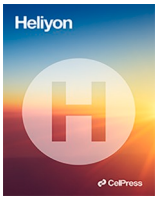
Heliyon 10 (2024) e36313

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Systematic review and meta-analysis on the classification metrics of machine learning algorithm based radiomics in hepatocellular carcinoma diagnosis

Nurin Syazwina Mohd Haniff [a], Kwan Hoong Ng [b], Izdihar Kamal [a,c,*], Norhayati Mohd Zain [c], Muhammad Khalis Abdul Karim [a]

[a] Department of Physics, Faculty of Science, Universiti Putra Malaysia, UPM, 43400 Serdang, Selangor, Malaysia
[b] Department of Biomedical Imaging, Universiti Malaya, 50603 Kuala Lumpur, Malaysia
[c] Research Management Centre, KPJ Healthcare University, 71800, Nilai, Negeri Sembilan, Malaysia

ARTICLE INFO

Keywords:
ystematic review
PRISMA
Radiomics
Machine learning
Hepatocellular carcinoma
Meta-analysis

ABSTRACT

The aim of this systematic review and meta-analysis is to evaluate the performance of classification metrics of machine learning-driven radiomics in diagnosing hepatocellular carcinoma (HCC). Following the PRISMA guidelines, a comprehensive search was conducted across three major scientific databases—PubMed, ScienceDirect, and Scopus—from 2018 to 2022. The search yielded a total of 436 articles pertinent to the application of machine learning and deep learning for HCC prediction. These studies collectively reflect the burgeoning interest and rapid advancements in employing artificial intelligence (AI)-driven radiomics for enhanced HCC diagnostic capabilities. After the screening process, 34 of these articles were chosen for the study. The area under curve (AUC), accuracy, specificity, and sensitivity of the proposed and basic models were assessed in each of the studies. Jamovi (version 1.1.9.0) was utilised to carry out a meta-analysis of 12 cohort studies to evaluate the classification accuracy rate. The risk of bias was estimated, and Logistic Regression was found to be the most suitable classifier for binary problems, with least absolute shrinkage and selection operator (LASSO) as the feature selector. The pooled proportion for HCC prediction classification was high for all performance metrics, with an AUC value of 0.86 (95 % CI: 0.83–0.88), accuracy of 0.83 (95 % CI: 0.78–0.88), sensitivity of 0.80 (95 % CI: 0.75–0.84) and specificity of 0.84 (95 % CI: 0.80–0.88). The performance of feature selectors, classifiers, and input features in detecting HCC and related factors was evaluated and it was observed that radiomics features extracted from medical images were adequate for AI to accurately distinguish the condition. HCC based radiomics has favourable predictive performance especially with addition of clinical features that may serve as tool that support clinical decision-making.

1. Introduction

Hepatocellular carcinoma (HCC) is a form of liver cancer that is among the main causes of cancer-related fatalities globally [1]. Despite the availability of hepatectomy surgery, liver transplantation, radiofrequency ablation, and chemotherapy, the survival rate of

* Corresponding author. Department of Physics, Faculty of Science, Universiti Putra Malaysia, UPM, 43400 Serdang, Selangor, Malaysia.
E-mail address: izdiharkamal@upm.edu.my (I. Kamal).

https://doi.org/10.1016/j.heliyon.2024.e36313
Received 21 September 2023; Received in revised form 13 August 2024; Accepted 13 August 2024
Available online 14 August 2024

HCC patients can be improved significantly through early detection and treatment [2,3]. With the emergence of precision medicine, the diagnosis of HCC and other types of cancer has become much more accurate and efficient. Diagnosing and treating cancer requires medical imaging, and radiologists use a selection of modalities including US, CT, and MRI to identify the issue based on their visual assessment of the images. CT and MRI have been demonstrated to be more sensitive than US, however US is still beneficial as recent research has uncovered the potential of radiomic analysis of US images for early diagnosis, prognosis, and prediction of HCC [4].

Radiomics is a rapidly developing field which involves the extraction and analysis of numerical image features derived from medical imaging data [5–7]. It incorporates state-of-the-art image analysis techniques to extract and, together, adapt the machine learning (ML) models to analyse a broad range of imaging features that quantify tumour phenotypic characteristics [8]. These features can furnish valuable information about the tumour's structure, texture, and spatial relationships, which can be used to enhance diagnostic, prognostic, and predictive accuracy in a variety of medical setting. It has been employed for diagnosis, characterization, and treatment planning of different types of cancer, including lung cancer [9], breast cancer [10], pancreatic cancer and hepatocellular carcinoma [7,10–13]. This extraction method growing in popularity as they extract information that is not visible to the naked eye, providing more detailed information about a specific disease and accounting for tumour heterogeneity [14–16]. Several researchers has demonstrated the reproducibility and repeatability of radiomics across various methods, as well as its ability to improve diagnostic accuracy through machine learning [17–22]. Radiomics has also been shown to be effective in constructing prediction models for early recurrence of HCC and in distinguishing HCC from non-HCC with higher accuracy using a combined model incorporating clinical factors [23,24].

