Abunahel *et al. European Radiology Experimental* (2026)10:186
https://doi.org/10.1186/s41747-026-00837-5

European Radiology EXPERIMENTAL

SYSTEMATIC REVIEW Open Access

MRI radiomics of pancreatic tumours: a systematic review of methodological quality and imaging reproducibility

Bassam M. Abunahel[1,2,3*], Robba Rai[2,3,4], Yves De Deene[2,3,5,6], Majid Alizadeh[1], Bowen Xin[2,7], William Rae[8], Trang Pham[2,3], Peter Metcalfe[1], Anatoly B. Rosenfeld[1], Susanna Guatelli[1] and Lois Holloway[1,2,3,4]

Abstract

Objectives Radiomics holds potential for personalised treatment, but reproducibility is hindered by inconsistent reporting of imaging parameters and methodological details. This study aims to identify and evaluate key MR imaging parameters to support radiomics reproducibility and to systematically evaluate the quality of published MRI radiomics studies focused on pancreatic tumours.

Materials and methods The literature search was conducted in the Scopus, PubMed, and Web of Science databases. Human studies investigating MRI radiomics in pancreatic tumours were included. An imaging reporting checklist was developed and used to identify and evaluate imaging parameters in eligible studies. The methodological quality score (METRICS) and checklist for evaluation of radiomics research (CLEAR) score were calculated for each study.

Results Seventy-eight studies, encompassing 241 MRI pulse sequences (143 anatomical and 98 functional), were included. The imaging checklist identified ten core items as key parameters for ensuring reproducible MRI radiomics. Across all studies, none provided complete or consistent reporting of these parameters, with frequent omissions in imaging acquisition and reconstruction details. Methodological and reporting quality remained suboptimal, reflected by median METRICS and CLEAR scores of 52.4% (45–60%) and 62.3% (56–69%), respectively. Reproducibility and quality of radiomics studies would benefit from clear reporting of imaging parameters, consistent documentation of radiomics settings, and control of confounding variables. Additionally, transparency in model validation, preprocessing, and explanation methods would further enhance reproducibility and clinical relevance.

Conclusion A focused imaging reproducibility checklist has been developed, and it has been demonstrated that no pancreatic tumour MRI radiomics study reported all necessary imaging parameters for reproducibility. Methodological and reporting quality on broader radiomics components was also poor, highlighting the potential value of such a checklist for improving standardised reporting in radiomics studies.

Key Points

Question There is inconsistent reporting of imaging parameters across MRI protocols used for pancreatic tumours radiomics analysis.

Findings The developed MR imaging reporting checklist provides structured guidelines to promote consistent documentation of MR imaging parameters and support reproducible radiomics research.

Relevance statement By identifying how key MR imaging acquisition and reconstruction parameters are currently presented in MRI radiomics studies of pancreatic tumours, this study highlights inconsistencies and omissions in

*Correspondence:
Bassam M. Abunahel
bmma729@uowmail.edu.au
Full list of author information is available at the end of the article

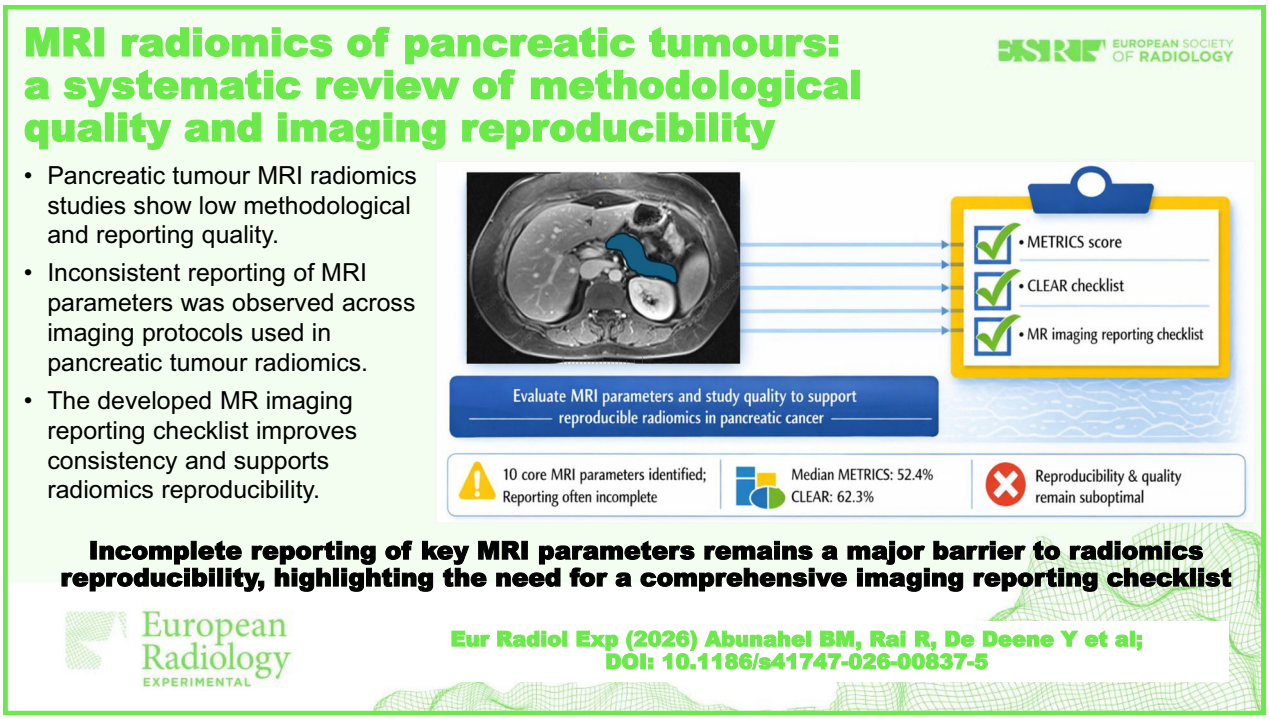
Springer

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/.

reporting, underscoring the need for more transparent reporting and adherence to guidelines to support radiomics reproducibility.

Keywords Magnetic resonance imaging, Pancreatic neoplasms, Radiomics

Graphical Abstract



Background

Pancreatic cancer is the most fatal disease of the pancreas, with an annual mortality rate of 6.9 per 100,000 individuals and a 5-year survival rate of less than 10% [1]. Among the available imaging techniques, magnetic resonance imaging (MRI) is the preferred modality for the clinical management of pancreatic tumours because of its superior soft-tissue contrast [2, 3]. Traditionally, the diagnosis of a pancreatic tumour has relied on subjective interpretation of imaging studies [4, 5]. However, medical images can also undergo mathematical analysis to extract quantitative features that serve as imaging biomarkers, a process known as radiomics [6, 7]. Radiomics-based biomarkers have the potential to support clinical decision-making by providing quantitative insights into tumour characterisation, prognosis, and treatment response [8, 9]. In pancreatic tumours, MRI-based radiomics biomarkers have been investigated for differential diagnosis, prognostic stratification, treatment response prediction, assessment of tumour aggressiveness and metastatic potential, and non-invasive characterisation of histological and molecular features [10]. However, the clinical translation of these radiomics biomarkers remains significantly impacted by their reproducibility and repeatability [11]. While reproducibility represents the robustness of biomarkers across variations in imaging conditions, such as differences in MRI scanners, acquisition and reconstruction parameters, and clinical sites, repeatability reflects their stability under consistent imaging conditions, commonly assessed through test–retest studies using the same scanner and imaging protocol. For instance, variations in imaging parameters, such as differences in scanner type, magnetic field strength, or slice thickness, can significantly change the extracted radiomics features, thereby impacting their reproducibility across institutions and patient cohorts [12]. Ensuring reproducibility in radiomics requires comprehensive and transparent reporting across the entire radiomics pipeline [13]. This pipeline comprises a multi-step workflow, including image acquisition, segmentation, feature extraction, feature selection and reduction, and modelling [14]. To strengthen reproducibility, several reporting

guidelines have been introduced. For example, the Imaging Biomarker Standardisation Initiative (IBSI) provides a framework for standardising radiomics by defining feature mathematics, specifying pre-processing steps (such as image resampling), and offering digital phantoms for validation [15]. In addition, quality assessment tools such as the methodological radiomics score (METRICS), the Radiomics Quality Score (RQS), and the Checklist for Evaluation of Radiomics Research (CLEAR) have been developed to ensure high-quality, reproducible radiomics [16, 17]. Without comprehensive reporting, developed radiomics models cannot be independently validated, thereby restricting their reliable use in clinical practice [18]. Moreover, MRI sequence generation methods and image reconstruction must also be clearly detailed to ensure repeatability and enable external validation and potential clinical translation of radiomics models [19, 20].

Radiomics has demonstrated promising potential in pancreatic tumours, with several models developed to assist in outcome prediction, differential diagnosis, and tumour classification [9]. However, the quality of these models and their suitability for clinical translation remains underexplored. To address this gap, this systematic review aims to propose key MR imaging acquisition and reconstruction parameters that should be detailed to ensure imaging reproducibility, an aspect insufficiently explored in previous radiomics reviews, and to assess the presentation of these details in the included studies. It also evaluates the methodological and reporting quality of MRI radiomics studies in pancreatic tumours, using both the METRICS score and the CLEAR checklist.

Methods

Search strategy

This systematic review protocol has been registered and is publicly available on the Open Science Framework (OSF) database [21]. A systematic literature search was conducted in Scopus, PubMed, and Web of Science databases, in collaboration with a subject librarian, to identify studies on MRI-based radiomics in pancreatic tumours. No time restriction was applied. The search followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [22], with titles and abstracts screened prior to full-text eligibility assessment. Reference lists of included articles were manually reviewed to define additional studies. Two authors (B.A. and R.R.) were involved in the eligibility assessment process. Detailed search strings and keyword combinations used in the three databases are provided in the Supplementary Material.

Eligibility and exclusion criteria

Studies were eligible if they investigated quantitative image analysis of pancreatic tumours (benign, precancerous, or malignant) using a radiomics pipeline that extracted handcrafted quantitative features, including shape, histogram features, second-order texture features and high-order texture (*i.e.*, transformed) features. MRI and hybrid MRI modalities (*e.g.*, positron emission tomography (PET)/MRI, MRI-guided linear accelerator (MRI-Linac)) were considered eligible. Studies were excluded if they: involved non-human subjects; employed fully end-to-end deep learning approaches without feature extraction; were not published in English; or were not original research articles, such as reviews, editorials, or book chapters.

Data extraction

The following data were extracted from the included studies, if available: authors, country, year of publication, study design, cohort size, clinical application, type of MRI sequence (*e.g.*, anatomical, functional), image preprocessing techniques employed, segmentation methods, radiomics software utilised, type of extracted features (*e.g.*, original, transformed, or filtered), feature extraction method (single slice or 3D), applied image filters, feature selection and feature fusion techniques (if applicable), the number and types of radiomics features incorporated into model development, the machine learning model employed (including its name and type), feature or model explainability methods (*e.g.*, SHAP (Shapley Additive Explanations), LIME (local interpretable model agnostic explanation)) and the validation method utilised (in accordance with the TRIPOD (Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis) validation statement [23]), where applicable.

Imaging acquisition and reconstruction parameters checklist

A structured checklist was developed by an MRI physicist (Y.D.) and an MRI radiographer (R.R.), with reference to relevant literature, and reviewed by a machine learning expert (B.X.) and a radiation oncologist (T.P.). The checklist was used to identify and evaluate key parameters that should be detailed to ensure imaging reproducibility and to assess how these details were presented in the eligible studies. The proposed reporting items, and their subitems, were compared with the current IBSI imaging reporting checklist with the goal of highlighting any differences in imaging parameters suggested for reporting to ensure imaging reproducibility, as necessary to ensure reproducibility of radiomics.

The development of a comprehensive checklist for reporting MR imaging parameters is motivated by the need to improve protocol transparency and facilitate full reproducibility of MRI protocols across clinical sites, which is essential for the clinical translation of radiomics

beyond research settings [24]. The current IBSI imaging reporting checklist provides important recommendations for radiomics studies; however, it primarily focuses on fundamental pulse-sequence parameters (*e.g.*, repetition time and flip angle) and does not capture additional acquisition parameters needed to replicate MR protocols across scanners and institutions [15]. These include software-related parameters, such as image reconstruction algorithms, as well as hardware-related parameters, such as coil configuration, gradients, and magnetic field strength. These parameters are known to influence MR image properties, including contrast, noise, image quality, and spatial resolution [25].

Given that radiomics quantifies image-derived patterns, including intensity distributions (*e.g.*, histogram-based features), spatial relationships (*e.g.*, texture-based features), and transformed image representations (*e.g.*, Laplacian of Gaussian- and wavelet-based features), variations in MRI parameters can substantially impact extracted radiomics features [26]. Therefore, comprehensive reporting of MR imaging parameters is critical for improving radiomics reproducibility and supporting their eventual clinical translation [17]. The ten core imaging parameters proposed in this study were defined based on evidence demonstrating their influence in shaping both qualitative and quantitative MR image properties [27–73]. Full details of these parameters are summarised in Table 1.

Methodological and reporting quality of included studies

The methodological quality of the studies was assessed using the Methodological Radiomics Score (METRICS) and the CLEAR [16, 17], with calculation details provided in the supplementary material. To minimise subjectivity in the evaluation process, two studies were initially selected and independently scored by five authors [74, 75]. Discrepancies in scoring were discussed in meetings to establish consensus on the scoring guidelines, which were then applied consistently to all included studies by two authors (B.A. and M.A.). In case of disagreements on individual papers, a third reviewer (B.X., W.R., or L.H.) was consulted.

Statistical analysis

Continuous variables are presented as the mean and standard deviation (SD) for normally distributed data, and the median with interquartile range (IQR) for non-normally distributed data. Normality of METRICS and CLEAR scores was assessed using the Shapiro–Wilk test and visual inspection of histograms and Q–Q plots. The relationship between methodological and reporting quality was assessed by correlating METRICS and CLEAR scores. Changes in study quality over time were evaluated by correlating METRICS and CLEAR scores with publication year. Pearson's correlation coefficient was used when normality assumptions were met; otherwise, Spearman's rank-order correlation was applied. A p-value < 0.05 was considered statistically significant. Statistical analyses and data visualisation were performed using Python (3.11.9) with matplotlib (v3.10.0) and seaborn (v0.13.2) libraries, and IBM SPSS Statistics (v29.0.2.0).

Results

The results of the review are organised as follows. The general characteristics of the included studies are first summarised. Subsequently, each component of the radiomics pipeline, including imaging acquisition, preprocessing, image segmentation, feature extraction, feature selection and dimensionality reduction and modelling, is detailed in separate sections (Fig. 1). Finally, the methodological and reporting quality of the included studies is assessed using the METRICS and CLEAR scores.

Characteristics of the included studies

The total number of retrieved studies was 765 (Fig. 2). Seventy-eight studies were eligible and included in the review [74–151]. These studies incorporated a total of 9,887 participants. The cohort size ranged from 16 to 597 participants, with a median of 104. The studies primarily implemented a retrospective design, with only two studies conducted prospectively [93, 132]. A total of 57 studies were in Asia [74, 76–81, 83, 84, 86–101, 103, 109, 110, 112–114, 117–125, 127–129, 132–141, 145, 147, 149–151], 13 studies in Europe [75, 82, 84, 85, 102, 107, 108, 116, 130, 142, 143, 146, 148], and eight studies in North America [104–106, 111, 115, 126, 131, 144]. A total of 50 studies applied a multiparametric MR imaging protocol for radiomics analysis, while the remaining 28 studies relied on a single MR image type [75, 78, 80, 82–84, 97, 100, 102, 104–106, 109–111, 115, 116, 124, 126, 127, 129, 131, 133, 137, 140, 142, 148, 149]. Extracted data were recorded in Supplementary Table S1.

