Diagnostic and Interventional Imaging 104 (2023) 435–447

Review/Research, New developments & Artificial Intelligence

A primer on artificial intelligence in pancreatic imaging

Taha M. Ahmed[a], Satomi Kawamoto[a], Ralph H. Hruban[b], Elliot K. Fishman[a], Philippe Soyer[c], Linda C. Chu[a,*]

[a] *The Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins Hospital, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA*
[b] *Sol Goldman Pancreatic Research Center, Department of Pathology, Johns Hopkins Hospital, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA*
[c] *Université Paris Cité, Faculté de Médecine, Department of Radiology, Hôpital Cochin-APHP, 75014, 75006, Paris, France, 7501475006*

ARTICLE INFO

Keywords:
Artificial intelligence
Deep learning
Radiomics
Pancreas

ABSTRACT

Artificial Intelligence (AI) is set to transform medical imaging by leveraging the vast data contained in medical images. Deep learning and radiomics are the two main AI methods currently being applied within radiology. Deep learning uses a layered set of self-correcting algorithms to develop a mathematical model that best fits the data. Radiomics converts imaging data into mineable features such as signal intensity, shape, texture, and higher-order features. Both methods have the potential to improve disease detection, characterization, and prognostication. This article reviews the current status of artificial intelligence in pancreatic imaging and critically appraises the quality of existing evidence using the radiomics quality score.

1. Introduction

The use of artificial intelligence (AI) in radiology has rapidly gained attention in recent years. Radiology, being a digital image-based specialty, serves as the ideal testing ground for medical applications of AI. This, coupled with the increasing demand for clinical imaging and a shortage of radiologists, has led to a surge in AI-based presentations and articles in radiology conferences and journals and has necessitated the establishment of dedicated AI journals within the field. The majority of current research in AI in pancreatic imaging can be classified as either deep learning-based approaches or radiomics based approaches. Much of this novel research is already being integrated into clinical practice, with radiology leading all medical specialties in the number of federally approved AI tools with over 200 radiology-related AI programs approved by the Federal Food and Drug Administration as of the end of 2022 [1]. Accompanying the optimism surrounding this growth, however, are concerns about reproducibility, lack of generalizability and uncharted clinical translation of models reported in the literature. In addition, the relative infancy of AI, coupled with its "black-box" complexity, which demands multidisciplinary expertise in the peer review process, has led to a varied quality of results being published.

This purpose of this article was to review the current status of AI in pancreatic imaging and critically appraises the quality of existing evidence using the radiomics quality score.

2. Deep learning

Deep learning is a subfield of machine learning in which algorithms are designed to automatically identify patterns in data by learning from multiple examples rather than being explicitly programmed. It utilizes deep neural networks, which are inspired by biological neural networks, to mimic the function of neurons in the brain. These networks consist of interconnected nodes that process input data and adjust the network weights to minimize prediction errors. The concepts of deep learning were first developed in the 1950s, but it was not until the advent of powerful parallel computing hardware, vast amounts of training data, and improved training techniques and network architectures that deep learning approaches were able to reach their full potential for clinical application [2]. In a deep neural network, the intermediate layers are referred to as hidden layers, as they process the input data to generate intermediate representations that lead to the final output. This final output is then compared to ground truths, and errors are used as feedback to adjust the weights in the network to minimize error in subsequent iterations [3].

Abbreviations: AI, Artificial intelligence; AIP, Autoimmune pancreatitis; AUC, Area under the curve; CT, Computed tomography; EUS, Endoscopic ultrasound; EUS-FNA, Endoscopic ultrasound-guided fine needle aspiration; IPMN, Intraductal papillary mucinous neoplasm; MCN, Mucinous cystic neoplasm; MRI, Magnetic resonance imaging; PDAC, Pancreatic ductal adenocarcinoma; pNET, Pancreatic neuroendocrine tumor; SCN, Serous cystic neoplasm; SPN, Solid pseudopapillary neoplasms; RQS, Radiomics quality score

* Corresponding author:
E-mail address: lchu1@jhmi.edu (L.C. Chu).

https://doi.org/10.1016/j.diii.2023.03.002
2211-5684/

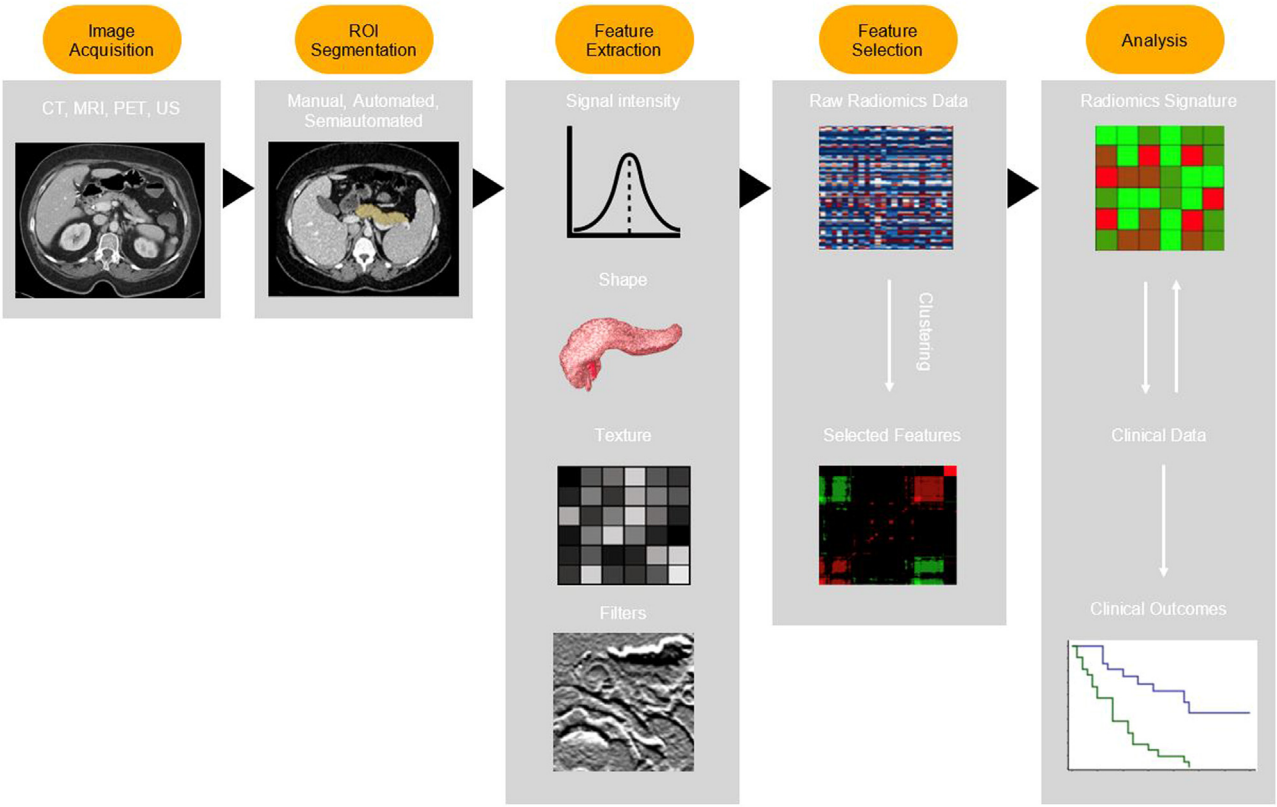

Fig. 1. Schematic illustration of the radiomics workflow.

3. Radiomics

Radiomics converts imaging data into high-dimensional mineable features that can be used to characterize the inherent spatial heterogeneity of neoplasms and other non-neoplastic disease processes [4]. This process begins with the conversion of imaging data into high-dimensional mineable features, including signal intensity, shape, texture and higher-order features (Fig. 1) [5–7]. Signal intensity features are derived from histograms of individual voxel signal intensities, providing measures of central tendency and shape of the distribution. Shape features are extracted from the three-dimensional surface of the region of interest, and texture features are calculated in three dimensions, taking into account the correlation of the signal intensities of the surrounding voxels. Higher-order features are features that are extracted following the application of a secondary filter, such as a wavelet or Gaussian filter [5,6]. The number of features extracted during this image analysis process can vary widely, depending on the software package and filters used. However, a high number of features and a low number of cases in a group for a classification task can result in overfitting of the model. To mitigate this risk, it is essential to perform feature selection or dimension reduction to reduce the number of features and increase the validity and generalizability of the results. Once appropriate features have been selected, they are subsequently analyzed with advanced machine learning algorithms, such as random forest or support vector machine, to perform specific classification tasks that can be used to help answer clinical questions.

4. Disease detection

Imaging features of pancreatic ductal adenocarcinoma (PDAC) early in the disease process can be subtle and potentially be overlooked during initial reads [8,9]. Literature on the applications of radiomics and deep learning in the detection of disease has continued to grow in recent years, with radiomics tools increasingly being investigated for their role as "second readers" to decrease the number of these misses that may represent early curable cancers. Numerous computed tomography (CT) and endoscopic ultrasound (EUS) based studies have reported on the utility of radiomic features in distinguishing pancreatic cancer from noncancerous pancreas (Table 1) [10,11]. CT radiomic analysis in a recent study with 436 patients with PDAC and 479 healthy controls demonstrated 94.7% sensitivity and 95.4% specificity in detecting PDAC from amongst healthy controls [12]. A deep learning-based study that was validated on an external test set of 1473 CT examinations also reported positive results with a sensitivity of 89.7% and a specificity of 92.8% [13].

While most prior studies had applied radiomics as a "second reader" to catch a diagnosis that may be missed due to human error, several more recent studies have reported that radiomics models may be able to detect PDAC before it is even discernable to the human eye on imaging [14–17]. During the development of PDAC, the pancreas undergoes various morphological changes. PDACs may arise from detectable precancerous lesions such as intraductal papillary mucinous neoplasms (IPMN), and the pancreatic parenchyma upstream from a subtle cancer may show focal parenchymal atrophy and changes of chronic pancreatitis. Each of these can gradually increase the heterogeneity of the pancreatic tissue and result in detectable morphological and textural changes [18–20]. These alterations may be difficult to interrogate on visual assessment, making AI the ideal tool to analyze them. Current studies on this have extracted radiomics features from prediagnostic CT examinations (3–36 months prior to the diagnosis of PDAC) and developed radiomics models that could predict the risk of developing PDAC with an accuracy ranging from 89.3% to100% [14–17].

AI has also made advancements in the detection of a variety of solid and cystic pancreatic neoplasms aside from PDAC. In a recent study, a deep learning model was able to detect PDAC, pancreatic neuroendocrine tumors (pNET), solid pseudopapillary neoplasms (SPN), mucinous cystic neoplasms (MCN), serous cystic neoplasms,

Table 1
Representative peer-reviewed studies of machine learning in pancreatic imaging.

