

ORIGINAL ARTICLE

Open Access



An overview of meta-analyses on radiomics: more evidence is needed to support clinical translation

Jingyu Zhong^{1†}, Junjie Lu^{2†}, Guangcheng Zhang³, Shiqi Mao⁴, Haoda Chen⁵, Qian Yin⁶, Yangfan Hu¹, Yue Xing¹, Defang Ding¹, Xiang Ge¹, Huan Zhang^{7*} and Weiwu Yao^{1*} 

Abstract

Objective To conduct an overview of meta-analyses of radiomics studies assessing their study quality and evidence level.

Methods A systematic search was updated via peer-reviewed electronic databases, preprint servers, and systematic review protocol registers until 15 November 2022. Systematic reviews with meta-analysis of primary radiomics studies were included. Their reporting transparency, methodological quality, and risk of bias were assessed by PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) 2020 checklist, AMSTAR-2 (A Measurement Tool to Assess systematic Reviews, version 2) tool, and ROBIS (Risk Of Bias In Systematic reviews) tool, respectively. The evidence level supporting the radiomics for clinical use was rated.

Results We identified 44 systematic reviews with meta-analyses on radiomics research. The mean \pm standard deviation of PRISMA adherence rate was $65 \pm 9\%$. The AMSTAR-2 tool rated 5 and 39 systematic reviews as low and critically low confidence, respectively. The ROBIS assessment resulted low, unclear and high risk in 5, 11, and 28 systematic reviews, respectively. We reperformed 53 meta-analyses in 38 included systematic reviews. There were 3, 7, and 43 meta-analyses rated as convincing, highly suggestive, and weak levels of evidence, respectively. The convincing level of evidence was rated in (1) T2-FLAIR radiomics for IDH-mutant vs IDH-wide type differentiation in low-grade glioma, (2) CT radiomics for COVID-19 vs other viral pneumonia differentiation, and (3) MRI radiomics for high-grade glioma vs brain metastasis differentiation.

Conclusions The systematic reviews on radiomics were with suboptimal quality. A limited number of radiomics approaches were supported by convincing level of evidence.

Clinical relevance statement The evidence supporting the clinical application of radiomics are insufficient, calling for researches translating radiomics from an academic tool to a practicable adjunct towards clinical deployment.

[†]Jingyu Zhong and Junjie Lu contributed equally to this work and share first authorship

*Correspondence:

Huan Zhang

huanzhangy@163.com; Zh10765@rjh.com.cn

Weiwu Yao

yaoweiwuhuan@163.com; YWW4142@shtrhospital.com

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

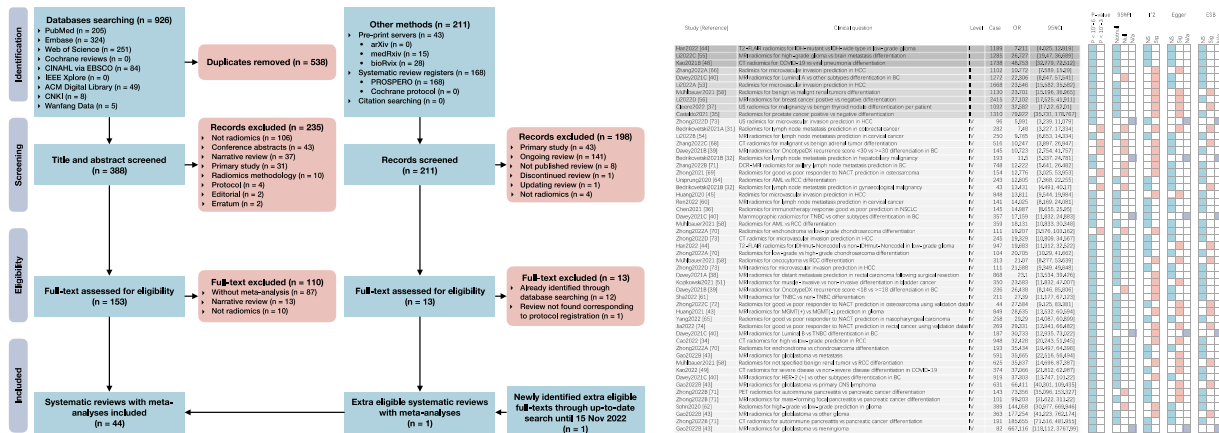
Key points

- The systematic reviews on radiomics studies were insufficient in reporting, suboptimal in methodological quality, and with high risk of bias.
- The meta-analyses covered a wide range of clinical questions, while only three of them were rated as convincing level of evidence.
- More radiomics investigation is necessary to allow clinical translation of radiomics to a practicable adjunct toward clinical deployment.

Keywords Radiomics, Quality improvement, Systematic review, Meta-analysis

Graphical Abstract

An overview of meta-analyses on radiomics: more evidence is needed to support clinical translation



The evidence supporting the clinical application of radiomics are insufficient, calling for researches translating radiomics from an academic tool to a practicable adjunct towards clinical deployment.

Insight Imaging (2023) Zhong J, Lu J et al. DOI: 10.1186/s13244-023-01437-2

Introduction

A decade has passed since the concept of radiomics was raised [1]. The concept of radiomics is based on an assumption that medical images contain information of disease-specific processes that are undetectable to naked eye [2]. Radiomics, a high-throughput methodology that extract large amounts of imaging biomarkers from medical images, is believed to be one of the most promising approaches for enhancing the existing images into deeper mineable data to support clinical decision-making [1–6]. The rapidly evolving field of

radiomics has attracted considerable interest, with a plethora of primary radiomics studies being published [7, 8]. Radiomics seems to potentially have a huge impact on clinical routine at first sight, but so far, little to none of these encouraging findings have served as evidence supporting these research tools translating into clinical application [9–13].

Primary radiomics studies are the sources of information for clinical evidence, while the systematic reviews and meta-analyses provide integration or synthesis of evidence with higher precision from conflicting results,

and address questions that cannot be asked in individual studies [14]. Although an increasing number of systematic reviews and meta-analyses are published in various medical fields, including radiomics [15], it is still unclear how far radiomics is from current research to clinical application [9–13]. There were systematic reviews attempt to cover a wide range of topics in radiomics [16, 17]. However, the number of published primary radiomics studies was too large to summarize in one single systematic review [16], and the evidence level rating of current radiomics was out of the aim of a methodological systematic review [17]. Nevertheless, the overview of systematic reviews is a relatively new type of publication that attempts to provide a broader evidence synthesis highlighting the knowledge gaps, biases, and priorities for future research, which helps clinical practitioners and policy-makers interpret the results of higher-level pieces of evidence in radiomics [18–20].

Therefore, our overview of systematic reviews of primary radiomics studies is aimed at assessing the study quality and the evidence level supporting radiomics application in clinical settings.

Methods

Protocol and registration

This overview of meta-analysis has been prospectively registered on PROSPERO (CRD42021272746), and the review protocol is available as Additional file 1: Note S1. The ethical approval was not required due to the nature of the study. The overview of meta-analysis was conducted as per guidelines [19–22]. The corresponding checklists are supplied in Additional file 1. The literature search, study selection, data extraction, and quality assessment were duplicated by two independent reviewers (J.Z. and either Y.H., Y.X., X.G., or D.D.). The disagreements were resolved by consults with a third independent reviewer (G.Z., S.M., H.C., Q.Y., G.Y., H.Z. or W.Y.). The data analysis was performed by a reviewer (J.Z.) under supervision of a statistical expert (J.L.).

Literature search and study selection

A systematic search was performed to identify systematic reviews with meta-analysis concerning on the radiomics applications for diagnostic, predictive, or prognostic purposes. The search strategy was tested for feasibility with the variations of the terms “radiomics”, “systematic review” and “meta-analysis”. The full formal search was performed until 30 September 2022 and was updated until 15 November 2022. We searched the peer-reviewed electronic databases (PubMed, Embase, Web of Science, Cochrane reviews via Cochrane Central, EBSCO Cumulative Index to Nursing and Allied Health Literature,

Institute of Electrical and Electronics Engineers and Institution of Engineering and Technology Xplore, Association for Computing Machinery Digital Library, China National Knowledge Infrastructure, Wanfang Data), preprint servers (arXiv, medRxiv, bioRxiv), and systematic review protocol registers (PROSPERO and Cochrane protocol via Cochrane Central). To identify additionally eligible systematic reviews, the reference lists of all included articles were screened, and radiomics experts were consulted.

