

REVIEW

Open Access



Artificial intelligence in pancreatic cancer histopathology and diagnostics - implications for clinical decisions and biomarker discovery?

Petra Weselá^{1,2}, Michal Eid^{3,4}, Petr Moravčík⁵, Jakub Vlažný^{4,6}, Jan Hlavsa⁵, Vladimír Procházka⁵, Zdeněk Kala⁵ and Petr Vaňhara^{1,7*}

Abstract

Artificial intelligence (AI) and machine learning (ML) are rapidly advancing fields within computer science, driving significant progress in cancer diagnostics. Various ML models have been developed to assist diagnosis, guide therapy decisions, and facilitate early disease detection. In this review, we discuss diverse AI and ML approaches and critically evaluate their applications and limitations in pancreatic cancer histopathology, diagnostics, and biomarker discovery.

Keywords Artificial intelligence, Machine learning, Multimodal learning, Pancreatic cancer, Pancreatic ductal adenocarcinoma, Biomarker discovery

Introduction

Artificial intelligence (AI) and machine learning (ML) have been introduced into various clinical fields over the past decades to enhance data evaluation and interpretation, as well as to manage increasing volumes of data [1]. AI is a field of computer science focused on developing systems that can mimic cognitive functions associated with human intelligence, such as the ability to identify hidden patterns in data, extract them and provide decisions based on prediction modeling [2]. ML is a subset of AI that allows a computational system to learn, self-correct and improve accuracy based on training and experience. ML uses mathematical and statistical models to obtain decisions autonomously, without direct instructions. High-quality data is essential for effective training

of ML models. The more accurate and relevant the training data are, the better the model performs [3].

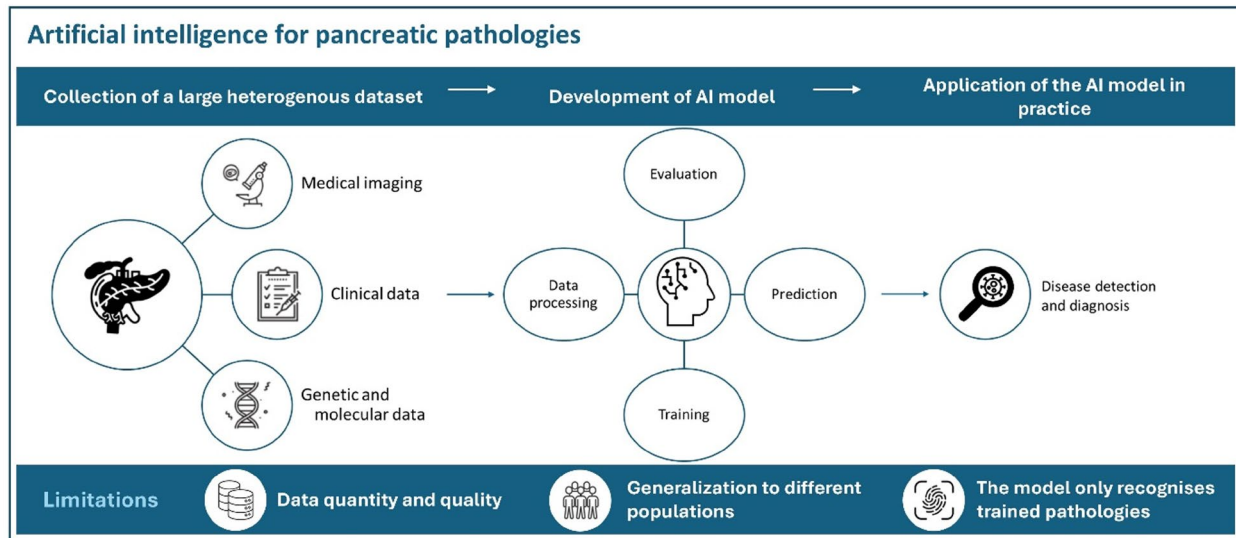
Generally, ML models offer robust learning ability and unique advantages in processing complex patterns in diverse types of input data. Therefore, they are ideally suited for image analysis tasks, especially in the field of digital pathology [4]. They can also combine a large amount of inherently highly diverse data, such as demographic and clinical parameters, biochemical values, radiology and histopathology image information and molecular-pathological data from sequence analyses. The AI and ML can support diagnosis and enhance the selection of the most appropriate treatment according to the patient's needs, bringing the concept of personalized medicine closer to the clinical practice [1]. In addition, the AI and ML prediction capabilities can improve the modeling of disease-free survival (DFS) and overall survival (OS) by integrating various input data and identifying relevant parameters related to patient outcomes [5, 6].

*Correspondence:
Petr Vaňhara
pvanhara@med.muni.cz

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Graphical abstract**Pancreatic cancer**

Pancreatic cancer (PC) covers a broad range of malignancies that develop from acinar, ductal or neuroendocrine cells within the pancreatic parenchyma. For a detailed review of pancreatic cancer pathology, classification, and staging, see Haeberle and Esposito [7]. The most common malignant form, pancreatic ductal adenocarcinoma (PDAC), is a highly aggressive disease with a median survival of 11.7 months [8] and increasing frequency in population. The absence of early symptoms frequently leads to the diagnosis of PDAC at advanced stages, significantly limiting therapeutic response and patient survival. PDAC exhibits significant morphological diversity, molecular heterogeneity, and complexity, which directly impact prognosis and treatment efficacy. Morphological and molecular variations are observed not only among individual PDAC patients but also within a single tumor. Sántha et al. demonstrated that the morphological heterogeneity of pancreatic cancer correlates with its structural and functional diversity [9]. This substantial heterogeneity significantly contributes to treatment resistance and failure [10]. Currently, surgical resection is a potential curative option available for only 15–20% of patients [11].

Prognosis of PDAC patients can be estimated based on overall clinical condition, previous medical history, nutritional status and eligibility to neoadjuvant therapy and/or surgical resection. Other high-risk biological and clinical features include, but are not limited to, the size of primary pancreatic tumor lesions exceeding 2–3 cm, presence of invaded lymph nodes, elevated Carbohydrate Antigen 19–9 (CA 19–9) levels, significant unintentional

weight loss, and relevant comorbidities. For patients with borderline resectable or locally advanced adenocarcinomas, where immediate radical resection is not recommended, neoadjuvant chemotherapy is typically administered as a primary intervention, with subsequent consideration for resection. In cases of metastatic PDAC, first-line systemic chemotherapy is initiated. Approximately half of these patients will subsequently receive second-line chemotherapy [12–14]. Despite recent progress in understanding the pathobiology of pancreatic cancer and radiology imaging-based diagnostics [15] there is a need for new computational approaches and biomarkers that would credibly predict disease progression or reveal early cases in population screenings.

While AI and ML approaches have already been introduced in radiology image analysis, they offer the potential to integrate also other types of input data and enhance clinical decision-making or providing support for preventive programs. Park et al. evaluated the significance of early pancreatic cancer detection by developing a model that simulates incidence and mortality data in the U.S. The model included a hypothetical screening program that could reduce mortality. These findings suggest that screening for cancer 4–6 years before clinical diagnosis could lead to an average gain of 2.4 life-years [16].

The concepts and applications of AI in pancreatic cancer have attracted significant attention and have already been extensively reviewed [17–21]. In this work, we evaluated studies published from 2018 to April 2025. Inclusion criteria required the studies to report detailed information on ML model performance, involve at least 30 patients, and use PC/PDAC histological data as

input for model training. Here, we aim to discuss recent advances in AI- and ML-based analyses of variable data inputs in PC, with a particular emphasis on histopathological diagnostics.

