0931-89430_Auto_Edited_R3.docx

WORD COUNT 4956
TIME SUBMITTED 05-JAN-2024 04:55PM
PAPER ID 105941104

Name of Journal: *World Journal of Gastroenterology*
Manuscript NO: 89430
Manuscript Type: SYSTEMATIC REVIEWS

Current status of magnetic resonance imaging radiomics in hepatocellular carcinoma: A quantitative review with radiomics quality score

Valentina Brancato, Marco Cerrone, Nunzia Garbino, Marco Salvatore, Carlo Cavaliere

Abstract

BACKGROUND

Radiomics is a promising tool that may increase the value of Magnetic Resonance Imaging (MRI) for different tasks related to the management of patients with hepatocellular carcinoma (HCC). However, its implementation in clinical practice is still far, with many issues related to the methodological quality of radiomic studies.

AIM

To systematically review the current status of MRI radiomic studies concerning HCC using the radiomics quality score (RQS).

METHODS

A systematic literature search of PubMed, Google Scholar, and Web of Science databases was performed to identify original articles focusing on the use of MRI radiomics for HCC management published between 2017 and 2023. The methodological quality of radiomic studies was assessed using the RQS tool. Spearman's correlation (ρ) analysis was performed to explore if RQS was correlated with journal metrics and characteristics of the studies. The level of statistical significance was set at $p < 0.05$.

RESULTS

127 articles were included, of which 43 focused on HCC prognosis, 39 on prediction of pathological findings, 16 on prediction of the expression of molecular markers outcomes, 18 had a diagnostic purpose, and 11 had multiple puropses. The mean RQS was 8 ± 6.22, and the corresponding percentage was 24.15% ± 15.25% (ranging from 0.0 to 58.33%). RQS was positively correlated with journal Impact Factor (ρ = 0.36, p =2.98×10-5), 5-years Impact Factor (ρ = 0.33, p =1.56×10-4), number of patients included in the study (ρ = 0.51, p < 9.37×10-10) and number of radiomics features extracted in the study (ρ = 0.59, p < 4.59×10-13), and time of publication (ρ = -0.23, p = 0.0072).

CONCLUSION

Although MRI radiomics in HCC represents a promising tool to develop adequate personalized treatment as a noninvasive approach in HCC patients, our study revealed that studies in this field still lack the quality required to allow its introduction into clinical practice.

Key Words: Hepatocellular carcinoma; Systematic Review; Magnetic Resonance Imaging; Radiomics; Radiomics Quality Score

Brancato V, Cerrone M, Garbino N, Salvatore M, Cavaliere C. Current status of magnetic resonance imaging radiomics in hepatocellular carcinoma: A quantitative review with radiomics quality score. *World J Gastroenterol* 2023; In press

Core Tip: This systematic review aimed at evaluating the status of Magnetic Resonance Imaging (MRI) radiomic studies related to hepatocellular carcinoma (HCC) using the Radiomics Quality Score (RQS) to assess methodological quality. A systematic literature search identified 127 articles covering various steps of HCC management. The mean

RQS was 8 ± 6.22, with significant variation. RQS was significantly correlated with journal Impact Factor, 5-year Impact Factor, the number of patients involved, the number of radiomic features extracted, and the publication year. Despite the potential of MRI radiomics in HCC, its clinical implementation is hindered by a lack of quality in studies in this field.

INTRODUCTION

Medical imaging has progressed over the last few decades from a simple diagnostic tool for diseases to a massive supply of quantitative data free of the normal subjective interpretation that characterizes conventional clinical practice. The introduction of technological advances and the quest for precision medicine have given rise to a new potential branch of research known as "radiomics". Radiomics is a quantitative technique that turns digitized medical pictures into high-dimensional mineable features that may be correlated with clinical endpoints such as pathological findings, treatment response, and survival. Radiomics can also be integrated with other quantitative data, such as genomics and pathomics data, to provide a comprehensive approach to disease[1–4]. As a quantitative analysis of digital images, radiomics has the potential to reveal specific disease characteristics that are otherwise inaccessible to the naked eye using conventional imaging modalities. This method may increase the quantity of clinically relevant data that may be extracted from medical images, offering the possibility of discovering innovative imaging biomarkers for the diagnosis, characterization, and prediction of outcomes in a wide range of diseases, including oncologic diseases[5]. In the field of oncology, the rationale behind radiomics is that biological tumor characteristics might be mirrored by quantifying medical image heterogeneity using extracted radiomic features, encompassing aspects of tumor progression, response to therapeutic interventions, and clinical outcomes. Quantitative imaging has garnered significant interest in the non-invasive detection of tumor heterogeneity, and recent radiomics studies across various oncological fields have