Magnetic resonance imaging protocols

Of the 78 included studies, 67 utilised a conventional MRI scanner, whereas hybrid MRI-linear accelerator and MRI-PET systems were used in the remaining 11 studies [75, 82, 84, 85, 100, 101, 104–106, 115, 132]. A total of 241 MRI sequences were applied for radiomics analysis across the 78 included studies. Of these, 143 were anatomical and 98 were functional sequences. T2-weighted imaging (T_{2WI}) ($n = 75$) and T1-weighted imaging (T_{1WI}; $n = 60$) were the most frequently reported anatomical sequences, while diffusion-weighted imaging (DWI) ($n = 53$) and

Table 1 Defined and evaluated key imaging acquisition and reconstruction parameters in the included studies

	Core imaging items	Sub-items	IBSI addressed	Reference	Reported impact on radiomics
1	Image type	T1WI, T2WI, mixed T1WI, T2WI, DWI, or SWI	Yes	[15]	Yes [20, 27]
2	Sequence name	SE, GRE, IR, or EPI	Yes	[15]	Yes [20]
3	Scanner parameters (hardware)	Field strength, vendor name, scanner type, RF pattern, and pulse numbers	Yes	[15]	Yes [28–31]
4	Acquisition parameters	IT, TR, TE, FOV, ETL, NEX, slice thickness, slice gap, matrix size, spatial resolution, flip angle, bandwidth, acquisition time, and phase oversampling	Yes	[15]	Yes [28, 29, 31, 32]
5	Signal suppression	STIR, FLAIR, CHESS, DIXON, SPAIR, SPIR	No	[33, 34]	Yes [72, 73]
6	Motion management	Breath-hold, trigger, navigator	Yes	[15]	Yes [35]
7	Acceleration imaging	Technique (SENSE, compressed SENSE, GRAPPA, SMASH)	No	[36]	Yes [37–39]
		Acceleration direction	No	[40]	Yes [37]
		Acceleration factor	No	[41]	Yes [37, 38]
8	K-space trajectory and image reconstruction	Cartesian or non-Cartesian (e.g., single or multi-shot, partial Fourier, radial, spiral, random)	No	[15, 42]	Yes [31, 43]
9	DWI parameters	*b*-value (nominal and actual)	No	[44, 45]	Yes, for nominal b-value [46, 47]
		Diffusion time (small and large delta)	No	[48, 49]	No
		Diffusion gradient waveform (trapezoid, sine, monopolar, bipolar, or oscillating)	No	[50–52]	No
		Diffusion pulse building block (e.g., STEAM or PFGSE)	No	[53, 54]	No
		Diffusion-sensitising gradient direction (phase encoding, frequency encoding, slice selection, MDDW, Q-Ball imaging, trace encoding (1D or 3D trace) or orthogonal)	No	[55–57]	Yes [31]
		Diffusion model (Gaussian (e.g., DTI) or non-Gaussian (e.g., IVIM, DKI), along with derived metrics such as FA, MD, AD, RD and ADK.	No	[58–60]	No
		Fitting models (mono, bi-, or multi-exponential or kurtosis)	No	[61, 62]	No
		Techniques to mitigate ADC estimation inaccuracy due to EPI susceptibility artifacts, static magnetic field inhomogeneity, eddy currents, and motion artifacts	No	[63, 64]	No
10	DCE-MRI	Plain, arterial, venous, Porto-venous, or delayed	Yes	[15]	Yes [65, 66]
		Arterial input function	No	[67–69]	No
		Pharmacokinetic model	No	[70, 71]	No

AD Axial diffusivity, *ADC* Apparent diffusion coefficient, *ADK* Apparent diffusion kurtosis, *CHESS* Chemical shift selective, *DCE* Dynamic contrast enhancement, *DKI* Diffusion kurtosis imaging, *DTI* Diffusion tensor imaging, *DWI* Diffusion weighted imaging, *EPI* Echo planar imaging, *ETL* Echo train length, *FA* Fractional anisotropy, *FLAIR* Fluid attenuated inversion recovery, *FOV* Field of view, *GRAPPA* Generalised auto calibrating partially parallel acquisitions, *GRE* Gradient echo, *IBSI* Image biomarker standardisation initiative, *IR* Inversion recovery, *IT* Inversion time, *IVIM* Intravoxel incoherent motion, *MD* Mean diffusivity, *MDDW* Multidimensional diffusion weighted imaging, *NEX* Number of excitation, *PFGSE* Pulsed field gradient spin echo, *RD* Radial diffusivity, *RF* Radiofrequency, *SE* Spin echo, *SENSE* Sensitivity encoding, *SMASH* Simultaneous acquisition of spatial harmonics, *SPAIR* Spectral attenuated inversion recovery, *SPIR* Spectral pre-saturation inversion recovery, *STEAM* Stimulated echo acquisition mode, *STIR* Short tau inversion recovery, *SWI* Susceptibility weighted imaging, *TE* Echo time, *TR* Repetition time

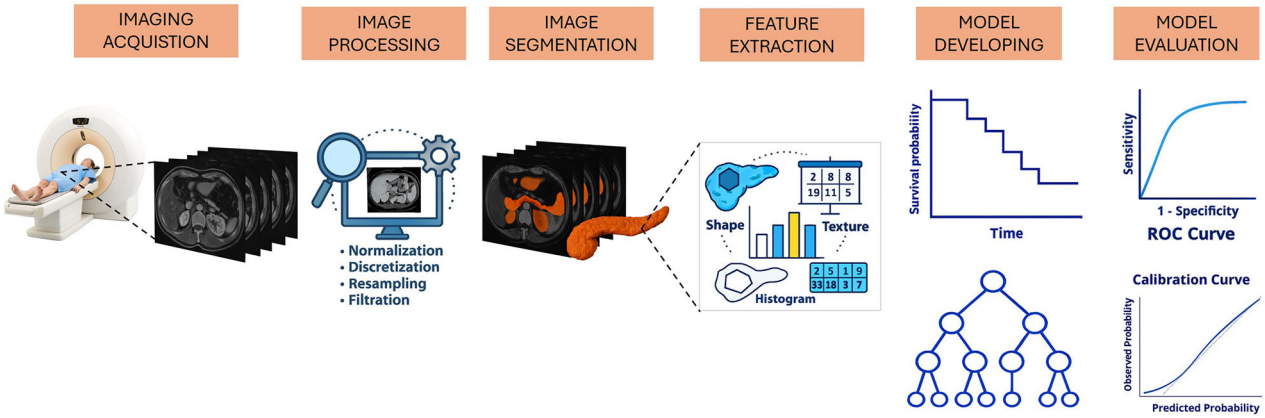


Fig. 1 An overview of the radiomics pipeline, illustrating the main steps, including image acquisition, image preprocessing, feature extraction, feature selection, model development, and evaluation

dynamic contrast-enhanced T1-weighted imaging (DCE-T_{1WI}) ($n = 45$) were the predominant functional sequences. T_{2WI} acquisitions were mainly performed using fast spin echo (FSE)-based sequences, whereas T_{1WI} acquisitions predominantly used gradient echo (GRE)-based sequences. All DWI sequences were acquired using echo planar imaging (EPI), and all combined T_{1WI}/T_{2WI} acquisitions used balanced steady-state free precession (SSFP). The specific MRI sequence used for radiomics analysis was not reported for 93 images, including 31 T_{2WI}, 32 T_{1WI}, and 30 DCE-T_{1WI} sequences (Table 2).

Image Preprocessing

Four image preprocessing techniques were applied prior to radiomics feature extraction. This included image filtration, intensity normalisation, intensity discretisation, and image resampling. Different types of filters were applied in 45 studies, while the remaining 33 studies relied only on original images [79, 84, 86, 87, 93, 100–108, 111, 112, 114, 116, 117, 120, 121, 128, 129, 131, 133, 138, 139, 141, 142, 144–146]. Intensity normalisation was performed in 35 studies, of which 30 specified the normalisation method [75, 76, 81, 89–92, 94, 96–98, 101, 104, 106, 114, 115, 117, 119, 121, 122, 126, 134–137, 140, 146, 147, 149, 151], while the remaining 5 studies did not report the method applied [112, 123, 125, 133, 144]. Similarly, intensity discretisation was applied only in 14 studies [74, 75, 82, 85, 93, 101, 102, 104–107, 119, 122, 148]. Finally, different resampling techniques were applied in 26 studies, with 13 studies specifying the interpolation method [74, 75, 82, 85, 93, 94, 97, 98, 107, 112, 134, 135, 148], while the remaining 13 did not report it [96, 101, 102, 114, 119, 122–126, 137, 149, 151]. Figure 3 illustrates the four image preprocessing techniques, detailing their specific types and frequency across the included studies.

Image segmentation, feature extraction, and feature processing

Manual segmentation was applied in 69 studies, whereas semi-automated and fully automated segmentation were applied in only five and four studies [123, 125, 144, 150], respectively. Of these manual segmentation studies, only 34 studies addressed inter-observer radiomics reproducibility [76, 77, 79–81, 85, 87–91, 95–98, 101, 107, 112–115, 118–120, 122, 124, 126, 127, 135–137, 139–141, 145–147, 149].

Most studies used 3D volumes of tumours for radiomics extraction, while nine studies only used the largest slice for analysis [80, 85, 107, 118, 121, 128–130, 146]. Twenty-nine studies employed in-house developed radiomics software [76, 77, 80, 84, 99, 105, 109–112, 114–116, 120, 128, 129, 131–133, 138–140, 142–145, 147, 150, 151], whereas the remaining studies used an IBSI-compliant radiomics software. All the 78 included studies extracted handcrafted radiomics features, with only one study incorporating deep learning-derived features [144]. Finally, feature normalisation was implemented in nine studies, with feature scaling used in one study [97] and Z-score in five studies [74, 87, 114, 135, 137], while three studies did not report the type of feature normalisation [95, 123, 145].

Feature selection and dimensionality reduction

Redundant (*i.e.*, non-informative) radiomics features were eliminated in 69 studies, while nine studies developed models without prior feature elimination [80, 105, 116, 129, 131, 132, 138, 142, 143]. A range of feature selection

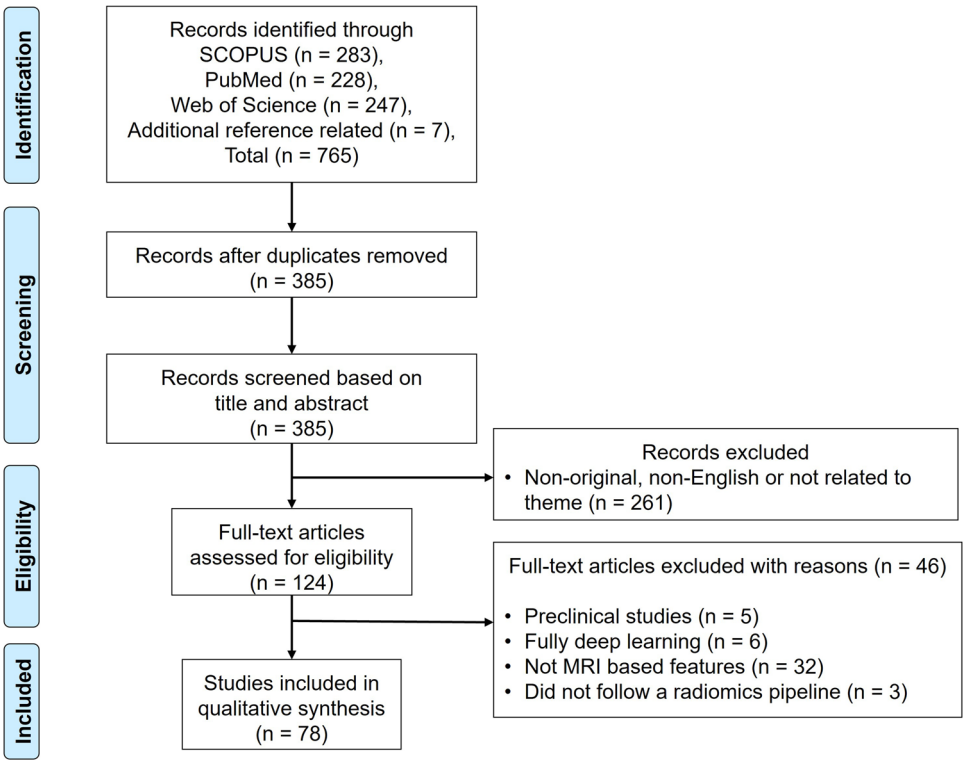


Fig. 2 Flowchart of the eligibility assessment workflow, including the identification, screening, and final inclusion or exclusion of studies

and dimensionality reduction techniques was applied across the studies, including embedded, filter, and wrapper-based methods. The most used embedded method was least absolute shrinkage and selection operator regression (LASSO) (n = 38), followed by the Mann–Whitney *U*-test ($n = 17$) for filter-based methods, and recursive feature elimination ($n = 5$) for wrapper-based methods. Other dimensionality reduction techniques, such as principal component analysis (PCA) ($n = 3$), were also applied (Fig. 4). Out of the 50 studies that utilised multiparametric MRI for feature extraction, only 10 studies identified informative features from single images [85, 86, 103, 118, 119, 121, 138, 141, 143, 146].

Statistical and machine learning modelling in radiomics

Thirteen different model types were applied across the studies, with logistic regression being the most frequently applied across all three radiomics applications: prediction ($n = 15$), differentiation ($n = 7$), and classification ($n = 7$) (Fig. 5). Only one study applied a neural network for the prediction of differential diagnosis [133]. Of the 78 included studies, 59 studies validated their models, with only 15 studies (19.2%) performing external validation [74, 83, 95, 97, 99, 101, 113–115, 118, 122, 127, 134, 144, 145]. The remaining 21 studies lack any form of model validation [80, 85, 86, 100, 105, 108–111, 116, 128, 129, 131, 132, 138, 140–143, 146]. Notably, none of the included studies evaluated the predictive performance of an existing published radiomics model on an independent dataset. Lastly, only 16 studies (20.6%) reported the use of feature or model explanation methods, including feature maps [81, 105, 123, 135, 137, 146], case-study analysis [88–91, 93, 102, 110, 143], and SHAP [99, 149]. Applications of MRI-derived radiomics for pancreatic tumours are illustrated in Fig. 6.