Learning schemes in computational histopathology of PDAC

The development of ML models for use in histopathology involves three principal approaches: supervised learning, unsupervised learning, and reinforcement learning. Currently, ML-assisted histopathological decision-making relies on supervised learning, which uses high-quality images that have been annotated by experienced pathologists to identify regions of normal and tumor tissue and typical dysmorphologies. In PDAC, morphological parameters such as normal tissue regions, invasive ductal structures, and the extent of desmoplastic fibroinflammatory stroma are defined within the training dataset. Additionally, the presence of PDAC histopathological variants, such as foamy cells, enlarged ducts, or vacuolated cells, can also be incorporated. Currently, various ML algorithms are designed to address specific problems in medical diagnosis. The concepts and categories of ML are illustrated in a Venn diagram (Fig. 1). AI methods for the analysis of histopathological images can be broadly categorized into two groups: traditional ML methods and deep learning methods. The biggest advantage of using deep learning methods over traditional methods is that they can process large and unstructured data such as

images, text, and speech without the need to extract features manually. For a detailed description of the different ML models, their training and validation, see Greener et al. [22].

Convolutional neural networks (CNNs) as a subset of Deep neural networks (DNN) are the often-used ML model in pancreatic histopathology to evaluate image data. CNN can recognize patterns directly from raw pixel images, requiring minimal preprocessing. CNNs can identify patterns with robustness to variability, distortion, and simple geometric transformations, such as rotations, translations, and scaling. This ability makes them highly effective for biological image-related tasks, e.g. object detection, image classification and segmentation, where the input data may vary in position, size, or orientation [23–25] or for analysis of whole slide images (WSI) [26]. This makes CNNs particularly valuable for analyzing complex, high-resolution tissue images, enabling more accurate and efficient diagnoses in pathology and the discovery of predictive biomarkers.

Applications of ML in diagnostics and prediction PDAC

Several studies have been published recently describing the use of ML for the early detection of pancreatic cancer. For example, Cao et al. used deep learning methods to detect and classify pancreatic lesions through non-contrast computer tomography. The model achieved the area under curve (AUC) value of 0.986–0.996 in detecting

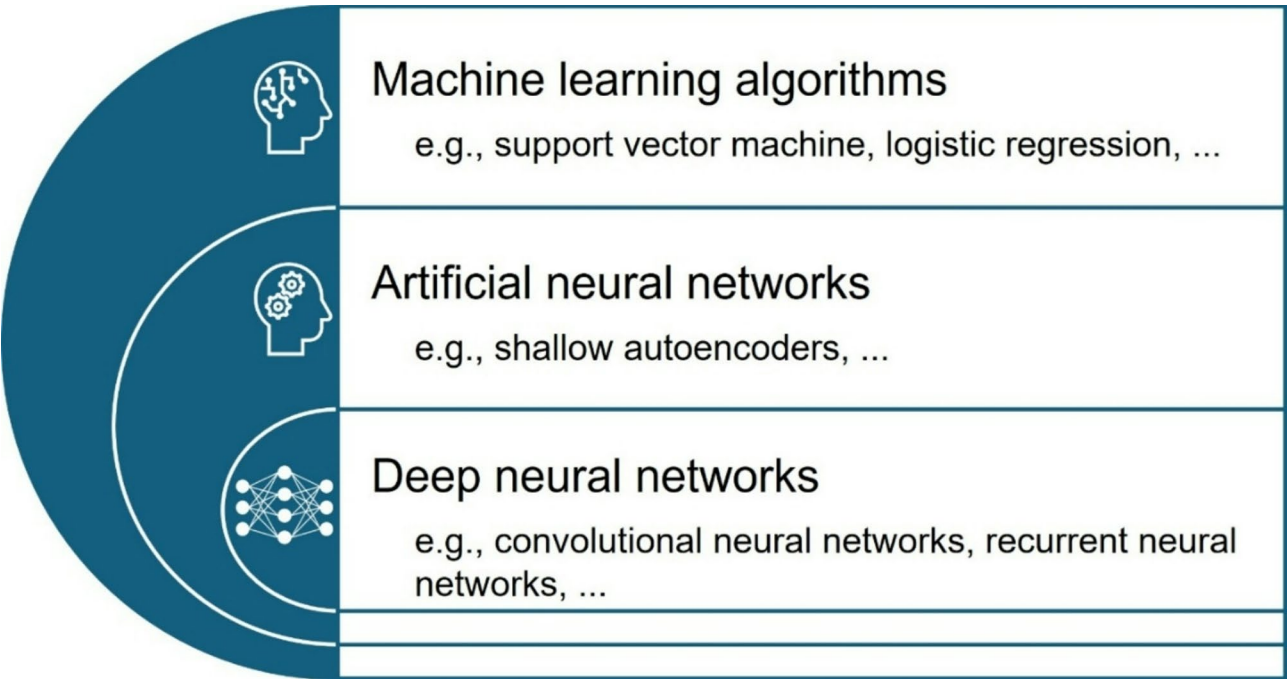


Fig. 1 Venn diagram demonstrating the principal hierarchy of AI models represented by machine learning algorithms, artificial neural networks and deep neural networks

lesions and was not evaluated inferior to the average results of human radiologists [27]. Alternative approach was used by Placido et al. in their publication, who used longitudinal clinical records to predict the risk of pancreatic cancer. The AUC of the model for the training data was 0.88 for predicting the occurrence of cancer within 36 months and 0.83 for 3 months prior to diagnosis. The results for the validation set were lower, with AUC values of 0.78 and 0.76, respectively [28].

To date, only a limited number of publications have investigated the detection of PDAC using ML coupled with standard histopathology techniques (Table 1). The ML models achieved high reliability above 90%. In all studies, the metrics that are appropriate for unbalanced cohorts, F1 score and balanced accuracy, have been used to evaluate reliability. However, direct comparison of the models is limited, as the studies included different numbers of groups that the classification models were trained for, variability in how the groups were defined or a lack of clear group definitions, differences in the number of data sources and/or scanners used for training, and variations in the diagnostic evaluation of images by pathologists.

In the pilot studies, models are typically trained using distinct groups, such as cancer and normal cohorts, due to the clear separation between these groups. Similarly, in publications by Fu et al. and Naito et al. two groups, PDAC patients, and healthy donors, were analyzed. However, in both publications, the characteristics of the control group were not provided in detail, and the extent of physiological heterogeneity was not considered as well. Due to the class imbalance in the study by Fu et al., F1 scores provide a more appropriate metric for evaluating

model performance. Although Naito et al. utilized a dataset more than twice as large, the reported model reliability is comparable. This similarity may be explained by greater heterogeneity within the control group in the Naito et al. study, which is described as ‘non-adenocarcinoma’, in contrast to the Fu et al. study, where the control group consisted of normal pancreatic tissue [29, 30]. The greater heterogeneity within the dataset may explain why a substantially larger number of images was required to achieve an F1 score exceeding 0.95. This also emphasized the need for external dataset validation, which neither of these studies included.

Kriegsmann et al. successfully applied ML for classification of patients into three categories: PDAC, benign pancreatic conditions, and pancreatic intraepithelial neoplasia. Patients in the study were recruited from two different centers, each using distinct types of scanners. Importantly, patients from both centers were equally distributed between the training and test datasets [31]. Despite analyzing a smaller patients’ cohort, classifying into three groups, utilizing data from varied scanners across two centers, and including PDAC patients at various disease stages, this model achieved results comparable to two earlier publications.