shown a strong association between imaging heterogeneity and the characteristics of solid tumors[6].

Hepatocellular carcinoma (HCC) is one of the leading causes of cancer-related deaths worldwide and poses serious challenges for screening, early diagnosis and treatment firstly because most HCC is diagnosed at an advanced stage when curative treatment options are limited, and also because of its complex heterogeneity at multiple levels: heterogeneity between tumor nodules from the same patient (intertumor heterogeneity), within the same tumor nodule (intratumor heterogeneity) and between patients (interpatient heterogeneity)[7,8]. Furthermore, current clinical practice based on single bioptic or tumor tissue section fails to discover useful biomarkers, and many existing staging systems for HCC are based on postoperative pathological examinations, which cannot aid in preoperative decision-making[9]. In contrast to numerous other solid tumors, HCC can be diagnosed by using distinctive enhancement patterns on dynamic multiphasic CT or MRI, without additional histopathologic confirmation[10,11]. Although imaging plays an important role in the screening, early identification, and management of HCC patients, the imaging evaluation of HCC is still based on subjective interpretation of qualitative imaging descriptors and tumor size estimate, both of which are prone to variability[10,12,13]. Of note, although CT is more generally available, faster, and needs less experience to administer and interpret pictures than MRI, its downsides include radiation exposure and low soft tissue contrast, which demands the use of iodinated contrast agents. The increased soft tissue contrast of MRI, on the other hand, enables for the examination of a range of tissue features that may be relevant in HCC therapy[14,15]. In this context, recent advantages in MRI radiomics can potentially address the urgent need for noninvasive, radiation-free strategies that can aid in the early detection of HCC and preoperative prediction of tumor behavior, as well as address the inherent variability of qualitative imaging descriptors and provide previously unavailable information to obtain a better stratification of HCC patients for a more precise treatment decision making.

Over the last decade, there has been a significant increase in radiomics studies in the field of HCC. Many of these studies have demonstrated the effectiveness of radiomic features for differential diagnosis, grading, predicting microvascular invasion, overall survival, recurrence, and treatment response[16–19]. Nevertheless, radiomics is presently limited to academic literature in the context of HCC, as physicians question its utility due to the absence of a translation from research studies to clinical application. This is attributed, at least in part, to the overall deficiency of streamlined and productive methods for integrating imaging biomarkers into clinical practice[20–22]. Lambin *et al*[2] developed the Radiomics Quality Score (RQS) to provide a standardized evaluation of the radiomics performance, reproducibility, and clinical. The RQS metric system determines the validity and comprehensiveness of radiomics investigations. This tool is modality-independent tool and was designed to assess the methodological quality of radiomics studies. The methodology and analyses of a radiomics study are evaluated based on 16 criteria that reward or penalize, promoting the best scientific practice[2]. Recent research tried to examine the current state of the art in HCC radiomics, stressing the major concepts, clinical applications, and limitations[23–25]. However, it is clear from these research that the bulk of radiomic investigations on HCC have been conducted on CT, with only a few looking into MRI. Furthermore, the quality of science and reporting in HCC MRI radiomics research investigations is mainly unknown.

Hence, the objective of this study was to provide a comprehensive overview of the existing state of MRI radiomic investigations related to HCC. Simultaneously, we aimed to evaluate the methodological quality of each study using the Radiomics Quality Score (RQS) to assess the radiomics analyses conducted in prior publications. The study's goal is to promote the quality of MRI radiomics research studies in HCC as a diagnostic, prognostic, and/or predictive tool, to allow radiomics to become an appropriate medical decision-making tool by facilitating the combined analysis of clinical data and high-throughput imaging features, while taking advantage of the benefits arising from the MRI technique.