Radiomics as a prediction tool for pancreatic tumours

A total of 45 studies applied radiomics for prediction of prognosis [75–85, 149], treatment response [103–106], gene mutation status [92, 93], molecular expression profiles [86–91], metastasis [74, 94–102], malignant potential [108–111], early recurrence [114], postoperative complications [107, 151], invasive behaviour [112, 113], mural nodules [116], and necessity of online adaptive radiotherapy [115]. Across these studies, Cox regression remained the primary model for survival analysis, while logistic regression and support vector machine were mainly used to model gene mutations and metastatic risk. For estimating malignancy risk, invasiveness, and early recurrence, random forest, gradient boosting (XGBoost),

Table 2 Overview of 241 anatomical and functional MRI pulse sequences reported in the included studies. Numbers in parentheses indicate the number of sequences

Sequences	Category	Type	Sequence name and frequency
241 MRI sequences	Anatomical ($n = 143$)	T2WI ($n = 75$)	Half-Fourier, single-shot turbo spin echo ($n = 7$)
			Single-shot turbo spin echo ($n = 12$)
			Fast spin echo ($n = 25$)
			Unspecified sequences ($n = 31$)
		T1WI ($n = 60$)	Fast spin echo ($n = 2$)
			Gradient echo ($n = 26$)
			Unspecified sequences ($n = 32$)
			Gradient echo sequences were reported with different names, including LAVA, VIBE, THRIVE
		T1/T2WI ($n = 6$)	Balanced steady state free precession ($n = 6$)
			All were acquired using low-field strength (0.35 T) MRI scanners
		Magnetic resonance cholangio-pancreatography (MRCP) ($n = 2$)	Spin echo ($n = 2$)
	Functional ($n = 98$)	DWI ($n = 53$)	All were echo planar imaging (EPI) ($n = 53$)
		DCE_T1WI ($n = 45$)	Spoil gradient echo ($n = 15$)
			Unspecified sequences ($n = 30$)

DCE Dynamic contrast-enhanced, *DWI* Diffusion weighted imaging, *LAVA* Liver acquisition with volume acceleration, *MRI* Magnetic resonance imaging, *MRCP* Magnetic resonance cholangio-pancreatography, *THRIVE* T1 high resolution isotropic volume examination, *T1WI* T1 weighted imaging, *T2WI* T2 weighted imaging, *VIBE* Volumetric interpolated Breath hold examination

and multilayer perceptron (MLP) models were commonly applied. Only nine models were externally validated [74, 83, 95, 97, 99, 101, 113–115], whereas 11 remained unvalidated. Furthermore, none of these studies attempted to validate an existing model using an independent dataset. Finally, prospective data were used in only one study [93].

Regarding MRI protocols used in this radiomics application, both quantitative and qualitative sequences were applied, with balanced SSFP being the predominant sequence applied to model radiotherapy treatment response from inter-fraction radiomics variations [104–106]. However, none of these models evaluated whether the observed variations reflected technical factors —such as patient setup variability—rather than true biological response.

Radiomics as a differential diagnosis tool for pancreatic tumours

Seventeen studies applied radiomics for differentiation of pancreatic lesions [117–133]. Only three included external validation [118, 122, 127], whereas 11 models remained unvalidated. None of these studies attempted to validate an existing model using an independent cohort. Furthermore, prospective data were used in only one study [132]. Quantitative MRI were the predominant pulse sequence (64%) applied in this radiomics application.

Radiomics as a classification tool for pancreatic tumours

The remaining 16 studies were for classification of histologic grades and molecular subtypes [134–148, 150]. Only three included external validation [134, 144, 145], while six models remained unvalidated. None of these studies attempted to validate an existing published model using an independent cohort. All models were developed using retrospective data. Finally, both quantitative and qualitative MRI sequences were applied in this radiomics application.

Reported imaging acquisition and reconstruction parameters

Using our modified IBSI-based reporting checklist (Table 1), all the 78 studies reported the image type and scanner related parameters; however, none reported the RF pulse flip angle pattern or the number of RF pulses. Of the 241 reported sequences, 133 (55.2%) specified the sequence name (Table 2). Reporting of sequence acquisition

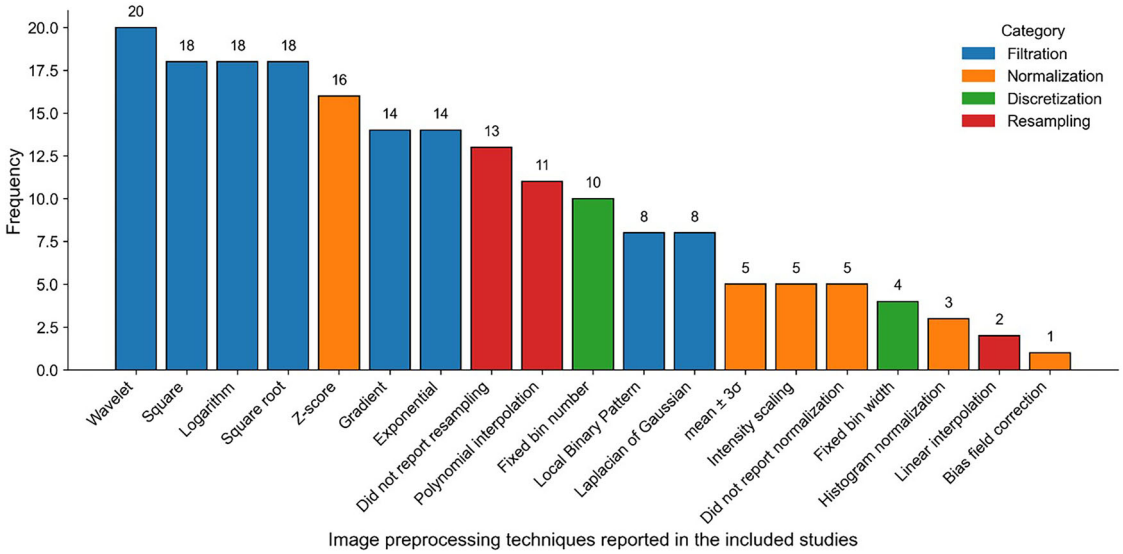


Fig. 3 An overview of the image preprocessing techniques applied across the included studies, including image filtration, normalisation, resampling, and discretisation

parameters was inconsistent, and no study reported the use of phase oversampling techniques, the k-space sampling trajectory, the reconstruction algorithm, or any distortion compensation methods for static magnetic field inhomogeneity. Of the three studies that applied SENSE parallel imaging with EPI, none reported the acceleration factor or direction of acceleration [93, 129, 142]. Fat/water suppression techniques were reported in only six studies, using SPAIR [128], CHESS [129], or DIXON [99, 107, 110, 130]. Finally, only two studies reported the use of motion compensation methods, including periodically rotated overlapping parallel lines with enhanced reconstruction (PROPELLER) and BLADE [85, 122].

None of the 53 DWI sequences reported the diffusion time, the gradient waveform type, or the diffusion preparation module. Only two studies reported diffusion models (both non-Gaussian) [120, 133], and four reported the diffusion gradient direction [120, 129, 133, 142]. Furthermore, none of the studies accounted for potential inaccuracies in ADC estimation arising from EPI-related susceptibility or static magnetic-field inhomogeneity. Although all DWI sequences reported the nominal b-value, none reported the actual b-value. For the DCE-MRI, only one study generated four quantitative perfusion maps, yet neither the applied pharmacokinetic model nor the arterial input function used to derive these maps was reported [93]. The reported MR imaging parameters across the included studies are summarised in Fig. 7, and the full extracted MRI pulse sequence data are provided in Supplementary Table S2.

Methodological quality of the included studies

Variation in METRICS scores across five authors ranged from 52.30–78.60% for the first study [74] and 53.40–82.50% for the second study [75]. Heatmaps illustrating the distribution of METRICS scores between the authors are provided in Supplementary Figs. S1 and S2, respectively. Following the consensus scoring guidelines (Supplementary Table S3), METRICS scores ranged from 27.2% to 78.6%, with a median (IQR) of 52.4% (45.1–60.0%), indicating that most studies were clustered within the moderate-to-low-quality range, with only a limited number of studies achieving higher scores (Supplementary Fig. S3). The five most frequently reported METRICS characteristics were defining clear study eligibility criteria, using high-quality reference standard or outcome measure, eliminating redundant radiomics features, using appropriate performance evaluation metrics, and accounting for uncertainty in radiomics analysis (Supplementary Fig. S3). The five least frequently reported METRICS characteristics were clinical translation of the MRI acquired imaging data, transparent reporting of feature extraction parameters, including default configuration, handling of confounding factors, external validation of the developed model using an independent dataset from other clinics and open science practices (data, code or model availability) (Supplementary Fig. S3). Item 17 of METRICS was not applicable for any study as it evaluates only fully end-to-end deep learning models, which is beyond the scope of this review. Other details are shown in Supplementary Table S4.

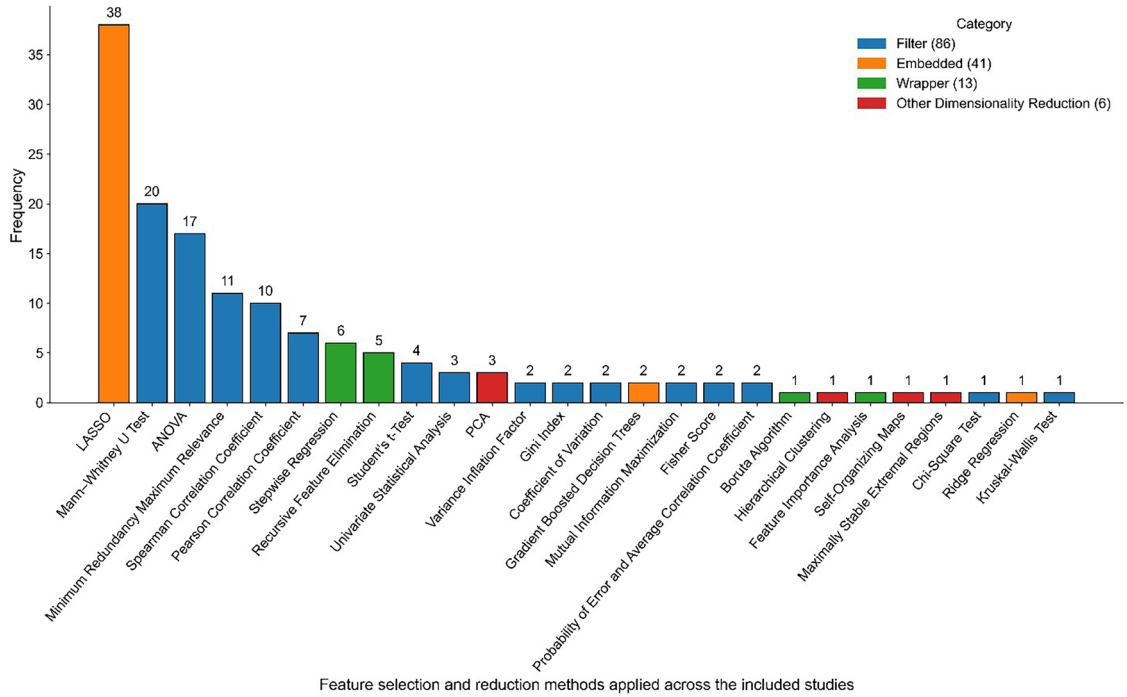


Fig. 4 A summary of the feature selection and dimensionality reduction techniques applied across the included studies, including embedded, filter, and wrapper-based methods. PCA, Principal component analysis; ANOVA, Analysis of variance; LASSO, Least absolute shrinkage and selection operator

Reporting quality of the included studies

Variation in CLEAR scores across five authors ranged from 45.70–82.20% for the first study [74] and 50.00–94.40% for the second study [75]. Heatmaps illustrating the distribution of CLEAR scores between the authors are provided in Supplementary Figs. S4 and S5, respectively. Following the consensus scoring guidelines (Supplementary Table S5), CLEAR scores ranged from 30.2% to 80.8%, with a median (IQR) of 62.3% (56.4–69.1%), indicating variability in reporting quality across the included studies, with most studies clustered within the moderate-quality range (Supplementary Fig. S6). The most consistently reported items of the CLEAR checklist were a clear description of the reference standard or outcome measure, a detailed listing of the number and classes of extracted radiomics features, feature selection and dimensionality reduction techniques, and clear reporting of the final radiomics features included in the predictive model (Supplementary Fig. S6). In contrast, many items of the CLEAR were underreported. None of the studies detailed use of the CLEAR checklist, and only eight studies referred to non-radiomics reporting guidelines such as TRIPOD. Only four studies described how class imbalance was managed [107, 130, 150, 151]. Access to image preprocessing, source or model code was not reported in any study and only 18 studies reported that their data were available upon request. None of these studies accounted for confounding factors raised from using multi-centre or multi-scanner data heterogeneity. None of the studies clearly reported the radiomics extraction parameters or reported using default configuration parameters. Only 16 studies clearly reported the use of feature or model explanation methods (Supplementary Fig. S6). Finally, most studies (74%) reported their imaging protocols; however, none of them had comprehensively reported all the imaging acquisition and reconstruction parameters specified in our checklist. Other details are presented in Supplementary Table S6.

A statistically significant positive correlation was observed between METRICS and CLEAR scores (Spearman $r = 0.728$, $p < 0.001$). Publication year also showed significant positive correlation with METRICS ($r = 0.406$, $p = 0.0002$) and CLEAR scores ($r = 0.291$, $p = 0.0097$) (Fig. 8). Other details are provided in Supplementary Figs. S7 and S8. The distributions of METRICS and CLEAR scores across the 78 included studies are presented in Supplementary Fig. S9.

Discussion

This systematic review aimed to identify and evaluate key MRI imaging parameters likely to impact imaging and thus radiomics reproducibility. It also assessed the

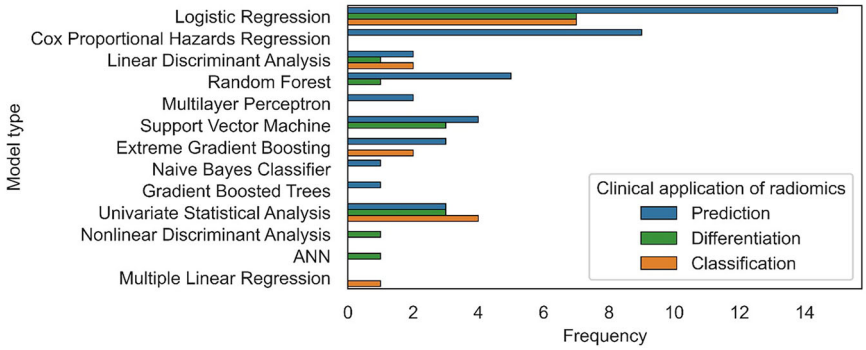


Fig. 5 A summary of the statistical and machine learning models applied across the included studies, illustrating the diversity of modelling approaches used for radiomics analysis. ANN, artificial neural network

methodological and reporting quality of MRI radiomics studies focused on pancreatic tumours. A total of 78 studies were included, yielding 241 MRI pulse sequences. Of these, 143 represented anatomical imaging, while 98 corresponded to functional imaging. A total of 10 core MR imaging and reconstruction parameters were identified and evaluated across the included studies, yet none of the included studies reported them comprehensively. Thus, none of these studies could be confidently validated and thus used in an independent dataset or clinical practice significantly limiting potential for clinical translation. Furthermore, the methodological and reporting quality of the included studies was also limited, with median METRICS and CLEAR scores of 52.4% and 62.3%, retrospectively.

It is worth noting that anatomical sequences were the predominant imaging types used for radiomics analysis of pancreatic tumours, whereas functional sequences were largely limited to DWI and DCE. Even though other functional sequences—such as T_1 and T_2 mapping—have demonstrated clinical significance in pancreatic tumours [152], yet they have not been incorporated into radiomics. Moreover, given the variability in reported methods for image preprocessing, feature processing, feature selection, and modelling, establishing robust radiomics features for pancreatic tumours remains essential, as these factors have been shown to influence radiomics stability in other tumour types [153–155]. A recent systematic review evaluating the reporting quality of radiomics software and key image preprocessing steps, including intensity discretization and image resampling, found that many radiomics studies fail to fully report preprocessing parameters despite their potential influence on radiomics reproducibility [156]. Finally, it should be noted that noise profiles and image intensities differ between manufacturers, which may not be completely addressed by image preprocessing. This lack of transparency may potentially limit the clinical translation of radiomics.