Clinical problem	Imaging modality	Machine learning technique	Validation	Accuracy, AUC, sensitivity, specificity or C-statistic of best model	Ref #
Detection of PDAC	CT	Radiomics	External validation	Accuracy: 86.5%	[12]
	CT	Deep learning CT	External validation	AUC: 0.95	[13]
	EUS	Deep learning	Cross validation	Accuracy: 87.5%	[10]
	CT	Radiomics	External validation (specificity only)	Accuracy: 92.2%	[14]
	CT	Radiomics	External validation	Accuracy: 89.3%	[16]
	CT	Radiomics	External validation	Accuracy:86.0%	[15]
	CT	Radiomics	Split sample	Accuracy: 100% (non-CP patients) 95% (CP patients)	[17]
	CT	Radiomics	Split sample	Accuracy: 99.2%	[11]
Detection of non-PDAC malignancy	CT	Deep learning	Cross sample	AUC: 0.91	[21]
	CT	Deep learning	Split sample	Accuracy: 82.7%	[23]
	CT	Deep learning	Cross validation	Sensitivity:78.8%	[22]
Classification of pancreatic neoplasm and mass vs. pancreatitis	CEUS	Radiomics + Deep learning	External	AUC: 0.967	[25]
	MRI	Radiomics	Cross validation	AUC: 0.942	[26]
	MRI	Radiomics	Cross validation	AUC: 0.962	[27]
	CT	Radiomics	Cross validation	AUC: 0.93	[28]
	CT	Radiomics	Cross validation	AUC: 0.97	[31]
	CT	Radiomics	Split sample	AUC: 0.975	[32]
	EUS	Deep learning	Cross validation	AUC: 0.940	[34]
	EUS	Deep learning	Split sample	AUC: 0.930	[35]
	EUS	Deep learning	None	AUC: 0.932	[36]
	EUS	Deep learning	Cross validation	AUC: 0.940	[37]
	CT	Radiomics	Cross validation	Accuracy: 82.1%	[33]
Classification of pancreatic neoplasms	CT	Radiomics	None	AUC: 0.77	[43]
	CT	Radiomics	None	AUC: 0.821	[42]
	CT	Radiomics	None	Sensitivity: 79%, Specificity: 71%	[39]
	CT	Radiomics	Cross validation	AUC: 0.93	[40]
	CT	Radiomics	Split sample	AUC: 0.85	[41]
	CT	Radiomics	Cross validation + External	AUC: 0.873	[38]
	CT	Radiomics	Cross validation	AUC: 0.86	[44]
	CT	Radiomics	Cross validation	AUC: 0.94	[45]
	CT	Radiomics	Split sample	AUC: 0.75	[46]
	CT	Radiomics	Cross validation	AUC: 0.837	[47]
	CT	Radiomics	None	AUC: 0.989	[48]
	CT	Radiomics	Split sample	AUC: 0.817	[49]
	CT	Deep learning	Cross validation	AUC: 0.973	[53]
	CT	Deep learning	Cross validation	AUC: 0.98	[50]
	CT	Deep learning	Cross validation	Accuracy: 72.8%	[52]
	CT	Deep learning	Cross validation	Accuracy: 83.6%	[51]
	MRI & CT	Radiomics	Cross validation	AUC: 0.94 (MRI), 0.864 (CT)	[56]
Classification of IPMN dysplasia grade	MRI	Radiomics	External validation	AUC: 0.822	[57]
	CT	Radiomics	Cross validation	AUC: 0.76	[58]
	CT	Radiomics	Cross validation	AUC: 0.77	[59]
	CT	Radiomics	Cross validation	AUC: 0.87	[60]
	CT	Radiomics	Cross validation +External validation	AUC: 0.71	[61]
	EUS	Deep learning	Split sample	Accuracy: 99.6%	[63]
	EUS	Deep learning	Cross validation	Accuracy: 94.0%	[64]
Prediction of pNET Grade	CT & MRI	Radiomics	Split sample	AUC: 0.85	[66]
	CT	Radiomics	Split Sample	Sensitivity: 54%, Specificity: 80%	[67]
	CT	Radiomics	Cross validation + External validation	AUC: 0.881	[69]
	CT	Radiomics	Split sample	AUC: 0.82	[77]
	PET/MRI	Radiomics	Split sample	AUC: 0.72	[70]
	MRI	Radiomics	Split sample	AUC: 0.736	[71]
	MRI	Radiomics	None	AUC: 0.695	[72]
	CT	Radiomics	Cross validation	AUC: 0.876	[75]
	CT	Radiomics	None	AUC: 0.79	[73]
	CT	Radiomics	Cross validation	AUC: 0.90	[74]
	CT	Radiomics	Cross validation	AUC: 0.837	[76]
Prediction of survival in patients with PDAC	CT	Radiomics	External validation	C-index: 0.554 (DFS), 0.545 (OS)	[81]
	CT	Radiomics	Cross validation	AUC: 0.715	[80]
	CT	Radiomics	Cross validation + External validation	AUC: 0.722 (internal), 0.689 (external)	[78]
	CT	Radiomics	Cross validation	AUC: 0.72	[79]
	CT	Radiomics	None	AUC: 0.756	[83]
	CT	Radiomics	Split sample	AUC: 0.74	[84]
	CT	Radiomics	None	AUC: 0.716	[93]
	CT	Deep learning	Cross validation + External validation	AUC: 0.723 (external), 0.734 (internal)	[82]
	CT	Radiomics	Cross validation + External validation	AUC: 0.94	[97]

(*continued*)

Table 1 (*Continued*)

Clinical problem	Imaging modality	Machine learning technique	Validation	Accuracy, AUC, sensitivity, specificity or C-statistic of best model	Ref #
Prediction of treatment response in patients with PDAC	MRI	Radiomics	Split sample	AUC: 0.845	[98]
	CT	Radiomics	None	AUC: 0.851	[99]
	CT	Radiomics	None	AUC: 0.71	[100]
	CT	Radiomics	Cross validation	AUC: 0.81	[101]
	CT	Radiomics	Cross validation	AUC: 0.75	[102]
	CT	Deep learning	Split sample	AUC: 0.74	[103]

AUC indicates area under the receiver operating characteristics curve; CT indicates computed tomography; DFS indicates disease-free survival; EUS indicates endoscopic ultrasound; IPMN indicates intraductal papillary mucinous neoplasm; OS indicates overall survival; PDAC indicates pancreatic ductal adenocarcinoma; PET indicates positron emission tomography; pNET indicates pancreatic neuroendocrine tumor; MRI indicates magnetic resonance imaging.

and IPMNs with a sensitivity of 98%–100% for solid lesions and 92%–93% for cystic lesions larger than 1.0 cm across two test sets consisting of 1192 patients [21]. The performance of this model was not significantly different from that of radiologists (95–100% for solid lesions and 93–98% for cystic lesions > 1.0 cm) [21]. Similar prior deep learning studies have reported the sensitivity of detecting PDAC, pNET and pancreatic cystic lesions ranging from 78.8% to 87.6% [22,23].

5. Classification of pancreatic masses

Pancreatic malignancies such as PDAC and pNET can often resemble benign conditions like pancreatitis or benign tumors such as serous cystic neoplasm, or non-invasive tumors with malignant potential like MCNs and IPMNs. When imaging reveals a pancreatic mass with uncertain or suspicious features, endoscopic ultrasound is often used to determine the nature of the mass. EUS, however, is an invasive procedure with associated complications and costs. Radiomics and deep learning may make it possible to classify these pancreatic masses more accurately, and, thus more appropriately triage patients based on who will receive the greatest benefit from endoscopic ultrasound.

A growing area of radiomic research has been on differentiating the lesions of chronic pancreatitis, mass-forming pancreatitis and autoimmune pancreatitis (AIP) from those of PDAC. These disease processes share overlapping imaging features, making differentiating them challenging even for the experienced radiologist. Additionally, autoimmune and other forms of pancreatitis can be difficult for even experienced pathologists to diagnose on the limited samples generated from endoscopic ultrasound-guided fine needle aspiration (EUS-FNA). The presence of chronic pancreatitis can also significantly lower the sensitivity of EUS-FNA for evaluating focal pancreatic masses from 89% to 54% [24]. In a recent study with 414 patients with PDAC and 144 patients with chronic pancreatitis, an ultrasound-based radiomics model achieved an area under the curve (AUC) of 0.967 in differentiating PDAC lesions from chronic pancreatitis lesions [25]. This model also achieved a higher than or not different sensitivity and specificity than radiologists in the same task. Similar studies have reported not different results in radiomics-based differentiation of PDAC from mass-forming pancreatitis lesions on both CT and MRI [26–28]. While EUS-FNA is quite effective at detecting pancreatic malignancy with a sensitivity of 85%, its sensitivity for identifying AIP is much lower, varying from 7.9% to 58.2% [29,30]. Li et al. developed a CT radiomics model that could differentiate AIP from PDAC with an accuracy of 94% [31]. Prior studies have demonstrated similar results and have demonstrated radiomics models to have outperformed radiologists in the challenging diagnosis of AIP [32]. Radiomic features have also shown efficacy in differentiating recurrent acute pancreatitis from chronic pancreatitis [33]. Deep learning studies based on EUS images have also shown excellent performance in differentiating PDAC from chronic pancreatitis with an AUC ranging from 0.93 to 0.94 [34–37]

Once a mass has been detected, radiomics and deep learning can be used to differentiate and classify various pancreatic tumors. Among literature on solid pancreatic tumors, most studies have reported on radiomics models distinguishing PDAC from pNET [38–41]. Some studies have also reported on differentiation of PDAC from lymphoma and pNET from renal cell carcinoma metastases to the pancreas [42,43]. Radiomic features have also been reported to be helpful in differentiating pNET from SPN on magnetic resonance imaging (MRI) [44].

Studies on classification of cystic pancreatic tumors have also reported excellent results. These studies have predominantly employed three different strategies in cyst classification: 1), A multiclass method to distinguish each category of pancreatic cyst. 2), A binary approach to separate benign cysts from those with malignant potential. 3), A binary classification of mucin-producing cysts into high-grade or low-grade dysplasia. A recent multiclass study consisting of 214 patients reported a radiomics model to perform on par with experienced academic radiologists at classifying various cystic tumors (IPMN, MCN, serous cystadenoma, SPN, PNET) with an AUC of 0.940 for the radiomics model compared to an AUC of 0.895 for the radiologists [45]. Binary studies on classification of cystic tumors have also demonstrated successful discriminatory capacity between benign pancreatic serous cystadenomas from other cystic neoplasms with malignant potential (MCNs, IPMNs and SPNs) with prior radiomics based studies reporting an AUC of 0.75–0.989 in differentiation of SCNs from MCNs [46–49]

Yang et al. reported a radiomics based model distinguishing SCNs from MCNs with an AUC of 0.98 in a study with 63 SCNs and 47 MCNs [50]. The overarching rationale for these binary classification studies is that ruling in a diagnosis of a serous cystadenoma can avoid unnecessary clinical and imaging follow-ups and reduce healthcare burden. Deep learning has also shown excellent results in classifying cystic pancreatic masses with a pooled accuracy ranging from 72.8% to 92.5% [50–53].