Janseen et al. [32] used ML algorithms to evaluate presence or absence of residual disease in PDAC patients after neoadjuvant therapy (Table 1). This multicentric study used data from 14 centers using four different scanners. The aim of the model was to distinguish residual pancreatic cancer from nonneoplastic pancreatic ducts in routine Hematoxylin and Eosin-stained WSI. The model achieved the overall F1 score of 0.78, and the best

Table 1 Application of ML in computational histopathology of pancreatic cancer

Aim of study	Data source	Algorithm	Sample size	Result	Reference
PDAC diagnostics	WSI	CNN	231 cases	Accuracy 100% WSI level Accuracy 95.3% patch-level F1 score 0.9525 patch level Normal class F1 score 0.9542 patch level Cancerous class	Fu et al. [29]
PDAC diagnostics	WSI	CNN	532 images	Accuracy 94.17% AUC 0.9836 F1 score 0.9581 Sensitivity 0.9302 Specificity 0.9706	Naito et al. [30]
PDAC diagnostics	WSI	CNN	201 cases	BACC aggregated 92.12%	Kriegsmann et al. [31]
RPC detection	WSI	CNN	528 cases	F1 score 0.78	Janseen et al. [32]
PC grading	WSI	CNN	138 images	F1 score 0.96 H&E stained images F1 score 0.88 MGG stained images F1 score 0.89 H&E and MGG stained images	Sehmi et al. [33]
PAAD grading	WSI	BCNN	138 images	F1 score 0.86 H&E stained images F1 score 0.78 MGG stained images F1 score 0.79 H&E and MGG stained images	Ghoshal et al. [34]
Molecular subtype of PDAC	WSI	CNN	207 images 139 cases	BAAC 96.19% classical subtype BAAC 83.03% basal subtype	Ahmadvand et al. [35]

AUC, Area under the curve, BACC, Balanced accuracy, BCNN, Bayesian convolutional neural network, CNN, Convolutional Neural Network, PAAD, Pancreatic adenocarcinoma, PC, Pancreatic cancer, PDAC, Pancreatic ductal adenocarcinoma, RPC, Residual pancreatic carcinoma, WSI, Whole slide image

reliability calculated for each type of scanner separately was 0.81 [32]. Reliability is lower than in the above publications, but in this case the dataset was composed from multiple sources and scan types, color augmentation and normalization was needed, and the goal of classification was also different.

In addition to PDAC diagnosis, ML models were also used to identify grades (Table 1). In the publication, Sehmi et al. classified images into four groups - Normal, Grade I, Grade II, and Grade III. Histological images were stained with May-Grünwald-Giemsa and Haematoxylin and Eosin. For learning, 14 pre-trained CNNs were used and then the model with the best reliability was selected, having an F1 score of 0.96 [33]. Ghoshal et al. used the same dataset as Sehmi et al. but aimed to determine the uncertainty in the prediction of PC stage by the

model using Bayesian CNNs [34]. The different objective in developing the model is probably the reason why the results of reliability for Ghoshal et al.'s model differ from those obtained by Sehmi et al.

AI was also used to determine the molecular subtype of PDAC, specifically the classical or basal subtype. The model achieved a balanced accuracy of 96.19% for classical subtype of PDAC and 83.03% for basal molecular subtype [35]. The major advantage of this study is that reliability has been validated using an external dataset. The reliability for classifying the basal subtype is lower, but this is most likely due to the fact that this group was less represented in the data (19.4% cases and 25.1% images).

These examples demonstrate how successful and reliable ML models can be. When the network is properly trained on a sufficiently large heterogeneous dataset, its reliability can be very high for both the validation and test sets. Such algorithms can play a significant role in clinical practice, assisting with the efficient evaluation of large numbers of samples. The significant heterogeneity among tumors can considerably impact model performance. This is evidenced by cases where even experienced pathologists cannot reach a consensus on the precise diagnosis [36].

Table 2 AI in the biomarker discovery for PC diagnosis

Algorithm	Sample size	Data source	Result	Reference
RF	489 samples	RNA seq	AUC 0.945	Mahawan et al. [40]
SVM	107 cases	MS	Accuracy 97.4% AUC 0.997 Sensitivity 100% Specificity 95.0%	Iwano et al. [41]
1D CNN + LSTM	590 cases	Proteomics	Accuracy 97% AUC 0.98	Karar et al. [42]
GLM	208 cases	Proteomics	AUC 0.95 Sensitivity 84% Specificity 95%	Athanasίου et al. [43]
GBM	1023 cases	ELISA	AUC 0.919 PDAC vs. non-PDAC AUC 0.925 PDAC vs. healthy controls	Firpo et al. [44]
SVM	1504 cases	mRNA	AUC 0.985 resectable PC AUC 0.967 all PDAC stages	Lee et al. [45]
SVM	501 cases	RNAseq	AUC of 0.936 Sensitivity of 93.68% Specificity of 91.57%	Yu et al. [46]
RF	100 cases	SMS	AUC 0.977 Sensitivity 0.94 Specificity 0.94	Chen et al. [47]
OPLS-DA	830 samples	Lipidomic MS	Accuracy 94.18% AUC 0.983 Sensitivity 95.97% Specificity 90.46%	Worlab et al. [48]

1D CNN, one-dimensional convolutional neural network, AUC, Area under the curve, DL, Deep learning, ELISA, Enzyme-linked immunosorbent assay, GBM, Generalized boosted regression model, GLM, Generalized linear model, LSTM, Long short-term memory, MS, Mass spectrometry, OPLS-DA, Orthogonal partial least squares discriminant analysis, RF, Random forest, SMS, Shotgun metagenomic sequencing, SVM, Support vector machine

ML-assisted biomarker discovery in pancreatic cancer

Timely diagnosis of PDAC allows earlier initiation of therapy, more accurate monitoring, and ultimately leads to improved patient survival. CA 19–9 is the most widely used biomarker associated with PC and PDAC [37, 38]; however, it lacks sufficient sensitivity and specificity for effective PDAC screening [39]. Therefore, there is a critical need to develop reliable, preferably non-invasive biomarkers for early detection of PDAC. Several studies reported biomarker development using different types of ML models, input data or sample source, reaching AUC values higher than 0.9 (Table 2).

Mahawan et al. proposed a *in silico* ML pipeline for constituting composite biomarker signatures from public gene expression databases, and demonstrated its efficacy on 15 selected genes for prediction of risk of PDAC metastasis [40]. Iwano et al. introduced collective biomarkers, based on 36 peripheral blood serum metabolites, identified by support vector machine-driven analysis of liquid chromatography-mass spectrometry (MS) spectra [41]. Karar et al. proposed a combination model composed of one-dimensional CNN and long short-term memory that identifies patterns in urinary biomarkers to automatically diagnose PDAC [42] and Worlab et al. used lipidomic profiling to discriminate between PDAC and healthy controls [48]. Multicentric proteomic analysis by Athanasiou et al. revealed serum

Table 3 ML models combining multiple types of data for PAAD

Algorithm	Sample size	Data source	Result	Reference
MMF– both datasets	166 cases	WSI	C index=0.653 MMF	Chen et al. [49]
AMIL– WSI		Molecular profile	C index=0.580 AMIL	
SNN– molecu- lar profile			C index=0.593 SNN	
			Survival AUC 0.701 MMF	
			Survival AUC 0.594 AMIL	
			Survival AUC 0.643 SNN	

AMIL, Attention-based multiple-instance learning, AUC, Area under the curve, MMF, Deeplearning-based multimodal fusion, SNN, Self-normalizing network, WSI, Whole slide image

protein signatures correlating with early stage PDAC and its subtypes [43]. Chen et al. created a model that integrates microbial signatures and CA 19–9 levels [47]. Firpo et al. [44] used several separate models for training in the first step and subsequently combined them into a single model using generalized boosted regression.