Notably, manual segmentation remains the predominant image segmentation approach in MRI radiomics of pancreatic tumours, with limited adoption of semi-automated or fully automated segmentation. Despite the influence of segmentation variability on radiomics stability [157], inter-observer reproducibility was assessed in only 34 of the 69 studies (49.3%) employing manual segmentation, highlighting the need for further evaluation of the effect of inter-observer segmentation variability on radiomics stability in pancreatic tumours. Furthermore, the three studies that applied inter-fraction radiomics variations (*i.e.*, delta radiomics) to predict radiotherapy response interpreted longitudinal changes in radiomics as indicators of tumour response [104–106]. However, these changes may also be influenced by technical factors, such as variations in patient positioning between treatment fractions [11]. Therefore, the contribution of technical variability to the observed radiomics changes remains unclear, as none of these studies incorporated additional assessments, such as test–retest scan, to distinguish setup-related variability from true tumour response.

Notably, several MRI acquisition and reconstruction parameters not included in the current IBSI checklist—but incorporated into our reporting checklist (Table 1)—are known to influence radiomics stability. These include the k-space sampling trajectory and reconstruction algorithm [31, 43], acceleration related parameters [37–39], DWI-specific settings including nominal b-values and phase-encoding direction [45, 46], and signal suppression techniques [72, 73]. Additional parameters considered in our evaluation but absent from the IBSI checklist affect image quality and contrast, yet their impact on radiomics stability remains not investigated. These include the choice of DCE arterial input function and

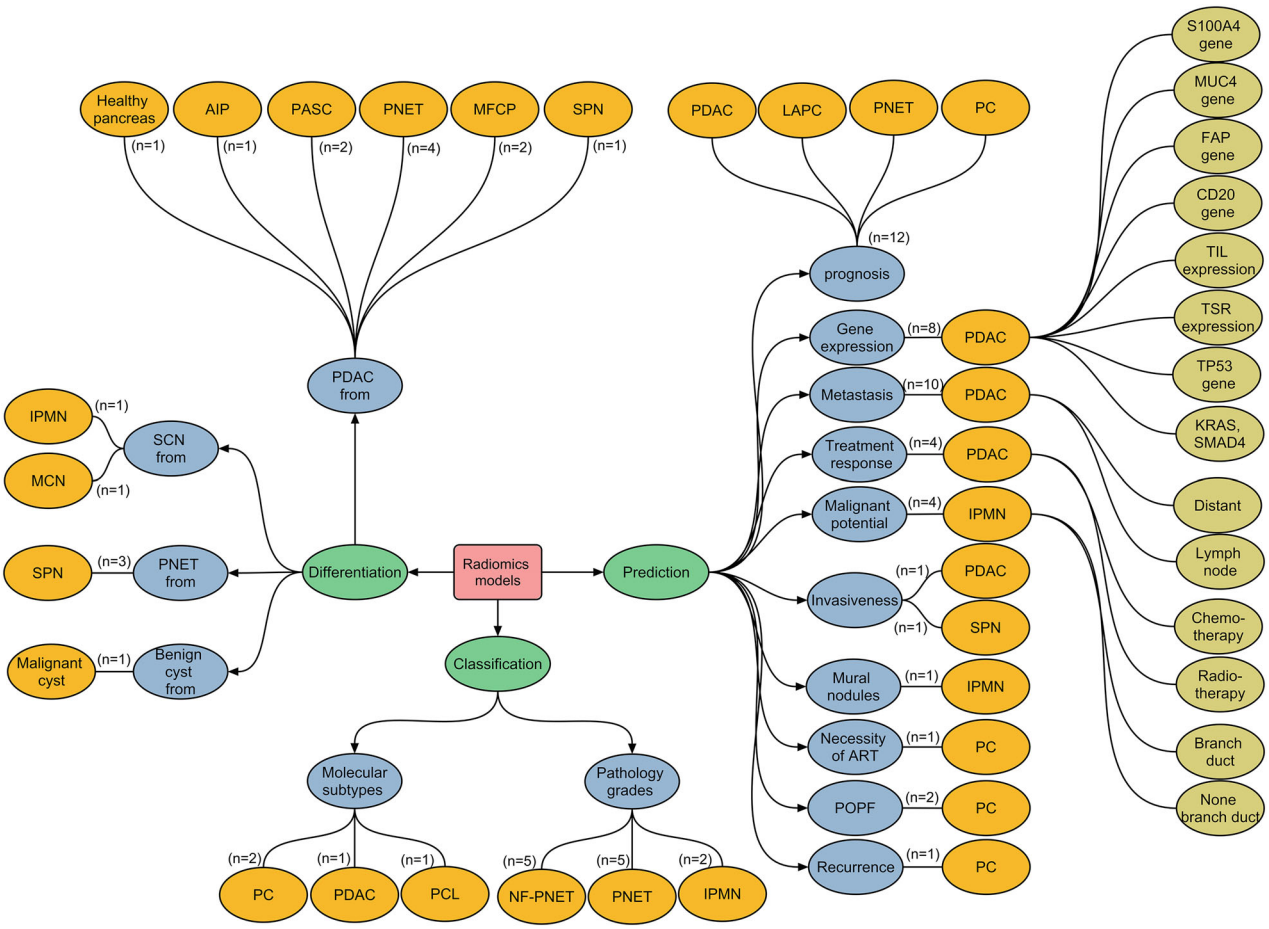


Fig. 6 Applications of MRI-derived radiomics of pancreatic tumours. AIP, Autoimmune pancreatitis; ART, Adaptive radiation therapy; FAP, Fibroblast activation protein; IPMN, intraductal papillary mucinous neoplasm; LAPC, Locally advanced pancreatic cancer; MCN, Mucinous cystic neoplasm; MFCP, Mass forming chronic pancreatitis; MUC4, Mucin-4 protein; NF-PNET, Non-functioning pancreatic neuroendocrine tumour; PASC, Pancreatic adenosquamous carcinoma; PC, Pancreatic cancer; PCL, Pancreatic cystic lesions; PDAC, Pancreatic ductal adenocarcinoma; PNET, Pancreatic neuroendocrine tumour; POPF, Post-operative pancreatic fistula; SCN, Serous cystic neoplasm; SPN, Solid-pseudopapillary neoplasm; TIL, Tumour infiltrating lymphocyte; TSR, Tumour-stroma ratio

pharmacokinetic model, diffusion time, diffusion sequence components, gradient waveform design, diffusion model (Gaussian or non-Gaussian), fitting approach (mono-, bi-, or multi-exponential), and correction strategies for ADC inaccuracies arising from EPI susceptibility, static-field inhomogeneity, eddy currents, and motion. Collectively, the relevance of both the IBSI and our additional parameters to radiomics stability remains insufficiently defined and warrants further investigation to support the development of a more comprehensive MR imaging reporting checklist for future radiomics research.

Using our modified imaging reporting checklist, none of the 241 MRI sequences demonstrated complete or consistent reporting, with frequent missing of reporting of essential acquisition and reconstruction parameters. Across the 78 included studies, the missing parameters included acceleration parameters [94, 129, 142], signal suppression techniques (reported only in six studies) [99, 107, 110, 128–130], k-space sampling trajectory, image reconstruction algorithm, RF-pulse flip-angle pattern, number of RF pulses, diffusion time, actual b-value, gradient encoding direction, diffusion-gradient waveform, diffusion-gradient directions, arterial input function, and pharmacokinetic perfusion model.

Considering and reporting these parameters in radiomics research is essential, as multiple studies have highlighted that acquisition and reconstruction variability affect image quality and radiomics stability. For instance, lower acceleration factors with SENSE yield more stable radiomics than compressed SENSE [38], and GRAPPA can result in more robust radiomics than SENSE [39]. Additionally, partial Fourier acquisition significantly alters

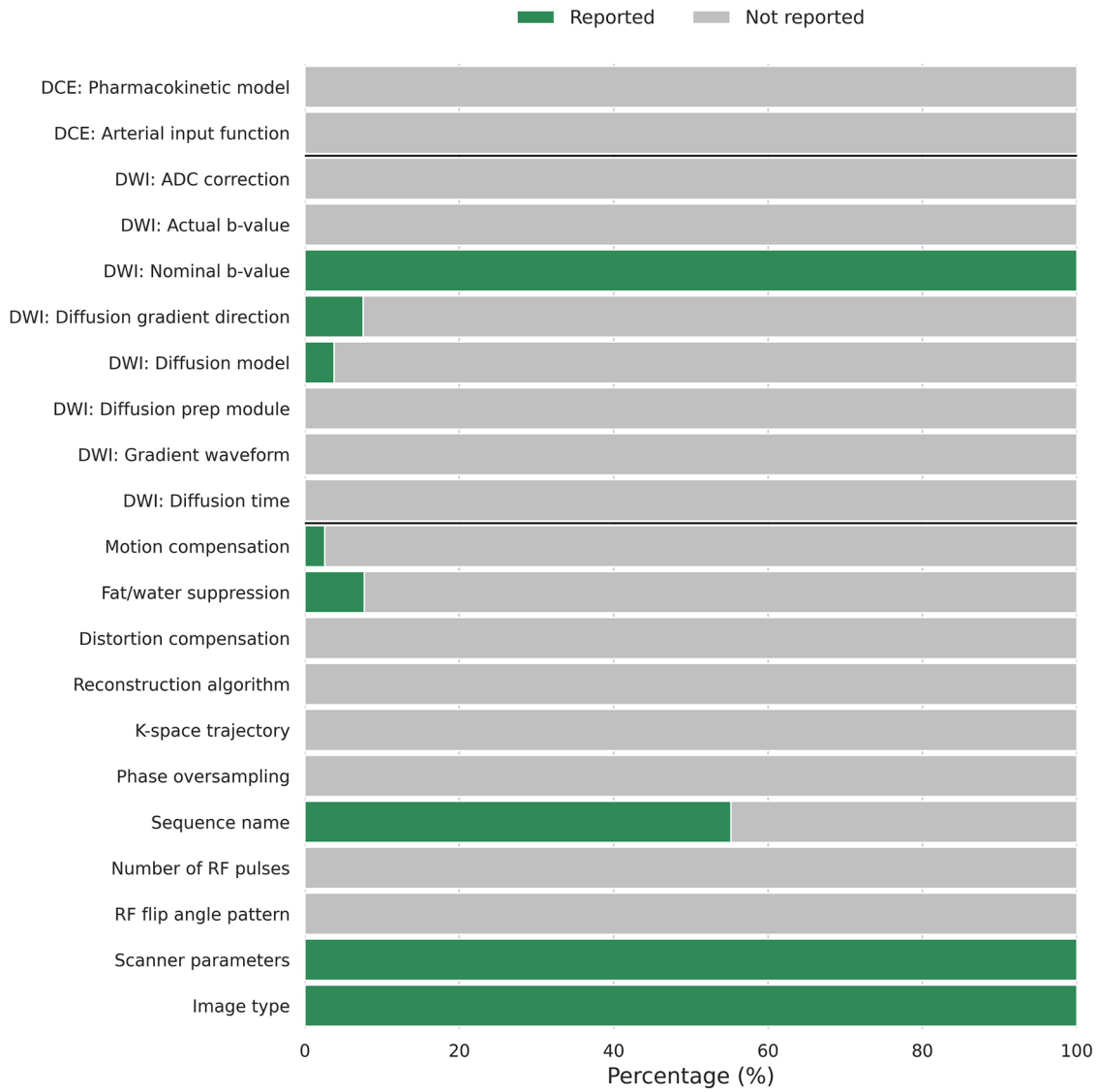


Fig. 7 Frequency of reported MR imaging acquisition and reconstruction parameters across the 241 anatomical and functional MRI pulse sequences identified across the included studies. DCE, Dynamic contrast enhancement; ADC, Apparent diffusion coefficient; DWI, Diffusion weighted imaging; RF, Radiofrequency pulse

radiomics values, whereas radial k-space sampling trajectory shows minimal deviation from Cartesian sampling [31]. Changes in phase-encoding direction also introduce significant radiomics variability [31]. In terms of image reconstruction, total-variation and wavelet-based methods generate more stable radiomics features than inverse Fourier transform or conjugate-gradient methods [43]. Moreover, variations in RF flip angle patterns and the number of RF pulses likewise affect radiomics stability [31]. Similarly, signal suppression techniques differ in their effectiveness for signal suppression [158, 159], yet their influence on radiomics feature stability remains insufficiently explored [72]. Finally, although some of the reported sequences shared identical acquisition parameters, defining robust and reproducible radiomics features of pancreatic tumours remains warranted, as these sequences were acquired on different MRI scanners from various vendors—a factor known to affect radiomics feature stability [160]. For example, among the 60 T1WI sequences, 15 had identical acquisition parameters; however, they were obtained using scanners from different vendors or with varying field strengths [81, 86, 88–91, 100–102, 119, 123, 125, 127, 135, 137], highlighting the need for future studies to define radiomics features that

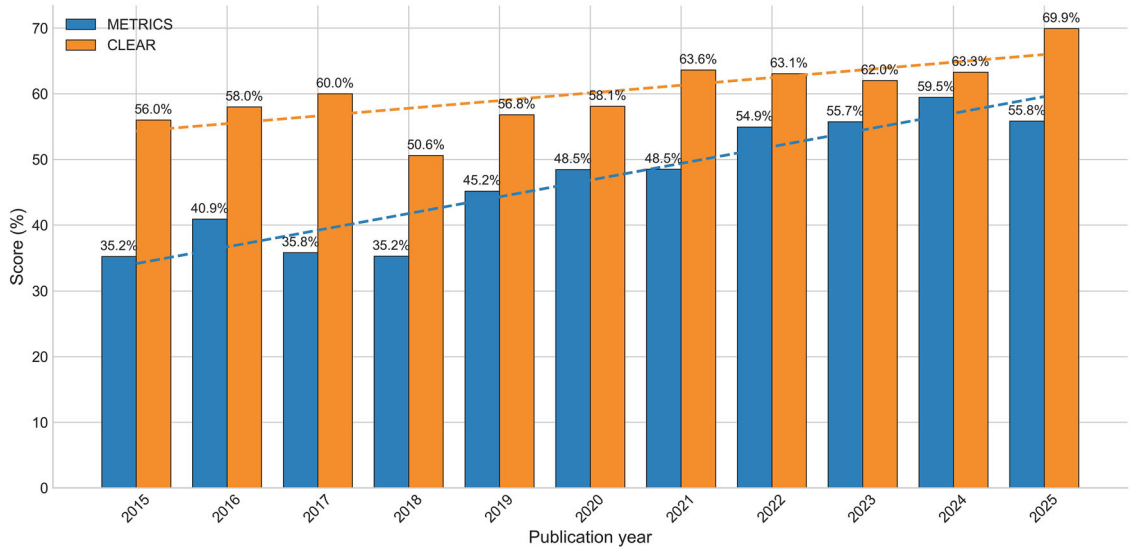


Fig. 8 Bar chart showing the mean METRICS and CLEAR scores (percentage) by year of publication. The dashed lines indicate the trend of METRICS and CLEAR scores over time. A significant positive correlation with publication year is observed for both METRICS (Spearman $r = 0.406$, $p = 0.0002$) and CLEAR (Spearman $r = 0.291$, $p = 0.0097$) scores

are reproducible across these different imaging protocols and scanner settings.

For the DWI, diffusion time is a key parameter influencing ADC measurement [161, 162], as longer diffusion times allow more water molecules to interact with tissue barriers, resulting in lower ADC values [163]. The pancreatic tumour microenvironment is highly complex, with multiple compartments and diffusion-restricting structures, leading to time-dependent, non-Gaussian (*i.e.*, bi-exponential) diffusion behaviour [49, 164, 165]. Thus, both diffusion time and fitting model strongly influences the estimated diffusion coefficient [62, 166], yet their impact on radiomics remains underexplored. Other factors are known to influence quantitative ADC estimation, yet their impact on radiomics has not been investigated. These include diffusion gradient waveform [167], diffusion gradient directions or the number of directions acquired [168], actual b-value [169], and static magnetic field inhomogeneity and EPI-related susceptibility [170]. Techniques such as segmented EPI, shimming, dielectric pads, and field-mapping have been proposed to mitigate these distortions [171], yet their role in improving radiomics robustness has not been evaluated.