6. Risk stratification of IPMN

The 2017 International Consensus Guidelines for the treatment of IPMNs, referred to as the Fukuoka criteria, suggest that surgery is necessary for all main duct IPMNs for patients who are fit for the procedure [54]. However, the management of branch duct IPMNs is more uncertain as only a small percentage (14.4%–47.9%) of resected branch duct IPMNs have high-grade dysplasia or an associated invasive carcinoma and, with a prior study demonstrating that among patients who underwent surgical resection only 25% of branch duct IPMN and 66% of main-duct IPMN harbored high-grade dysplasia or an associated invasive carcinoma [54,55] . This highlights the importance of finding more accurate markers of high-risk disease to aid clinical decision-making. The third category of radiomics-based cyst

classification studies have attempted to fill this gap through applying radiomics to risk-stratify patients with IPMNs. Studies have found that radiomics features extracted from CT, MRI and ultrasound have superior ability in identifying high-risk vs low-risk disease compared to clinical features and/or Fukuoka criteria. The most recent of such studies enrolled 66 patients and compared both MRI and CT radiomics models. In this study, the MRI model outperformed the CT radiomics model and achieved an AUC of 0.940 in preoperatively predicting malignant potential of IPMNs [56]. Numerous similar prior studies with MRI or CT radiomics model have been conducted, and despite methodological variation, these studies have reported comparably strong results (AUC range, 0.71–0.96) [57–61]. Prior studies have also integrated clinical models based on the Fukuoka guidelines with radiomics models and have demonstrated superior performance of the combined models [57,58]. Future integration of additional multidimensional data, such as novel radiogenomic features of cyst fluid DNA, into machine learning models has potential to further improve upon the performance of existing models [62]. Several deep learning studies have also explored malignancy prediction of IPMNs using EUS and CT images and have reported an accuracy range of 94.0–99.6% [63,64]. Given the shortcomings of existing risk assessment criteria, future clinical adoption of these tools has the potential to significantly impact clinical management.

7. Prediction of tumor grade

Surgical resection is currently the only curative treatment option for patients with pNETs and is recommended for all patients with functioning pNETs and pNETs larger than 2 cm [65]. However, it is unclear whether surgical resection or conservative watchful waiting is the best approach for managing small (≤ 2 cm) non-functioning pNETs. The histologic grade of a pNET is a crucial prognostic factor for predicting patient survival and radiomics features have shown promise in improving the classification of pNET tumor grades compared to traditional qualitative imaging features (ill-defined margins, main pancreatic duct dilatation, heterogeneous enhancement, low-level enhancement and vascular involvement). Liu et al. recently showed that a multimodality radiomics model combining both MRI and CT radiomics with clinical data could distinguish between non-functional G1 and G2/G3 pNET tumors with an AUC of 0.85 [66]. In another study, Pulvirenti et al. presented a CT radiomics model with a sensitivity and specificity at predicting pNET grade similar to EUS-FNA, which is the current gold standard for invasive diagnostic testing [67,68]. While earlier studies on this topic reported strong results (AUC range, 0.695–0.90), they were limited by relatively small sample sizes given the rarity of pNETs [69–75]. However, more recently published literature has predominantly consisted of larger cohorts of patients and has reported comparably strong results as reported in earlier studies (AUC range, 0.82–0.876), suggesting that incorporating radiomics features with traditional imaging features may indeed improve accuracy of pNET grade prediction [66,76,77].

8. Survival prediction

A unique application of radiomics beyond diagnostics has been to predict preoperatively a patient's survival should they undergo surgical resection of PDAC. Survival prediction of PDAC currently relies mainly on postoperative features such as the TNM stage and margins. This precludes preoperative survival prediction which could separate patients who will benefit from surgery from those who will not. Although surgical resection remains the only cure for PDAC, it is associated with significant complications and carries a small mortality risk. The development of an accurate preoperative survival prediction model could allow for a quantitative risk-benefit analysis prior to pancreatic resection and allow for individualized triaging of patients for surgery based on overall anticipated benefit from resection. In the current literature, four studies directly compared the prognostic performance of radiomics models with the clinical TNM staging criteria, with all four studies reporting that the radiomics models outperformed clinical criteria in predicting overall survival [78–81]. In one of these studies however, utility of the model at external validation was clinically insignificant, highlighting the inherent limitations of radiomics when deployed in real-world settings [81]. Deep learning models have also been applied to predict overall survival among patients undergoing resection [82]. Similar existing research has also shown that radiomics features can predict overall survival in patients with locally advanced or unresectable PDAC. Addition of carbohydrate antigen 19–9, an established prognostic serum biomarker, to these models has been demonstrated to improve further survival prediction [83,84]. Broadly, two themes have also emerged from current literature on radiomics based prognostication of PDAC: tumors with lower attenuation (represented by first-order median and first-order minimum) and tumors with more heterogenous texture (represented by first-order entropy) were predictive of worse survival. These findings are in line with previous studies which have shown that tumors with lower enhancement on CT, representing lower vascularity, have a worse response to treatment and shorter overall survival [81,84]. Lower attenuation is also linked to higher grade tumors, a higher risk of lymph node metastasis, and worse progression-free survival [81,83]. It has been previously suggested that lower attenuation may represent areas of hypoxic necrosis or venous invasion within tumors, both of which indicate aggressive tumor biology [85–88]. In extreme cases, these findings may be picked up by the human eye; however, for most cases, automated quantification is necessary to detect these changes at an earlier stage. Heterogeneity at a genomic level has also previously been associated with worse survival, which is also consistent with radiomics data in lung, esophageal and brain tumors reporting a similar association between tumor heterogeneity and survival [89–92]. However, the exact relationship between heterogeneity and survival in PDAC is not yet clear, as PDACs are genetically homogenous, and an earlier radiomic study reported the opposite association [93,94]. This highlights the larger issue of the limitations of drawing clinical conclusions from radiomic features in the absence of robust explanations of the underlying biological and pathophysiological processes those features are surrogates for.

9. Treatment response and surgical resectability

Neoadjuvant therapy for PDAC is associated with lower rates of post-operative nodal involvement and perineural invasion, and higher rates of negative margin resection [95,96]. However, determining the response to neoadjuvant treatment and resectability can be difficult. Radiomics have demonstrated the potential to identify a rapid response to chemotherapy and early down-staging to a surgically resectable tumor by evaluating the longitudinal evolution of radiomic features over chemoradiation cycles, termed delta radiomic features. Nasief et al. notably showed that CT delta radiomics features, particularly skewness and kurtosis (a measure of the shape of the distribution), could differentiate good responders from poor responders of chemoradiation therapy in a validation cohort of 40 patients with an AUC of 0.94 [97]. While prior studies had also reported these same features to be most predictive, other recent MRI and CT radiomics studies have reported similar associations but with different radiomics features [98,99]. This heterogeneity among radiomic features found to be significantly associated with treatment response makes reproducing results challenging and highlights the importance of feature selection in radiomic analyses. Radiomics features have also been found to be predictive of resection margin positivity and preoperative risk of lymph node metastases, and have been reported to outperform a multidisciplinary assessment of preoperative risk of superior mesenteric artery involvement by PDAC [100–102]. While more research is needed to understand and validate

the true capability of these radiomics models, if optimized, they offer the potential to allow for personalized tailoring of surgery, chemotherapy, or radiation therapy regimens to a patient's unique tumor imaging characteristics.

Beyond radiomics, a deep learning model applied to pre-chemotherapy CT examinations from patients with PDAC yielded an AUC of 0.74 for distinguishing pathological responders from non-responders [103]. Addition of semantic features, such as a 10% decrease in carbohydrate antigen 19–9 serum level, further improved the model AUC to 0.79 [103]. These results, however, have yet to be validated in an independent dataset.

10. Radiomics quality score

The Radiomics Quality Score (RQS) is a 16-item checklist designed to evaluate the reproducibility and validity of radiomics studies. Since first proposed by Lambin et al., the RQS has been applied to numerous radiomics studies to assess their quality [104]. Over time, the RQS has gradually continued to garner interest as the number of radiomics studies continues to grow, and the need for objective quality control measures in this burgeoning field becomes increasingly more pressing. While other quality assessment tools for radiomics studies have been proposed, their use is limited, and the RQS has become the de facto standard for radiomics quality assessment [105]. The first four RQS items primarily pertain to imaging and protocol quality and feature reproducibility. These assess image protocol quality/documentation, presence of multiple segmentations, incorporation of phantom studies to detect inter-scanner differences, and imaging at different time points to evaluate feature robustness to temporal variabilities. The next three RQS items focus on radiomics feature-related aspects and assess inclusion of feature reduction or adjustment, multivariable analysis with non-radiomics features and discussion of the biological correlates of selected features. The three subsequent RQS items address statistical quality checks, including determination of risk groups through cut-off analysis, inclusion of discrimination statistics, and reporting of calibration statistics. The eleventh RQS item checks for registry in a prospective trial database and is the most highly scored item in the checklist. The next two RQS items include performance-related checks through assessing the studies validation protocols and comparison to a gold standard. RQS items fourteen and fifteen evaluate the clinical transferability of results through assessing whether the study discussed potential clinical applications and conducted a cost-effectiveness analysis of clinical adoption. The final item evaluates whether the imaging data or radiomics extraction code used was open source. Not all items within the RQS checklist are weighed equally with some items assigned more points when fulfilled. Failure to meet the criteria of certain items also penalizes studies and can result in the detraction of points. The overall score is calculated out of a total of 36 and reported as a percentage.

The RQS has previously been classified into six domains [106]. Domain 1, protocol quality and stability in image and segmentation, consists of items 1–4. Domain 2, feature selection and validation, consists of items 5 and 12. Domain 3, biologic/clinical validation and utility consists of items 6, 7, 13 and 14. Domain 4, model performance index consists of items 8, 9, and 10. Domain 5, high level of evidence, consists of items 11 and 15. Domain 6, open science and data, consists of item 16.

In this review the RQS was applied to assess all 54 included radiomics studies of pancreatic imaging (Table 2). The mean RQS score across all studies was quite low (32.1%) (Table 3).

In domain 1, 94.4% of studies documented image acquisition protocols or the use of publicly available imaging database. Automatic segmentation or segmentation by more than one reader was carried out in 61.1% of studies. Only 9.3% of studies detected and adjusted for intra-scanner differences or employed a phantom. Extraction of radiomics features at different time points, known as delta radiomics, was only employed by two studies (3.7%) that were longitudinally evaluating treatment response using radiomics.