All ML models demonstrated promising efficacy; however, their direct evaluation is limited. Besides the type of input data and ML-model, the studies also differed in patient inclusion criteria. Some publications included patients in all stages of PDAC [40, 46–48] whereas others did not specify the stage [42, 45] or focused only on selected stages, namely T2 and T3 [41] or stages I and II [43, 44]. The definition of the control group also varied in studies, ranging from “healthy controls” [40–42, 45–48] which were not fully described in all cases, or including patients with chronic pancreatitis, family history of PDAC or genetic predispositions [43] and patients with chronic pancreatitis and intraductal papillary mucinous neoplasm [44]. Karar et al. stratified the dataset into three classes for ML model development– PDAC, healthy controls, and benign hepatobiliary disease [42]. A valid experimental design is therefore a critical prerequisite for reliable, reproducible and comparative outputs.

In summary, ML represents a powerful approach for identifying novel biomarkers to facilitate the detection of pancreatic cancer. As illustrated by the examples above, ML models can attain high accuracy despite considerable differences in model design or methodology.

Integration of input data from multiple modalities

A promising approach to achieve even more accurate predictions or prognoses in PDAC uses a combination of different data sources as inputs to an AI model. In cancer, histological data can be integrated with immunohistochemical, metabolic, genetic, and radiographic data. This integration can increase the effectiveness of clinical workflows, improve diagnostic accuracy, and ultimately lead to personalized diagnosis and treatment plans for patients. Chen et al. reported using used weakly supervised, multimodal deep-learning algorithm for successful

prediction of prognosis of 14 cancer types including PDAC based on WSI (Table 3).

This study highlighted the importance of using multiple data sources, specifically a combination of histology images and molecular profile data, to predict outcomes and reveal prognostic features that correlate with poor and favorable outcomes in 14 cancer types. Data from 5720 patients was used to construct the model for all evaluated cancer types. Models for each single type of cancer have also been established. To create the model for PDAC, 166 patients were used. The C index of the whole dataset was 0.645, compared with 0.585 for the model using histological data alone and 0.607 for the model using molecular features alone. Similarly, the survival AUC improved from 0.616 for the model using only molecular features and 0.598 for histological images to 0.662 when both were combined. These results are similar to the model evaluations for PDAC, which are presented in Table 4 [49]. For both models, a combination of both datasets gave the best results.

In other types of cancer, Hou et al. reported a similar conclusion, using a multimodal model to predict the survival of hepatocellular carcinoma patients. Dataset used in the study was constructed using WSI and mRNA expression levels from publicly available databases. The prediction of OS based on both datasets achieved a C index of 0.746 and surpassed that of models that were based only on either histopathological data (0.714) or genetic information (0.666) [50]. Mobadersany et al., used a combination of genetic biomarkers and histology images from 796 samples as inputs to a CNN to predict the survival of patients with gliomas. The best model performance was achieved by combining data from different sources leading to improved prognostic prediction probability [51]. In contrast, Höhn et al., developed a skin cancer classifier to distinguish between melanomas and benign nevi, and reported that its performance was comparable or even better when trained on a single dataset versus a combination of two datasets. Here, 430 WSI and basic patient data (age, sex, anatomical site of lesion) were used for construction of CNN. The AUC of the patient data model was 0.770, and the balanced accuracy (BA) was 0.720, while the WSI model achieved an AUC of 0.923 and a BA of 0.832. Using both datasets, the best approach achieved an AUC of 0.896 and a BA of 0.830 [52]. Although the differences in performance between using WSIs and both datasets are minimal, the study underscores that incorporating multiple datasets does not necessarily enhance classifier accuracy.

Therefore, including low- or non-informative variables from different data sources does not necessarily improve the accuracy of a classifier. In that case, a decrease in model reliability usually results from overfitting, particularly in models built with numerous uncorrelated

variables. Initially, adding features can enhance accuracy, but beyond a certain point, it may lead to a decline. In fact, including too many features can cause the model to fit noise rather than meaningful patterns. This is especially relevant in supervised learning, where the number of samples should be proportional to the number of parameters [53, 54].

Challenges and limitations

A major challenge in training ML algorithms is the requirement for large volumes of precisely annotated (labeled) data [55]. Moreover, annotations are difficult to obtain, as they are time-consuming and must be performed and validated by expert pathologists. A possible solution to this problem is the use pretrained networks, e.g., ImageNet by Stanford Vision and Learning Lab [56] followed by additional training on a smaller set of annotated images [57]. This approach streamlines and accelerates the training process, although it does not always result in better performance than a network that was trained from the beginning [58]. Additionally, the features representations learned by pretrained ML models may not be well-suited for all histopathological images. To overcome this limitation, disease-specific digital pathology datasets have been proposed and made publicly available [59–62].

Most ML prediction models typically distinguish PDAC patients from either healthy individuals or those with other types of lesions. However, it's crucial to consider various pancreatic changes that differ from PDAC [20, 63]. Therefore, a computational tool capable of discriminating not only PDAC but also other distinct pancreatic diseases is highly needed. Unbalanced cohorts within the datasets then pose another major challenge, potentially leading to biased accuracy estimates if an appropriate evaluation metrics are not properly applied. Furthermore, the histopathological heterogeneity of PDAC likely contributes to the variability in treatment responses [64]. Consequently, improving the performance of ML models will require a large, diverse dataset representing different PDAC subtypes and forms. Future development ML models should prioritize patient staging, rigorously correlating it with histopathology and OS. Additionally, investigating the response to neoadjuvant/adjuvant chemo(radio)therapy and exploring multiple treatment options is highly relevant [18].

Generalization remains a challenge for ML models, which often perform well on training data but poorly on data from different sources. This issue represents one of the greatest challenges to the widespread implementation of AI in medicine. Carrillo-Perez et al. used two datasets in their publication, where one contained histological images of PDAC patients and the other a control set, and showed that the model trained on the dataset created this

way achieved accuracy of 99.04%, and when the model was evaluated on a different dataset, achieved an accuracy 67.20% [65]. To minimize this problem, it is crucial to create a training set that includes as much variability as possible, incorporating data from different centers, instruments, and varying staining protocols. In addition, Carrillo-Perez showed that the model trained on dataset integrating multiple sources does not preferentially discriminate between the classes, but rather between the datasets, indicating that stronger normalization would be needed when using datasets from various sources [65]. A similar problem was successfully addressed by Pečinka et al. for MS data recorded from whole cancer cells [66]. Since the problem of model generalization can often result from overfitting, it is necessary for publications to clearly state how this issue has been addressed. Authors should explicitly describe the methods and strategies implemented to prevent overfitting, such as cross-validation techniques, regularization methods, or the use of external validation datasets. Additionally, it is important to outline the specific measures implemented to minimize overfitting, and therefore ensuring the robustness and reliability of the reported performance of the ML model. Another technical issue significantly affecting AI analysis, but usually not the interpretation by a human pathologist is the color variability of histological samples. Some studies have shown that deep learning algorithms' ability to segment or classify histological images decreases when there is high color variability or inconsistency within the training set [67, 68].

Predictive algorithms are inherently limited to recognizing only the patterns they have been trained on. While ideally an algorithm would be trained to identify all potential pathological conditions, this approach is often extremely difficult or even impossible. A practical alternative is to develop AI methods that, in addition to providing predictions, also generate a certainty score for each prediction [69]. When WSI are used to train algorithms, their immense size limits computational efficiency, as these gigapixel images are too large to fit into a Graphics Processing Unit's (GPU) memory. Several approaches address this issue; a common method assumes that all patches—small tiles into which the image is divided—contain morphological information correlated with the label of the entire image [70].