For the DCE-MRI, reporting the pharmacokinetic perfusion model is essential, as model selection can substantially change the estimated perfusion parameters [172, 173], yet its impact on radiomics stability remains insufficiently investigated. It is also notable that none of the studies investigated radiomics stability across different contrast enhancement phases. Evidence from prostate, lung, and colorectal cancers shows that radiomics stability is highly dependent on contrast agent phase [65, 66], underscoring the importance of phase selection when interpreting DCE-derived radiomics.

According to the METRICS and CLEAR criteria, the included studies demonstrated several reporting limitations. None of the studies specified how radiomics features were calculated or reported their configuration parameters, despite evidence that variations in texture-feature computation can change feature values and their interpretation [174]. Confounding factors related to multi centre design, scanner variability, and population heterogeneity were also unaddressed, raising concerns about radiomics generalisability [175]. While harmonisation techniques such as ComBat aim to reduce scanner related variability, their impact on radiomics reproducibility is still inadequately assessed [176]. Also, patient bias adjustment strategies established in clinical modelling [177, 178], still underutilised in radiomics modelling. External validation was similarly limited, as only 15 studies (19.2%) evaluated their models on independent datasets, underscoring ongoing barriers to data sharing and the necessity of external testing to confirm radiomics generalisability [179]. Transparency was similarly limited, with no study providing preprocessing or model code and only 18 providing data upon request, a notable concern for federated learning where clear reporting is essential for identifying site-specific differences and interpreting model behaviour [180]. Finally, explainability was rarely addressed, with only 16 studies (20.5%) using feature or

model explanation techniques, despite their potential to enhance model interpretability [181].

The statistically significant correlation observed between METRICS and CLEAR scores (Spearman $r = 0.728$, $p < 0.001$), highlighting that radiomics studies with more rigorous methodologies tended to report their procedures and analyses more transparently [182]. This combination (*i.e.*, methodological rigour and reporting transparency) is essential for producing reproducible and meaningful radiomics, which will support the integration of radiomics into clinical workflow [182–184]. Additionally, the significant positive correlation observed between publication year and both METRICS and CLEAR scores confirms that radiomics study quality has improved over time, consistent with earlier RQS-based reviews [185–187]. However, despite this progress, radiomics remains not yet fully mature for clinical translation [183]. A key limitation identified in our study is inconsistent reporting of MR imaging parameters, which varied substantially across the applied MRI protocols. This variability, and lack of all required information to implement sequences independently, will likely limit radiomics reproducibility and remains a major barrier to the clinical translation of radiomics.

It is worth noting that the RQS has been widely used to evaluate the quality of radiomics studies [188]. However, some of its evaluation criteria (such as prospective design) are often not applicable in radiomics research which often results in low scores, even for well conducted studies [185]. In contrast, the METRICS tool was developed to address the increasing use of machine learning in radiomics. It evaluates the entire radiomics pipeline, from image acquisition and feature extraction to model development and validation [16]. As a result, studies that received low RQS scores have often achieved considerably higher scores when assessed using METRICS [189]. This pattern was also observed in the subset of 11 studies included in our review that matched with an earlier RQS-based systematic review [190]. Within this shared group, METRICS consistently assigned higher scores than RQS (Supplementary Table S7). These findings suggest that METRICS provides a more comprehensive assessment of radiomics study quality by recognising methodological strengths that may not be fully captured by RQS. Therefore, METRICS may be better a more suitable tool for evaluating the quality of radiomics studies, while comparison with the updated RQS version should be explored in future studies [191].

There are several limitations that need to be acknowledged when interpreting the findings of this review. First, the quality of the studies was assessed using the METRICS and CLEAR checklists. Although other tools—such as the RQS—were also applicable, different radiomics quality evaluation tools can yield varying interpretations of methodological rigour. However, METRICS, was developed on the foundation of RQS, offers a more comprehensive framework that incorporates additional components of the radiomics pipeline, including deep radiomics. Second, the checklist used to extract MRI parameters was developed by an MRI physicist (Y.D.) and MRI radiographer (R.R.) based on literature evidence on factors impacting images. Although the checklist was informed by multidisciplinary input, it could be further strengthened and adopted with greater confidence through a formal, structured process involving a broader international, multidisciplinary panel of experts. To enhance reproducibility and broader applicability, future work should aim to establish a globally accepted checklist, ideally developed through consensus methods such as Delphi and incorporated into IBSI reporting checklist. Third, no statistical conclusion could be made regarding the extracted MRI parameters due to the incomplete and inconsistent reporting across the studies. The wide variety between the reported imaging and reconstruction parameters underscores the need for standardising imaging protocols. Fourth, it is important to note that the METRICS and CLEAR checklist serve as guidelines for assessing radiomics studies. At this stage, there is no universally preferred methodology for developing radiomics based models. Finally, it is important to note that this review focused on MRI radiomics studies applying handcrafted radiomics for pancreatic tumours, with the primary objective of evaluating reproducibility across the radiomics pipeline, including imaging acquisition, image preprocessing, tumour segmentation, feature extraction, feature selection, and modelling. End-to-end deep learning approaches represent an emerging alternative that can minimise variability associated with several stages of the radiomics workflow, particularly tumour segmentation and manual feature engineering, through automated feature learning directly from imaging data, potentially enabling more streamlined and consistent workflows. As the field continues to advance, challenges related to the interpretability and explainability of deep learning approaches have emerged, highlighting the need for robust strategies to support their clinical translation. Unlike handcrafted radiomics, where features are mathematically defined, including shape, first-order, and texture descriptors, many of which have been standardised by the Image Biomarker Standardisation Initiative (IBSI), deep learning models learn latent image representations whose biological meaning and clinical significance remain largely unknown, highlighting the need for robust interpretability and explainability approaches to support clinical translation [192].

Conclusion

Radiomics holds potential promise for personalised management of pancreatic tumours. However, current MRI-based studies exhibit suboptimal methodological and reporting quality, with frequent omissions in reporting imaging acquisition and reconstruction parameters. These gaps hinder meaningful comparison of radiomics models across different clinics and patient populations, and thus clinical implementation. The proposed MRI reporting checklist provides a structured framework to improve the ability to validate radiomics reproducibility. The need for a globally accepted checklist—ideally through consensus methods such as Delphi and integration with the current IBSI checklist—will be essential for shifting MRI radiomics from exploratory research toward a reliable precision-medicine tool suitable for routine clinical decision making.

Abbreviations

CLEAR	Checklist for evaluation of radiomics research
EPI	Echo planar imaging
GRE	Gradient echo
METRICS	Methodological quality score
SHAP	Shapley additive explanations
SSFP	Steady-state free precession
TRIPOD	Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis

Supplementary information

The online version contains supplementary material available at https://doi.org/10.1186/s41747-026-00837-5.

Additional information Table S1: General characteristics of the 78 included studies. **Figure S1:** Heatmap showing the distribution of METRICS scores across five authors for the first study [22], illustrating reproducibility in scoring patterns among the authors. **Figure S2:** Heatmap showing the distribution of METRICS scores across five authors for the second study [11], illustrating reproducibility in scoring patterns among the authors. **Table S3:** Consensus scoring guidelines used for evaluating all 78 included studies with the METRICS checklist. **Figure S3:** METRICS score of the 78 included studies. Each column represents one of the 30 METRICS items, and each row related to an individual study. Green cells indicate that the criterion was met, red cell; not met, and grey cells; not applicable. The percentage of studies meeting each item is shown in parentheses. The final column shows the METRICS score as a percentage of the total for each study. **Table S4:** Methodological Radiomics Score (METRICS) quality scoring tool of the 78 included studies. **Figure S4:** Heatmap showing the distribution of CLEAR scores across five authors for the first study [22], illustrating reproducibility in scoring patterns among the authors. **Figure S5:** Heatmap showing the distribution of CLEAR scores across five authors for the second study [11], illustrating reproducibility in scoring patterns among the authors. **Table S5:** Scoring guidelines used for evaluating all 78 included studies with the CLEAR checklist. **Figure S6:** CLEAR score of the 78 included studies. Each column represents one of the 58 CLEAR items, and each row related to an individual study. Green cells indicate that the criterion was reported, red cell; not reported, and grey cells; not applicable. The percentage of studies meeting each item is shown in parentheses. The final column shows the CLEAR score as a percentage of the total for each study. **Table S6:** Reporting quality of the 78 included studies based on CLEAR checklist. **Figure S7:** Scatter plot showing the relationship between METRICS and CLEAR scores. Each point represents a study. A significant positive correlation is observed (Spearman $r = 0.728$, $p < 0.001$), indicating that studies with higher METRICS score tend to have higher CLEAR score. **Figure S8:** Distribution of METRICS and CLEAR scores across the 78 included studies. Histograms depict score distributions, and Q–Q plots assess normality. Deviations from the reference line in the Q-Q plot indicate non-normality distribution. **Figure S9:** Distribution of methodological and reporting quality scores across the 78 included studies. Boxplots illustrate the distribution of METRICS and CLEAR scores, with individual study scores shown as data points. **Table S7:** Comparison of methodological quality across selected radiomics studies evaluated with two quality scoring tools; RQS and METRICS.

Supplementary Table A.2

Authors contributions

Bassam M. Abunahel: Writing—review & editing, Writing—original draft, Methodology, Formal analysis, Data curation, Data analysis, Investigation, Software, Validation, Visualisation, Conceptualisation. Robba Rai: Writing—review & editing, Methodology, Data curation, Conceptualisation. Yves De Deene: Writing—review & editing, Methodology, Data curation, Conceptualisation. Majid Alizadeh: Writing— review & editing, Methodology, Data curation. Bowen Xin: Writing—review &editing, Methodology, Conceptualisation. William Rae: Writing—review &editing, Methodology, Conceptualisation. Trang Pham: Writing—review &editing, Methodology. Peter Metcalfe: Writing—review &editing, Supervision. Anatoly B. Rosenfeld: Writing—review &editing, Supervision. Susanna Guatelli: Writing— review &editing, Supervision. Lois Holloway: Writing—review &editing, Supervision, Methodology, Data curation, Conceptualisation.

Funding

Bassam M. Abunahel reports financial support from the Australian Government Department of Education, which facilitated the resources for this study. All other authors declare no competing financial interests or personal relationships that could have influenced the results reported in this study. Open Access funding enabled and organised by CAUL and its Member Institutions.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable, as this study was a systematic review, it used only previously published data and did not involve human participants or animal subjects.

Consent for publication

Not applicable.

Competing interests

Not applicable.

Author details

[1]Centre for Medical Radiation Physics, University of Wollongong, Wollongong, NSW 2522, Australia. [2]Ingham Institute for Applied Medical Research, Medical Physics, Liverpool, NSW 2170, Australia. [3]Liverpool and Macarthur Cancer Therapy Centre, Liverpool, NSW 2170, Australia. [4]South West Sydney Clinical Campuses, School of Clinical Medicine, Faculty of Medicine and Health, UNSW, Sydney, NSW 2170, Australia. [5]School of Science, Western Sydney University, Locked Bag, Penrith 1797 NSW, Australia. [6]Department of Human Structure and Repair, Ghent University, Corneel Heymanslaan 10, 9000 Gent, Belgium. [7]Australian e-Health Research Centre, CSIRO, Sydney, NSW 2015, Australia. [8]Medical Imaging Department, Prince of Wales Hospital, Randwick, NSW 2031, Australia.

Received: 6 April 2026 Revised: 6 August 2026 Accepted: 10 September 2026
Published online: 20 November 2026