We include all the systematic reviews with meta-analysis concerning on the radiomics applications for diagnostic, predictive, or prognostic purposes in humans. There was no restriction for publication period, target population, study setting, or comparator group, while only articles in English, Chinese, Japanese, German, and French were available. We excluded with following criteria: (a) primary study systematic review without meta-analysis, and article with insufficient information for assessment; (b) systematic review purely assessed artificial intelligence, machine learning or deep learning; (c) systematic review focused on methodology or robustness issue other than clinical-relevant questions. After excluding duplicates, we screened the titles and abstract for potentially available systematic reviews and then, confirmed their eligibility by reading the full-texts, supplementary materials, and related review protocols. The detailed search strategy and study selection process are provided in Additional file 1: Note S2.

Data extraction and study assessment

The data were extracted according to a predefined data extraction sheet (Additional file 1: Table S1). This sheet includes bibliographical information, study characteristics, and effect metrics at level of meta-analyses and those at level of individual primary studies. The contingency tables at the level of individual studies were extracted or reconstructed for repeating meta-analysis.

The PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) 2020 checklist [22], the AMSTAR-2 (A MeaSurement Tool to Assess systematic Reviews, version 2) tool [23], and the ROBIS (Risk Of Bias In Systematic reviews) tool [24] for reporting quality, methodological quality, and risk of bias assessment, respectively. The operational definitions of these three tools can be found in Additional file 1: Tables S2 to S5. The PRISMA 2020 checklist is updated to guide systematic reviewers for transparently reporting with a checklist for abstract of twelve items and a checklist for full-text of twenty-seven items. The AMSTAR-2 tool is developed and modified for critically appraising systematic reviews with sixteen questions to assess their methodological quality.

The three-phase ROBIS tool is specifically designed to assess the risk of bias in systematic reviews covering four domains: study eligibility criteria, identification and selection of studies, data collection and study appraisal, and synthesis and findings. These tools have been tailored to radiological systematic reviews and applied for identifying overlooked reporting items, insufficient methodology, and potential risk of bias, respectively [25–27].

A training phase was introduced to test and modify the tools to reach an operational definition of each item and make sure that all reviewers have a shared understanding. Reached consensus during data extraction and quality assessments is available in Additional file 1: Note S3.

Data analysis and strength of evidence

The statistical analysis was performed with R language version 4.1.3 within using relevant packages [28, 29]. The differences of PRISMA adherence rate, AMSTAR-2 rating, and ROBINS assessment were compared by (a) Journal Citation Report quartile (Q1 or Q2-Q4), (b) journal type (imaging or non-imaging), (c) first authorship (radiologist or non-radiologist), (d) biomarker (diagnostic, predictive, or prognostic), and (e) publication year (2020, 2021, or 2022), using student's *t* test, one-way analysis of variance, and Chi-square test. A two-tailed $p < 0.05$ was recognized as statistical significance, unless specified otherwise.

The meta-analyses were re-performed with R language version 4.1.3 using relevant packages to allow evidence rating [30, 31]. The diagnostic odds ratio (OR) and the corresponding 95% confidence interval (CI) were pooled as summary effect size using random-effect models, and corresponding *p* values were calculated. The sensitivity, specificity, area under curve and index of concordance were not included for analysis, because the corresponding methodology has not been well established so far [21]. The I^2 statistic was used to assess heterogeneity among primary studies. The 95% prediction intervals (PI) were calculated to facilitate more conservative prediction for potential application of radiomics models. The Egger's test was conducted for small-study effects and publication bias. Excess significance bias was evaluated by a Chi-square test comparing the actual observed number of primary studies with a $p < 0.05$ with the expected number of primary studies with statistical significance.

The strength of evidence supporting radiomics for clinical use was categorized into five levels: convincing, highly suggestive, suggestive, weak, and not suggestive (Additional file 1: Table S6), based on the results of a series of aforementioned analyses [21]. The detailed data analysis process is available in Additional file 1: Note S4.

Results

Literature search

The flow diagram of selection process is shown in Fig. 1. Our primary literature search resulted in 926 records, in which 43 systematic reviews were included. No extra available systematic review was identified through preprint servers, systematic review protocol registers, or citation searching. The up-to-date search identified 1 extra eligible systematic review. Finally, 44 systematic reviews were included for the current overview [32–75]. The lists of the included systematic reviews with meta-analyses, and the excluded articles with justifications are provided in Additional file 1: Note S5.

Study characteristics

The characteristics of the included systematic reviews were summarized (Table 1 and Additional file 1: Tables S7 and S8). The systematic reviews most frequently evaluated the application of radiomics in breast cancer ($n=7$), followed by glioma ($n=5$) and liver cancer ($n=4$). The systematic reviews evaluating non-oncological diseases were less common. Only 2 and 1 systematic review investigated the radiomics in COVID-19 and pancreatitis, respectively.

The quality assessment tools used in included systematic reviews varied (Table 1). The Radiomics Quality Score (RQS) rating was the most employed tool ($n=31$), followed by the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) checklist ($n=4$), the Image Biomarker Standardization Initiative (IBSI) checklist ($n=2$), and the CheckList for Artificial Intelligence in Medical imaging (CLAIM) ($n=1$). The revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool, and Prediction model Risk Of Bias ASsessment Tool (PROBAST) were applied by 30 and 2 systematic reviews for risk of bias assessment.

Quality and risk of bias assessment

The result of quality assessment is presented in Fig. 2 and Table 2. The evaluation results for individual systematic reviews are available in Additional file 1: Tables S9 to S11. The overall mean \pm standard deviation (median, range) of PRISMA adherence rate was $65 \pm 9\%$ (64%, 48%–83%) for reporting quality. The AMSTAR-2 rated 5 and 39 systematic reviews as low and critically low confidence in methodological quality assessment, respectively. The overall risk of bias assessment by ROBIS tool resulted 5, 11, and 28 systematic reviews as having a low, unclear, and high risk of bias, respectively. The PRISMA adherence rate of systematic reviews with a first authorship of radiologist was higher than those without (68 ± 7 vs 62 ± 10 , $p=0.023$).