Artificial Intelligence (AI) has been extensively utilised in radiomics, thus enhancing its capabilities and potential uses. AI has been developing over time and is renowned for its capacity to refine tumour assessment and treatment planning in oncology. In recent years, a range of techniques have been developed, including those based on ML and deep learning (DL) models. Feature selection is a method used for classification models to reduce data dimensionality and eliminate redundant features, which can improve the model's performance. According to Shan et al. least absolute shrinkage and selection operator (LASSO) is one of feature selection that can increase the model performance [23]. Dimensions reduction is one of the methods used to avoid overfitting of constructed model. In 2021, Liu et al. generate reliable prediction model by using Principal Component Analysis (PCA) to reduce input features from 1419 to 20 principal components [25]. Dai et al. (2021) also highlight that combination feature selection is method can be considered improves the performance of model [26]. Liao et al. (2020) provided a new apporach to detect HCC using both stratified 5-fold cross validation mthod and genetic algorithm [3]. A fully automated ML also shown to be efficient in diagnosing HCC and predicts patients' survival outcomes.

The potential of ML and DL techniques for the diagnosis of HCC has garnered increasing attention in recent years [16,25,27,28]. However, a comprehensive evaluation of the current state of knowledge on the use of these techniques for HCC diagnosis and identification of areas for future research is lacking. In this review, we evaluated the application of ML and DL algorithms for the diagnosis of HCC and evaluated their performance.

2. Methodology

2.1. Literature search strategy

We conducted a literature search to identify articles published in English that pertain to algorithms and radiomics used for classifying hepatocellular carcinoma (HCC) by searching three databases: PubMed, ScienceDirect, and Scopus. The publication range for this search was between 2018 and 2023. The literature search was executed and adhered to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines and recommendations [29]. A search for relevant articles was conducted online using description and three main keywords, 'algorithms', 'Hepatocellular Carcinoma', and 'radiomics'. After filtering by title and abstract, studies irrelevant to the research topic, systematic reviews and incomplete articles were removed. The full text was then filtered for the types of data used on the predictive models. Studies that used biopsy or genotypes as their features were excluded to focus only on studies using radiomic features as input data.

2.2. Inclusion and exclusion criteria

Data were extracted based on the following characteristics in each study: year of publications; demographics such as imaging techniques, extracted features and algorithms used to classify HCC including types of feature selections. Image processing techniques was also included in this study. Inclusion criteria for quantitative meta-analysis include studies that provide: (i) patients underwent US, CT and MRI examination prior diagnosed with HCC; (ii) implementation of HCC segmentation using either manual, semi- or automatic segmentation;

(iii) application of radiomics features models in the prediction of HCC; (iv) presence of feature selector and usage of machine learning and deep learning for classifying HCC; (v) performance metrics can be directly or indirectly extracted from the reported results to evaluate the predictive model's performance. Nevertheless, exclusion criteria include studies with: (i) features extracted from biopsy (genomic features); (ii) phantom study; (iii) manuscript written in other language than English; (v) studies with unextractable data.

Table 1
List of the relevant information and element for the systematic review and meta-analysis.

Extraction Element	Contents	Type
1	Author	The authors of articles
2	Algorithm	Types of machine learning and feature selection used
3	Datasets	Sample size and amount of data used for each training, test and validation sets
4	Types of modalities	The types of imaging modalities used in the dataset.
5	Features extracted	Types of features extracted.
6	Performance evaluation	Accuracy, specificity, sensitivity and AUC of the model used in the articles