In domain 2, 94.4% used appropriate methods (LASSO, random forest, recursive feature elimination, support vector machines, false discovery rate with univariate logistic regression) for feature reduction or adjustment for multiple testing. While 81.5% of studies validated their results, the overall RQS for validation was low (19.3%) with only 22.2% of studies conducting external validation with an even smaller subset (7.4%) validating their model based on datasets from two distinct institutes. The RQS awards five full points for validation using data from three datasets from distinct institutes or validation of a previously established signature, and none of the included studies met these criteria.

In domain 3, most studies (70.7%) conducted multivariate analysis including non-radiomics features, to permit for inferencing between radiomics features and non-radiomics variables. Meanwhile, 31.5% of studies compared radiomics models with gold standards (e.g. radiologist's read for detection/classification or TNM staging for survival prediction). Only 13.0% of studies detected and discussed potential biological correlates of radiomics features.

In domain 4, all included studies reported discrimination statistics (e.g. AUC, C-statistic), with 53.7% reporting the additional application of a resampling method such as cross-validation or bootstrapping. Fewer studies conducted model calibration (33.3%) and cut-off analysis (18.5%).

Domain 5 reported the poorest performance with none of the included studies reporting a prospective study design or conducting cost effectiveness analysis.

Within domain 6, 44.4% reported utilization of an open-source feature extraction code (e.g. Pyradiomics, IBEX). However, only one study reported using open-source scans.

When RQS was stratified by domain, domain 3, biologic/clinical validation and utility, scored the highest (53.2%) while domain 5, high level of evidence, scored the lowest (0%) (Fig. 2). When RQS was stratified according to clinical problem addressed, studies on classification of the degree of dysplasia in an IPMN scored highest (39%) while studies on prediction of treatment response scored lowest (27%) (Fig. 3).

There were several key findings from our RQS analysis. Overall, our review found that we are currently doing well in terms of studies reporting on potential clinical utility, feature reduction or adjustment of multiple testing and discrimination statistics with these being the three highest ranked RQS criteria (items 14, 5, 9). This suggests that the radiomics feature extraction workflow of most studies is appropriate and that most studies provide some commentary on clinical applications of results. Studies within our review, however, were identified and classified based on their impact and relevance in addressing key clinical problems within pancreatic imaging. Therefore, we acknowledge the potential for selection bias among studies reporting clinical utility and associated discrimination statistics within this review. Another key takeaway was that literature is limited in terms of studies reporting prospective design, cost effectiveness analysis and imaging at multiple time points, with these being the 3 lowest ranked criteria. Prospectively comparing the stability of radiomics features over time and evaluating the cost effectiveness of AI tools is essential in determining clinical applicability, and future studies need to focus on fulfilling these criteria. Greater emphasis on validating results more robustly is also needed. While 81.5% of studies validated their results in some way, the overall RQS for validation was only 19.3%, which is indicative of the low overall quality of current validation approaches. External multicenter validation of future studies is needed to truly gauge the clinical generalizability of current studies.

Table 2
Radiomics quality score of individual studies.

Study [Ref#]	Image protocol quality (0,+1,+2)	Multiple segmentations (0,+1)	Phantom study (0,+1)	Imaging at multiple time points (0,+1)	Feature reduction or adjustment for multiple testing (−3 or +3)	Multivariable analysis with non-radiomics features (0,+1)	Detect and discuss biological correlates (0,+1)	Cut-off analyses (0,1)	Discrimination statistics (0, +1, +2)	Calibration statistics (0, +1, +2)	Prospective study registered in a trial database (0,+7)	Validation (0 to +5)	Comparison to 'gold standard' (0,+2)	Potential clinical utility (0,+2)	Cost-effectiveness analysis (0,+1)	Open science and data (0 - +4)	**Score (n/36)**
Chen PT et al. [12]	+1	+1	0	0	+3	0	0	0	+1	0	0	+4	0	+2	0	+2	14
Mukherjee et al. [14]	+1	+1	0	0	+3	0	0	0	+2	0	0	+4	+2	+2	0	+1	16
Javed et al. [16]	0	0	0	0	+3	0	0	0	+1	0	0	+3	0	+2	0	0	9
Qureshi et al. [15]	0	0	0	0	+3	0	0	0	+1	0	0	+3	0	+2	0	0	9
Chen et al. [17]	+1	+1	0	0	+3	0	0	0	+1	0	0	+2	+2	+2	0	0	12
Chu et al. [11]	+1	+1	0	0	+3	0	0	0	+1	0	0	+2	0	+2	0	0	13
Liu et al. [26]	+1	+1	0	0	+3	+1	0	0	+2	+1	0	+2	0	+2	0	+1	14
Deng et al. [27]	+1	+1	0	0	+3	+1	0	0	+2	+1	0	+3	+2	+2	0	+1	17
Zhang et al. [28]	+1	+1	0	0	+3	+1	0	0	+2	0	0	+2	0	+2	0	+1	13
Li et al. [31]	+1	+1	0	0	+3	+1	0	0	+2	0	0	+2	0	+2	0	+1	13
Park et al. [32]	+1	+1	0	0	+3	0	0	0	+1	0	0	+2	0	+2	0	0	11
Mashayekhi et al. [33]	+1	+1	0	+1	+3	0	0	0	+2	+1	0	+2	0	+2	0	0	13
Pol et al. [43]	+1	+1	0	0	+3	+1	0	0	+1	0	0	−5	0	+2	0	0	4
Huang et al. [42]	+1	0	0	0	+3	+1	0	0	+1	0	0	−5	+2	+2	0	0	5
Reinert et al. [39]	+1	0	+1	0	+3	+1	0	+1	+1	0	0	−5	0	+2	0	+1	6
Zhang et al. [40]	+1	+1	0	0	+3	+1	0	0	+2	0	0	+2	0	+2	0	0	12
Wang et al. [41]	+1	+1	0	0	−3	+1	0	0	+1	0	0	+2	0	+2	0	0	5
He et al. [38]	+1	0	0	0	+3	+1	0	0	+2	+1	0	+2	+2	+2	0	0	14
Shi et al. [44]	+1	+1	0	0	+3	+1	0	+1	+2	+1	0	+2	+2	+2	0	+1	17
Chu et al. [45]	+1	+1	0	0	+3	+1	0	0	+2	0	0	+2	+2	+2	0	0	14
Yang et al. [46]	+1	+1	0	0	+3	0	0	0	+1	0	0	+2	0	+2	0	0	10
Wei et al. [47]	+1	+1	0	0	+3	+1	+1	0	+2	0	0	+2	0	+2	0	0	13
Xie et al. [48]	+1	0	0	0	+3	+1	0	+1	+1	+2	0	−5	+2	+2	0	+1	9
Chen et al. [49]	+1	+1	0	0	+3	+1	0	0	+1	+1	0	+2	0	+2	0	0	12
Cheng et al. [56]	+1	+1	0	0	+3	+1	0	+1	+2	0	0	+2	+2	+2	0	0	15
Cui et al. [57]	+1	+1	0	0	+3	+1	+1	+1	+1	+1	0	+4	0	+2	0	+1	17
Attiyeh et al. [58]	+1	0	0	0	+3	+1	0	+1	+2	0	0	+2	0	+2	0	0	12
Chakraborty et al. [59]	+1	0	0	0	+3	+1	+1	+1	+2	0	0	+2	0	+2	0	0	13
Polk et al. [60]	+1	0	0	0	+3	+1	0	0	+2	0	0	+2	+2	+2	0	0	13
Tobaly et al. [61]	+1	0	0	0	+3	+1	0	+1	+2	0	0	+3	0	+2	0	+1	14
Liu et al. [66]	+1	+1	0	0	+3	+1	0	0	+1	+1	0	+2	0	+2	0	+1	13
Pulvirenti et al. [67]	0	0	0	0	+3	+1	0	0	+1	0	0	+2	+2	+2	0	0	11
Gu et al. [69]	+1	+1	+1	0	+3	+1	0	+1	+2	+1	0	+3	0	+2	0	+1	17
Chiti et al. [77]	+1	0	0	0	+3	0	0	0	+1	0	0	+2	0	+2	0	+1	10
Mori et al. [70]	+1	0	+1	0	+3	+1	0	0	+1	0	0	+2	0	+2	0	0	11
Bian et al. [71]	+1	+1	+1	0	+3	+1	0	0	+2	+1	0	+2	0	+2	0	+1	15
Li et al. [72]	+1	0	0	0	+3	0	+1	0	+1	0	0	−5	0	+2	0	+1	4
Zhao et al. [75]	+1	+1	0	0	+3	+1	+1	+1	+2	+2	0	+2	0	+2	0	0	16
Benedetti et al. [73]	+1	0	0	0	+3	+1	0	0	+1	0	0	−5	+2	+2	0	+1	6
Bevilacqua et al. [74]	+1	0	0	0	+3	0	0	+1	+2	0	0	+2	0	+2	0	0	11
Wang et al. [76]	+1	+1	0	0	+3	+1	0	0	+2	+1	0	+2	0	+2	0	0	11
Healy et al. [81]	+1	+1	0	0	+3	+1	+1	+1	+2	+1	0	+3	+2	+2	0	+1	19
Xie et al. [80]	+1	0	0	0	+3	+1	0	+1	+2	+1	0	+2	+2	+2	0	+1	16
Wang et al. [78]	+1	+1	0	0	+3	+1	0	+1	+2	+1	0	+4	+2	+2	0	+1	19
Shi et al. [79]	+1	+1	0	0	+3	+1	0	+1	+2	+1	0	+2	+2	+2	0	0	16
Cheng et al. [83]	+1	0	0	0	−3	+1	0	+1	+1	0	0	−5	0	+2	0	0	0
Attiyeh et al. [84]	+1	+1	0	0	+3	+1	+1	0	+1	0	0	+2	0	0	0	+1	11

(*continued on next page*)

Table 2 (*Continued*)

Study [Ref#]	Image protocol quality (0,+1,+2)	Multiple segmentations (0,+1)	Phantom study (0,+1)	Imaging at multiple time points (0,+1)	Feature reduction or adjustment for multiple testing (−3 or +3)	Multivariable analysis with non-radiomics features (0,+1)	Detect and discuss biological correlates (0,+1)	Cut-off analyses (0,1)	Discrimination statistics (0, +1, +2)	Calibration statistics (0, +1, +2)	Prospective study registered in a trial database (0,+7)	Validation (0 to +5)	Comparison to 'gold standard' (0,+2)	Potential clinical utility (0,+2)	Cost-effectiveness analysis (0,+1)	Open science and data (0 - +4)	**Score (n/36)**
Eilaghi et al. [93]	+1	0	0	0	−3	0	0	0	+1	0	0	−5	0	+2	0	0	0
Nasief et al. [97]	+1	+1	+1	+1	+3	0	0	0	+2	0	0	+2	0	+2	0	+1	14
Simpson et al. [98]	+1	0	0	+1	+3	0	0	0	+2	0	0	+2	0	+2	0	0	11
Gregucci et al. [99]	+1	0	0	0	−3	0	0	0	+1	0	0	−5	0	+2	0	+1	0
Rigorili et al. [100]	+1	+1	0	0	+3	+1	0	0	+1	0	0	−5	+2	+2	0	0	0
Bian et al. [101]	+1	+1	0	0	+3	+1	0	0	+2	+1	0	+2	0	+2	0	+1	14
Bian et al. [102]	+1	+1	0	0	+3	+1	0	+1	+2	0	0	+2	0	+2	0	+1	14

11. Opportunities and challenges

Ongoing advances in AI within pancreatic imaging continue to make yearly strides and demonstrate increasingly promising results. By the end of 2022, there were over 200 federally approved radiology-related AI tools, constituting a near 300% increase since 2020 [1]. However, despite recent advancements, only 30% of radiology practices reported using AI tools, highlighting certain challenges that precede widespread incorporation of AI into clinical practice [107]. These challenges can broadly be categorized into those pertaining to the quality and generalizability of existing evidence and those pertaining to the practical barriers to clinical adoption.