MATERIALS AND METHODS

Search strategy and selection criteria

A systematic search was conducted for all published studies exploring the role of MRI radiomics in the field of HCC. PubMed, Web of Science and Google Scholar electronic databases were comprehensively explored and used to build the search. Only studies published in the last six years were selected. The last search was performed on 1st June 2023. The search terms consisted in: ("radiomics" OR "texture" OR "histogram") AND ("MRI" OR "Magnetic Resonance Imaging") AND ("Hepatocellular Carcinoma" OR "HCC"). The literature search was limited to English language publications and studies of human subjects. Two reviewers, after having independently screened the identified titles and abstracts, assessed the full text of articles aiming at exploring MRI radiomics in the field of HCC and that were not review articles. For articles meeting these criteria with full text available, the following further selection criteria had to be fulfilled: involvement of adult patients (age > 18 years); involvement of patients with HCC confirmed by pathology and/or surgery and/or overall analysis combined with medical history, clinical symptoms, and imaging data; presence of information about MRI protocol. Moreover, studies were excluded if they performed analyses on mixed patients (e.g., groups of patients with multiple hepatic malignant diseases) that did not allow conclusions to be drawn only about HCC patients; if they did not evaluate an outcome measure; if they were focused only on semantic imaging features (radiologist-dependent). After selecting the studies that met the inclusion and exclusion criteria, reference lists of these studies were also searched in order to recruit any potential eligible studies. In addition, pre-existing reviews/systematic reviews/meta-analyses were also searched in order to recruit any other potentially eligible studies from their reference lists.

Planning and conducting the review

After the above-mentioned selection procedure, selected articles were analysed by two reviewers, and data useful for conducting the systematic review were collected in a predesigned sheet. Extracted data will include the following: first author name,

publication year, Journal name, scientometric indexes (Impact Factor (IF), 5-years IF, CiteScore, H-index, first author IF with and without self-citations), study design, in particular prospective/retrospective, clinical purpose, specific output measured in the study, number and type of patients, imaging modalities used for radiomic feature extraction, information on ROI placement (segmentation technique and ROI type), software used for radiomic feature extraction, number and features type, feature selection methods (if used), classification methods, information on if models were applied to a separate dataset, highest accuracy/most important results and main findings.

This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement[26] (see Supplementary Materials for PRISMA Checklist).

Quality assessment with radiomics quality score

The methodological quality of each radiomics study was assessed by two reviewers using the RQS tool[2]. The assessment was performed independently, and any disagreement was resolved by consensus. RQS tool is composed of 16 items structured to assess various crucial steps in the workflow of radiomics analyses (see Supplementary Table S1). In particular, a maximum of 36 points can be assigned to each study: up to 2 points for the first RQS checkpoint (a single item, namely "Image protocol quality"), up to 3 points for the second RQS checkpoint (3 items, specifically on multiple segmentation strategies, the use of phantoms and multiple imaging time points) and up to 31 points for the third RQS checkpoint (12 items, encompassing feature extraction, exploratory analysis design as well as model building and validation). The total score ranges between −8 and 36 and can be translated into a final 0–100 RQS percentage, with −8 to 0 defined as 0%, indicating the lowest quality, and 36 as 100%, indicating the highest quality in terms of the methodology and reporting standards of the radiomics study[2].

Correlation analysis between RQS and journal metrics

Spearman's correlation (ρ) analysis was performed to explore if there was a correlation between RQS and journal metrics, comprising Impact Factor (IF) of the journal at the year of publication, 5-Year IF, CiteScore, and H-index at the year of publication. Additionally, Spearman's correlation was used to explore the correlation between RQS and H-index of the first author at the year of publication of the study (both with and without self-citations), time of publication (calculated as time between the publication date and the date of last literature research, in months), as well as the association with the number of patients involved and the number of radiomic features extracted in the study. Finally, to explore if there was a difference in RQS according to clinical purpose of the study, a subgroup analysis using Kruskal-Wallis H test was performed. In case of significance, Wilcoxon rank-sum post hoc tests with Bonferroni correction were carried out on each pair of groups. The significance level was set at 0.05. All statistical analysis was performed using SPSS (version 27).