The integration of ML models into clinical practice for PDAC remains in the early stages. Baxi et al. outline a phased framework for the clinical implementation of AI technologies. The initial phase involves validating the model's reliability, a process overseen by a pathologist to ensure diagnostic accuracy. Contingent upon successful validation, ML applications may subsequently be introduced as adjunctive diagnostic tools. This would be followed by prospective clinical trials designed to evaluate

model performance in real-world settings. In the final phase, the model would undergo continuous refinement and optimization through the incorporation of expanding datasets [71].

Regulatory and ethical consideration

The application of AI in medical practice raises important ethical considerations, particularly in relation to data confidentiality, patient privacy, data security, and informed consent. Since a large amount of sensitive information needs to be collected from patients and physicians to train and validate the model, it is necessary to consider the ethical aspects of data access from the perspective of public interest but also from the perspective of preserving the right to privacy of an individual [21, 72]. At the same time, data security should be considered to minimize the risk of data loss or hacking, as this could cause harm to patients [73, 74]. Another problem in the implementation of AI in medical practice is the absence of legal standards that address the issue of responsibility in case of misdiagnosis by ML-based tools, as they can have fatal consequences [19, 75]. Furthermore, bias in the model development process may also be an important obstacle. If the training set is not sufficiently diverse, biases in terms of demographic and socio-economic factors may occur. Over or under representation of some groups can cause poor model performance, incorrect predictions and even harm a patient [73, 76].

Conclusion

In recent years, various AI and ML based strategies have been developed to support diagnostic evaluations. With the growing availability of data, these models have begun to match—and in some instances exceed—the diagnostic performance of human experts. The integration of multimodal data sources has further enhanced the predictive accuracy and classification capabilities of ML models, offering improved clinical decision support. These advancements have the potential to reduce diagnostic turnaround times and assist in complex cases where definitive diagnoses are difficult to establish. While substantial barriers to routine clinical implementation remain, AI and ML represent a promising innovation with the potential to transform PDAC histopathology, streamline identification of novel biomarkers and improve prognosis of cancer patients.

Author contributions

PW and PV conceptualized the manuscript, PW, ME, PM, JV, ZK, VP and PV wrote the main manuscript text, PW, JH and PV, prepared the figures, all authors reviewed the manuscript.

Funding

This work was supported by the Ministry of Health of the Czech Republic, grant no. NU23-08-00241, and funds from the Faculty of Medicine, Masaryk University, grant. no. MUNI/A/1738/2024 and MUNI/A/1685/2024, and by

the CREATIC project funded by the European Union (grant agreement no. 101059788).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethical approval

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Histology and Embryology, Faculty of Medicine, Masaryk University, Kamenice 3, Brno 625 00, Czech Republic

²Department of Biostatistics, St. Anne's University Hospital Brno, Brno, Czech Republic

³Department of Internal Medicine, Hematology and Oncology, University Hospital Brno, Brno, Czech Republic

⁴Center for Precision Medicine, University Hospital Brno, Brno, Czech Republic

⁵Surgery Clinic, University Hospital Brno, Brno, Czech Republic

⁶Department of Pathology, University Hospital Brno, Brno, Czech Republic

⁷International Clinical Research Center, St. Anne's University Hospital Brno, Brno, Czech Republic

Received: 14 March 2025 / Accepted: 2 June 2025

Published online: 17 June 2025

References

1. van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nat Med*. 2021;27(5):775–84. <https://doi.org/10.1038/s41591-021-01343-4>.
2. Förtsch S, Klauschen F, Hufnagel P, Roth W. Artificial intelligence in pathology. *Dtsch Arzteblatt Int*. 2021;118(12):194–204. <https://doi.org/10.3238/arztebl.m.2021.0011>.
3. Burzykowski T, Rousseau A-J, Geubbelmans M, Valkenburg D. Introduction to machine learning. *Am J Orthod Dentofac Orthop*. 2023;163(5):732–4. <https://doi.org/10.1016/j.jajodo.2023.02.005>.
4. Salvi M, Acharya UR, Molinari F, Meiburger KM. The impact of Pre- and Post-Image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Comput Biol Med*. 2021;128:104129. <https://doi.org/10.1016/j.combiomed.2020.104129>.
5. Kosaraju S, Park J, Lee H, Yang JW, Kang M. Deep Learning-Based framework for Slide-Based histopathological image analysis. *Sci Rep*. 2022;12(1):19075. <https://doi.org/10.1038/s41598-022-23166-0>.
6. Mun SK, Wong KH, Lo S-CB, Li Y, Bayarsaikhan S. Artificial Intelligence for the Future Radiology Diagnostic Service. *Front. Mol. Biosci*. 2021, 7.
7. Haeberle L, Esposito I. Pathology of pancreatic Cancer. *Transl Gastroenterol Hepatol*. 2019;4:50. <https://doi.org/10.21037/tgh.2019.06.02>.
8. Groot VP, Rezaee N, Wu W, Cameron JL, Fishman EK, Hruban RH, Weiss MJ, Zheng L, Wolfgang CL, He J, Patterns. Timing, and predictors of recurrence following pancreatotomy for pancreatic ductal adenocarcinoma. *Ann Surg*. 2018;267(5):936–45. <https://doi.org/10.1097/SLA.0000000000002234>.
9. Sántha P, Lenggenhager D, Finstadsveen A, Dorg L, Tøndel K, Amrutkar M, Gladhaug IP, Verbeke C. Morphological heterogeneity in pancreatic Cancer reflects structural and functional divergence. *Cancers*. 2021;13(4):895. <https://doi.org/10.3390/cancers13040895>.
10. Verbeke C. Morphological heterogeneity in ductal adenocarcinoma of the Pancreas—Does It Matter? *Pancreatol*. 2016;16(3):295–301. <https://doi.org/10.1016/j.pan.2016.02.004>.
11. Varadhachary GR, Tamm EP, Abbruzzese JL, Xiong HQ, Crane CH, Wang H, Lee JE, Pisters PWT, Evans DB, Wolff RA. Borderline resectable pancreatic cancer: definitions, management, and role of preoperative therapy. *Ann Surg Oncol*. 2006;13(8):1035–46. <https://doi.org/10.1245/ASO.2006.08.011>.