When discussing the quality and generalizability of current evidence, perhaps the biggest challenge that needs to be overcome is the consolidation of large public annotated datasets. Radiomics and deep learning need to be trained on large datasets, and model performance and generalizability are critically dependent on the quality and size of these datasets. While the existence of such datasets had been previously limited, efforts to grow them are underway, with a notable dataset such as Imagenet and the National Cancer Institute's The Cancer Imaging Archive already being used by one included study to externally validate their model [108,109]. In addition, efforts to synthetically augment datasets through deep learning methods such as neural style transfer and generative adversarial networks, which garnered public interest due to its application in "deepfake" media, have also demonstrated potential, but their current utility is uncertain [110–112]. Although efforts to develop these databases and augment datasets are ongoing, their utilization in the development and validation of recent radiomics studies remains limited, as highlighted by the low overall RQS for validation amongst included studies. Another unanticipated issue with the growth of these public datasets is that of "off label" usage of public datasets, where data published for one task are used for another, which can lead to inflated results of deep learning algorithms. Such usage, labelled as implicit "data crime", arises due to public data being processed with hidden processing pipelines that alter data features and has been demonstrated to produce significantly biased results with up to 48% artificial improvement in models when applied to public datasets [113]. The enormous variation of human anatomy and pathology also present a problem. Studies of radiomics and deep learning often use pathologically confirmed “extremes” of abnormal and normal. The large gray zone of variation in between the extremes is an enormous challenge for all AI-driven tools.

Another notable challenge uncovered by our study was the low overall RQS of existing radiomics literature. This result was also in line with prior RQS studies across various radiological subspecialties, indicating a systemic lack of compliance with quality metrics within AI research in radiology [105]. The low overall RQS of included studies represents the significant methodological heterogeneity of existing literature. This heterogeneity makes it challenging to compare results and draw meaningful conclusions. While validation of general observations with local datasets is valuable, repeated publications of these observations without upholding robust reporting standards add little to our understanding. Publication of such studies is particularly problematic within the context of a prior review reporting that up to 94% of published literature on radiomics has reported overwhelmingly positive findings, which is highly suggestive of a publication bias within the field of AI in radiology [114]. Publication of studies that report less than optimal and perhaps even negative results is equally valuable to tease out the true associations between radiomic features and tumor biology. This is especially important given that our understanding of biological correlates of radiomic features is very limited with only 13% of included studies detecting and discussing potential biologic correlates of selected features.

Another concerning quality-related finding of our review was how none of the included studies employed a prospective study

Table 3
Radiomics quality score and associated adherence rate of all included radiomics-based studies of pancreatic imaging.

Domain	Adherence rate	RQS (%)
1: Protocol quality and stability in image and segmentation	**(51) 94.4%**	**30.8%**
Protocol quality	(51) 94.4%	47.2%
Multiple segmentation	(33) 61.1%	61.1%
Phantom study	(5) 9.3%	9.3%
Imaging at multiple time points	(2) 3.7%	3.7%
2: Feature selection and validation	**(51) 94.4%**	**65.2%**
Feature reduction or adjustment of multiple testing	(50) 92.6%	85.2%
Validation	(44) 81.5%	19.3%
3: Biologic/clinical validation and utility	**(54) 100%**	**53.2%**
Multivariate analysis with non-radiomics features	(38) 70.4%	70.4%
Biological correlates	(7) 13.0%	13.0%
Comparison to 'gold standard'	(17) 31.5%	31.5%
Potential clinical utility	(53) 98.1%	98.1%
4: Model performance index	**(54) 100%**	**42.2%**
Discrimination statistics	(54) 100%	76.9%
Calibration statistics	(18) 33.3%	18.5%
Cut-off analysis	(17) 31.5%	31.5%
5: High level of evidence	**0%**	**0%**
Prospective study	(0) 0%	0%
Cost-effective analysis	(0) 0%	0%
6: Open science and data	**(25) 46.3%**	**24.1%**
Total	**72.5%**	**32.1%**

RQS indicates radiomics quality score.
The numbers in parentheses (n) indicate numbers of the studies adherent to each category.

design. While this is perhaps reflective of the general scarcity of prospectively designed studies within radiology compared to other specialties, large prospective studies remain essential for the clinical validation of AI tools.

Beyond quality and generalizability issues, several practical barriers also exist, with one being the determination of the added clinical value of these models. To truly gauge a model's clinical utility, it must be compared against existing gold standards. For studies reporting models that autonomously detect and classify pancreatic lesions, the gold standard remains the radiologists' reads, while among studies reporting models that predict patient survival, the gold standard is TNM staging. Amongst included studies, only 31.5% reported a comparison with these gold standards making interpretation of the net clinical benefit of most current models questionable. While the vast majority (98.1%) of reviewed studies discussed potential clinical applications, simply suggesting prospective methods in which AI models may be of value is no longer adequate and future studies should report objective metrics such as incremental value over gold standards or decision curve analyses alongside their models.

Another practical barrier to implementation is the financial sustainability of AI tools. None of the included studies conducted a cost effectiveness analysis. An oft-cited criticism of radiomics has been that it can demonstrate excellent results "in vivo" but lacks meaningful commentary on a framework for how these tools will actually translate to day-to-day clinical practice. While the lack of cost effectiveness analysis of included studies highlights this problem, this is not surprising given the economic complexity of developing a framework for how AI will realistically fit into clinical practice and billing. For AI tools to be financially sustainable they need to demonstrate significantly superior diagnostic performance compared to radiologists, greatly enhance radiologists' productivity or secure separate reimbursement to offset their added cost. This is all while being seamlessly assimilated into picture archiving and communication systems. Marginally improved diagnostic accuracy without direct cost or time savings attributable to AI tools may not be sufficient to justify clinical adoption. Although computer aided detection to standard breast imaging can be billed at higher rate in the United States, a similar framework does not yet exist for use in abdominal imaging. Future radiomics studies would therefore be prudent to put forth frameworks that incorporate cost effectiveness analyses and adhere to RQS criteria.

The final and arguably most formidable barrier to AI adoption is that of the legal hurdles associated with AI use in healthcare. The current legal framework for computer aided detection tools and for AI in triaging (e.g., AIdoc) is cloudy, and how this approach translates to the next generation of increasingly autonomous and diagnostic AI is uncertain [115]. Mistakes are inevitable and consideration should be given to the difficult questions that will arise when these mistakes happen.

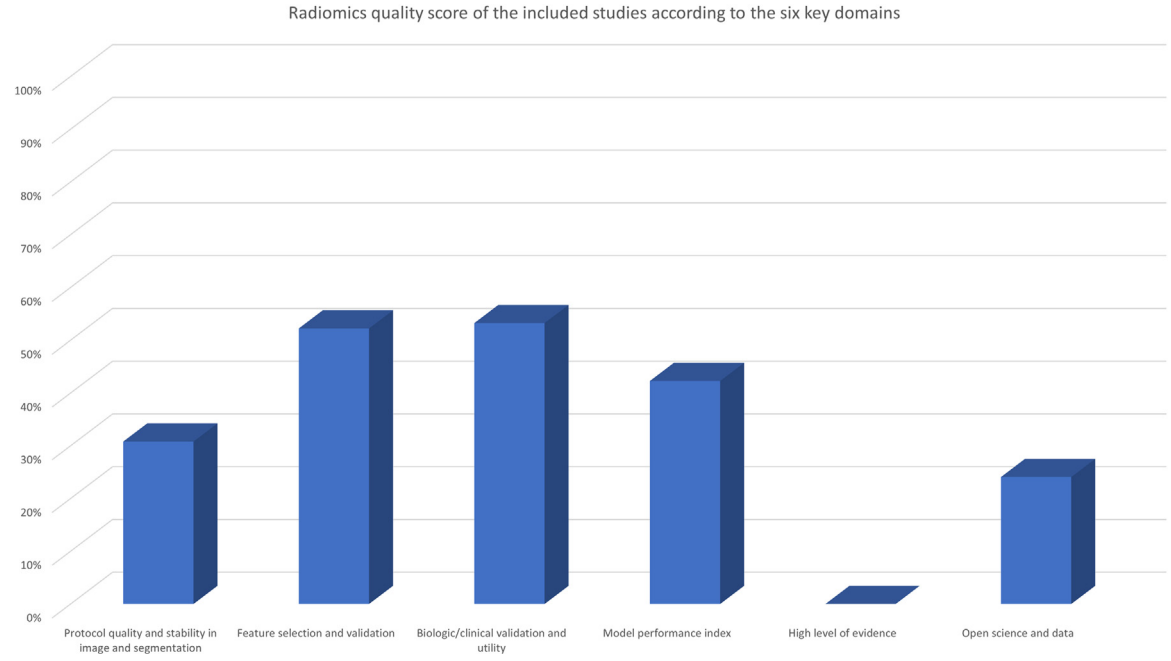

Fig. 2. Radiomics quality score of the included studies according to the six key domains.