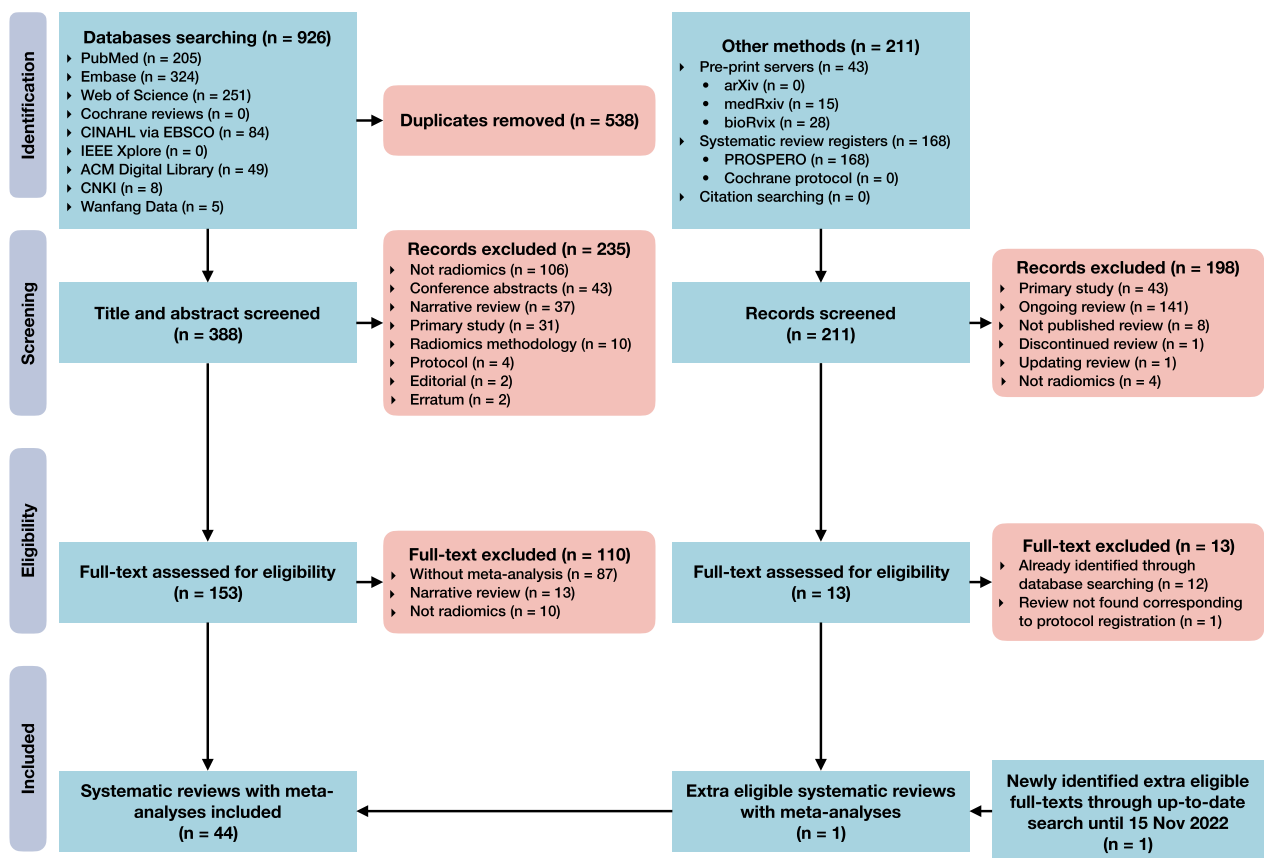


Fig. 1 Flow diagram of systematic review search and selection

Meta-analysis and strength of evidence

There were 53 meta-analyses in 38 systematic reviews re-conducted based on extracted or reconstructed data, covering 497 primary studies, 65,955 subjects, and 29,408 events [32, 33, 35–41, 44–47, 49–59, 61–63, 65–75] (Fig. 3). The meta-analyses in 6 systematic reviews were excluded due to unavailable data [34, 42, 43, 48, 60, 64]. Up to 47 meta-analyses reached a stringent p -value of less than 10^{-6} , and 6 meta-analyses presented p -values $< 10^{-3}$. None of the meta-analyses was deemed as non-significant. Twenty-eight meta-analyses presented an $I^2 > 50\%$. There were 5 meta-analyses conducting with less than three primary studies. For those performing with three or more primary studies, the 95%PI excluded the null value in 37 meta-analyses. Egger’s test of 28 meta-analyses reached $p > 0.05$ for indicating no small-study effects or publication bias. The excess significance bias was not presented in 35 meta-analyses.

Accordingly, there were 3, and 7 meta-analyses rated as convincing, and highly suggestive level of evidence, respectively (Table 3). The radiomics has been rated as convincing level of evidence in (1) T2-FLAIR radiomics for IDH-mutant vs IDH-wide type differentiation in

low-grade glioma (diagnostic OR 7.2, 95%CI 4.0 to 12.9; $p = 3.13 \times 10^{-11}$), (2) CT radiomics for COVID-19 vs other viral pneumonia differentiation (OR 26.7, 95%CI 19.5 to 36.7; $p = 4.54 \times 10^{-82}$), and (3) MRI radiomics for high-grade glioma vs brain metastasis differentiation (OR 48.8, 95%CI 32.8 to 72.5; $p = 7.25 \times 10^{-92}$). The meta-analyses were rated as highly suggestive mainly due to high heterogeneity, significant small-study effects and publication bias. In spite of these dramatic statistical significances, 43 meta-analyses were rated as weak pieces of evidence. The reason for them failed to reach a higher level of evidence was mainly inadequate number of participants.

Discussion

An increasing number of studies are investigating the potential of radiomics as a diagnostic, predictive, or prognostic tool in multiple clinical scenarios, while none of the radiomics academic research has been successfully translated into daily clinical practice. Our overview of systematic reviews with meta-analyses identified 44 systematic reviews and reperformed 53 meta-analyses. The radiomics seemed to be convincing tools in answering three clinical questions including: (1) differentiation

Table 1 Characteristics of included systematic reviews

Characteristics	Data
Included primary studies, mean ± standard deviation, median (range)	22.4 ± 22.7, 15 (5 to 133)
Impact factor, mean ± standard deviation, median (range)	5.33 ± 1.63, 5.74 (2.37 to 10.06)
JCR quartile, <i>n</i>	<i>N</i> = 42*
Q1	18
Q2-Q4	24
Not applicable	2
Journal type, <i>n</i>	<i>N</i> = 44
Imaging	22
Non-imaging	22
First authorship, <i>n</i>	<i>N</i> = 44
Radiologist	21
Non-radiologist	23
Imaging modality, <i>n</i>	<i>N</i> = 81**
CT	25
MRI	35
PET	11
US	8
MMG	2
Biomarker, <i>n</i>	<i>N</i> = 44
Diagnostic	25
Predictive/Prognostic	13
Diagnostic and Predictive/Prognostic	6
Topics	<i>N</i> = 44
Oncologic	
Breast	7
Chest	4
Gastrointestinal	9
Genitourinary	6
Head and Neck	3
Musculoskeletal	3
Neuro	7
Gynecologic	2
Non-oncologic	
COVID-19	2
Pancreatitis	1
Quality assessment tool, <i>n</i>	<i>N</i> = 70**
CLAIM	1
IBSI	2
PROBAST	2
QUADAS-2	30
RQS	31
TRIPOD	4

*Two systematic reviews have no impact factor [35, 44]. ** Systematic reviews with multiple imaging modalities, or multiple quality assessment tools were counted, respectively. CLAIM = Checklist for Artificial Intelligence in Medical Imaging, IBSI = Image Biomarker Standardization Initiative, Journal Citation Report, PROBAST = Prediction Model Risk of Bias Assessment Tool, QUADAS = Quality Assessment of Diagnostic Accuracy Studies, RQS = Radiomics Quality Score rating, TRIPOD = Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

of IDH-mutant vs IDH-wide type in low-grade glioma, (2) differentiation of COVID-19 vs other viral pneumonia, and (3) differentiation of high-grade glioma vs brain metastasis. However, the included systematic reviews were insufficient in reporting, suboptimal in methodological quality, and with high risk of bias. The suboptimal study quality might lead to insufficient confidence in radiomics application and thereby hinder the clinical translation of radiomics even there was high-level of supporting evidence.

The radiomics were most frequently employed in oncological field with a representing example of breast cancer which accounting for seven of included systematic reviews, resulting only three non-oncological radiomics systematic reviews. Sollini et al. [16] declared that the number of oncological image minding studies was six-times of those in non-oncological field. Spadarella et al. [17] found that more than nine tenths of their included systematic reviews focused on oncological radiomics. It is not surprising because the concept of radiomics was raised to mine the medical images for extra deeper information related to oncological genomics [1]. However, the radiomics-biological correlation is more than radio-genomics but covers the diverse clinical, imaging, and molecular profile data, which allow understanding of complex diseases to achieve accurate diagnosis in order to provide the best possible treatment [12, 76]. The radiomics investigations are encouraged to expand to the non-oncological field for wider potential applications.