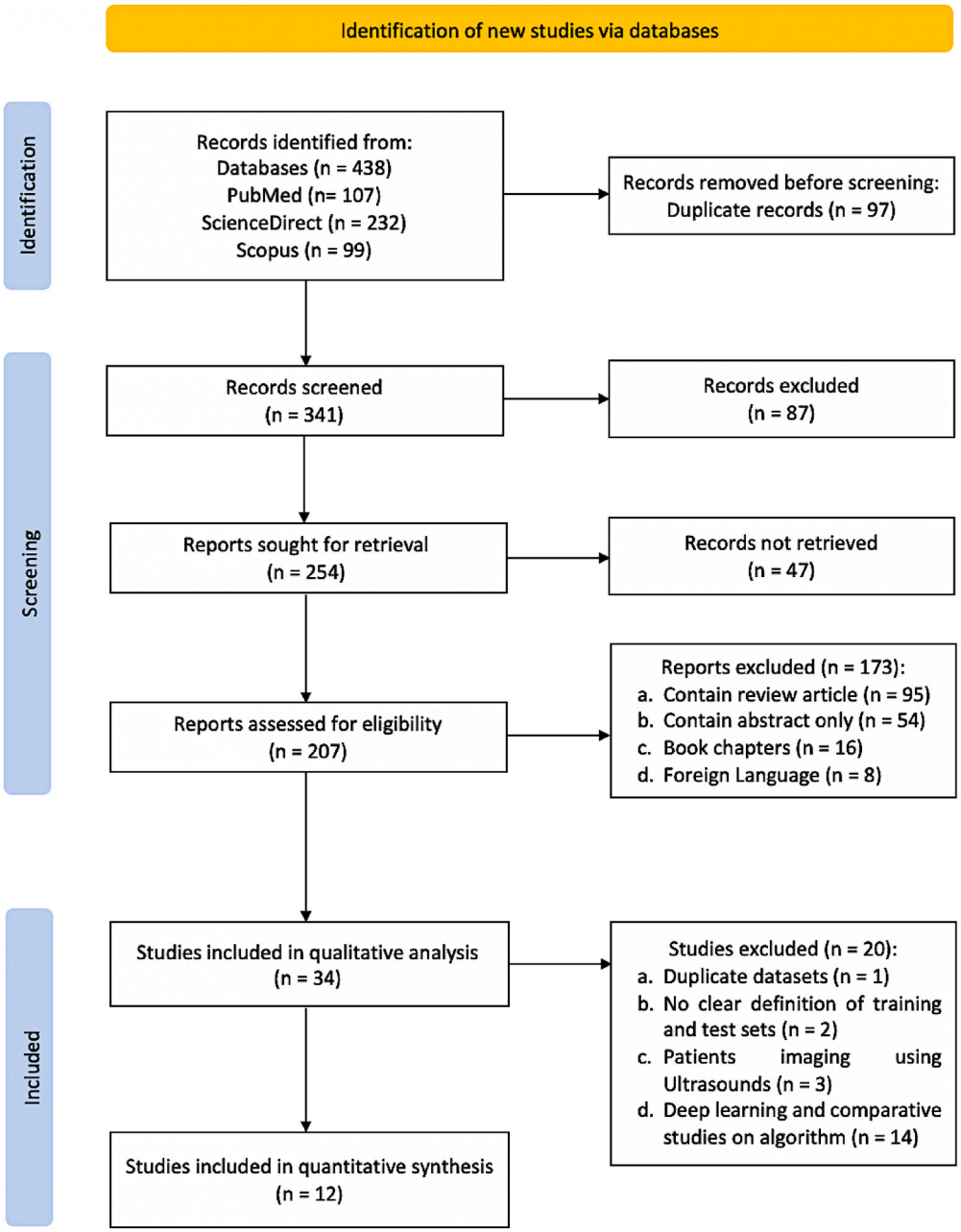


Fig. 1. Flow chart shows the approach used to identify eligible studies based on the PRISMA strategy.

Table 2
Algorithm, datasets, modalities and classification metrics performance of the included studies.

Author	Year	Algorithm	Datasets				Types of Modalities	Types of Radiomic features extracted	Performance Evaluation
			Sample size	Training	Test	Validation			
Nie et al. [1]	2020	Multivariate Logistic Regression (LR) -LASSO	156	119	37	–	CT [Triphasic contrast CT image]	i. Intensity statistics features ii. Shape features iii. Texture features iv. Filter & Wavelet features	**Radiomics Signature** Accuracy = 86.49 Sensitivity = 76.92 Specificity = 91.67 AUC = 0.865 **Radiomics Nomogram** Accuracy = 89.19 Sensitivity = 92.31 Specificity = 87.50 AUC = 0.917
Dong et al. [4]	2019	Support Vector Machine (SVM) -Sparse Representation	42	–	–	–	US	i. Ultrasound feature map: • DEA feature map (time-domain feature) • SDSD feature map (frequency-domain feature) • OND feature map (statistical feature) ii. Texture features iii. Wavelet features	**DOSM** Accuracy = 0.9286 Sensitivity = 0.8571 Specificity = 100 AUC = 0.9501 **DOM** Accuracy = 0.881 Sensitivity = 0.8095 Specificity = 0.9524 AUC = 0.9184 **DM** Accuracy = 0.8571 Sensitivity = 0.8095 Specificity = 0.9048 AUC = 0.9093 **GM** Accuracy = 0.8333 Sensitivity = 0.8095 Specificity = 0.8571 AUC = 0.8594
Mokrane et al. [20]	2020	i. K- Nearest Neighbour (KNN) ii. Support Vector Machine (SVM) iii. Random Forest (RF) -SMOTE Course-to-fine strategy (feature selection)	178	106	36	36	CT (Triphasic CT)	i. Standardized subtraction (delta 1) ii. Direct subtraction (delta 2) iv. Relative subtraction (delta 3)	Sensitivity = 0.7 Specificity = 0.54 AUC = 0.66 (95%CI 0.64–0.84)
Ding et al. [24]	2021	i. Max-Relevance and Min-Redundancy (mRMR) ii. Random Forests (RF) iii. LASSO	224	156	68	–	MRI [Contrast-enhanced]	i. First Order Statistics ii. Gray-level dependence matrix (GLDM) iii. Gray-level co-occurrence matrix (GLCM)	**Radiomics** Accuracy = 86.8 % Sensitivity = 88.9 % Specificity = 82.6 % AUC = 0.931