References

1. Kleeff J, Michalski C, Friess H, Büchler MW (2006) Pancreatic cancer: from bench to 5-year survival. Pancreas 33:111–118. https://doi.org/10.1097/01.mpa.0000229010.62538.f2
2. Tummala P, Junaidi O, Agarwal B (2011) Imaging of pancreatic cancer: an overview. J Gastrointest Oncol 2:168–174. https://doi.org/10.3978/j.issn.2078-6891.2011.036
3. Kalra M, Maher M, Mueller P, Saini S (2003) State-of-the-art imaging of pancreatic neoplasms. Br J Radiol 76:857–865. https://doi.org/10.1259/bjr/16642775
4. Schima W, Ba-Ssalamah A, Goetzinger P, Scharitzer M, Koelblinger C (2007) State-of-the-art magnetic resonance imaging of pancreatic cancer. Top Magn Reson Imaging 18:421–429. https://doi.org/10.1097/rmr.0b013e31816459e0
5. Fusari M, Maurea S, Imbriaco M et al (2010) Comparison between multislice CT and MR imaging in the diagnostic evaluation of patients with pancreatic masses. Radiol Med 115:453–466. https://doi.org/10.1007/s11547-010-0490-7
6. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563–577. https://doi.org/10.1148/radiol.2015151169
7. Kumar V, Gu Y, Basu S et al (2012) Radiomics: the process and the challenges. J Magn Reson Imaging 30:1234–1248. https://doi.org/10.1016/j.mri.2012.06.010
8. Koçak B, Durmaz EŞ, Ateş E, Kılıçkesmez Ö (2019) Radiomics with artificial intelligence: a practical guide for beginners. Diagn Interv Radiol 25:485–495. https://doi.org/10.5152/dir.2019.19321
9. Abunahel BM, Pontre B, Kumar H, Petrov MS (2021) Pancreas image mining: a systematic review of radiomics. Eur Radiol 31:3447–3467. https://doi.org/10.1007/s00330-020-07376-6
10. Marti-Bonmati L, Cerdá-Alberich L, Pérez-Girbés A et al (2022) Pancreatic cancer, radiomics and artificial intelligence. Br J Radiol 95: 20220072. https://doi.org/10.1259/bjr.20220072
11. Park JE, Park SY, Kim HJ, Kim HS (2019) Reproducibility and generalizability in radiomics modeling: possible strategies in radiologic and statistical perspectives. Korean J Radiol 20:1124–1137. https://doi.org/10.3348/kjr.2018.0070
12. Baeßler B, Weiss K, Dos Santos DP (2019) Robustness and reproducibility of radiomics in magnetic resonance imaging: a phantom study. Invest Radiol 54:221–228. https://doi.org/10.1097/RLI.0000000000000530
13. Zhao B (2021) Understanding sources of variation to improve the reproducibility of radiomics. Front Oncol 11: 633176. https://doi.org/10.3389/fonc.2021.633176
14. McCague C, Ramlee S, Reinius M et al (2023) Introduction to radiomics for a clinical audience. Clin Radiol 78:83–98. https://doi.org/10.1016/j.crad.2022.08.149
15. Zwanenburg A, Vallières M, Abdalah MA et al (2020) The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 295:328–338. https://doi.org/10.1148/radiol.2020191145
16. Kocak B, Akinci D'Antonoli T, Mercaldo N et al (2024) METhodological RadiomICs Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMII. Insights Imaging 15:8. https://doi.org/10.1186/s13244-023-01572-w
17. Kocak B, Baessler B, Bakas S et al (2023) CheckList for EvaluAtion of Radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMII. Insights Imaging 14:75. https://doi.org/10.1186/s13244-023-01415-8
18. Huang EP, O'Connor JP, McShane LM et al (2023) Criteria for the translation of radiomics into clinically useful tests. Nat Rev Clin Oncol 20:69–82. https://doi.org/10.1038/s41571-022-00707-0
19. Virostko J, Craddock RC, Williams JM et al (2021) Development of a standardized MRI protocol for pancreas assessment in humans. PLoS One 16:e0256029. https://doi.org/10.1371/journal.pone.0256029
20. Ford J, Dogan N, Young L, Yang F (2018) Quantitative radiomics: impact of pulse sequence parameter selection on MRI-based textural features of the brain. Contrast Media Mol Imaging 2018: 1729071. https://doi.org/10.1155/2018/1729071
21. Abunahel BM, Rai R, De Deen Y et al (2025) Evaluating methodological and reporting quality of studies considering magnetic resonance imaging (MRI)-derived radiomics features in relation to pancreatic tumours: a systematic review using METRICS score and CLEAR checklist. Open Science Framework. https://doi.org/10.17605/OSF.IO/F3VWA
22. Sarkis-Onofre R, Catalá-López F, Aromataris E, Lockwood C (2021) How to properly use the PRISMA Statement. Syst Rev 10: 117. https://doi.org/10.1186/s13643-021-01671-z
23. Collins GS, Reitsma JB, Altman DG, Moons KG (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Br J Surg 102:148–158. https://doi.org/10.1002/bjs.9736
24. Van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B (2020) Radiomics in medical imaging—"how-to" guide and critical reflection. Insights Imaging 11:91. https://doi.org/10.1186/s13244-020-00887-2
25. Brown RW, Cheng Y-CN, Haacke EM, Thompson MR, Venkatesan R (2014) Magnetic resonance imaging: physical principles and sequence design. John Wiley & Sons
26. Lambin P, Rios-Velazquez E, Leijenaar R et al (2012) Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 48:441–446. https://doi.org/10.1016/j.ejca.2011.11.036
27. Mitchell-Hay RN, Ahearn TS, Murray AD, Waiter GD (2022) Investigation of the Inter-and Intrascanner Reproducibility and Repeatability of Radiomics Features in T1-Weighted Brain MRI. J Magn Reson Imaging 56:1559–1568. https://doi.org/10.1002/jmri.28191
28. Buch K, Kuno H, Qureshi MM, Li B, Sakai O (2018) Quantitative variations in texture analysis features dependent on MRI scanning parameters: a phantom model. J Appl Clin Med Phys 19:253–264. https://doi.org/10.1002/acm2.12482
29. Lee J, Steinmann A, Ding Y et al (2021) Radiomics feature robustness as measured using an MRI phantom. Sci Rep 11: 3973. https://doi.org/10.1038/s41598-021-83593-3
30. Xue C, Yuan J, Poon DM et al (2021) Reliability of MRI radiomics features in MR-guided radiotherapy for prostate cancer: Repeatability, reproducibility, and within-subject agreement. Med Phys 48:6976–6986. https://doi.org/10.1002/mp.15232
31. Yuan J, Xue C, Lo G et al (2021) Quantitative assessment of acquisition imaging parameters on MRI radiomics features: a prospective anthropomorphic phantom study using a 3D-T2W-TSE sequence for MR-guided-radiotherapy. Quant Imaging Med Surg 11:1870. https://doi.org/10.21037/qims-20-865
32. Mayerhoefer ME, Szomolanyi P, Jirak D, Materka A, Trattnig S (2009) Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: an application-oriented study. Med Phys 36:1236–1243. https://doi.org/10.1118/1.3081408
33. Higashi M, Tanabe M, Okada M, Furukawa M, Iida E, Ito K (2020) Influence of fat deposition on T1 mapping of the pancreas: evaluation by dual-flip-angle MR imaging with and without fat suppression. Radiol Med 125:1–6. https://doi.org/10.1007/s11547-019-01087-9
34. Delfaut EM, Beltran J, Johnson G, Rousseau J, Marchandise X, Cotten A (1999) Fat suppression in MR imaging: techniques and pitfalls. Radiographics 19:373–382. https://doi.org/10.1148/radiographics.19.2.g99mr03373
35. Sun M, Baiyasi A, Liu X et al (2022) Robustness and reproducibility of radiomics in T2 weighted images from magnetic resonance image guided linear accelerator in a phantom study. Phys Med 96:130–139. https://doi.org/10.1016/j.ejmp.2022.03.002
36. Kojima S, Shinohara H, Hashimoto T, Suzuki S (2018) Undersampling patterns in k-space for compressed sensing MRI using two-dimensional Cartesian sampling. Radiol Phys Technol 11:303–319. https://doi.org/10.1007/s12194-018-0469-y
37. Xue C, Zhou Y, Lo GG et al (2021) Reliability of radiomics features due to image reconstruction using a standardized T2-weighted pulse sequence for MR-guided radiotherapy: An anthropomorphic phantom

study. Magn Reson Med 85:3434–3446. https://doi.org/10.1002/mrm.28650
38. Kim M, Jung SC, Park JE, Park SY, Lee H, Choi KM (2021) Reproducibility of radiomic features in SENSE and compressed SENSE: impact of acceleration factors. Eur Radiol 31:6457–6470. https://doi.org/10.1007/s00330-021-07760-w
39. Hu K, Deng W, Li N et al (2022) Impact of parallel acquisition technology on the robustness of magnetic resonance imaging radiomic features. J Comput Assist Tomogr 46:906–913. https://doi.org/10.1097/RCT.0000000000001344
40. Forster J, Sieverding L, Breuer J, Lutz O, Schick F (2001) Fast acceleration-encoded magnetic resonance imaging. Med Phys 28:28–35. https://doi.org/10.1118/1.1333413
41. Pruessmann KP (2006) Encoding and reconstruction in parallel MRI. NMR Biomed 19:288–299. https://doi.org/10.1002/nbm.1042
42. Moummad I, Jaudet C, Lechervy A et al (2021) The impact of resampling and denoising deep learning algorithms on radiomics in brain metastases MRI. Cancers (Basel) 14:36. https://doi.org/10.3390/cancers14010036
43. Yang F, Dogan N, Stoyanova R, Ford JC (2018) Evaluation of radiomic texture feature error due to MRI acquisition and reconstruction: a simulation study utilizing ground truth. Phys Med 50:26–36. https://doi.org/10.1016/j.ejmp.2018.05.017
44. Grant KB, Agarwal HK, Shih JH et al (2015) Comparison of calculated and acquired high b value diffusion-weighted imaging in prostate cancer. Abdom Imaging 40:578–586. https://doi.org/10.1007/s00261-014-0246-2
45. DelPriore MR, Biswas D, Hippe DS et al (2021) Breast cancer conspicuity on computed versus acquired high b-value diffusion-weighted MRI. Acad Radiol 28:1108–1117. https://doi.org/10.1016/j.acra.2020.03.011
46. He Y, Rong Y, Chen H et al (2020) Impact of different b-value combinations on radiomics features of apparent diffusion coefficient in cervical cancer. Acta Radiol 61:568–576. https://doi.org/10.1177/0284185119870157
47. Zhang J, Qiu Q, Duan J et al (2019) Variability of radiomic features extracted from multi-b-value diffusion-weighted images in hepatocellular carcinoma. Transl Cancer Res 8:130–140. https://doi.org/10.21037/tcr.2019.01.14
48. Zhou IY, Gao DS, Chow AM et al (2014) Effect of diffusion time on liver DWI: An experimental study of normal and fibrotic livers. Magn Reson Med 72:1389–1396. https://doi.org/10.1002/mrm.25035
49. Reynaud O (2017) Time-Dependent Diffusion MRI in Cancer: Tissue Modeling and Applications. Front Phys 5: 58. https://doi.org/10.3389/fphy.2017.00058
50. Kiselev VG (2017) Fundamentals of diffusion MRI physics. NMR Biomed 30: e3602. https://doi.org/10.1002/nbm.3602
51. Huang SY, Nummenmaa A, Witzel T et al (2015) The impact of gradient strength on *in vivo* diffusion MRI estimates of axon diameter. Neuroimage 106:464–472. https://doi.org/10.1016/j.neuroimage.2014.12.008
52. Kennis M, van Rooij SJH, Kahn RS, Geuze E, Leemans A (2016) Choosing the polarity of the phase-encoding direction in diffusion MRI: Does it matter for group analysis? Neuroimage Clin 11:539–547. https://doi.org/10.1016/j.nicl.2016.03.022
53. Paulsen JL, Özarslan E, Komlosh ME, Basser PJ, Song Y-Q (2015) Detecting compartmental non-Gaussian diffusion with symmetrized double-PFG MRI. NMR Biomed 28:1550–1556. https://doi.org/10.1002/nbm.3363
54. Gardier R, Savoy A, Haro JLV et al (2024) Comparing steam and PGSE diffusion MRI signal of RAT lymph nodes using in-silico simulations. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp 1–5. https://doi.org/10.1109/ISBI56570.2024.10635470
55. Hasan KM, Parker DL, Alexander AL (2001) Comparison of gradient encoding schemes for diffusion-tensor MRI. J Magn Reson Imaging 13:769–780. https://doi.org/10.1002/jmri.1107
56. Schilling K, Nath V, Blaber J et al (2017) Effects of b-value and number of gradient directions on diffusion MRI measures obtained with Q-ball imaging. SPIE Med Imag 10133 https://doi.org/10.1117/12.2254545
57. Gunbey Karabekmez L (2022) The role of multidirectional diffusion weighted imaging in the diagnosis of breast carcinoma in magnetic resonance imaging: DW-MRI in breast carcinoma diagnosis. Arch Breast Cancer 9:488–496. https://doi.org/10.32768/abc.202294488-496
58. De Santis S, Gabrielli A, Palombo M, Maraviglia B, Capuani S (2011) Non-Gaussian diffusion imaging: a brief practical review. J Magn Reson Imaging 29:1410–1416. https://doi.org/10.1016/j.mri.2011.04.006
59. Wochner P, Schneider T, Stockmann J, Lee J, Sinkus R (2023) Diffusion phase-imaging in anisotropic media using non-linear gradients for diffusion encoding. PLoS One 18:e0281332. https://doi.org/10.1371/journal.pone.0281332
60. Ma C, Li Y, Wang L et al (2017) Intravoxel incoherent motion DWI of the pancreatic adenocarcinomas: monoexponential and biexponential apparent diffusion parameters and histopathological correlations. Cancer Imaging 17: 12. https://doi.org/10.1186/s40644-017-0114-8
61. Ogura A, Hatano I, Osakabe K, Yamaguchi N, Koyama D, Watanabe H (2016) Importance of fractional b value for calculating apparent diffusion coefficient in DWI. AJR Am J Roentgenol 207:1239–1243. https://doi.org/10.2214/AJR.15.15945
62. Kartalis N, Manikis GC, Loizou L et al (2016) Diffusion-weighted MR imaging of pancreatic cancer: a comparison of mono-exponential, bi-exponential and non-Gaussian kurtosis models. Eur J Radiol Open 3:79–85. https://doi.org/10.1016/j.ejro.2016.04.002
63. Mesri HY, David S, Viergever MA, Leemans A (2020) The adverse effect of gradient nonlinearities on diffusion MRI: from voxels to group studies. Neuroimage 205: 116127. https://doi.org/10.1016/j.neuroimage.2019.116127
64. Bruce IP, Petty C, Song AW (2018) Simultaneous and inherent correction of B0 and eddy-current induced distortions in high-resolution diffusion MRI using reversed polarity gradients and multiplexed sensitivity encoding (RPG-MUSE). Neuroimage 183:985–993. https://doi.org/10.1016/j.neuroimage.2018.09.055
65. Watzenboeck ML, Beer L, Kifjak D et al (2024) Contrast agent dynamics determine radiomics profiles in oncologic imaging. Cancers (Basel) 16:1519. https://doi.org/10.3390/cancers16081519
66. McHugh DJ, Porta N, Little RA et al (2021) Image contrast, image preprocessing, and T1 mapping affect mri radiomic feature repeatability in patients with colorectal cancer liver metastases. Cancers (Basel) 13:240. https://doi.org/10.3390/cancers13020240
67. Huang W, Chen Y, Fedorov A et al (2016) The impact of arterial input function determination variations on prostate dynamic contrast-enhanced magnetic resonance imaging pharmacokinetic modeling: a multicenter data analysis challenge. Tomography 2:56–66. https://doi.org/10.18383/j.tom.2015.00184
68. Huang W, Chen Y, Fedorov A et al (2019) The impact of arterial input function determination variations on prostate dynamic contrast-enhanced magnetic resonance imaging pharmacokinetic modeling: a multicenter data analysis challenge, part II. Tomography 5:99–109. https://doi.org/10.18383/j.tom.2018.00027
69. Fluckiger JU, Schabel MC, DiBella EV (2010) Toward local arterial input functions in dynamic contrast-enhanced MRI. J Magn Reson Imaging 32:924–934. https://doi.org/10.1002/jmri.22339
70. Zhao W, Liu C, Huan Y et al (2024) Reproducibility and reliability of pancreatic pharmacokinetic parameters derived from dynamic contrast-enhanced magnetic resonance imaging. Acta Radiol 65:681–688. https://doi.org/10.1177/02841851241246364
71. Harrington KA, Shukla-Dave A, Paudyal R, Do RK (2021) MRI of the Pancreas. J Magn Reson Imaging 53:347–359. https://doi.org/10.1002/jmri.27148
72. Wichtmann BD, Harder FN, Weiss K et al (2023) Influence of image processing on radiomic features from magnetic resonance imaging. Invest Radiol 58:199–208. https://doi.org/10.1097/RLI.0000000000000921
73. Fields BK, Demirjian NL, Hwang DH et al (2021) Whole-tumor 3D volumetric MRI-based radiomics approach for distinguishing between benign and malignant soft tissue tumors. Eur Radiol 31:8522–8535. https://doi.org/10.1007/s00330-021-07914-w
74. Huang Y, Zhou S, Luo Y et al (2023) Development and validation of a radiomics model of magnetic resonance for predicting liver metastasis in resectable pancreatic ductal adenocarcinoma patients. Radiat Oncol 18: 79. https://doi.org/10.1186/s13014-023-02273-w
75. Michalet M, Valenzuela G, Nougaret S, Tardieu M, Azria D, Riou O (2025) Development of Multiparametric Prognostic Models for Stereotactic Magnetic Resonance Guided Radiation Therapy of Pancreatic Cancers.