12. Garajová I, Peroni M, Gelsomino F, Leonardi FA. Simple overview of pancreatic Cancer treatment for clinical oncologists. *Curr Oncol*. 2023;30(11):9587–601. <https://doi.org/10.3390/curroncol30110694>.
13. Halbrook CJ, Lyssiotis CA, Magliano MP di, Maitra A. Pancreatic Cancer: Advances and Challenges. *Cell* 2023, 186 (8), 1729–1754. <https://doi.org/10.1016/j.cell.2023.02.014>
14. Krishna V, Tiu E, Krishna V, Vrabac D, Shah K, Abuzeid W, Smith K, Davelaar J, Nuesca C, Larson BK, Fountzilias C, Rajpurkar P, Hendifar AE, Collisson EA, Joshi A, Singhi AD. Development of artificial Intelligence–Derived histological biomarkers for First-Line treatment selection in metastatic pancreatic ductal adenocarcinoma (mPDAC). *J Clin Oncol*. 2023;41(4suppl):743–743. https://doi.org/10.1200/JCO.2023.41.4_suppl.743.
15. Elbanna KY, Jang H-J, Kim TK. Imaging diagnosis and staging of pancreatic ductal adenocarcinoma: A comprehensive review. *Insights Imaging*. 2020;11:58. <https://doi.org/10.1186/s13244-020-00861-y>.
16. Park J, Lim F, Prest M, Ferris JS, Aziz Z, Agyekum A, Wagner S, Gulati R, Hur C. Quantifying the potential benefits of early detection for pancreatic Cancer through a counterfactual simulation modeling analysis. *Sci Rep*. 2023;13(1):20028. <https://doi.org/10.1038/s41598-023-46751-3>.
17. Huang B, Huang H, Zhang S, Zhang D, Shi Q, Liu J, Guo J. Artificial intelligence in pancreatic Cancer. *Theranostics*. 2022;12(16):6931–54. <https://doi.org/10.7150/thno.77949>.
18. Liu W, Zhang B, Liu T, Jiang J, Liu Y. Artificial intelligence in pancreatic image analysis: A review. *Sensors*. 2024;24(14):4749. <https://doi.org/10.3390/s24144749>.
19. Hameed BS, Krishnan UM. Artificial Intelligence-Driven diagnosis of pancreatic Cancer. *Cancers*. 2022;14(21):5382. <https://doi.org/10.3390/cancers14215382>.
20. Schuurmans M, Alves N, Vendittelli P, Huisman H, Hermans J. Setting the research agenda for clinical artificial intelligence in pancreatic adenocarcinoma imaging. *Cancers*. 2022;14(14):3498. <https://doi.org/10.3390/cancers14143498>.
21. Daher H, Punchayil SA, Ismail AAE, Fernandes RR, Jacob J, Algazzar MH, Mansour M. Advancements in pancreatic Cancer detection: integrating biomarkers, imaging technologies, and machine learning for early diagnosis. *Cureus* 16 (3), e56583. <https://doi.org/10.7759/cureus.56583>
22. Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. *Nat Rev Mol Cell Biol*. 2022;23(1):40–55. <https://doi.org/10.1038/s41580-021-00407-0>.
23. LeCun Y, Bengio Y, Hinton G. Deep Learning. *Nature*. 2015;521(7553):436–44. <https://doi.org/10.1038/nature14539>.
24. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, Cui C, Corrado G, Thrun S, Dean J. A guide to deep learning in healthcare. *Nat Med*. 2019;25(1):24–9. <https://doi.org/10.1038/s41591-018-0316-z>.
25. Liu J, Fu F. Convolutional neural network model by deep learning and teaching robot in keyboard musical instrument teaching. *PLoS ONE*. 2023;18(10):e0293411. <https://doi.org/10.1371/journal.pone.0293411>.
26. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J Pathol Inf*. 2016;7(1):29. <https://doi.org/10.4103/2153-3539.186902>.
27. Cao K, Xia Y, Yao J, Han X, Lambert L, Zhang T, Tang W, Jin G, Jiang H, Fang X, Nogueira I, Li X, Guo W, Wang Y, Fang W, Qiu M, Hou Y, Kovarnik T, Vocka M, Lu Y, Chen Y, Chen X, Liu Z, Zhou J, Xie C, Zhang R, Lu H, Hager GD, Yuille AL, Lu L, Shao C, Shi Y, Zhang Q, Liang T, Zhang L, Lu J. Large-Scale pancreatic Cancer detection via Non-Contrast CT and deep learning. *Nat Med*. 2023;29(12):3033–43. <https://doi.org/10.1038/s41591-023-02640-w>.
28. Placido D, Yuan B, Hjaltekin JX, Zheng C, Haue AD, Chmura PJ, Yuan C, Kim J, Umerton R, Antell G, Chowdhury A, Franz A, Brais L, Andrews E, Marks DS, Regev A, Ayandeh S, Brophy MT, Do NV, Kraft P, Wolpin BM, Rosenthal MH, Fillmore NR, Brunak S, Sander C. A deep learning algorithm to predict risk of pancreatic Cancer from disease trajectories. *Nat Med*. 2023;29(5):1113–22. <https://doi.org/10.1038/s41591-023-02332-5>.
29. Fu H, Mi W, Pan B, Guo Y, Li J, Xu R, Zheng J, Zou C, Zhang T, Liang Z, Zou J, Zou H. Automatic pancreatic ductal adenocarcinoma detection in whole slide images using deep convolutional neural networks. *Front Oncol*. 2021;11:665929. <https://doi.org/10.3389/fonc.2021.665929>.
30. Naito Y, Tsuneki M, Fukushima N, Koga Y, Higashi M, Notohara K, Aishima S, Ohike N, Tajiri T, Yamaguchi H, Fukumura Y, Kojima M, Hirabayashi K, Hamada Y, Norose T, Kai K, Omori Y, Sukeda A, Noguchi H, Uchino K, Itakura J, Okabe Y, Yamada Y, Akiba J, Kanavati F, Oda Y, Furukawa T, Yano H. A deep learning model to detect pancreatic ductal adenocarcinoma on endoscopic Ultrasound-Guided Fine-Needle biopsy. *Sci Rep*. 2021;11(1):8454. <https://doi.org/10.1038/s41598-021-87748-0>.