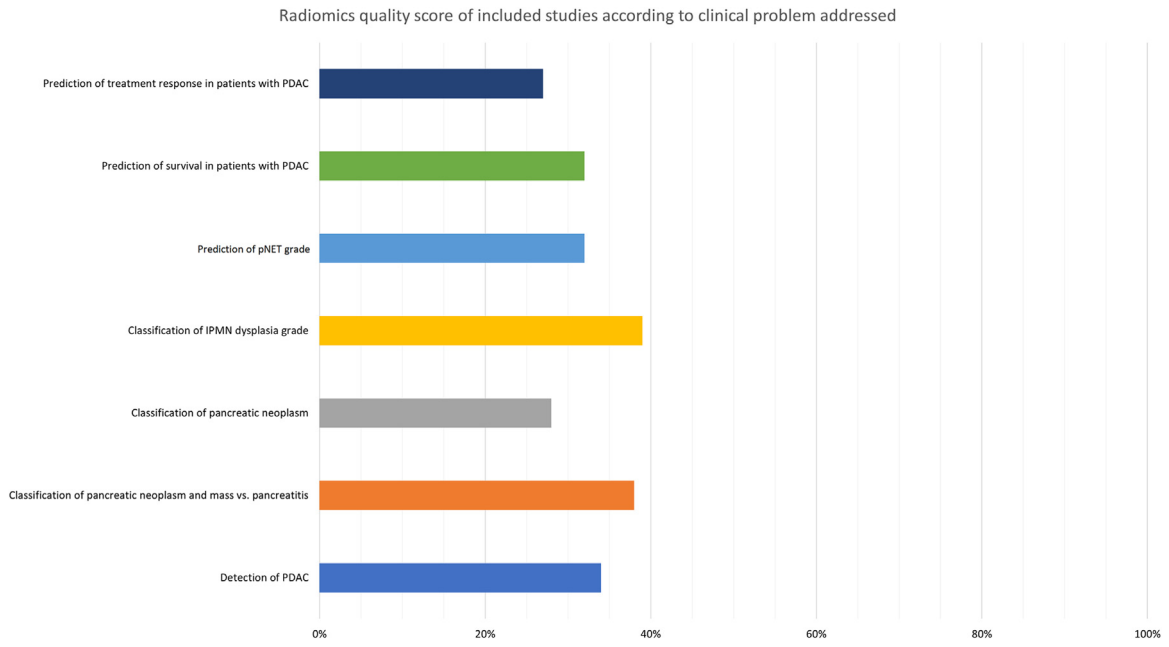

Fig. 3. Radiomics quality score of included studies according to clinical problem addressed. PDAC indicates pancreatic ductal adenocarcinoma. pNET indicates pancreatic neuroendocrine tumor.

Who is responsible for the accuracy of an autonomous AI system when it makes an error? How do we factor in the radiologist's liability when using AI tools? What is the liability of the health system that purchases an AI product? It has been suggested that analysis of how lawsuits involving autonomous cars, which share certain similarities with medical AI tools, have been handled by the courts could be instructive in providing a legal framework for medical AI [116]. However, autonomous cars are not entirely analogous to medical AI, and law governing AI remains in its infancy with substantial uncertainty regarding the manner in which liability for mistakes will be allocated and the extent to which that will impact and stunt future development of AI. Despite the various current shortcomings of AI, AI holds tremendous potential. Steps must be taken to ensure that this potential is not sidelined before it can be realized due to underlying legal issues.

12. Conclusion

AI has made significant progress in the detection, classification, and prognostication of pancreatic lesions, through techniques such as radiomics and deep learning [117]. Despite this promise, the quality of existing literature is far from robust. While we acknowledge that the potential value of existing literature may extend beyond what may be formally evaluated through the RQS tool, to fully realize the benefit of these advancements, current results need to be validated through higher quality studies and multicenter trials that include the full spectrum of normal and abnormal. Fundamental questions still need addressing before clinical adoption, and efforts to establish sound evidence for future studies is warranted [118]. Given the rate of discovery of AI in abdominal imaging however, we optimistically believe that these challenges will inevitably be overcome and that a future in which synergy between radiologists and machines will become the norm is not a matter of 'if' but only a matter of 'when' [119].

Human rights

The authors declare that the work described has been performed in accordance with the Declaration of Helsinki of the World Medical Association revised in 2013 for experiments involving humans.

Informed consent and patient details

The authors declare that this article does not contain any personal information that could lead to the identification of the patients.

Disclosure of interest

The authors do not have any conflict of interest to disclose in relation with this article.

Funding

This study was supported by the Lustgarten Foundation (recipient, Taha M. Ahmed)

Author contributions

All authors attest that they meet the current International Committee of Medical Journal Editors (ICMJE) criteria for Authorship. The author(s) declare(s) that they had full access to all of the data in this study and the author(s) take(s) complete responsibility for the integrity of the data and the accuracy of the data analysis.

References

[1] AI Central n.d. https://aicentral.acrdsi.org/ (accessed November 30, 2022).
[2] Cheng PM, Montagnon E, Yamashita R, Pan I, Cadrin-Chênevert A, Perdigón Romero F, et al. Deep learning: an update for radiologists. Radiographics 2021;41:1427–45.
[3] Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. Radiographics 2017;37:505–15.
[4] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563–77.
[5] Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. Magn Reson Imaging 2012;30:1234–48.
[6] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumor phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
[7] Boeken T, Feydy J, Lecler A, Soyer P, Feydy A, Barat M, et al. Artificial intelligence in diagnostic and interventional radiology: where are we now? Diagn Interv Imaging 2023;104:1–5.
[8] Gangi S, Fletcher JG, Nathan MA, Christensen JA, Harmsen WS, Crownhart BS, et al. Time interval between abnormalities seen on CT and the clinical diagnosis

