10.1016/S2589-7500(19)30123-2].
36. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024. [PMID: 38626948, PMCID: PMC11019967, doi:10.1136/bmj-2023-078378].
37. Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019. [PMID: 31665002, PMCID: PMC6821018, doi:10.1186/s12916-019-1426-2].
38. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early stage clinical evaluation of decision support systems driven by

artificial intelligence: DECIDE-AI. BMJ. 2022. [PMID: 35584845, PMCID: PMC9116198, doi:10.1136/bmj-2022-070904].

39. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. Nat Med. 2019;25(9):1337-40. [PMID: 31427808, doi:10.1038/s41591-019-0548-6].