RESULTS

Study selection

A total of 537 articles were identified from scientific electronic scientific databases. Only 211 articles were retained after the removal of duplicates.

We reviewed the titles and abstracts of these records, excluding 59 due to non-compliance with inclusion criteria (29 unrelated to the topic, 16 were reviews, 5 conducted analyses on mixed patients, and 9 did not assess an outcome measure). The full text of 149 articles was assessed, leading to the exclusion of 16 off-topic articles. Additionally, four studies were excluded for not evaluating an outcome measure, and two for analyzing mixed patients. Thirteen more articles were found through references in selected articles or existing reviews/systematic reviews/meta-analyses, and seven of these were incorporated into the review. A total of 127 data sets were included in the review. Figure 1 shows the PRISMA flow diagram of the included studies based on the inclusion and exclusion criteria.

Characteristics of included studies

The details regarding the characteristics of the 127 studies chosen for this review are presented in Table 1. Approximately half of these studies (51 out of 127) were published in the last two years, and only 9 studies deviated from a retrospective design. Most of the selected studies (43 out of 127) explored radiomic approaches for HCC prognosis after surgical, radiofrequency ablation and/or trans-arterial chemo embolization (TACE) treatment. 40 studies investigated the ability of radiomics in predicting pathological findings (e.g. microvascular invasion – MVI –, VECT, histologic grade), of which 27 aimed at investigating the performance of radiomics analysis for MVI prediction. 16 studies aimed at exploring if MRI radiomics could infer the expression of molecular markers (e.g. CK19, Ki67, GPC3) outcomes. Among the remaining studies, 24/127 aimed to evaluate the power of radiomics for distinguishing HCC from other solid hepatic lesions, while 11 had multiple aims.

The number of total included patients was 18.949, with a sample size varying from 17 to 602 patients (median: 309.5). Most studies (96 out of 127) explored more than one phase/sequence to perform radiomic analysis. Most studies (106 out of 127) performed 3D segmentation. In 114 of them, segmentation was manually performed, while in the remaining studies was used a semiautomatic (12 studies) or automatic (2 study) segmentation approach. Concerning software used for feature extraction, PyRadiomics was the most popular (used in 42 out of 127 studies), followed by AK software (used in 23 out of 127 studies) and Matlab (used in 19 out of 127studies). The number of radiomics features extracted from each phase/sequence ranged from 3 to 3144 (mean: 68 ± 206). Shape features were extracted in 55 out of 127 studies, first-order features in all but three studies, textural features in 82/127 studies, and features from filtered images (e.g. wavelet, Laplacian of gaussian) in 34 out of 127 studies. Concerning feature selection algorithms, the Least Absolute Shrinkage and Selection Operator (LASSO) regression was the most widely used (used in 55 out of 127 studies). Other frequently used algorithms for feature selection were ICC (used in 25 studies), correlation (used in 12 studies) and minimum redundancy maximum relevancy – mRMR – (used in 9 studies). The performance metrics of the studies, when present, corresponded to

accuracy in 9 out of 127 studies, AUC in 99 out of 127 studies and to C-index in 12 out of 127 studies. Most studies involved machine learning techniques for radiomic analysis, of which 51 splitted the subjects into training and test cohort to test the prediction models performance. Further details on these characteristics can be found in Table 1 and Supplementary Table S2.