of pancreatic cancer: retrospective review of CT scans obtained before diagnosis. AJR Am J Roentgenol 2004;182:897–903.
[9] Gonoi W, Hayashi TY, Okuma H, Akahane M, Nakai Y, Mizuno S, et al. Development of pancreatic cancer is predictable well in advance using contrast-enhanced CT: a case-cohort study. Eur Radiol 2017;27:4941–50.
[10] Ozkan M, Cakiroglu M, Kocaman O, Kurt M, Yilmaz B, Can G, et al. Age-based computer-aided diagnosis approach for pancreatic cancer on endoscopic ultrasound images. Endosc Ultrasound 2016;5:101–7.
[11] Chu LC, Park S, Kawamoto S, Fouladi DF, Shayesteh S, Zinreich ES, et al. Utility of CT radiomics features in differentiation of pancreatic ductal adenocarcinoma from normal pancreatic tissue. AJR Am J Roentgenol 2019;213:349–57.
[12] Chen PT, Chang D, Yen H, Liu K-L, Huang SY, Roth H, et al. Radiomic features at CT can distinguish pancreatic cancer from noncancerous pancreas. Radiol Imaging Cancer 2021;3:e210010.
[13] Chen PT, Wu T, Wang P, Chang D, Liu K-L, Wu M-S, et al. Pancreatic cancer detection on CT scans with deep learning: a nationwide population-based study. Radiology 2023;306:172–82.
[14] Mukherjee S, Patra A, Khasawneh H, Korfiatis P, Rajamohan N, Suman G, et al. Radiomics-based machine-learning models can detect pancreatic cancer on prediagnostic CT at a substantial lead time before clinical diagnosis. Gastroenterology 2022;163:1435–46 e3.
[15] Qureshi TA, Gaddam S, Wachsman AM, Wang L, Azab L, Asadpour V, et al. Predicting pancreatic ductal adenocarcinoma using artificial intelligence analysis of pre-diagnostic computed tomography images. Cancer Biomark 2022;33:211–7.
[16] Javed S, Qureshi TA, Gaddam S, Wang L, Azab L, Wachsman AM, et al. Risk prediction of pancreatic cancer using AI analysis of pancreatic subregions in computed tomography images. Front Oncol 2022;12:1007990.
[17] Chen W, Zhou Y, Asadpour V, Parker RA, Puttock EJ, Lustigova E, et al. Quantitative radiomic features from CT can predict pancreatic cancer up to 36 months before diagnosis. Clin Transl Gastroenterol 2023;14:e00548.
[18] Rezaee N, Barbon C, Zaki A, He J, Salman B, Hruban RH, et al. Intraductal papillary mucinous neoplasm (IPMN) with high-grade dysplasia is a risk factor for the subsequent development of pancreatic ductal adenocarcinoma. HPB 2016;18:236–46.
[19] Nakahodo J, Kikuyama M, Fukumura Y, Horiguchi S, Chiba K, Tabata H, et al. Focal pancreatic parenchyma atrophy is a harbinger of pancreatic cancer and a clue to the intraductal spreading subtype. Pancreatology 2022;22:1148–58.
[20] Raimondi S, Lowenfels AB, Morselli-Labate AM, Maisonneuve P, Pezzilli R. Pancreatic cancer in chronic pancreatitis: etiology, incidence, and early detection. Best Pract Res Clin Gastroenterol 2010;24:349–58.
[21] Park HJ, Shin K, You MW, Kyung SG, Kim SY, Park SH, et al. Deep learning-based detection of solid and cystic pancreatic neoplasms at contrast-enhanced CT. Radiology 2023;306:140–9.
[22] Abel L, Wasserthal J, Weikert T, Sauter AW, Nesic I, Obradovic M, et al. Automated detection of pancreatic cystic lesions on CT using deep learning. Diagnostics 2021;11:901.
[23] Si K, Xue Y, Yu X, Zhu X, Li Q, Gong W, et al. Fully end-to-end deep-learning-based diagnosis of pancreatic tumors. Theranostics 2021;11:1982–90.
[24] Fritscher-Ravens A, Brand L, Knöfel WT, Bobrowski C, Topalidis T, Thonke F, et al. Comparison of endoscopic ultrasound-guided fine needle aspiration for focal pancreatic lesions in patients with normal parenchyma and chronic pancreatitis. Am J Gastroenterol 2002;97:2768–75.
[25] Tong T, Gu J, Xu D, Song L, Zhao Q, Cheng F, et al. Deep learning radiomics based on contrast-enhanced ultrasound images for assisted diagnosis of pancreatic ductal adenocarcinoma and chronic pancreatitis. BMC Med 2022;20:74.
[26] Liu J, Hu L, Zhou B, Wu C, Cheng Y. Development and validation of a novel model incorporating MRI-based radiomics signature with clinical biomarkers for distinguishing pancreatic carcinoma from mass-forming chronic pancreatitis. Transl Oncol 2022;18:101357.
[27] Deng Y, Ming B, Zhou T, Wu J-L, Chen Y, Liu P, et al. Radiomics model based on MR images to discriminate between pancreatic ductal adenocarcinoma and mass-forming chronic pancreatitis. Front Oncol 2021;11:620981.
[28] Zhang H, Meng Y, Li Q, Yu J, Liu F, Fang X, et al. Two nomograms for differentiating mass-forming chronic pancreatitis from pancreatic ductal adenocarcinoma in patients with chronic pancreatitis. Eur Radiol 2022;32:6336–47.
[29] Morishima T, Kawashima H, Ohno E, Yamamura T, Funasaka K, Nakamura M, et al. Prospective multicenter study on the usefulness of endosocopic ultrasound-guided fine-needle biopsy for the diagnosis of autoimmune pancreatitis. Gastrointest Endosc 2016;84:241–8.
[30] Ishikawa T, Kawashima H, Ohno E, Suhara H, Hayashi D, Hiramatsu T, et al. Usefulness of endoscopic ultrasound-guided fine-needle biopsy for the diagnosis of autoimmune pancreatitis using a 22-gauge Franseen needle: a prospective multicenter study. Endoscopy 2020;52:978–85.
[31] Li J, Liu F, Fang X, Cao K, Meng Y, Zhang H, et al. CT radiomics features in differentiation of autoimmune pancreatitis from pancreatic ductal adenocarcinoma: a propensity score analysis. Acad Radiol 2022;29:358–66.
[32] Park S, Chu LC, Hruban RH, Vogelstein B, Kinzler KW, Yuille AL, et al. Differentiating autoimmune pancreatitis from pancreatic ductal adenocarcinoma with CT radiomics features. Diagn Interv Imaging 2020;101:555–64.
[33] Mashayekhi R, Parekh VS, Faghih M, Singh VK, Jacobs MA, Zaheer A. Radiomic features of the pancreas on CT imaging accurately differentiate functional abdominal pain, recurrent acute pancreatitis, and chronic pancreatitis. Eur J Radiol 2020;123:108778.
[34] Săftoiu A, Vilmann P, Gorunescu F, Janssen J, Hocke M, Larsen M, et al. Efficacy of an artificial neural network-based approach to endoscopic ultrasound elastography in diagnosis of focal pancreatic masses. Clin Gastroenterol Hepatol 2012;10:84–90 e1.
[35] Das A, Nguyen CC, Li F, Li B. Digital image analysis of EUS images accurately differentiates pancreatic cancer from chronic pancreatitis and normal tissue. Gastrointest Endosc 2008;67:861–7.
[36] Săftoiu A, Vilmann P, Gorunescu F, Gheonea DI, Gorunescu M, Ciurea T, et al. Neural network analysis of dynamic sequences of EUS elastography used for the differential diagnosis of chronic pancreatitis and pancreatic cancer. Gastrointest Endosc 2008;68:1086–94.
[37] Tonozuka R, Itoi T, Nagata N, Kojima H, Sofuni A, Tsuchiya T, et al. Deep learning analysis for the detection of pancreatic cancer on endosonographic images: a pilot study. J Hepatobiliary Pancreat Sci 2021;28:95–104.
[38] He M, Liu Z, Lin Y, Wan J, Li J, Xu K, et al. Differentiation of atypical non-functional pancreatic neuroendocrine tumor and pancreatic ductal adenocarcinoma using CT based radiomics. Eur J Radiol 2019;117:102–11.
[39] Reinert CP, Baumgartner K, Hepp T, Bitzer M, Horger M. Complementary role of computed tomography texture analysis for differentiation of pancreatic ductal adenocarcinoma from pancreatic neuroendocrine tumors in the portal-venous enhancement phase. Abdom Radiol 2020;45:750–8.
[40] Tao Z, Yu X, Hang W, Hong Y, Yiu L, Xing W, et al. Multiple machine learning in differentiating pancreatic ductal adenocarcinoma from pancreatic neuroendocrine tumor. Clin Med 2022;11:6789.
[41] Wang Z, Chen X, Wang J, Cui W, Ren S, Wang Z. Differentiating hypovascular pancreatic neuroendocrine tumors from pancreatic ductal adenocarcinoma based on CT texture analysis. Acta Radiol 2020;61:595–604.
[42] Huang Z, Li M, He D, Wei Y, Yu H, Wang Y, et al. Two-dimensional texture analysis based on CT images to differentiate pancreatic lymphoma and pancreatic adenocarcinoma: a preliminary study. Acad Radiol 2019;26:e189–95.
[43] van der Pol CB, Lee S, Tsai S, Larocque N, Alayed A, Williams P, et al. Differentiation of pancreatic neuroendocrine tumors from pancreas renal cell carcinoma metastases on CT using qualitative and quantitative features. Abdom Radiol 2019;44:992–9.
[44] Shi YJ, Zhu HT, Liu YL, Wei YY, Qin XB, Zhang XY, et al. Radiomics analysis based on diffusion kurtosis imaging and T2-weighted imaging for differentiation of pancreatic neuroendocrine tumors from solid pseudopapillary tumors. Front Oncol 2020;10:1624.
[45] Chu LC, Park S, Soleimani S, Fouladi DF, Shayesteh S, He J, et al. Classification of pancreatic cystic neoplasms using radiomic feature analysis is equivalent to an experienced academic radiologist: a step toward computer-augmented diagnostics for radiologists. Abdom Radiol 2022;47:4139–50.
[46] Yang J, Guo X, Ou X, Zhang W, Ma X. Discrimination of pancreatic serous cystadenomas from mucinous cystadenomas with CT textural features based on machine learning. Front Oncol 2019;9:494.
[47] Wei R, Lin K, Yan W, Guo Y, Wang Y, Li J, et al. Computer-aided diagnosis of pancreas serous cystic neoplasms: a radiomics method on preoperative MDCT images. Technol Cancer Res Treat 2019;18:1533033818824339.
[48] Xie H, Ma S, Guo X, Zhang X, Wang X. Preoperative differentiation of pancreatic mucinous cystic neoplasm from macrocystic serous cystic adenoma using radiomics: preliminary findings and comparison with radiological model. Eur J Radiol 2020;122:108747.
[49] Chen S, Ren S, Guo K, Daniels MJ, Wang Z, Chen R. Preoperative differentiation of serous cystic neoplasms from mucin-producing pancreatic cystic neoplasms using a CT-based radiomics nomogram. Abdom Radiol 2021;46:2637–46.
[50] Yang R, Chen Y, Sa G, Li K, Hu H, Zhou J, et al. CT classification model of pancreatic serous cystic neoplasms and mucinous cystic neoplasms based on a deep neural network. Abdom Radiol 2022;47:232–41.
[51] Dmitriev K, Kaufman AE, Javed AA, Hruban RH, Fishman EK, Lennon AM, et al. Classification of pancreatic cysts in computed tomography images using a random forest and convolutional neural network ensemble. Med Image Comput Comput Assist Interv 2017;10435:150–8.
[52] Li H, Shi K, Reichert M, Lin K, Tselousov N, Braren R, et al. Differential diagnosis for pancreatic cysts in CT scans using densely-connected convolutional networks. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:2095–8.
[53] Liang W, Tian W, Wang Y, Wang P, Wang Y, Zhang H, et al. Classification prediction of pancreatic cystic neoplasms based on radiomics deep learning models. BMC Cancer 2022;22:1237.
[54] Tanaka M, Fernández-Del Castillo C, Kamisawa T, Jang JY, Levy P, Ohtsuka T, et al. Revisions of international consensus Fukuoka guidelines for the management of IPMN of the pancreas. Pancreatology 2017;17:738–53.
[55] Pulvirenti A, Margonis GA, Morales-Oyarvide V, McIntyre CA, Lawrence SA, Goldman DA, et al. Intraductal papillary mucinous neoplasms: have IAP consensus guidelines changed our approach? Results from a multi-institutional study. Ann Surg 2021;274:e980–7.
[56] Cheng S, Shi H, Lu M, Wang C, Duan S, Xu Q, et al. Radiomics analysis for predicting malignant potential of intraductal papillary mucinous neoplasms of the pancreas: comparison of CT and MRI. Acad Radiol 2022;29:367–75.
[57] Cui S, Tang T, Su Q, Wang Y, Shu Z, Yang W, 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;21:26.
[58] Attiyeh MA, Chakraborty J, Gazit L, Langdon-Embry L, Gonen M, Balachandran VP, et al. Preoperative risk prediction for intraductal papillary mucinous neoplasms by quantitative CT image analysis. HPB 2019;21:212–8.
[59] Chakraborty J, Midya A, Gazit L, Attiyeh M, Langdon-Embry L, Allen PJ, et al. CT radiomics to predict high-risk intraductal papillary mucinous neoplasms of the pancreas. Med Phys 2018;45:5019–29.