Int J Radiat Oncol Biol Phys 122:678–689. https://doi.org/10.1016/j.ijrobp.2025.03.039
76. Lu W, Wu G, Miao X et al (2024) The radiomics nomogram predicts the prognosis of pancreatic cancer patients with hepatic metastasis after chemoimmunotherapy. Cancer Immunol Immunother 73:87. https://doi.org/10.1007/s00262-024-03644-2
77. Shao C, Zhang J, Guo J et al (2023) A radiomics nomogram model for predicting prognosis of pancreatic ductal adenocarcinoma after high-intensity focused ultrasound surgery. Int J Hyperth 40: 2184397. https://doi.org/10.1080/02656736.2023.2184397
78. Xu X, Qu J, Zhang Y, Qian X, Chen T, Liu Y (2023) Development and validation of an MRI-radiomics nomogram for the prognosis of pancreatic ductal adenocarcinoma. Front Oncol 13: 1074445. https://doi.org/10.3389/fonc.2023.1074445
79. Zhang X, Gao A, Ma L, Yu N (2024) Integrating intratumoral and peritumoral radiomics with clinical risk factors for prognostic prediction in pancreatic ductal adenocarcinoma patients undergoing combined chemotherapy and HIFU ablation. Int J Hyperth 41: 2410342. https://doi.org/10.1080/02656736.2024.2410342
80. Choi MH, Lee YJ, Yoon SB, Choi J-I, Jung SE, Rha SE (2019) MRI of pancreatic ductal adenocarcinoma: texture analysis of T2-weighted images for predicting long-term outcome: MH Choi et al. Abdom Radiol (NY) 44:122–130. https://doi.org/10.1007/s00261-018-1681-2
81. Bian Y, Liu C, Li Q et al (2022) Preoperative radiomics approach to evaluating tumor-infiltrating CD8+ T cells in patients with pancreatic ductal adenocarcinoma using noncontrast magnetic resonance imaging. J Magn Reson Imaging 55:803–814. https://doi.org/10.1002/jmri.27871
82. Kaissis G, Ziegelmayer S, Lohöfer F et al (2019) A machine learning model for the prediction of survival and tumor subtype in pancreatic ductal adenocarcinoma from preoperative diffusion-weighted imaging. Eur Radiol Exp 3:41. https://doi.org/10.1186/s41747-019-0119-0
83. Qiu H, Xu M, Wang Y et al (2022) A novel preoperative MRI-based radiomics nomogram outperforms traditional models for prognostic prediction in pancreatic ductal adenocarcinoma. Am J Cancer Res 12:2032–2049
84. Cusumano D, Boldrini L, Yadav P et al (2021) Delta radiomics analysis for local control prediction in pancreatic cancer patients treated using magnetic resonance guided radiotherapy. Diagnostics (Basel) 11:72. https://doi.org/10.3390/diagnostics11010072
85. Mapelli P, Bezzi C, Palumbo D et al (2022) 68Ga-DOTATOC PET/MR imaging and radiomic parameters in predicting histopathological prognostic factors in patients with pancreatic neuroendocrine well-differentiated tumours. Eur J Nucl Med Mol Imaging 49:2352–2363. https://doi.org/10.1007/s00259-022-05677-0
86. Liang L, Luo R, Ding Y et al (2021) S100A4 overexpression in pancreatic ductal adenocarcinoma: imaging biomarkers from whole-tumor evaluation with MRI and texture analysis. Abdom Radiol (NY) 46:623–635. https://doi.org/10.1007/s00261-020-02676-3
87. Deng Y, Li Y, Wu J-L et al (2022) Radiomics models based on multi-sequence MRI for preoperative evaluation of MUC4 status in pancreatic ductal adenocarcinoma: a preliminary study. Quant Imaging Med Surg 12:5129. https://doi.org/10.21037/qims-22-112
88. Meng Y, Zhang H, Li Q et al (2021) Noncontrast magnetic resonance radiomics and multilayer perceptron network classifier: An approach for predicting fibroblast activation protein expression in patients with pancreatic ductal adenocarcinoma. J Magn Reson Imaging 54:1432–1443. https://doi.org/10.1002/jmri.27648
89. Li Q, Yu J, Zhang H et al (2022) Prediction of tumor-infiltrating CD20+ B-cells in patients with pancreatic ductal adenocarcinoma using a multilayer perceptron network classifier based on non-contrast MRI. Acad Radiol 29:e167–e177. https://doi.org/10.1016/j.acra.2021.11.013
90. Bian Y, Liu YF, Jiang H et al (2021) Machine learning for MRI radiomics: a study predicting tumor-infiltrating lymphocytes in patients with pancreatic ductal adenocarcinoma. Abdom Radiol (NY) 46:4800–4816. https://doi.org/10.1007/s00261-021-03159-9
91. Meng Y, Zhang H, Li Q et al (2022) Magnetic resonance radiomics and machine-learning models: an approach for evaluating tumor-stroma ratio in patients with pancreatic ductal adenocarcinoma. Acad Radiol 29:523–535. https://doi.org/10.1016/j.acra.2021.08.013
92. Gao J, Chen X, Li X et al (2021) Differentiating TP53 mutation status in pancreatic ductal adenocarcinoma using multiparametric MRI-derived radiomics. Front Oncol 11: 632130. https://doi.org/10.3389/fonc.2021.632130
93. Choi MH, Yoon SB, Lee YJ et al (2024) Rim enhancement of pancreatic ductal adenocarcinoma: investigating the relationship with DCE-MRI-based radiomics and next-generation sequencing. Front Oncol 14: 1304187. https://doi.org/10.3389/fonc.2024.1304187
94. Yuan Z, Shu Z, Peng J et al (2024) Prediction of postoperative liver metastasis in pancreatic ductal adenocarcinoma based on multiparametric magnetic resonance radiomics combined with serological markers: a cohort study of machine learning. Abdom Radiol (NY) 49:117–130. https://doi.org/10.1007/s00261-023-04047-0
95. Xie N, Fan X, Chen D et al (2023) Peritumoral and intratumoral texture features based on multiparametric MRI and multiple machine learning methods to preoperatively evaluate the pathological outcomes of pancreatic cancer. J Magn Reson Imaging 58:379–391. https://doi.org/10.1002/jmri.28538
96. Shi L, Wang L, Wu C, Wei Y, Zhang Y, Chen J (2022) Preoperative prediction of lymph node metastasis of pancreatic ductal adenocarcinoma based on a radiomics nomogram of dual-parametric MRI imaging. Front Oncol 12: 927077. https://doi.org/10.3389/fonc.2022.927077
97. Shi Z, Ma C, Huang X, Cao D (2022) Magnetic resonance imaging radiomics-based nomogram from primary tumor for pretreatment prediction of peripancreatic lymph node metastasis in pancreatic ductal adenocarcinoma: a multicenter study. J Magn Reson Imaging 55:823–839. https://doi.org/10.1002/jmri.28048
98. Zeng P, Qu C, Liu J et al (2023) Comparison of MRI and CT-based radiomics for preoperative prediction of lymph node metastasis in pancreatic ductal adenocarcinoma. Acta Radiol 64:2221–2228. https://doi.org/10.1177/02841851221142552
99. Shen J, Li Q, Li L et al (2025) Contrast-enhanced MRI-based intratumoral heterogeneity assessment for predicting lymph node metastasis in resectable pancreatic ductal adenocarcinoma. Insights Imaging 16:76. https://doi.org/10.1186/s13244-025-01956-0
100. Gao J, Huang X, Meng H et al (2020) Performance of multiparametric functional imaging and texture analysis in predicting synchronous metastatic disease in pancreatic ductal adenocarcinoma patients by hybrid PET/MR: initial experience. Front Oncol 10:198. https://doi.org/10.3389/fonc.2020.00198
101. Gao J, Bai Y, Miao F et al (2023) Prediction of synchronous distant metastasis of primary pancreatic ductal adenocarcinoma using the radiomics features derived from 18F-FDG PET and MRI. Clin Radiol 78:746–754. https://doi.org/10.1016/j.crad.2023.06.011
102. De Robertis R, Tomaiuolo L, Pasquazzo F et al (2022) Correlation between ADC histogram-derived metrics and the time to metastases in resectable pancreatic adenocarcinoma. Cancers (Basel) 14:6050. https://doi.org/10.3390/cancers14246050
103. Liang L, Ding Y, Yu Y et al (2021) Whole-tumour evaluation with MRI and radiomics features to predict the efficacy of S-1 for adjuvant chemotherapy in postoperative pancreatic cancer patients: a pilot study. BMC Med Imaging 21: 75. https://doi.org/10.1186/s12880-021-00605-4
104. Simpson G, Jin W, Spieler B et al (2022) Predictive value of delta-radiomics texture features in 0.35 Tesla magnetic resonance setup images acquired during stereotactic ablative radiotherapy of pancreatic cancer. Front Oncol 12: 807725. https://doi.org/10.3389/fonc.2022.807725
105. Tomaszewski M, Latifi K, Boyer E et al (2021) Delta radiomics analysis of magnetic resonance guided radiotherapy imaging data can enable treatment response prediction in pancreatic cancer. Radiat Oncol 16: 237. https://doi.org/10.1186/s13014-021-01957-5
106. Simpson G, Spieler B, Dogan N et al (2020) Predictive value of 0.35 T magnetic resonance imaging radiomic features in stereotactic ablative body radiotherapy of pancreatic cancer: A pilot study. Med Phys 47:3682–3690. https://doi.org/10.1002/mp.14200
107. Skawran SM, Kambakamba P, Baessler B et al (2021) Can magnetic resonance imaging radiomics of the pancreas predict postoperative pancreatic fistula? Eur J Radiol 140: 109733. https://doi.org/10.1016/j.ejrad.2021.109733
108. Flammia F, Innocenti T, Galluzzo A et al (2023) Branch duct-intraductal papillary mucinous neoplasms (BD-IPMNs): An MRI-based radiomic

model to determine the malignant degeneration potential. Radiol Med 128:383–392. https://doi.org/10.1007/s11547-023-01609-6
109. Cheng S, Shi H, Lu M et al (2022) Radiomics analysis for predicting malignant potential of intraductal papillary mucinous neoplasms of the pancreas: comparison of CT and MRI. Acad Radiol 29:367–375. https://doi.org/10.1016/j.acra.2021.04.013
110. Jeon SK, Kim JH, Yoo J, Kim J-E, Park SJ, Han JK (2021) Assessment of malignant potential in intraductal papillary mucinous neoplasms of the pancreas using MR findings and texture analysis. Eur Radiol 31:3394–3404. https://doi.org/10.1007/s00330-020-07425-0
111. Hoffman DH, Ream JM, Hajdu CH, Rosenkrantz AB (2017) Utility of whole-lesion ADC histogram metrics for assessing the malignant potential of pancreatic intraductal papillary mucinous neoplasms (IPMNs). Abdom Radiol (NY) 42:1222–1228. https://doi.org/10.1007/s00261-016-1001-7
112. Liang X, He W, Huang C et al (2022) Preoperative prediction of invasive behavior of pancreatic solid pseudopapillary neoplasm by MRI-based multiparametric radiomics models. Abdom Radiol (NY) 47:3782–3791. https://doi.org/10.1007/s00261-022-03639-6
113. Xie N, Fan X, Xie H et al (2023) Preoperative extrapancreatic extension prediction in patients with pancreatic cancer using multiparameter MRI and machine learning-based radiomics model. Acad Radiol 30:1306–1316. https://doi.org/10.1016/j.acra.2022.09.017
114. Tang TY, Li X, Zhang Q et al (2020) Development of a novel multiparametric MRI radiomic nomogram for preoperative evaluation of early recurrence in resectable pancreatic cancer. J Magn Reson Imaging 52:231–245. https://doi.org/10.1002/jmri.27024
115. Nasief HG, Parchur AK, Omari E et al (2022) Predicting necessity of daily online adaptive replanning based on wavelet image features for MRI guided adaptive radiation therapy. Radiother Oncol 176:165–171. https://doi.org/10.1016/j.radonc.2022.10.001
116. D'Onofrio M, Tedesco G, Cardobi N et al (2021) Magnetic resonance (MR) for mural nodule detection studying Intraductal papillary mucinous neoplasms (IPMN) of pancreas: Imaging-pathologic correlation. Pancreatology 21:180–187. https://doi.org/10.1016/j.pan.2020.11.024
117. Liu J, Hu L, Zhou B, Wu C, Cheng Y (2022) 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 18: 101357. https://doi.org/10.1016/j.tranon.2022.101357
118. Deng Y, Ming B, Zhou T et al (2021) Radiomics model based on MR images to discriminate pancreatic ductal adenocarcinoma and mass-forming chronic pancreatitis lesions. Front Oncol 11: 620981. https://doi.org/10.3389/fonc.2021.620981
119. Song T, Zhang Q-W, Duan S-F et al (2021) MRI-based radiomics approach for differentiation of hypovascular non-functional pancreatic neuroendocrine tumors and solid pseudopapillary neoplasms of the pancreas. BMC Med Imaging 21: 36. https://doi.org/10.1186/s12880-021-00563-x
120. Shi Y-J, Zhu H-T, Liu Y-L et al (2020) Radiomics analysis based on diffusion kurtosis imaging and T2 weighted imaging for differentiation of pancreatic neuroendocrine tumors from solid pseudopapillary tumors. Front Oncol 10:1624. https://doi.org/10.3389/fonc.2020.01624
121. Li X, Zhu H, Qian X, Chen N, Lin X (2020) MRI texture analysis for differentiating nonfunctional pancreatic neuroendocrine neoplasms from solid pseudopapillary neoplasms of the pancreas. Acad Radiol 27:815–823. https://doi.org/10.1016/j.acra.2019.07.012
122. Shen K, Su W, Liang C, Shi D, Sun J, Yu R (2024) Differentiating small (< 2 cm) pancreatic ductal adenocarcinoma from neuroendocrine tumors with multiparametric MRI-based radiomic features. Eur Radiol 34:7553–7563. https://doi.org/10.1007/s00330-024-10837-x
123. Li Q, Zhou Z, Chen Y et al (2023) Fully automated magnetic resonance imaging-based radiomics analysis for differentiating pancreatic adenosquamous carcinoma from pancreatic ductal adenocarcinoma. Abdom Radiol (NY) 48:2074–2084. https://doi.org/10.1007/s00261-023-03801-8
124. Shiraishi M, Igarashi T, Hiroaki F, Oe R, Ohki K, Ojiri H (2022) Radiomics based on diffusion-weighted imaging for differentiation between focal-type autoimmune pancreatitis and pancreatic carcinoma. Br J Radiol 95: 20210456. https://doi.org/10.1259/bjr.20210456
125. Li Q, Li X, Liu W et al (2023) Non-enhanced magnetic resonance imaging-based radiomics model for the differentiation of pancreatic adenosquamous carcinoma from pancreatic ductal adenocarcinoma. Front Oncol 13: 1108545. https://doi.org/10.3389/fonc.2023.1108545
126. Ansari G, Mirza-Aghazadeh-Attari M, Afyouni S, Mohseni A, Shahbazian H, Kamel IR (2024) Utilization of texture features of volumetric ADC maps in differentiating between serous cystadenoma and intraductal papillary neoplasms. Abdom Radiol (NY) 49:1175–1184. https://doi.org/10.1007/s00261-024-04187-x
127. Fang X, Zhang Q, Liu F et al (2023) T2-Weighted image radiomics nomogram to predict pancreatic serous and mucinous cystic neoplasms. Acad Radiol 30:1562–1571. https://doi.org/10.1016/j.acra.2022.10.008
128. Wang Y, Zhang X, Wang B et al (2019) Value of texture analysis of intravoxel incoherent motion parameters in differential diagnosis of pancreatic neuroendocrine tumor and pancreatic adenocarcinoma. Chin Med Sci J 34:1–9. https://doi.org/10.24920/003531
129. Shindo T, Fukukura Y, Umanodan T et al (2016) Histogram analysis of apparent diffusion coefficient in differentiating pancreatic adenocarcinoma and neuroendocrine tumor. Medicine (Baltimore) 95:e2574. https://doi.org/10.1097/MD.0000000000002574
130. Sarac D, Badza Atanasijevic M, Mitrovic Jovanovic M et al (2025) Applicability of radiomics for differentiation of pancreatic adenocarcinoma from healthy tissue of pancreas by using magnetic resonance imaging and machine learning. Cancers (Basel) 17:1119. https://doi.org/10.3390/cancers17071119
131. Taffel MT, Luk L, Ream JM, Rosenkrantz AB (2019) Exploratory study of apparent diffusion coefficient histogram metrics in assessing pancreatic malignancy. Can Assoc Radiol J 70:416–423. https://doi.org/10.1016/j.carj.2019.07.001
132. Cheng S-H, Liu D, Hou B et al (2020) PET-MR imaging and mr texture analysis in the diagnosis of pancreatic cysts: a prospective preliminary study. Acad Radiol 27:996–1005. https://doi.org/10.1016/j.acra.2019.09.001
133. Malagi AV, Shivaji S, Kandasamy D et al (2023) Pancreatic mass characterization using IVIM-DKI MRI and machine learning-based multiparametric texture analysis. Bioengineering 10:83. https://doi.org/10.3390/bioengineering10010083
134. Zhu H-B, Zhu H-T, Jiang L et al (2024) Radiomics analysis from magnetic resonance imaging in predicting the grade of nonfunctioning pancreatic neuroendocrine tumors: a multicenter study. Eur Radiol 34:90–102. https://doi.org/10.1007/s00330-023-09957-7
135. Bian Y, Zhao Z, Jiang H et al (2020) Noncontrast radiomics approach for predicting grades of nonfunctional pancreatic neuroendocrine tumors. J Magn Reson Imaging 52:1124–1136. https://doi.org/10.1002/jmri.27176
136. Bian Y, Li J, Cao K et al (2021) Magnetic resonance imaging radiomic analysis can preoperatively predict G1 and G2/3 grades in patients with NF-pNETs. Abdom Radiol (NY) 46:667–680. https://doi.org/10.1007/s00261-020-02706-0
137. Liu C, Bian Y, Meng Y et al (2022) Preoperative prediction of G1 and G2/3 grades in patients with nonfunctional pancreatic neuroendocrine tumors using multimodality imaging. Acad Radiol 29:e49–e60. https://doi.org/10.1016/j.acra.2021.05.017
138. Kulali F, Semiz-Oysu A, Demir M, Segmen-Yilmaz M, Bukte Y (2018) Role of diffusion-weighted MR imaging in predicting the grade of nonfunctional pancreatic neuroendocrine tumors. Diagn Interv Imaging 99:301–309. https://doi.org/10.1016/j.diii.2017.10.012
139. Guo C-g, Ren S, Chen X et al (2019) Pancreatic neuroendocrine tumor: prediction of the tumor grade using magnetic resonance imaging findings and texture analysis with 3-T magnetic resonance. Cancer Manag Res 10:1933–1944. https://doi.org/10.2147/CMAR.S195376
140. Ohki K, Igarashi T, Ashida H et al (2021) Usefulness of texture analysis for grading pancreatic neuroendocrine tumors on contrast-enhanced computed tomography and apparent diffusion coefficient maps. Jpn J Radiol 39:66–75. https://doi.org/10.1007/s11604-020-01038-9
141. Li W, Xu C, Ye Z (2021) Prediction of pancreatic neuroendocrine tumor grading risk based on quantitative radiomic analysis of MR. Front Oncol 11: 758062. https://doi.org/10.3389/fonc.2021.758062