Quality assessment with radiomics quality score

Supplementary Table S3 provides the RQS details of all included studies. The average total RQS score was 8 ± 6.22, corresponding to a percentage of 24.15% ± 15.25%, with a range from 0.0 to 58.33% (refer to Figure 2). Concerning the first RQS checkpoint, nearly all studies, excluding ten, provided thorough documentation of the imaging protocol, yet none achieved the maximum points for utilizing a public protocol. In relation to the second RQS checkpoint (items 2 to 4), a majority of studies (84.25%, 107 out of 127) employed multiple segmentation, mainly by different radiologists, but none of the articles met the requirement for 'imaging at multiple time points' and only one article met the requirement for a 'phantom study'. With respect to the third RQS checkpoint (items 5 to 16), feature reduction techniques were applied in all but 15 studies (88.28%). Multivariable analysis with non-radiomics features was performed in 85 studies (66.92%) of the 128 included articles. However, only 40 (31.25%) identified and discussed biological correlates and only 50 (39.06%) provided cut-off analysis.

Of the 127 studies included, almost all (123) reported discrimination statistics and their statistical significance. About a quarter of these studies used resampling techniques. However, only 58 studies reported calibration statistics, and none of them applied resampling techniques.

A significant proportion (39.37%) of the studies (50 out of 127) did not provide any validation of their results. Only three studies validated their results using one external validation cohort and five studies used two external validation cohorts. Furthermore, only 47 out of 127 research examined the clinical utility of the produced model using decision curve analysis, while 42 out of 127 studies compared radiomics models with the particular gold standard (based on the study purpose).

Lastly, no study disclosed code and data to the public or performed a cost-effectiveness analysis.

Correlation analysis between RQS and journal metrics

A significant positive correlation was found between RQS and journal Impact Factor ($\rho = 0.36$, $p = 2.98\times10^{-5}$), 5-years Impact Factor ($\rho = 0.33$, $p = 1.56\times10^{-4}$), number of patients involved ($\rho = 0.51$, $p < 9.37\times10^{-10}$) and number of radiomics features ($\rho = 0.59$, $p < 4.59\times10^{-13}$) extracted in the study. On the other hand, there was a significant negative correlation between RQS and time between the publication and the performed literature research ($\rho = -0.23$, $p = 0.0072$) and there were no statistically significant differences identified in the RQS among studies with different objectives.

DISCUSSION

In this systematic review, we aimed at summarizing the current status of the fast-growing research on MRI radiomics for the management of HCC. We explored whether it could offer diagnostic, prognostic, and predictive information about pathological outcomes and molecular expression. Additionally, we assessed the quality of the science and reporting across the studies using the RQS tools. 127 studies from November 2017 onwards were examined in our study. Despite promising results obtained from each of them (with best AUC and C-indexes reaching 0.98 and 0.94, respectively), our study revealed that the methodological variability of the research is considerable, and the reporting quality is insufficient.

Mean RQS was 8 out of 36, with a mean percentage RQS of 24.15%. These results are consistent with previously published data on a variety of tumors, including prostate, breast, lung, renal, and brain cancer[152–156]. Recent studies evaluating research quality in HCC radiomics also align with our findings[25,157]. However, direct comparison with our study is not possible due to differences in purpose and inclusion criteria.

The results of our analysis showed that the poor RQS scores of the included studies were mostly caused by the absence of rigorous procedures pertaining to radiomics workflow.

Regarding RQS checkpoint 1, practically all investigations have a thorough documentation of the imaging methodology. Nevertheless, the lack of public image methods in the investigations negatively impacts the radiomic studies' repeatability and reproducibility. Notably, the CE-T1WI MRI sequence emerged as the most extensively explored, given its primary role in preoperative HCC assessment. Nevertheless, there exists variability in MRI acquisition due to differences in manufacturers, scanning protocols, contrast media, and phases employed. A significant diversity across the included studies was also noted in terms of RQS checkpoint 2. Specifically, 108 out of 127 studies adopted multiple segmentations to mitigate bias arising from segmentation variability. It's crucial to highlight, however, the lack of consistency among studies regarding the type of ROI (2D/3D) and the segmentation method used (manual, semi-automatic, automatic). It is worth mentioning that the majority of studies used manual or semi-automated image segmentation with manual correction, which restricts the studies included. Both manual and semi-automatic segmentation can introduce significant observation bias, which may affect studies on intra- and interobserver variation in ROI/VOI delineation[1].