31. Kriegsmann M, Kriegsmann K, Steinbuss G, Zgorzelski C, Kraft A, Gaida MM. Deep learning in pancreatic tissue: identification of anatomical structures, pancreatic intraepithelial neoplasia, and ductal adenocarcinoma. *Int J Mol Sci*. 2021;22(10):5385. <https://doi.org/10.3390/ijms22105385>.
32. Janssen BV, Oteman B, Ali M, Valkema PA, Adsay V, Basturk O, Chatterjee D, Chou A, Crobach S, Doukas M, Drillenburger P, Esposito I, Gill AJ, Hong S-M, Jansen C, Kliffen M, Mittal A, Samra J, van Velthuysen M-LF, Yavas A, Kazemier G, Verheij J, Steyerberg E, Besselink MG, Wang H, Verbeke C, Fariña A, de Boer OJ, Pathologists. (ISGPP), for the I. S. G. of P; consortium, the P. and H. A. I. R. (PHAIR); pathologists (ISGPP), for the I. S. G. of P; consortium, the P. and H. A. I. R. (PHAIR). Artificial Intelligence-Based segmentation of residual pancreatic Cancer in resection specimens following neoadjuvant treatment (ISGPP-2): international improvement and validation study. *Am J Surg Pathol*. 2024;48(9):1108. <https://doi.org/10.1097/PAS.0000000000002270>.
33. Sehmi MNM, Fauzi MFA, Ahmad WSHMW, Chan EWL. Pancreatic Cancer grading in pathological images using deep learning convolutional neural networks. *F1000Research* October 18, 2021. <https://doi.org/10.12688/f1000research.73161.1>
34. Ghoshal B, Ghoshal B, Tucker A. Leveraging uncertainty in deep learning for pancreatic adenocarcinoma grading. In: Yang G, Aviles-Rivero A, Roberts M, Schönlieb C-B, editors. *Medical image Understanding and analysis*. Springer International Publishing: Cham; 2022. pp. 565–77. https://doi.org/10.1007/978-3-031-12053-4_42.
35. Ahmadvand P, Farahani H, Farnell D, Darbandsari A, Topham J, Karasinska J, Nelson J, Naso J, Jones SJM, Renouf D, Schaeffer DF, Bashashati A. A deep learning approach for the identification of the molecular subtypes of pancreatic ductal adenocarcinoma based on whole slide pathology images. *Am J Pathol*. 2024;194(12):2302–12. <https://doi.org/10.1016/j.ajpath.2024.08.006>.
36. Wilmink JW, Besselink MG, Brosens LAA, Wang H, Verbeke CS, Verheij J. Amsterdam international consensus meeting: tumor response scoring in the pathology assessment of resected pancreatic Cancer after neoadjuvant therapy. *Mod Pathol*. 2021;34(1):4–12. <https://doi.org/10.1038/s41379-020-00683-9>.
37. Velthuysen M-L, Basturk O, Campbell F, Doglioni C, Esposito I, Feakins R, Fukushima N, Gill A, J, Hruban R, H, Kaplan J, Koerkamp B, G, Hong S-M, Krasinskas A, Luchini C, Offerhaus J, Sarasqueta A, F, Shi C, Singhi A, Stoop T, F, Soer E, C, Thompson E, Tienhoven G, van. Farnell D, Darbandsari A, Topham J, Karasinska J, Nelson J, Naso J, Jones SJM, Renouf D, Schaeffer DF, Bashashati A. A deep learning approach for the identification of the molecular subtypes of pancreatic ductal adenocarcinoma based on whole slide pathology images. *Am J Pathol*. 2024;194(12):2302–12. <https://doi.org/10.1016/j.ajpath.2024.08.006>.
38. Poruk KE, Gay DZ, Brown K, Mulvihill JD, Boucher KM, Scaife CL, Firpo MA, Mulvihill SJ. The clinical utility of CA 19–9 in pancreatic adenocarcinoma: diagnostic and prognostic updates. *Curr Mol Med*. 2013;13(3):340–51. <https://doi.org/10.2174/1566524011313030003>.
39. Majumder S, Taylor WR, Foote PH, Berger CK, Wu CW, Mahoney DW, Bamlet WR, Burger KN, Postier N, de la Fuente J, Doering KA, Lidgard GP, Allawi HT, Petersen GM, Chari ST, Ahlquist DA, Kisiel JB. High detection rates of pancreatic Cancer across stages by plasma assay of novel methylated DNA markers and CA19-9. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2021;27(9):2523–32. <https://doi.org/10.1158/1078-0432.CCR-20-0235>.
40. Mahawan T, Luckett T, Mielgo Iza A, Pornputtapong N, Caamaño Gutiérrez E. Robust and consistent biomarker candidates identification by a machine learning approach applied to pancreatic ductal adenocarcinoma metastasis. *BMC Med Inf Decis Mak*. 2024;24(Suppl 4):175. <https://doi.org/10.1186/s12911-024-02578-0>.
41. Iwano T, Yoshimura K, Watanabe G, Saito R, Kiritani S, Kawaida H, Moriguchi T, Murata T, Ogata K, Ichikawa D, Arita J, Hasegawa K, Takeda S. High-Performance collective biomarker from liquid biopsy for diagnosis of pancreatic Cancer based on mass spectrometry and machine learning. *J Cancer*. 2021;12(24):7477–87. <https://doi.org/10.7150/jca.63244>.
42. Karar ME, El-Fishawy N, Radad M. Automated classification of urine biomarkers to diagnose pancreatic Cancer using 1-D convolutional neural networks. *J Biol Eng*. 2023;17:28. <https://doi.org/10.1186/s13036-023-00340-0>.
43. Athanasiou A, Kureshi N, Wittig A, Sterner M, Huber R, Palma NA, King T, Schiess R. Biomarker discovery for early detection of pancreatic ductal adenocarcinoma (PDAC) using multiplex proteomics technology. *J Proteome Res*. 2024;24(1):315–22. <https://doi.org/10.1021/acs.jproteome.4c00752>.