The quality and risk of bias assessment tools for radiomics systematic reviews varied. The RQS and QUADAS-2 tool were the most used tool for study quality and risk of bias assessment, respectively. The RQS was most used for quality assessment in the included systematic reviews and has been long served a necessary role as the de facto reference tool for assessing radiomics studies [17]. However, the RQS was far from perfect. With an increasing trend of deep learning application in radiomics, RQS could not well identify the advantages and disadvantages in radiomics studies applying as the CLAIM [72]. The TRIPOD checklist might further identify room for improvement in radiomics studies, but some items were not suitable for radiomics studies [76]. The IBSI checklist has highly overlapped with other checklists and somehow too complicated to use [73]. Recently, Checklist for EvaluAtion of Radiomics research (CLEAR) has been developed as a single documentation standard for radiomics research that can guide authors and reviewers [77]. However, the reproducibility and effectiveness of this tool has not been fully investigated yet. The QUADAS-2 tool was employed repeatedly because most of the radiomics studies were diagnostic accuracy studies. The PROBAST

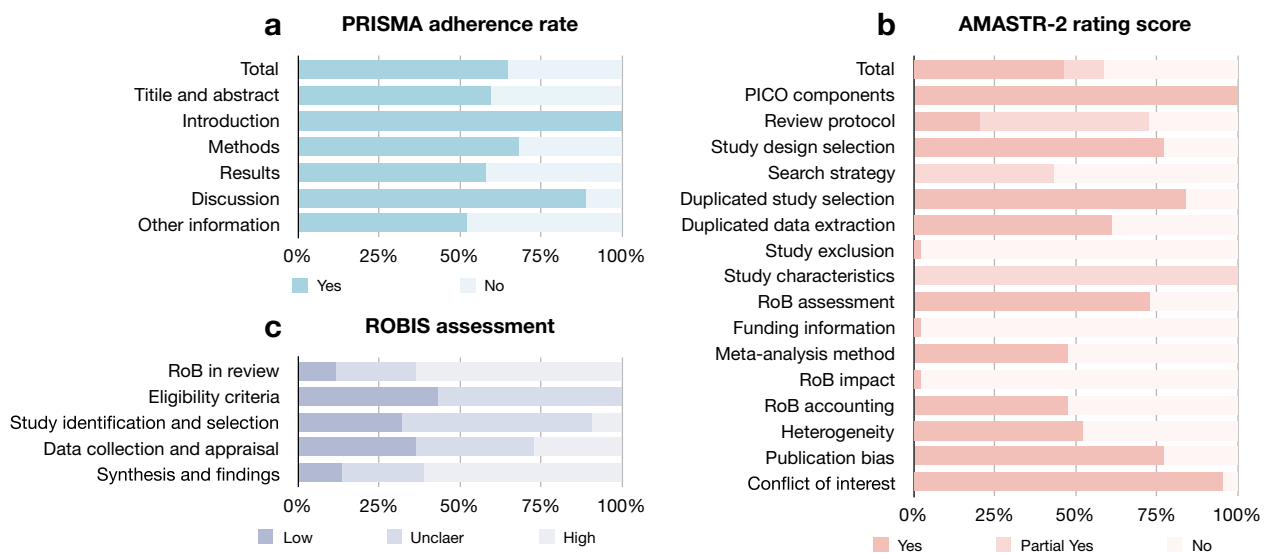


Fig. 2 Result of quality assessment of included systematic reviews. **a** PRISMA adherence rate, **b** AMSTAR-2 rating score, and **c** ROBIS assessment

Table 2 Result of quality assessment of included systematic reviews

Subgroup	No. of systematic reviews	PRISMA adherence rate Mean ± standard deviation, %	AMSTAR-2 rating		ROBINS assessment		
			Low, n	Critically low, n	Low, n	Unclear, n	High, n
Overall	44	65 ± 9	5	39	5	11	28
JCR quartile (N = 42), p-value		0.406		> 0.99			0.073
Q1	18	67 ± 8	2	16	0	4	14
Q2-Q4	24	64 ± 10	3	21	5	7	12
Journal type (N = 44), p-value		0.689		0.154			0.291
Imaging	22	64 ± 10	4	18	3	5	14
Non-imaging	22	65 ± 8	1	21	2	6	14
First authorship (N = 44), p-value		0.023		0.187			0.344
Radiologist	21	68 ± 7	1	20	3	5	13
Non-radiologist	23	62 ± 10	4	19	2	6	15
Biomarker (N = 44), p-value		0.124		0.117			0.132
Diagnostic	25	64 ± 9	5	20	4	9	12
Predictive/Prognostic	13	63 ± 9	0	13	1	2	10
Diagnostic and Predictive/Prognostic	6	72 ± 8	0	6	0	0	6
Publication year (N = 44), p-value		0.208		0.698			0.499
2020	3	65 ± 10	0	3	0	1	2
2021	20	62 ± 9	3	17	1	4	15
2022	21	67 ± 8	2	19	4	6	11

tool might be also suitable for most of radiomics studies because it is developed for predictive models of both diagnostic and prognostic purpose. Other guidelines and checklists are developed or under development for radiomics and artificial intelligence studies including artificial intelligence extensions for TRIPOD, QUADAS-2, and PROBAST [79].

Further validation is needed for their feasibility and efficiency in improving quality of radiomics studies.

The evidence rating highlighted three pieces of convincing evidence of radiomics approaches answering clinical questions. However, there were seven pieces of highly suggestive evidence hindered by the high