None of the studies determined scanner/manufacturer variability or collected images at multiple time points, making it difficult to detect potential feature variability between scanners and manufacturers, as well as temporal variability. Positively, all but twelve studies performed feature reduction, which is consistent with the third RQS checkpoint. In fact, excessive dimensionality of features can negatively affect model performance and lead to overfitting[158]. The RQS showed high variability in items 6, 7 and 8. However, it is important to note that these items are highly dependent on the aim of the study.

Another notable finding from our review was that only nine of the studies in the review were prospective studies, which is the highest weighting in the RQS tool. This constitutes a significant drawback in radiological studies since a meticulously planned prospective trial serves to diminish and control potential confounding factors, thereby offering a superior level of evidence regarding the trial's quality. This elucidates the

rationale behind assigning the highest weight (7 points) to studies with a prospective design in the RQS tool, representing approximately 20% of the total score. Thus, this limitation highlights the importance of conducting well-designed prospective studies.

It is noteworthy that nearly half of the examined papers lacked outcome validation which increases the risk of false-positive results and hinders the implementation of radiomics in clinical practice. However, approximately half of the studies that did not validate their results with an independent cohort chose to perform cross-validation.

The majority of the studies did not provide open access to their data sets, segmentations or codes, which limits the ability to verify and reproduce their results [159,160].

Cost-effectiveness analyses that evaluate radiomic prediction models from a health economic perspective when applied in clinical practice have the same limitation. The assumption is that a new predictor should be no more costly than existing predictors, given comparable accuracy. In addition, the health impact of a radiomics predictor is compared to a condition in which no radiomic predictor[2,157]. However, this criterion of RQS is not as important as the need to standardize and validate the models.

As far as we are aware, this is the first systematic review that looks into the possibility of employing MRI radiomics to gather information regarding the management of HCC and to assess studies using the RQS tool.

Previous studies evaluated the quality of radiomic analysis in different studies for different oncologic applications[152–156,161]. Similar to our study, Wakabasaky *et al*[25] assessed whether radiomics is a valuable and reproducible method for clinical management of HCC using RQS. However, their work included studies up to 2018 and was not focused on MRI modality. In addition, Wang *et al*[157] also aimed to assess the methodological quality of radiomics studies for HCC management. However, although similar findings with respect to our study were found (mean RQS of 10), their study was focused on the prediction of MVI in HCC patients and also included studies involving other imaging modalities than MRI.

I contrast to most studies that focus on assessing the quality of radiomic studies by means of RQS, our approach involved exploring the potential correlation between RQS

and scientometric indixes. Our findings revealed that publications that have higher RQS were published in journals with higher IF and 5-years IF. However, studies with high/Low RQS and low/high IF and 5-years IF were also found. Moreover, although no significant correlation was found, it was observed that RQS tended to increase with time (decreasing number of months passed from literature research). Interestingly, we discovered that the quality of included studies increased as the number of included patients and extracted attributes grew.

It is crucial to underscore that only 45 out of 127 studies referenced the Image Biomarker Standardization Initiative (IBSI) guidelines or utilized software for radiomic feature extraction compliant with IBSI standards (e.g., PyRadiomics). Emphasizing the importance of adhering to standardized radiomic features nomenclature and calculation according to IBSI, our study highlights the need for future research to align with these standards, thus enhancing the reproducibility of scientific researches[162].

Despite the insights gained, our study is not without limitations. The RQS scoring system, as acknowledged in prior research, is not a definitive standard for evaluating radiomics studies and requires ongoing refinement for widespread acceptance in radiology. The existing research is limited by issues including conducting phantom studies across all scanners, applying imaging at multiple time points, and lacking definition for a particular study purpose[163,164]. Additionally, the predominantly retrospective nature of the included studies introduces bias, compounded by the absence of external validation cohorts and comparisons with reference standards, hindering conclusive remarks on the efficacy of MRI radiomics in HCC[165,166]. Variability in sample size, inclusion criteria, and methodological settings across studies precluded a meta-analysis aligned with study objectives. Furthermore, the study did not explore specific shared radiomic features among different studies, considering the wide-ranging variability in imaging protocols and software for feature extraction.