44. Firpo MA, Boucher KM, Bleicher J, Khanderao GD, Rosati A, Poruk KE, Kamal S, Marzullo L, De Marco M, Falco A, Genovese A, Adler JM, De Laurenzi V, Adler DG, Affolter KE, Garrido-Laguna I, Scaife CL, Turco MC, Mulvihill SJ. Multianalyte serum biomarker panel for early detection of pancreatic adenocarcinoma. *JCO Clin Cancer Inf*. 2023;7:e2200160. <https://doi.org/10.1200/JCO.22.00160>.
45. Lee J, Kang SW, Sim E-J, Bae J-S, Koo S, Byoun M, Kwon S, Hong S, Kim Y, Youn Y, Jung K, Kim J, Jeong HH, Kim J, Hwang J-H. Novel mRNA Biomarker-Based liquid biopsy for the detection of resectable pancreatic Cancer. *BMC Cancer*. 2025;25:762. <https://doi.org/10.1186/s12885-025-14124-w>.
46. Yu S, Li Y, Liao Z, Wang Z, Wang Z, Li Y, Qian L, Zhao J, Zong H, Kang B, Zou W-B, Chen K, He X, Meng Z, Chen Z, Huang S, Wang P. Plasma extracellular vesicle long RNA profiling identifies a diagnostic signature for the detection of pancreatic ductal adenocarcinoma. 2020. <https://doi.org/10.1136/gutjnl-2019-318860>
47. Chen Y, Nian F, Chen J, Jiang Q, Yuan T, Feng H, Shen X, Dong L. Metagenomic microbial signatures for noninvasive detection of pancreatic Cancer. *Biomedicine*. 2025;13(4):1000. <https://doi.org/10.3390/biomedicine13041000>.
48. Wolrab D, Jirásko R, Cifková E, Höring M, Mei D, Chocholoušková M, Peterka O, Ildkowiak J, Hrnčiarová T, Kuchař L, Ahrends R, Brumarová R, Friedecký D, Vivo-Truyols G, Škrha P, Škrha J, Kučera R, Melichar B, Liebisch G, Burkhardt R, Wenk MR, Cazenave-Gassiot A, Karásek P, Novotný I, Greplová K, Hrstka R, Holčápek M. Lipidomic profiling of human serum enables detection of pancreatic Cancer. *Nat Commun*. 2022;13(1):124. <https://doi.org/10.1038/s41467-021-27765-9>.
49. Chen RJ, Lu MY, Williamson DFK, Chen TY, Lipkova J, Noor Z, Shaban M, Shady M, Williams M, Joo B, Mahmood F. Pan-Cancer integrative Histology-Genomic analysis via multimodal deep learning. *Cancer Cell*. 2022;40(8):865–e8786. <https://doi.org/10.1016/j.ccell.2022.07.004>.
50. Hou J, Jia X, Xie Y, Qin W. Integrative Histology-Genomic analysis predicts hepatocellular carcinoma prognosis using deep learning. *Genes*. 2022;13(10):1770. <https://doi.org/10.3390/genes13101770>.
51. Mobadersany P, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, Velázquez Vega JE, Brat DJ, Cooper L. A. D. Predicting Cancer Outcomes from Histology and Genomics Using Convolutional Networks. *Proc. Natl. Acad. Sci. U. S. A*. 2018; 115 (13), E2970–E2979. <https://doi.org/10.1073/pnas.1717139115>
52. Höhn J, Kriehoff-Henning E, Jutzi TB, Kalle C. Combining CNN-Based histologic whole slide image analysis and patient data to improve skin Cancer classification. *Eur J Cancer*. 2021;149:94–101. <https://doi.org/10.1016/j.ejca.2021.02.032>. Utikal J, S.Meier F, Gellrich F, F; Hobelsberger S, Hauschild, A; Schlager J, G; French, L; Heinzerling, L; Schlaak, M; Ghoreschi, K; Hilke, F. J; Poch, G; Kutzner, H; Heppt, M. V; Haferkamp, S; Sondermann, W; Schadendorf, D; Schilling, B; Goebeler, M; Hekler, A; Fröhling, S; Lipka, D. B; Kather, J. N; Krah, D; Ferrara, G; Hagggenmüller, S; Brinker, T. J.
53. Jabbar HK, Khan RZ. Methods to avoid Over-Fitting and Under-Fitting in supervised machine learning (Comparative Study). Computer science, communication and instrumentation devices. Research Publishing Services; 2014. pp. 163–72. https://doi.org/10.3850/978-981-09-5247-1_017.
54. Ying X. An overview of overfitting and its solutions. *J Phys Conf Ser*. 2019;1168:022022. <https://doi.org/10.1088/1742-6596/1168/2/022022>.
55. Komura D, Ishikawa S. Machine learning methods for histopathological image analysis. *Comput Struct Biotechnol J*. 2018;16:34–42. <https://doi.org/10.1016/j.csbj.2018.01.001>.
56. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Advances in neural information processing systems*. Volume 25. Curran Associates, Inc; 2012.
57. Srinidhi CL, Ciga O, Martel AL. Deep neural network models for computational histopathology: A survey. *Med Image Anal*. 2021;67:101813. <https://doi.org/10.1016/j.media.2020.101813>.
58. He K, Girshick R, Dollár PR. ImageNet Pre-Training. arXiv November 21, 2018. <https://doi.org/10.48550/arXiv.1811.08883>
59. Litjens G, Bandi P, Ehteshami Bejnordi B, Geessink O, Balkenhol M, Bult P, Halliövic A, Hermesen M, van de Loo R, Vogels R, Manson QF, Stathonikos N, Baidoshvili A, van Diest P, Wauters C, van Dijk M, van der Laak J. 1399 H&E-Stained Sentinel lymph node sections of breast Cancer patients: the CAMELYON dataset. *GigaScience*. 2018;7(6):giy065. <https://doi.org/10.1093/gigascience/giy065>.
60. Aresta G, Araújo T, Kwok S, Chennamsetty SS, Safwan M, Alex V, Marami B, Prastawa M, Chan M, Donovan M, Fernandez G, Zeineh J, Kohl M, Walz C, Ludwig F, Braunewell S, Baust M, Vu QD, To MNN, Kim E, Kwak JT, Galal S, Sanchez-Freire V, Brancati N, Frucci M, Riccio D, Wang Y, Sun L, Ma K, Fang J, Kone I, Boulmane L, Campilho A, Eloy C, Polónia A, Aguiar PBACH. Grand challenge on breast Cancer histology images. *Med Image Anal*. 2019;56:122–39. <https://doi.org/10.1016/j.media.2019.05.010>.
61. Bulten W, Kartasalo K, Chen P-HC, Ström P, Pinckaers H, Nagpal K, Cai Y, Steiner DF, van Boven H, Vink R, Hulsbergen-van de Kaa C, van der Laak J, Amin MB, Evans AJ, van der Kwast T, Allan R, Humphrey PA, Grönberg H, Samaratunga H, Delahunt B, Tsuzuki T, Häkkinen T, Egevad L, Demkin M, Dane S, Tan F, Valkonen M, Corrado GS, Peng L, Mermel CH, Ruusuvaori P, Litjens G, Eklund M. PANDA challenge consortium. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat Med*. 2022;28(1):154–63. <https://doi.org/10.1038/s41591-021-01620-2>.
62. Babaie M, Kalra S, Sriram A, Mitcheltree C, Zhu S, Khatami A, Rahnamayan S, Tizhoosh HR. Classification and retrieval of digital pathology scans. *A New Dataset*; 2017. pp. 8–16.
63. Schuurmans M, Alves N, Vendittelli P, Huisman H, Hermans J, Litjens G, Chang D, Verbeke C, Malats N, Löhr M. Artificial intelligence in pancreatic ductal adenocarcinoma imaging: A commentary on potential future applications. *Gastroenterology*. 2023;165(2):309–16. <https://doi.org/10.1053/j.gastro.2023.04.003>.
64. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, Cooc J, Weinkle J, Kim GE, Jakkula L, Feiler HS, Ko AH, Olshen AB, Danenberg KL, Tempero MA, Spellman PT, Hanahan D, Gray JW. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat Med*. 2011;17(4):500–3. <https://doi.org/10.1038/nm.2344>.
65. Carrillo-Perez F, Ortuno FM, Börjesson A, Rojas I, Herrera LJ. Performance comparison between Multi-Center histopathology datasets of a Weakly-Supervised deep learning model for pancreatic ductal adenocarcinoma detection. *Cancer Imaging*. 2023;23:66. <https://doi.org/10.1186/s40644-023-00586-3>.
66. Pečinka L, Moráň L, Kovačovicová P, Meloni F, Havel J, Pivetta T, Vaňhara P. Intact cell mass spectrometry coupled with machine learning reveals minute changes induced by single gene silencing. *Heliyon*. 2024;10(9):e29936. <https://doi.org/10.1016/j.heliyon.2024.e29936>.
67. Li X, Plataniotis KN. A complete color normalization approach to histopathology images using color cues computed from Saturation-Weighted statistics. *IEEE Trans Biomed Eng*. 2015;62(7):1862–73. <https://doi.org/10.1109/TBME.2015.2405791>.
68. Monaco J, Hipp J, Lucas D, Smith S, Balis U, Madabhushi A. Image segmentation with implicit color standardization using spatially constrained expectation maximization: detection of nuclei. In: Ayache N, Delingette H, Golland P, Mori K, editors. *Medical image computing and Computer-Assisted Intervention—MICCAI 2012*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer; 2012. pp. 365–72. https://doi.org/10.1007/978-3-642-33415-3_45.
69. Linmans J, Laak J. van der; Litjens, G. Efficient Out-of-Distribution Detection in Digital Pathology Using Multi-Head Convolutional Neural Networks. In *Proceedings of the Third Conference on Medical Imaging with Deep Learning*; PMLR, 2020; pp 465–478.
70. Pawłowski N, Bhooshan S, Ballas N, Ciompi F, Glocker B, Drozdal M. Needles in haystacks: on classifying tiny objects in large images. *ArXiv January*. 2020;6. <https://doi.org/10.48550/arXiv.1908.06037>.
71. Baxi V, Edwards R, Montalto M, Saha S. Digital pathology and artificial intelligence in translational medicine and clinical practice. *Mod Pathol*. 2022;35(1):23–32. <https://doi.org/10.1038/s41379-021-00919-2>.
72. Felländer-Tsai LAI, Ethics. Accountability, and sustainability: revisiting the hipocratic oath. *Acta Orthop*. 2019;91(1):1–2. <https://doi.org/10.1080/17453674.2019.1682850>.
73. Shreve JT, Khanani SA, Haddad TC. Artificial intelligence in oncology: current capabilities, future opportunities, and ethical considerations. *Am Soc Clin Oncol Educ Book* 2022, 42, 842–51. https://doi.org/10.1200/EDBK_350652
74. Kiener M. Artificial intelligence in medicine and the disclosure of risks. *AI Soc*. 2021;36(3):705–13. <https://doi.org/10.1007/s00146-020-01085-w>.
75. Loftus TJ, Tighe PJ, Filiberto AC, Efron PA, Brakenridge SC, Mohr AM, Rashidi P, Upchurch GR, Bihorac J. Artificial intelligence and surgical Decision-Making. *JAMA Surg*. 2020;155(2):148–58. <https://doi.org/10.1001/jamasurg.2019.4917>.
76. Abdullah YI, Schuman JS, Shabsigh R, Caplan A, Al-Aswad LA. Ethics of artificial intelligence in medicine and ophthalmology. *Asia-Pac J Ophthalmol*. 2021;10(3):289–98. <https://doi.org/10.1097/APO.0000000000000397>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.