142. Pereira JAS, Rosado E, Bali M, Metens T, Chao S-L (2015) Pancreatic neuroendocrine tumors: correlation between histogram analysis of apparent diffusion coefficient maps and tumor grade. Abdom Imaging 40:3122–3128. https://doi.org/10.1007/s00261-015-0524-7
143. De Robertis R, Maris B, Cardobi N et al (2018) Can histogram analysis of MR images predict aggressiveness in pancreatic neuroendocrine tumors? Eur Radiol 28:2582–2591. https://doi.org/10.1007/s00330-017-5236-7
144. Yao L, Zhang Z, Demir U et al (2023) Radiomics boosts deep learning model for IPMN classification. Mach Learn Med Imaging 14:134–143. https://doi.org/10.1007/978-3-031-45676-3_14
145. Cui S, Tang T, Su Q et al (2021) Radiomic nomogram based on MRI to predict grade of branching type intraductal papillary mucinous neoplasms of the pancreas: a multicenter study. Cancer Imaging 21: 26. https://doi.org/10.1186/s40644-021-00395-6
146. Ștefan PA, Lupean RA, Lebovici A et al (2022) Quantitative MRI of Pancreatic Cystic Lesions: a new diagnostic approach. Healthcare 10:1039. https://doi.org/10.3390/healthcare10061039
147. Gu H, Liang H, Zhong J, Wei Y, Ma Y (2022) How does the pancreatic solid pseudopapillary neoplasm confuse us: Analyzing from the point view of MRI-based radiomics? J Magn Reson Imaging 85:38–43. https://doi.org/10.1016/j.mri.2021.10.034
148. Kaissis G, Ziegelmayer S, Lohöfer F et al (2019) A machine learning algorithm predicts molecular subtypes in pancreatic ductal adenocarcinoma with differential response to gemcitabine-based versus FOLFIRINOX chemotherapy. PLoS One 14:e0218642. https://doi.org/10.1371/journal.pone.0218642
149. Dou Z, Lin J, Lu C et al (2025) The development of a multimodal prediction model based on CT and MRI for the prognosis of pancreatic cancer. BMC Gastroenterol 25:557. https://doi.org/10.1186/s12876-025-04119-z
150. Shi Y-J, Zhang H, Wang L-L et al (2025) Deep learning automatic segmentation and radiomics model for diagnosing pancreatic solid neoplasms in MRI. BMC Cancer 25: 1563. https://doi.org/10.1186/s12885-025-15021-y
151. Yan P, Tong K, Liu Z et al (2025) Machine learning model based on preoperative MRI and clinical data for predicting pancreatic fistula after pancreaticoduodenectomy. BMC Med Imaging 25: 410. https://doi.org/10.1186/s12880-025-01951-3
152. Beleù A, Canonico D, Morana G (2024) T1 and T2-mapping in pancreatic MRI: current evidence and future perspectives. Eur J Radiol Open 12: 100572. https://doi.org/10.1016/j.ejro.2024.100572
153. Trojani V, Bassi MC, Verzellesi L, Bertolini M (2024) Impact of preprocessing parameters in medical imaging-based radiomic studies: a systematic review. Cancers (Basel) 16:2668. https://doi.org/10.3390/cancers16152668
154. Abunahel BM, Pontre B, Ko J, Petrov MS (2022) Towards developing a robust radiomics signature in diffuse diseases of the pancreas: accuracy and stability of features derived from T1-weighted magnetic resonance imaging. J Med Radiat Sci 53:420–428. https://doi.org/10.1016/j.jmir.2022.04.002
155. Abunahel BM, Pontre B, Petrov MS (2022) Effect of gray value discretization and image filtration on texture features of the pancreas derived from magnetic resonance imaging at 3T. J Imaging 8:220. https://doi.org/10.3390/jimaging8080220
156. Kocak B, Yuzkan S, Mutlu S, Bulut E, Kavukoglu I (2023) Publications poorly report the essential RadiOmics ParametERs (PROPER): a meta-research on quality of reporting. Eur J Radiol 167: 111088. https://doi.org/10.1016/j.ejrad.2023.111088
157. Brown R, Lim K, Walker A et al (2025) Impact of number of observers on reproducibility of gynaecological cancer MRI radiomics to interobserver contour variation. Med Phys 52: e70085. https://doi.org/10.1002/mp.70085
158. Kirchgesner T, Perlepe V, Michoux N, Larbi A, Vande Berg B (2017) Fat suppression at 2D MR imaging of the hands: Dixon method versus CHESS technique and STIR sequence. Eur J Radiol 89:40–46. https://doi.org/10.1016/j.ejrad.2017.01.011
159. Joohyun L, Ho N (2023) A comparitive study of fat suppression image qualities according to shimming techniques in off-center MRI. J Med Radiat Sci 54:S7–S8. https://doi.org/10.1016/j.jmir.2023.06.027
160. Rai R, Holloway LC, Brink C et al (2020) Multicenter evaluation of MRI-based radiomic features: a phantom study. Med Phys 47:3054–3063. https://doi.org/10.1002/mp.14173
161. Stabinska J, Thiel TA, Zöllner HJ, Benkert T, Wittsack HJ, Ljimani A (2025) Investigation of diffusion time dependence of apparent diffusion coefficient and intravoxel incoherent motion parameters in the human kidney. Magn Reson Med 93:2020–2028. https://doi.org/10.1002/mrm.30396
162. Bourne R, Liang S, Panagiotaki E, Bongers A, Sved P, Watson G (2017) Measurement and modeling of diffusion time dependence of apparent diffusion coefficient and fractional anisotropy in prostate tissue ex vivo. NMR Biomed 30: e3751. https://doi.org/10.1002/nbm.3751
163. Ba R, Wang X, Zhang Z et al (2023) Diffusion-time dependent diffusion MRI: effect of diffusion-time on microstructural mapping and prediction of prognostic features in breast cancer. Eur Radiol 33:6226–6237. https://doi.org/10.1007/s00330-023-09623-y
164. Kiemen AL, Braxton AM, Grahn MP et al (2022) CODA: quantitative 3D reconstruction of large tissues at cellular resolution. Nat Methods 19:1490–1499. https://doi.org/10.1038/s41592-022-01650-9
165. Nissan N (2017) Modifications of pancreatic diffusion MRI by tissue characteristics: what are we weighting for? NMR Biomed 30: e3728. https://doi.org/10.1002/nbm.3728
166. Klaassen R, Gurney-Champion OJ, Engelbrecht MR et al (2018) Evaluation of six diffusion-weighted MRI models for assessing effects of neoadjuvant chemoradiation in pancreatic cancer patients. Int J Radiat Oncol Biol Phys 102:1052–1062. https://doi.org/10.1016/j.ijrobp.2018.04.064
167. Drobnjak I, Siow B, Alexander DC (2010) Optimizing gradient waveforms for microstructure sensitivity in diffusion-weighted MR. J Magn Reson 206:41–51. https://doi.org/10.1016/j.jmr.2010.05.017
168. Lebel C, Benner T, Beaulieu C (2012) Six is enough? Comparison of diffusion parameters measured using six or more diffusion-encoding gradient directions with deterministic tractography. Magn Reson Med 68:474–483. https://doi.org/10.1002/mrm.23254
169. Lee Y, Kettinger AO, Wilm BJ et al (2020) A comprehensive approach for correcting voxel-wise b-value errors in diffusion MRI. Magn Reson Med 83:2173–2184. https://doi.org/10.1002/mrm.28078
170. Hasler SW, Bernchou U, Behrens CP et al (2024) Impact of geometric correction on echo-planar imaging-based apparent diffusion coefficient maps for abdominal radiotherapy. Biomed Phys Eng Express 10:065010. https://doi.org/10.1088/2057-1976/ad7597
171. Manson EN, Inkoom S, Mumuni AN (2022) Impact of magnetic field inhomogeneity on the quality of magnetic resonance images and compensation techniques: a review. Rep Med Imaging 15:43–56. https://doi.org/10.2147/RMI.S369491
172. Keil VC, Mädler B, Gieseke J et al (2017) Effects of arterial input function selection on kinetic parameters in brain dynamic contrast-enhanced MRI. J Magn Reson Imaging 40:83–90. https://doi.org/10.1016/j.mri.2017.04.006
173. Debus C, Floca R, Nörenberg D, Abdollahi A, Ingrisch M (2017) Impact of fitting algorithms on errors of parameter estimates in dynamic contrast-enhanced MRI. Phys Med Biol 62:9322–9340. https://doi.org/10.1088/1361-6560/aa8989
174. Ardakani AA, Bureau NJ, Ciaccio EJ, Acharya UR (2022) Interpretation of radiomics features–a pictorial review. Comput Methods Programs Biomed 215: 106609. https://doi.org/10.1016/j.cmpb.2021.106609
175. Benkarim O, Paquola C, Park B-y et al (2022) Population heterogeneity in clinical cohorts affects the predictive accuracy of brain imaging. PLoS Biol 20:e3001627. https://doi.org/10.1371/journal.pbio.3001627
176. Hajianfar G, Hosseini SA, Bagherieh S, Oveisi M, Shiri I, Zaidi H (2024) Impact of harmonization on the reproducibility of MRI radiomic features when using different scanners, acquisition parameters, and image preprocessing techniques: a phantom study. Med Biol Eng Comput 62:2319–2332. https://doi.org/10.1007/s11517-024-03071-6
177. Yang J, Soltan AA, Eyre DW, Yang Y, Clifton DA (2023) An adversarial training framework for mitigating algorithmic biases in clinical machine learning. NPJ Digit Med 6:55. https://doi.org/10.1038/s41746-023-00805-y
178. Austin PC (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 46:399–424. https://doi.org/10.1080/00273171.2011.568786

179. Ramspek CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M (2021) External validation of prognostic models: what, why, how, when and where? Clin Kidney J 14:49–58. https://doi.org/10.1093/ckj/sfaa188
180. Anees A, Field M, Holloway L (2024) A neural network-based vertical federated learning framework with server integration. Eng Appl Artif Intell 138: 109276. https://doi.org/10.1016/j.engappai.2024.109276
181. Linardatos P, Papastefanopoulos V, Kotsiantis S (2020) Explainable AI: a review of machine learning interpretability methods. Entropy 23:18. https://doi.org/10.3390/e23010018
182. Kocak B, Ammirabile A, Ambrosini I et al (2025) Explanation and elaboration with examples for METRICS (METRICS-E3): an initiative from the EuSoMII Radiomics Auditing Group. Insights Imaging 16:175. https://doi.org/10.1186/s13244-025-02061-y
183. Kocak B, dos Santos DP, Dietzel M (2025) The widening gap between radiomics research and clinical translation: rethinking current practices and shared responsibilities. Eur J Radiol Artif Intell 1: 100004. https://doi.org/10.1016/j.ejrai.2025.100004
184. Kocak B, Ponsiglione A, Stanzione A et al (2024) CLEAR guideline for radiomics: early insights into current reporting practices endorsed by EuSoMII. Eur J Radiol 181: 111788. https://doi.org/10.1016/j.ejrad.2024.111788
185. Barry N, Kendrick J, Molin K et al (2025) Evaluating the impact of the Radiomics Quality Score: a systematic review and meta-analysis. Eur Radiol 35:1701–1713. https://doi.org/10.1007/s00330-024-11341-y
186. Kocak B, Keles A, Kose F, Sendur A (2025) Quality of radiomics research: comprehensive analysis of 1574 unique publications from 89 reviews. Eur Radiol 35:1980–1992. https://doi.org/10.1007/s00330-024-11057-z
187. Spadarella G, Stanzione A, Akinci D'Antonoli T et al (2023) Systematic review of the radiomics quality score applications: an EuSoMII Radiomics Auditing Group Initiative. Eur Radiol 33:1884–1894. https://doi.org/10.1007/s00330-022-09187-3
188. Lambin P, Leijenaar RT, Deist TM et al (2017) Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 14:749–762. https://doi.org/10.1038/nrclinonc.2017.141
189. Koçak B, D'Antonoli TA, Cuocolo R (2024) Exploring radiomics research quality scoring tools: a comparative analysis of METRICS and RQS. Diagn Interv Radiol 30:366–369. https://doi.org/10.4274/dir.2024.242793
190. Staal FC, Aalbersberg EA, van der Velden D et al (2022) GEP-NET radiomics: a systematic review and radiomics quality score assessment. Eur Radiol 32:7278–7294. https://doi.org/10.1007/s00330-022-08996-w
191. Lambin P, Woodruff HC, Mali SA et al (2025) Radiomics Quality Score 2.0: towards radiomics readiness levels and clinical translation for personalized medicine. Nat Rev Clin Oncol 22:831–846. https://doi.org/10.1038/s41571-025-01067-1
192. Cè M, Chiriac MD, Cozzi A et al (2024) Decoding radiomics: a step-by-step guide to machine learning workflow in hand-crafted and deep learning radiomics studies. Diagnostics (Basel) 14:2473. https://doi.org/10.3390/diagnostics14222473

Publisher's Note

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