Study [Reference]	Clinical question	Level	Case	OR	95%CI	P-value	95%NPI	I ²	Egger	ESB
						P < 10 ⁻⁶ P < 10 ⁻³	Nonnull Null N/a	NS Sig	NS Sig N/a	NS Sig N/a
Han2022 [44]	T2-FLAIR radiomics for IDH-mutant vs IDH-wide type in low-grade glioma	I	1189	7.211	[4.025, 12.919]					
Li2022C [55]	MRI radiomics for high-grade glioma vs brain metastasis differentiation	I	1285	26.727	[19.47, 36.689]					
Kao2021B [48]	CT radiomics for COVID-19 vs viral pneumonia differentiation	I	1738	48.753	[32.779, 72.512]					
Zhang2022A [66]	Radiomics for microvascular invasion prediction in HCC	II	1102	10.772	[7.589, 15.29]					
Davey2021C [40]	MRI radiomics for Luminal A vs other subtypes differentiation in BC	II	1272	22.306	[8.647, 57.541]					
Li2022A [53]	Radiomics for microvascular invasion prediction in HCC	II	1668	23.546	[15.582, 35.582]					
Mühlbauer2021 [58]	Radiomics for benign vs malignant renal tumors differentiation	II	1130	23.701	[15.196, 36.965]					
Li2022D [56]	MRI radiomics for breast cancer positive vs negative differentiation	II	2415	27.102	[17.525, 41.911]					
Cleere2022 [37]	US radiomics for malignancy vs benign thyroid nodule differentiation per patient	II	1092	32.582	[17.12, 62.01]					
Castaldo2021 [35]	Radiomics for prostate cancer positive vs negative differentiation	II	1310	79.922	[35.731, 178.767]					
Zhong2022D [73]	US radiomics for microvascular invasion prediction in HCC	IV	96	5.991	[3.239, 11.079]					
Bedrikovetski2021A [31]	Radiomics for lymph node metastasis prediction in colorectal cancer	IV	282	7.48	[3.227, 17.334]					
Li2022B [54]	MRI radiomics for lymph node metastasis prediction in cervical cancer	IV	250	9.765	[6.653, 14.334]					
Zhang2022C [68]	CT radiomics for malignant vs benign adrenal tumor differentiation	IV	516	10.247	[3.897, 26.947]					
Davey2021B [39]	MRI radiomics for OncotypeDX recurrence score <30 vs >=30 differentiation in BC	IV	145	10.723	[2.754, 41.757]					
Bedrikovetski2021B [32]	Radiomics for lymph node metastasis prediction in hepatobiliary malignancy	IV	193	11.5	[5.337, 24.781]					
Zhang2022B [71]	DCR-MRI radiomics for axillary lymph node metastasis prediction in BC	IV	748	12.222	[5.641, 26.482]					
Zhong2021 [69]	Radiomics for good vs poor responder to NACT prediction in osteosarcoma	IV	154	12.776	[3.025, 53.953]					
Ursprung2020 [64]	Radiomics for AML vs RCC differentiation	IV	243	12.805	[7.368, 22.255]					
Bedrikovetski2021B [32]	Radiomics for lymph node metastasis prediction in gynaecological malignancy	IV	43	13.431	[4.491, 40.17]					
Huang2020 [45]	Radiomics for microvascular invasion prediction in HCC	IV	848	13.811	[9.544, 19.984]					
Ren2022 [60]	MRI radiomics for lymph node metastasis prediction in cervical cancer	IV	141	14.025	[8.169, 24.081]					
Chen2021 [36]	Radiomics for immunotherapy response good vs poor prediction in NSCLC	IV	145	14.987	[8.655, 25.95]					
Davey2021C [40]	Mammographic radiomics for TNBC vs other subtypes differentiation in BC	IV	357	17.159	[11.832, 24.883]					
Mühlbauer2021 [58]	Radiomics for AML vs RCC differentiation	IV	359	18.131	[10.833, 30.348]					
Zhong2022A [70]	Radiomics for enchondroma vs low-grade chondrosarcoma differentiation	IV	111	19.207	[3.576, 103.162]					
Zhong2022D [73]	CT radiomics for microvascular invasion prediction in HCC	IV	245	19.329	[10.809, 34.567]					
Han2022 [44]	T2-FLAIR radiomics for IDHmut- Noncodel vs non-IDHmut- Noncodel in low-grade glioma	IV	947	19.683	[11.912, 32.522]					
Zhong2022A [70]	Radiomics for low-grade vs high-grade chondrosarcoma differentiation	IV	104	20.705	[10.29, 41.662]					
Mühlbauer2021 [58]	Radiomics for oncocytoma vs RCC differentiation	IV	313	21.07	[8.277, 53.639]					
Zhong2022D [73]	MRI radiomics for microvascular invasion prediction in HCC	IV	111	21.588	[9.349, 49.848]					
Davey2021A [38]	MRI radiomics for distant metastasis prediction in rectal carcinoma following surgical resection	IV	868	23.1	[13.534, 39.426]					
Kozikowski2021 [51]	MRI radiomics for muscle-invasive vs non-invasive differentiation in bladder cancer	IV	350	23.583	[11.832, 47.007]					
Davey2021B [39]	MRI radiomics for OncotypeDX recurrence score <18 vs >=18 differentiation in BC	IV	236	26.438	[8.146, 85.806]					
Sha2022 [61]	MRI radiomics for TNBC vs non-TNBC differentiation	IV	211	27.39	[11.177, 67.123]					
Zhong2022C [72]	Radiomics for good vs poor responder to NACT prediction in osteosarcoma using validation data	IV	44	27.584	[9.125, 83.381]					
Huang2021 [43]	MRI radiomics for MGMT(+) vs MGMT(-) prediction in glioma	IV	849	28.635	[13.532, 60.594]					
Yang2022 [65]	Radiomics for good vs poor responder to NACT prediction in nasopharyngeal carcinoma	IV	258	29.29	[14.087, 60.899]					
Jia2022 [74]	Radiomics for good vs poor responder to NACT prediction in rectal cancer using validation data	IV	269	29.331	[12.941, 66.482]					
Davey2021C [40]	MRI radiomics for Luminal B vs TNBC differentiation in BC	IV	187	30.733	[12.935, 73.022]					
Cao2022 [34]	CT radiomics for high vs low-grade prediction in RCC	IV	948	32.428	[20.243, 51.945]					
Zhong2022A [70]	Radiomics for enchondroma vs chondrosarcoma differentiation	IV	193	35.434	[19.497, 64.398]					
Gao2022B [43]	MRI radiomics for glioblastoma vs metastasis	IV	591	35.665	[22.516, 56.494]					
Mühlbauer2021 [58]	Radiomics for not specified benign renal tumor vs RCC differentiation	IV	625	35.837	[14.696, 87.387]					
Kao2022 [49]	CT radiomics for severe disease vs non-severe disease differentiation in COVID-19	IV	374	37.066	[21.812, 62.987]					
Davey2021C [40]	MRI radiomics for HER-2 (+) vs other subtypes differentiation in BC	IV	919	37.303	[13.747, 101.22]					
Gao2022B [43]	MRI radiomics for glioblastoma vs primary CNS lymphoma	IV	631	66.411	[40.301, 109.435]					
Zhong2022B [71]	PET radiomics for autoimmune pancreatitis vs pancreatic cancer differentiation	IV	143	73.356	[35.096, 153.327]					
Zhong2022B [71]	MRI radiomics for mass-forming focal pancreatitis vs pancreatic cancer differentiation	IV	101	99.203	[31.622, 311.22]					
Sohn2020 [62]	Radiomics for high-grade vs low-grade prediction in glioma	IV	389	144.058	[30.977, 669.946]					
Gao2022B [43]	MRI radiomics for glioblastoma vs other glioma	IV	363	177.254	[41.223, 762.174]					
Zhong2022B [71]	CT radiomics for autoimmune pancreatitis vs pancreatic cancer differentiation	IV	191	185.655	[71.516, 481.955]					
Gao2022B [43]	MRI radiomics for glioblastoma vs meningioma	IV	82	667.116	[118.112, 3767.99]					

Fig. 3 Summary of evidence rating. AML = angiomyolipoma, BC = breast cancer, HCC = hepatocellular carcinoma, NSCLC = non-small cell lung cancer, RCC = renal clear cell carcinoma, NACT = neoadjuvant chemotherapy, TNBC = triple negative breast cancer. NS = not significant, Sig = significant, N/a = not applicable

Table 3 Summary of converging and highly suggestive evidence

Study [References]	Clinical question	Evidence rating	PRISMA, %	AMSTAR	ROBINS
Han2022 [45]	T2-FLAIR radiomics for IDH-mutant vs IDH-wide type in low-grade glioma	Convincing	67	Critically low	Unclear
Kao2021B [56]	CT radiomics for COVID-19 vs viral pneumonia differentiation	Convincing	56	Critically low	High
Li2022C [49]	MRI radiomics for high-grade glioma vs brain metastasis differentiation	Convincing	67	Low	Unclear
Li2022A [54]	Radiomics in microvascular invasion prediction in hepatocellular carcinoma	Highly suggestive	77	Low	Unclear
Davey2021C [41]	MRI radiomics in Luminal A vs other subtypes differentiation in breast cancer	Highly suggestive	52	Critically low	High
Zhang2022A [67]	Radiomics in microvascular invasion prediction in hepatocellular carcinoma	Highly suggestive	77	Low	Low
Li2022D [57]	MRI radiomics in breast cancer positive vs negative differentiation	Highly suggestive	58	Critically low	High
Castaldo2021 [36]	Radiomics in prostate cancer positive vs negative differentiation	Highly suggestive	67	Critically low	Unclear
Mühlbauer2021 [59]	Radiomics in benign vs malignant renal tumors differentiation	Highly suggestive	60	Critically low	High
Cleere2022 [38]	US radiomics in malignancy vs benign thyroid nodule per patient	Highly suggestive	52	Critically low	High

heterogeneity. We did not investigate the potential source of heterogeneity due to the workload, but this should be explored in the individual systematic review to allow interruption of the results. Unfortunately, these systematic reviews did not perform related investigations. Indeed, less than a half of included systematic reviews conducted such an analysis. Another reason for failing to reach convincing level of evidence is significant small-study effects and publication bias. This was assessed by more than four fifths of the included systematic reviews. The radiomics were not rated as sufficient tools for other clinical applications. There were more pieces of weak evidence due to insufficient participants. This could not be solved by systematic reviews, but it might be overcome with more carefully designed prospective, multicenter, randomized controlled trials and data sharing [9, 11, 12]. Another concern on the systematic reviews and meta-analyses of radiomics was their relatively low study quality. Although our overview identified three potential application of radiomics with high-level of supporting evidence, they were all with suboptimal quality that should be taken into consideration when applying the evidence. A systematic approach is encouraged to establish to comprehensively evaluate the radiomics tool, in order to tell whether the tool can be used in the clinical practice. The GRADE (grading of recommendations assessment, development and evaluation) system can be used for diagnostic tests or strategies [80], but the feasibility of this approach for radiomics researches needs to be verified.

There are several limitations to address. First, we only included systematic reviews with meta-analyses to identify the most possible candidate to be supported by high-level evidence. We only included the primary studies mentioned in the meta-analyses for re-analysis, because updating the literature search may lead to a too heavy workload. As a rapidly developing field, our meta-analyses may not include all the eligible radiomics studies. Second, most of the meta-analyses were based on training or validation dataset, which potentially overestimated the results. The future analysis is encouraged to be conducted using testing dataset of the strictly designed studies. Third, the majority of the included primary studies were retrospective, single-center, small-scale studies and have been assessed as suboptimal quality. Further, the overall quality of included systematic reviews was also insufficient. Therefore, the aforementioned evidence level rating results should be cautiously interpreted. Lastly, the evidence rating criteria of diagnostic accuracy tests have not been well established. We only estimated the

diagnostic odds ratio as effect size, but not the corresponding sensitivity, specificity, and area under curve value for each meta-analysis, whose potential role in evidence rating needs further investigation.