[60] Polk SL, Choi JW, McGettigan MJ, Rose T, Ahmed A, Kim J, et al. Multiphase computed tomography radiomics of pancreatic intraductal papillary mucinous neoplasms to predict malignancy. World J Gastroenterol 2020;26:3458–71.
[61] Tobaly D, Santinha J, Sartoris R, Dioguardi Burgio M, Matos C, Cros J, et al. CT-based radiomics analysis to predict malignancy in patients with intraductal papillary mucinous neoplasm of the pancreas. Cancers 2020;12:3089.
[62] Springer S, Masica DL, Dal Molin M, Douville C, Thoburn CJ, Afsari B, et al. A multimodality test to guide the management of patients with a pancreatic cyst. Sci Transl Med 2019;11:eaav4772.
[63] Schulz D, Heilmaier M, Phillip V, Treiber M, Mayr U, Lahmer T, et al. Accurate prediction of histological grading of intraductal papillary mucinous neoplasia using deep learning. Endoscopy 2023;55:1971.
[64] Kuwahara T, Hara K, Mizuno N, Okuno N, Matsumoto S, Obata M, et al. Usefulness of deep learning analysis for the diagnosis of malignancy in intraductal papillary mucinous neoplasms of the pancreas. Clin Transl Gastroenterol 2019;10:1–8.
[65] Falconi M, Eriksson B, Kaltsas G, Bartsch DK, Capdevila J, Caplin M, et al. ENETS consensus guidelines update for the management of patients with functional pancreatic neuroendocrine tumors and non-functional pancreatic neuroendocrine tumors. Neuroendocrinology 2016;103:153–71.
[66] Liu C, Bian Y, Meng Y, Liu F, Cao K, Zhang H, et al. Preoperative prediction of G1 and G2/3 grades in patients with nonfunctional pancreatic neuroendocrine tumors using multimodality imaging. Acad Radiol 2022;29:e49–60.
[67] Pulvirenti A, Yamashita R, Chakraborty J, Horvat N, Seier K, McIntyre CA, et al. Quantitative CT image analysis to predict pancreatic neuroendocrine tumor grade. JCO Clin Cancer Inform 2021:679–94.
[68] Boutsen L, Jouret-Mourin A, Borbath I, van Maanen A, Weynand B. Accuracy of pancreatic neuroendocrine tumour grading by endoscopic ultrasound-guided fine needle aspiration: analysis of a large cohort and perspectives for improvement. Neuroendocrinology 2018;106:158–66.
[69] Gu D, Hu Y, Ding H, Wei J, Chen K, Liu H, et al. CT radiomics may predict the grade of pancreatic neuroendocrine tumors: a multicenter study. Eur Radiol 2019;29:6880–90.
[70] Mori M, Palumbo D, Muffatti F, Partelli S, Mushtaq J, Andreasi V, et al. Prediction of the characteristics of aggressiveness of pancreatic neuroendocrine neoplasms (PanNENs) based on CT radiomic features. Eur Radiol 2022. doi: 10.1007/s00330-022-09351-9.
[71] Bian Y, Zhao Z, Jiang H, Fang X, Li J, Cao K, et al. Noncontrast radiomics approach for predicting grades of nonfunctional pancreatic neuroendocrine tumors. J Magn Reson Imaging 2020;52:1124–36.
[72] Li W, Xu C, Ye Z. Prediction of pancreatic neuroendocrine tumor grading risk based on quantitative radiomic analysis of MR. Front Oncol 2021;11:758062.
[73] Benedetti G, Mori M, Panzeri MM, Barbera M, Palumbo D, Sini C, et al. CT-derived radiomic features to discriminate histologic characteristics of pancreatic neuroendocrine tumors. Radiol Med 2021;126:745–60.
[74] Bevilacqua A, Calabrò D, Malavasi S, Ricci C, Casadei R, Campana D, et al. A [68Ga] Ga-DOTANOC PET/CT radiomic model for non-invasive prediction of tumor grade in pancreatic neuroendocrine tumors. Diagnostics 2021;11:870.
[75] Zhao Z, Bian Y, Jiang H, Fang X, Li J, Cao K, et al. CT-radiomic approach to predict G1/2 nonfunctional pancreatic neuroendocrine tumor. Acad Radiol 2020;27:e272–81.
[76] Wang X, Qiu JJ, Tan CL, Chen YH, Tan QQ, Ren SJ, et al. Development and validation of a novel radiomics-based nomogram with machine learning to predict histologic grade in pancreatic neuroendocrine tumors. Front Oncol 2022;12:843376.
[77] Chiti G, Grazzini G, Flammia F, Matteuzzi B, Tortoli P, Bettarini S, et al. Gastroenteropancreatic neuroendocrine neoplasms (GEP-NENs): a radiomic model to predict tumor grade. Radiol Med 2022;127:928–38.
[78] Wang F, Zhao Y, Xu J, Shao S, Yu D. Development and external validation of a radiomics combined with clinical nomogram for preoperative prediction prognosis of resectable pancreatic ductal adenocarcinoma patients. Front Oncol 2022;12:1037672.
[79] Shi H, Wei Y, Cheng S, Lu Z, Zhang K, Jiang K, et al. Survival prediction after upfront surgery in patients with pancreatic ductal adenocarcinoma: radiomic, clinic-pathologic and body composition analysis. Pancreatology 2021;21:731–7.
[80] Xie T, Wang X, Li M, Tong T, Yu X, Zhou Z. Pancreatic ductal adenocarcinoma: a radiomics nomogram outperforms clinical model and TNM staging for survival estimation after curative resection. Eur Radiol 2020;30:2513–24.
[81] Healy GM, Salinas-Miranda E, Jain R, Dong X, Deniffel D, Borgida A, et al. Preoperative radiomics model for prognostication in resectable pancreatic adenocarcinoma with external validation. Eur Radiol 2022;32:2492–505.
[82] Yao J, Cao K, Hou Y, Zhou J, Xia Y, Nogues I, et al. Deep learning for fully automated prediction of overall survival in patients undergoing resection for pancreatic cancer: a retrospective multicenter study. Ann Surg 2022. doi: 10.1097/SLA.0000000000005465.
[83] Cheng SH, Cheng YJ, Jin ZY, Xue HD. Unresectable pancreatic ductal adenocarcinoma: role of CT quantitative imaging biomarkers for predicting outcomes of patients treated with chemotherapy. Eur J Radiol 2019;113:188–97.
[84] Attiyeh MA, Chakraborty J, Doussot A, Langdon-Embry L, Mainarich S, Gönen M, et al. Survival prediction in pancreatic ductal adenocarcinoma by quantitative computed tomography image analysis. Ann Surg Oncol 2018;25:1034–42.
[85] Hruban RH, Gaida MM, Thompson E, Hong S-M, Noë M, Brosens LA, et al. Why is pancreatic cancer so deadly? The pathologist's view. J Pathol 2019;248:131–41.
[86] Sandrasegaran K, Lin Y, Asare-Sawiri M, Taiyini T, Tann M. CT texture analysis of pancreatic cancer. Eur Radiol 2019;29:1067–73.
[87] Lee SJ, Sung Y-N, Kim SJ, Shin S, Cho H, Hruban RH, et al. Comprehensive histological evaluation with clinical analysis of venous invasion in pancreatic ductal adenocarcinoma: from histology to clinical implications. Pancreatology 2020;20:1486–94.
[88] Shin J, Wood LD, Hruban RH, Hong S-M. Desmin and CD31 immunolabeling for detecting venous invasion of the pancreatobiliary tract cancers. PLoS One 2020;15:e0242571.
[89] Shim KY, Chung SW, Jeong JH, Hwang I, Park CK, Kim TM, et al. Radiomics-based neural network predicts recurrence patterns in glioblastoma using dynamic susceptibility contrast-enhanced MRI. Sci Rep 2021;11:9974.
[90] Larue RTHM, Klaassen R, Jochems A, Leijenaar RTH, Hulshof MCCM, van Berge Henegouwen MI, et al. Pre-treatment CT radiomics to predict 3-year overall survival following chemoradiotherapy of esophageal cancer. Acta Oncol 2018;57:1475–81.
[91] Ganeshan B, Skogen K, Pressney I, Coutroubis D, Miles K. Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. Clin Radiol 2012;67:157–64.
[92] Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature 2013;501:338–45.
[93] Eilaghi A, Baig S, Zhang Y, Zhang J, Karanicolas P, Gallinger S, et al. CT texture features are associated with overall survival in pancreatic ductal adenocarcinoma - a quantitative analysis. BMC Med Imaging 2017;17:38.
[94] Yachida S, Jones S, Bozic I, Antal T, Leary R, Fu B, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature 2010;467:1114–7.
[95] Fong ZV, Verdugo FL, Fernandez-Del Castillo C, Ferrone CR, Allen JN, Blaszkowsky LS, et al. Tolerability, attrition rates, and survival outcomes of neoadjuvant FOLFIRINOX for non-metastatic pancreatic adenocarcinoma: intent-to-treat analysis. J Am Coll Surg 2023. doi: 10.1097/XCS.0000000000000499.
[96] Brown ZJ, Heh V, Labiner HE, Brock GN, Ejaz A, Dillhoff M, et al. Surgical resection rates after neoadjuvant therapy for localized pancreatic ductal adenocarcinoma: meta-analysis. Br J Surg 2022;110:34–42.
[97] Nasief H, Zheng C, Schott D, Hall W, Tsai S, Erickson B, et al. A machine learning based delta-radiomics process for early prediction of treatment response of pancreatic cancer. NPJ Precis Oncol 2019;3:25.
[98] Simpson G, Spieler B, Dogan N, Portelance L, Mellon EA, Kwon D, et al. Predictive value of 0.35 T magnetic resonance imaging radiomic features in stereotactic ablative body radiotherapy of pancreatic cancer: a pilot study. Med Phys 2020;47:3682–90.
[99] Gregucci F, Fiorentino A, Mazzola R, Ricchetti F, Bonaparte I, Surgo A, et al. Radiomic analysis to predict local response in locally advanced pancreatic cancer treated with stereotactic body radiation therapy. Radiol Med 2022;127:100–7.
[100] Rigiroli F, Hoye J, Lerebours R, Lafata KJ, Li C, Meyer M, et al. CT radiomic features of superior mesenteric artery involvement in pancreatic ductal adenocarcinoma: a pilot study. Radiology 2021;301:610–22.
[101] Bian Y, Guo S, Jiang H, Gao S, Shao C, Cao K, et al. Radiomics nomogram for the preoperative prediction of lymph node metastasis in pancreatic ductal adenocarcinoma. Cancer Imaging 2022;22:4.
[102] Bian Y, Jiang H, Ma C, Cao K, Fang X, Li J, et al. Performance of CT-based radiomics in diagnosis of superior mesenteric vein resection margin in patients with pancreatic head cancer. Abdom Radiol 2020;45:759–73.
[103] Watson MD, Baimas-George MR, Murphy KJ, Pickens RC, Iannitti DA, Martinie JB, et al. Pure and hybrid deep learning models can predict pathologic tumor response to neoadjuvant therapy in pancreatic adenocarcinoma: a pilot study. Am Surg 2021;87:1901–9.
[104] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62.
[105] Spadarella G, Stanzione A, Akinci D, Antonoli T, Andreychenko A, Fanni SC, Ugga L, et al. Systematic review of the radiomics quality score applications: an EuSoMII radiomics auditing group initiative. Eur Radiol 2023;33:1884–94.
[106] Park JE, Kim HS, Kim D, Park SY, Kim JY, Cho SJ, et al. A systematic review reporting quality of radiomics research in neuro-oncology: toward clinical utility and quality improvement using high-dimensional imaging features. BMC Cancer 2020;20:29.
[107] Allen B, Agarwal S, Coombs L, Wald C, Dreyer K. 2020 ACR data science institute artificial intelligence survey. J Am Coll Radiol 2021;18:1153–9.
[108] Welcome to the cancer imaging archive. The cancer imaging archive (TCIA). https://www.cancerimagingarchive.net/(accessed January 30.
[109] ImageNet. https://www.image-net.org/index.php (accessed January 30, 2023).
[110] Barat M, Marchese U, Pellat A, Dohan A, Coriat R, Hoeffel C, et al. Imaging of pancreatic ductal adenocarcinoma: an update on recent advances. Can Assoc Radiol J 2022. doi: 10.1177/08465371221124927.
[111] Wolterink JM, Mukhopadhyay A, Leiner T, Vogl TJ, Bucher AM, Išgum I. Generative adversarial networks: a primer for radiologists. Radiographics 2021;41:840–57.
[112] Couteaux V, Zhang C, Mulé S, Milot L, Valette PJ, Raynaud C, et al. Synthetic MR image generation of macrotrabecular-massive hepatocellular carcinoma using generative adversarial networks. Diagn Interv Imaging 2023. doi: 10.1016/j.diii.2023.01.003.

[113] Shimron E, Tamir JI, Wang K, Lustig M. Implicit data crimes: machine learning bias arising from misuse of public data. Proc Natl Acad Sci USA 2022;119: e2117203119.
[114] Buvat I, Orlhac F. The dark side of radiomics: on the paramount importance of publishing negative results. J Nucl Med 2019;60:1543–4.
[115] Healthcare AI | AIdoc always-on AI. AIdoc. https://www.aidoc.com/(accessed January 31.
[116] Mezrich JL. Is artificial intelligence (AI) a pipe dream? Why legal issues present significant hurdles to AI autonomy. AJR Am J Roentgenol 2022;219:152–6.
[117] Barat M, Chassagnon G, Dohan A, Gaujoux S, Coriat R, Hoeffel C, et al. Artificial intelligence: a critical review of current applications in pancreatic imaging. Jpn J Radiol 2021;39:514–23.
[118] Soyer P, Fishman EK, Rowe SP, Patlas MN, Chassagnon G. Does artificial intelligence surpass the radiologist? Diagn Interv Imaging 2022;103: 445–7.
[119] Rowe SP, Soyer P, Fishman EK. The future of radiology: what if artificial intelligence is really as good as predicted? Diagn Interv Imaging 2022;103: 385–6.