CONCLUSION

In summary, despite the potential of recent developments in MRI radiomics to fulfill the urgent requirement for noninvasive, radiation-free, and quantitative approaches to support decision-making in HCC treatment, the current studies in this domain lack the requisite quality for integration into clinical practice. Emphasizing the significance of external validation, addressing concerns related to feature reproducibility, conducting clinical utility analyses, and fostering scientific openness are crucial steps that need to be addressed. This endeavor aims to provide fresh perspectives and contribute to the establishment of a consensus regarding the application of the radiomic method in assessing HCC.

ARTICLE HIGHLIGHTS

Research background

Radiomics is a promising tool that may increase the value of Magnetic Resonance Imaging (MRI) for different tasks linked to the management of patients with hepatocellular carcinoma (HCC).

Research motivation

Over the last decade, there has been a substantial increase in radiomics studies in the field of HCC. Many of these studies have demonstrated the power of radiomic features for differential diagnosis, grading, predicting microvascular invasion, overall survival, recurrence, and treatment response. However, the use of radiomics in HCC is currently limited to academic literature, and no studies have yet been translated into clinical applications. This has led to doubts among clinicians about the radiomics validity. This is in part due to many issues related to the methodological quality of radiomic studies.

Research objectives

To summarize the status of MRI radiomic studies concerning HCC, using the radiomics quality score (RQS) to assess the quality of the methodology used in each study.

Research methods

We systematically reviewed PubMed, Google Scholar, and Web of Science databases to identify original articles focused on using MRI radiomics for HCC management published between 2017 and 2023. The RQS tool was employed to evaluate the methodological quality of radiomic studies. Spearman's correlation (ρ) analysis was conducted to investigate the association between RQS and journal metrics, as well as the characteristics of the studies. The threshold for statistical significance was established at $p < 0.05$.

Research results

127 articles were included, of which 43 focused on HCC prognosis, 39 on prediction of pathological findings, 16 on prediction of the expression of molecular markers outcomes, 18 had a diagnostic purpose, and 11 had multiple aims. Mean RQS was 8 ± 6.22, with the corresponding percentage of 24.15% ± 15.25% (ranging from 0.0 to 58.33%). RQS was positively correlated with journal Impact Factor ($\rho = 0.36$, $p = 2.98\times10^{-5}$), 5-years Impact Factor ($\rho = 0.33$, $p = 1.56\times10^{-4}$), number of patients involved ($\rho = 0.51$, $p < 9.37\times10^{-10}$) and number of radiomics features ($\rho = 0.59$, $p < 4.59\times10^{-13}$) extracted in the study, and time of publication ($\rho = -0.23$, $p = 0.0072$).

Research conclusions

Although the MRI radiomics in HCC represents an auspicious tool for developing adequate personalized treatment as a noninvasive approach in HCC patients, our study revealed that studies in this field still lack the quality required to allow its introduction in clinical practice.

Research perspectives

Although recent advantages in MRI radiomics can potentially satisfy the urgent need for noninvasive, radiation-free and quantitative strategies that can aid in HCC treatment decision making, studies in this field still lack the quality required to allow its

introduction in clinical practice. Future studies including external validation and adhering to the standardization of radiomics features are necessary. Moreover, limitations and challenges related to feature reproducibility, analysis of the clinical utility, and openness of science need to be addressed. This work may provide new insights and contribute to a common understanding of the use of radiomics in the assessment of HCC.

0931-89430_Auto_Edited_R3.docx

ORIGINALITY REPORT

5%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet	107 words — 2%
2	mdpi-res.com Internet	40 words — 1%
3	bmccancer.biomedcentral.com Internet	36 words — 1%
4	worldwidescience.org Internet	15 words — < 1%
5	www.mdpi.com Internet	15 words — < 1%
6	www.alnap.org Internet	14 words — < 1%
7	elight.springeropen.com Internet	12 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 12 WORDS
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 12 WORDS