Conclusion

In conclusion, our overview of systematic reviews and meta-analyses highlighted three convincing and seven highly suggestive level of evidence for radiomics in answering clinical questions, while the low study quality and high risk of bias might lead to insufficient confidence in clinical translation. Future research should provide more scientific base for those with low-level of evidence and seek to validate the radiomics algorithms in clinical settings for those with high-level of evidence. Systematic reviews and meta-analyses on radiomics researches could continuously help the stakeholder to identify knowledge gaps, biases, and priorities for future research to promote the radiomics translation from an academic tool for generating papers to a practicable adjunct toward clinical deployment.

Abbreviations

AMSTAR-2	A Measurement Tool to Assess systematic Reviews, version 2
CLAIM	CheckList for Artificial Intelligence in Medical imaging
IBSI	Image Biomarker Standardization Initiative
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PROBAST	Prediction model Risk Of Bias ASsessment Tool
QUADAS-2	Revised QUality Assessment of Diagnostic Accuracy Studies
ROBIS	Risk Of Bias In Systematic reviews
RQS	Radiomics Quality Score
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/s13244-023-01437-2>.

Additional file 1: Supplementary Materials, Supplementary Review Protocol, and Supplementary PRISMA Checklists. **Supplementary Note S1.** Review protocol. **Supplementary Note S2.** Study search strategy and study selection. **Supplementary Note S3.** Consensus reached during data extraction and quality assessment. **Supplementary Note S4.** Data synthesis and analysis methods. **Supplementary Note S5.** List of included full-texts and excluded full-texts with justifications. **Supplementary Table S1.** Data extraction sheet. **Supplementary Table S2.** PRISMA 2020 abstract checklist for reporting quality assessment. **Supplementary Table S3.** PRISMA 2020 checklist for reporting quality assessment. **Supplementary Table S4.** AMSTAR-2 tool for methodological quality assessment. **Supplementary Table S5.** ROBIS tool for risk of bias assessment. **Supplementary Table S6.** Category of five levels of evidence based on meta-analyses. **Supplementary Table S7.** Bibliographic information of included systematic reviews. **Supplementary Table S8.** Review topic of included systematic reviews. **Supplementary Table S9.** PRISMA adherence rate of included systematic reviews. **Supplementary Table S10.** AMSTAR-2 ratings of included systematic reviews. **Supplementary Table S11.** ROBIS tool assessments of included systematic reviews.

Acknowledgements

Dr. Jingyu Zhong would like to express his gratitude to Ms. Hongyan Huang for her support and kindness as a mother and as a friend, and wish her a happy 60th birthday.

Author contributions

JZ, GZ, SM, HC, QY, YH, YX, DD and XG performed the literature search, data extraction, and quality assessment. JZ and JL performed meta-analyses and visualized data. JZ drafted the manuscript. HZ and WY supervised the study. All authors read and approved the final manuscript.

Funding

This study has received funding by: National Natural Science Foundation of China (82271934); Yangfan Project of Science and Technology Commission of Shanghai Municipality (22YF1442400); Shanghai Science and Technology Commission Science and Technology Innovation Action Clinical Innovation Field (18411953000); Medicine and Engineering Combination Project of Shanghai Jiao Tong University (YG2019ZDB09); Research Fund of Tongren Hospital, Shanghai Jiao Tong University School of Medicine (TRKYRC-XX202204, TRYJ2021JC06, TRGG202101, 2020TRYJ(LB)06, 2020TRYJ(JC)07); and Guangci Innovative Technology Launch Plan of Ruijin Hospital, Shanghai Jiao Tong University School of Medicine (2022-13). They played no role in the study design, data collection or analysis, decision to publish, or manuscript preparation.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its Additional files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Imaging, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200336, China. ²Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA. ³Department of Orthopedics, Shanghai Sixth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200233, China. ⁴Department of Medical Oncology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai 200433, China. ⁵Department of General Surgery, Pancreatic Disease Center, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China. ⁶Department of Pathology, Shanghai Sixth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200233, China. ⁷Department of Radiology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China.

Received: 13 January 2023 Accepted: 14 April 2023

Published online: 19 June 2023

References

- 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(4):441–446. <https://doi.org/10.1016/j.ejca.2011.11.036>
- Tomaszewski MR, Gillies RJ (2021) The biological meaning of radiomic features. *Radiology* 298(3):505–516. <https://doi.org/10.1148/radiol.2021202553>
- Gillies RJ, Kinahan PE (2016) Hricak H (2016) Radiomics: images are more than pictures, they are data. *Radiology* 278(2):563–577. <https://doi.org/10.1148/radiol.2015151169>
- Lambin P, Leijenaar RTH, Deist TM et al (2017) Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 14(12):749–762. <https://doi.org/10.1038/nrclinonc.2017.141>
- O'Connor JP, Aboagye EO, Adams JE et al (2017) Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* 14:169–186. <https://doi.org/10.1038/nrclinonc.2016.162>
- 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(1):91. <https://doi.org/10.1186/s13244-020-00887-2>
- Song J, Yin Y, Wang H, Chang Z, Liu Z, Cui L (2020) A review of original articles published in the emerging field of radiomics. *Eur J Radiol* 127:108991. <https://doi.org/10.1016/j.ejrad.2020.108991>
- Volpe S, Mastroleo F, Krengli M, Jereczek-Fossa BA (2023) Quo vadis Radiomics? Bibliometric analysis of 10-year Radiomics journey. *Eur Radiol*. <https://doi.org/10.1007/s00330-023-09645-6>
- Pinto Dos Santos D, Dietzel M, Baessler B (2021) A decade of radiomics research: are images really data or just patterns in the noise? *Eur Radiol* 31(1):1–4. <https://doi.org/10.1007/s00330-020-07108-w>
- Dewey M, Bossert M, Dodd JD, Thun S, Kressel HY (2019) Clinical imaging research: higher evidence, global collaboration, improved reporting, and data sharing are the grand challenges. *Radiology* 291(3):547–552. <https://doi.org/10.1148/radiol.2019181796>
- Halligan S, Menu Y, Mallett S (2021) Why did European Radiology reject my radiomic biomarker paper? How to correctly evaluate imaging biomarkers in a clinical setting. *Eur Radiol* 31(12):9361–9368. <https://doi.org/10.1007/s00330-021-07971-1>
- Fournier L, Costaridou L, Bidaut L et al (2021) European Society of Radiology Incorporating radiomics into clinical trials: expert consensus endorsed by the European Society of Radiology on considerations for data-driven compared to biologically driven quantitative biomarkers. *Eur Radiol* 31(8):6001–6012. <https://doi.org/10.1007/s00330-020-07598-8>
- Moskowitz CS, Welch ML, Jacobs MA, Kurland BF, Simpson AL (2022) Radiomic analysis: study design, statistical analysis, and other bias mitigation strategies. *Radiology* 304(2):265–273. <https://doi.org/10.1148/radiol.211597>
- Higgins JPT, Thomas J, Chandler J et al (2022) Cochrane handbook for systematic reviews of interventions version 6.3, 2022. Cochrane Web site. <https://training.cochrane.org/handbook>. Updated February, 2022. Accessed 1 Jul 2022.
- Booth A, Clarke M, Ghera D, Moher D, Petticrew M, Stewart L (2010) An international registry of systematic-review protocols. *Lancet* 377:108–109. [https://doi.org/10.1016/S0140-6736\(10\)60903-8](https://doi.org/10.1016/S0140-6736(10)60903-8)
- Sollini M, Antunovic L, Chiti A, Kirienco M (2019) Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. *Eur J Nucl Med Mol Imaging* 46(13):2656–2672. <https://doi.org/10.1007/s00259-019-04372-x>
- Spadarella G, Stanzone A, Akinici D'Antonoli T et al (2022) Systematic review of the radiomics quality score applications: an EuSoMI Radiomics Auditing Group Initiative. *Eur Radiol*. <https://doi.org/10.1007/s00330-022-09187-3>
- Hartling L, Vandermeer B, Fernandes RM (2014) Systematic reviews, overviews of reviews and comparative effectiveness reviews: a discussion of approaches to knowledge synthesis. *Evid Based Child Health* 9:486–494. <https://doi.org/10.1002/ebch.1968>
- Pollock A, Campbell P, Brunton G, Hunt H, Estcourt L (2017) Selecting and implementing overview methods: implications from five exemplar overviews. *Syst Rev* 6(1):145. <https://doi.org/10.1186/s13643-017-0534-3>
- Bougioukas KI, Liakos A, Tsapas A, Ntzani E, Haidich AB (2018) Preferred reporting items for overviews of systematic reviews including harms checklist: a pilot tool to be used for balanced reporting of benefits and harms. *J Clin Epidemiol* 93:9–24. <https://doi.org/10.1016/j.jclinepi.2017.10.002>
- Fusar-Poli P, Radua J (2018) Ten simple rules for conducting umbrella reviews. *Evid Based Ment Health* 21(3):95–100. <https://doi.org/10.1136/ebmental-2018-300014>
- Page MJ, McKenzie JE, Bossuyt PM et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71. <https://doi.org/10.1136/bmj.n71>
- Shea BJ, Reeves BC, Wells G et al (2017) AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised

- studies of healthcare interventions, or both. *BMJ* 358:j4008. <https://doi.org/10.1136/bmj.j4008>
24. Whiting P, Savović J, Higgins JP et al (2016) ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 69:225–234. <https://doi.org/10.1016/j.jclinepi.2015.06.005>
 25. Park HY, Suh CH, Woo S, Kim PH, Kim KW (2022) Quality reporting of systematic review and meta-analysis according to PRISMA 2020 guidelines: results from recently published papers in the Korean Journal of Radiology. *Korean J Radiol* 23(3):355–369. <https://doi.org/10.3348/kjr.2021.0808>
 26. Dang Y, Hou Y (2021) The prognostic value of late gadolinium enhancement in heart diseases: an umbrella review of meta-analyses of observational studies. *Eur Radiol* 31(7):4528–4537. <https://doi.org/10.1007/s00330-020-07437-w>
 27. Al-Okshi A, Horner K, Rohlin M (2021) A meta-review of effective doses in dental and maxillofacial cone beam CT using the ROBIS tool. *Br J Radiol* 94(1123):20210042. <https://doi.org/10.1259/bjr.20210042>
 28. Mangiafico SS (2015) An R companion for the handbook of biological statistics, version 1.3.2, 2015. rcompanion.org/rcompanion/. Accessed 1 Sept 2022.
 29. Mangiafico SS (2016) Summary and analysis of extension program evaluation in R, version 1.19.10.2016. rcompanion.org/handbook/. Accessed 1 Sept 2022.
 30. Michael Dewey, Wolfgang Viechtbauer (2022) CRAN task view: meta-analysis, version 2022-06-21. <https://CRAN.R-project.org/view=MetaAnalysis>. Accessed 1 Sept 2022.
 31. Gosling CJ, Solanes A, Fusar-Poli P, Radua J (2022) metaumbrella: An R package for conducting umbrella reviews version 1.05, 2022. <https://CRAN.R-project.org/package=metaumbrella/>. Accessed 1 Sept 2022.
 32. Bedrikovetski S, Dudi-Venkata NN, Kroon HM et al (2021) Artificial intelligence for pre-operative lymph node staging in colorectal cancer: a systematic review and meta-analysis. *BMC Cancer* 21(1):1058. <https://doi.org/10.1186/s12885-021-08773-w>
 33. Bedrikovetski S, Dudi-Venkata NN, Maicas G et al (2021) Artificial intelligence for the diagnosis of lymph node metastases in patients with abdominopelvic malignancy: a systematic review and meta-analysis. *Artif Intell Med* 113:102022. <https://doi.org/10.1016/j.artmed.2021.102022>
 34. Bhandari AP, Liong R, Koppen J, Murthy SV, Lasocki A (2021) Noninvasive determination of IDH and 1p19q status of lower-grade gliomas using MRI radiomics: a systematic review. *AJNR Am J Neuroradiol* 42(1):94–101. <https://doi.org/10.3174/ajnr.A6875>
 35. Cao X, Zhu M, Yin L et al (2022) CT radiomics for predicting pathological grade of renal clear cell carcinoma: meta-analysis. *Chin J Med Imaging Technol* 38(8):1197–1202. <https://doi.org/10.13929/j.isn.1003-3289.2022.08.017>
 36. Castaldo R, Cavaliere C, Soricelli A, Salvatore M, Pecchia L, Franzese M (2021) Radiomic and genomic machine learning method performance for prostate cancer diagnosis: systematic literature review. *J Med Internet Res* 23(4):e22394. <https://doi.org/10.2196/22394>
 37. Chen Q, Zhang L, Mo X et al (2021) Current status and quality of radiomic studies for predicting immunotherapy response and outcome in patients with non-small cell lung cancer: a systematic review and meta-analysis. *Eur J Nucl Med Mol Imaging* 49(1):345–360. <https://doi.org/10.1007/s00259-021-05509-7>
 38. Cleere EF, Davey MG, O'Neill S et al (2022) Radiomic detection of malignancy within thyroid nodules using ultrasonography—a systematic review and meta-analysis. *Diagnostics (Basel)* 12(4):794. <https://doi.org/10.3390/diagnostics12040794>
 39. Davey MS, Davey MG, Ryan EJ, Hogan AM, Kerin MJ, Joyce M (2021) The use of radiomic analysis of magnetic resonance imaging in predicting distant metastases of rectal carcinoma following surgical resection: a systematic review and meta-analysis. *Colorectal Dis* 23(12):3065–3072. <https://doi.org/10.1111/codi.15919>
 40. Davey MG, Davey MS, Ryan EJ et al (2021) Is radiomic MRI a feasible alternative to OncotypeDX® recurrence score testing? A systematic review and meta-analysis. *BJS Open* 5(5):zrab081. <https://doi.org/10.1093/bjsopen/zrab081>
 41. Davey MG, Davey MS, Boland MR et al (2021) Radiomic differentiation of breast cancer molecular subtypes using pre-operative breast imaging—a systematic review and meta-analysis. *Eur J Radiol* 144:109996. <https://doi.org/10.1016/j.ejrad.2021.109996>
 42. Deantonio L, Garo ML, Paone G et al (2022) 18F-FDG PET radiomics as predictor of treatment response in oesophageal cancer: a systematic review and meta-analysis. *Front Oncol* 12:861638. <https://doi.org/10.3389/fonc.2022.861638>
 43. Gao Y, Cheng S, Zhu L et al (2022) A systematic review of prognosis predictive role of radiomics in pancreatic cancer: heterogeneity markers or statistical tricks? *Eur Radiol* 32(12):8443–8452. <https://doi.org/10.1007/s00330-022-08922-0>
 44. Gao Z, Xiao Y, Zhu F et al (2022) Diagnostic value of radiomics in glioblastoma: a meta-analysis. *Chin J Evid Based Med* 22(2):232–242. <https://doi.org/10.7507/1672-2531.202108134>
 45. Han Z, Chen Q, Zhang L et al (2022) Radiogenomic association between the T2-FLAIR mismatch sign and IDH mutation status in adult patients with lower-grade gliomas: an updated systematic review and meta-analysis. *Eur Radiol* 32(8):5339–5352. <https://doi.org/10.1007/s00330-022-08607-8>
 46. Huang J, Tian W, Zhang L et al (2020) Preoperative prediction power of imaging methods for microvascular invasion in hepatocellular carcinoma: a systemic review and meta-analysis. *Front Oncol* 10:887. <https://doi.org/10.3389/fonc.2020.00887>
 47. Huang H, Wang FF, Luo S, Chen G, Tang G (2021) Diagnostic performance of radiomics using machine learning algorithms to predict MGMT promoter methylation status in glioma patients: a meta-analysis. *Diagn Interv Radiol* 27(6):716–724. <https://doi.org/10.5152/dir.2021.21153>
 48. Kao YS, Hsu Y (2021) A meta-analysis for using radiomics to predict complete pathological response in esophageal cancer patients receiving neoadjuvant chemoradiation. *In Vivo* 35(3):1857–1863. <https://doi.org/10.21873/in vivo.12448>
 49. Kao YS, Lin KT (2021) A meta-analysis of computerized tomography-based radiomics for the diagnosis of COVID-19 and viral pneumonia. *Diagnostics (Basel)* 11(6):991. <https://doi.org/10.3390/diagnostics11060991>
 50. Kao YS, Lin KT (2022) A meta-analysis of the diagnostic test accuracy of CT-based radiomics for the prediction of COVID-19 severity. *Radiol Med* 127(7):754–762. <https://doi.org/10.1007/s11547-022-01510-8>
 51. Kothari G, Korte J, Lehrer EJ et al (2021) A systematic review and meta-analysis of the prognostic value of radiomics based models in non-small cell lung cancer treated with curative radiotherapy. *Radiother Oncol* 155:188–203. <https://doi.org/10.1016/j.radonc.2020.10.023>
 52. Kozikowski M, Suarez-Ibarrola R, Osiecki R et al (2021) Role of radiomics in the prediction of muscle-invasive bladder cancer: a systematic review and meta-analysis. *Eur Urol Focus* 8(3):728–738. <https://doi.org/10.1016/j.euf.2021.05.005>
 53. Lee S, Choi Y, Seo MK et al (2022) Magnetic resonance imaging-based radiomics for the prediction of progression-free survival in patients with nasopharyngeal carcinoma: a systematic review and meta-analysis. *Cancers (Basel)* 14(3):653. <https://doi.org/10.3390/cancers14030653>
 54. Li L, Wu C, Huang Y, Chen J, Ye D, Su Z (2022) Radiomics for the preoperative evaluation of microvascular invasion in hepatocellular carcinoma: a meta-analysis. *Front Oncol* 12:831996. <https://doi.org/10.3389/fonc.2022.831996>
 55. Li L, Zhang J, Zhe X et al (2022) A meta-analysis of MRI-based radiomic features for predicting lymph node metastasis in patients with cervical cancer. *Eur J Radiol* 151:110243. <https://doi.org/10.1016/j.ejrad.2022.110243>
 56. Li Y, Liu Y, Liang Y et al (2022) Radiomics can differentiate high-grade glioma from brain metastasis: a systematic review and meta-analysis. *Eur Radiol* 32(11):8039–8051. <https://doi.org/10.1007/s00330-022-08828-x>
 57. Li Z, Ye J, Du H et al (2022) Preoperative prediction power of radiomics for breast cancer: a systemic review and meta-analysis. *Front Oncol* 12:837257. <https://doi.org/10.3389/fonc.2022.837257>
 58. Liang X, Yu X, Gao T (2022) Machine learning with magnetic resonance imaging for prediction of response to neoadjuvant chemotherapy in breast cancer: a systematic review and meta-analysis. *Eur J Radiol* 150:110247. <https://doi.org/10.1016/j.ejrad.2022.110247>
 59. Mühlbauer J, Egen L, Kowalewski KF et al (2021) Radiomics in renal cell carcinoma—a systematic review and meta-analysis. *Cancers (Basel)* 13(6):1348. <https://doi.org/10.3390/cancers13061348>

60. Pesapane F, Agazzi GM, Rotili A et al (2022) Prediction of the pathological response to neoadjuvant chemotherapy in breast cancer patients with MRI-radiomics: a systematic review and meta-analysis. *Curr Probl Cancer* 46(5):100883. <https://doi.org/10.1016/j.cuprobincancer>
61. Ren J, Li Y, Liu XY et al (2022) Diagnostic performance of ADC values and MRI-based radiomics analysis for detecting lymph node metastasis in patients with cervical cancer: a systematic review and meta-analysis. *Eur J Radiol* 156:110504. <https://doi.org/10.1016/j.ejrad.2022.110504>
62. Sha YS, Chen JF (2022) MRI-based radiomics for the diagnosis of triple-negative breast cancer: a meta-analysis. *Clin Radiol* 77(9):655–663. <https://doi.org/10.1016/j.crad.2022.04.015>
63. Sohn CK, Bisdas S (2020) Diagnostic accuracy of machine learning-based radiomics in grading gliomas: systematic review and meta-analysis. *Contrast Media Mol Imaging* 2020:2127062. <https://doi.org/10.1155/2020/2127062>
64. Ugga L, Perillo T, Cuocolo R et al (2021) Meningioma MRI radiomics and machine learning: systematic review, quality score assessment, and meta-analysis. *Neuroradiology* 63(8):1293–1304. <https://doi.org/10.1007/s00234-021-02668-0>
65. Ursprung S, Beer L, Bruining A et al (2020) Radiomics of computed tomography and magnetic resonance imaging in renal cell carcinoma—a systematic review and meta-analysis. *Eur Radiol* 30(6):3558–3566. <https://doi.org/10.1007/s00330-020-06666-3>
66. Yang C, Jiang Z, Cheng T et al (2022) Radiomics for predicting response of neoadjuvant chemotherapy in nasopharyngeal carcinoma: a systematic review and meta-analysis. *Front Oncol* 12:893103. <https://doi.org/10.3389/fonc.2022.893103>
67. Zhang J, Huang S, Xu Y, Wu J (2022) Diagnostic accuracy of artificial intelligence based on imaging data for preoperative prediction of microvascular invasion in hepatocellular carcinoma: a systematic review and meta-analysis. *Front Oncol* 12:763842. <https://doi.org/10.3389/fonc.2022.763842>
68. Zhang J, Li L, Zhe X et al (2022) The diagnostic performance of machine learning-based radiomics of DCE-MRI in predicting axillary lymph node metastasis in breast cancer: a meta-analysis. *Front Oncol* 12:799209. <https://doi.org/10.3389/fonc.2022.799209>
69. Zhang H, Lei H, Pang J (2022) Diagnostic performance of radiomics in adrenal masses: a systematic review and meta-analysis. *Front Oncol* 12:975183. <https://doi.org/10.3389/fonc.2022.975183>
70. Zhong J, Hu Y, Si L et al (2021) A systematic review of radiomics in osteosarcoma: utilizing radiomics quality score as a tool promoting clinical translation. *Eur Radiol* 31(3):1526–1535. <https://doi.org/10.1007/s00330-020-07221-w>
71. Zhong J, Hu Y, Ge X et al (2022) A systematic review of radiomics in chondrosarcoma: assessment of study quality and clinical value needs handy tools. *Eur Radiol*. <https://doi.org/10.1007/s00330-022-09060-3>
72. Zhong J, Hu Y, Zhang G et al (2022) An updated systematic review of radiomics in osteosarcoma: utilizing CLAIM to adapt the increasing trend of deep learning application in radiomics. *Insights Imaging* 13(1):138. <https://doi.org/10.1186/s13244-022-01277-6>
73. Zhong J, Hu Y, Xing Y et al (2022) A systematic review of radiomics in pancreatitis: applying the evidence level rating tool for promoting clinical transferability. *Insights Imaging* 13(1):139. <https://doi.org/10.1186/s13244-022-01279-4>
74. Zhong X, Long H, Su L et al (2022) Radiomics models for preoperative prediction of microvascular invasion in hepatocellular carcinoma: a systematic review and meta-analysis. *Abdom Radiol (NY)* 47(6):2071–2088. <https://doi.org/10.1007/s00261-022-03496-3>
75. Jia LL, Zheng QY, Tian JH et al (2022) Artificial intelligence with magnetic resonance imaging for prediction of pathological complete response to neoadjuvant chemoradiotherapy in rectal cancer: a systematic review and meta-analysis. *Front Oncol* 12:1026216. <https://doi.org/10.3389/fonc.2022.1026216>
76. Park JE, Kim D, Kim HS et al (2020) Quality of science and reporting of radiomics in oncologic studies: room for improvement according to radiomics quality score and TRIPOD statement. *Eur Radiol* 30(1):523–536. <https://doi.org/10.1007/s00330-019-06360-z>
77. 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 EuSoMI. *Insights Imaging* 14(1):75. <https://doi.org/10.1186/s13244-023-01415-8>
78. Holzinger A, Haibe-Kains B, Jurisica I (2019) Why imaging data alone is not enough: AI-based integration of imaging, omics, and clinical data. *Eur J Nucl Med Mol Imaging* 46(13):2722–2730. <https://doi.org/10.1007/s00259-019-04382-9>
79. EQUATOR network (2022) <https://www.equator-network.org>. Accessed 1 Nov 2022.
80. Grading of Recommendations, Assessment, Development and Evaluation (GRADE) Working Group (2022) GRADE Handbook. <https://gdt.gradepro.org/app/handbook/handbook.html>. Accessed 1 Dec 2022.

Publisher's Note

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

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)