(*continued on next page*)

Table 2 (*continued*)

Author	Year	Algorithm	Datasets				Types of Modalities	Types of Radiomic features extracted	Performance Evaluation
			Sample size	Training	Test	Validation			
								iv. Gray-level run length matrix (GLRLM) v. Gray-level size zone matrix (GLSZM)	**Combined Model** Accuracy = 94.1 % Sensitivity = 93.3 % Specificity = 95.7 % AUC = 0.927
Liu et al. [25]	2021	Support Vector Machine (SVM) [kernel type: c and gamma] Principal Component Analysis (PCA) Feature reduction	85	–	–	–	CT & MRI	i. First Order Statistics ii. 2D Shape feature iii. Gray-level Co-occurrence matrix (GLCM) (Total 1419 features)	Sensitivity = 0.68 Specificity = 0.88 AUC = 0.81
Dai et al. [26]	2021	i. Support Vector Machine (SVM) ii. Logistic Regression (LR) iii. Random Forest (RF) iv. Gradient Boosting Decision Tree (GBDT) SVM-RFE, mRMR, LASSO, LASSSO-RFE	69	68	1	–	MRI	i. Shape feature ii. Texture feature Intensity feature	**LASSO-RFE LR** Accuracy = 79.7 % Sensitivity = 82.8 % Specificity = 77.5 % AUC = 0.85 **GBDT** Accuracy = 87 % Sensitivity = 93.1 % Specificity = 82.5 % AUC = 0.895
Bousabarah et al. [28]	2021	i. Deep Convolutional Neural Network (DCNN) ii. Cluster Threshold (TR) Random Forest (RF)	174	122	26	26	MRI	i. Arterial features ii. Venous features iii. Delayed features iv. Textural feature Shape features	**DCNN** Sensitivity = 0.73 Specificity = 0.55 **DCNN + TR + RF** Sensitivity = 0.75 Specificity = 0.66
Ji et al. [45]	2019	Multivariate Cox Regression Backward step-wise elimination with Akaike Information Criteria (AIC)	470	210	107	153	CT [Contrast-enhanced]	i. First order statistics ii. Texture feature Wavelet feature	**Internal Validation** AUC = 0.84 **External Validation** AUC = 0.803
Zhang et al. [46]	2020	Multivariable Logistic Regression (MLR) i. LASSO	637	451	111	75	CT [Contrast-enhanced]	i. Intensity feature ii. Texture feature iii. Wavelet feature	**Test** AUC = 0.803 **Validation** AUC = 0.796
Li et al. [32]	2021	Logistic Regression (LR) -LASSO	301	131	113	57	MRI	i. Gray-level histogram ii. Form factor parameters	**Radiomics** Sensitivity = 89.2 % Specificity = 62.3 % AUC = 0.817 **Radiomics-Clinical** (Nomogram): External Validation (n = 133) Sensitivity =

(*continued on next page*)

Table 2 (*continued*)

Author	Year	Algorithm	Datasets				Types of Modalities	Types of Radiomic features extracted	Performance Evaluation
			Sample size	Training	Test	Validation			
									86.5 % Specificity = 75 % AUC = 0.881
Shen et al. [42]	2021	Multivariate Logistic regression = classifier Random forest (RF) = feature selector Combined Model Radiomics with Changes of Serum AFP Level	114	80	34	–	CT	i. Difference-in-difference (DD) Features ii. Histogram parameters iii. Texture parameters iv. Form factor parameters v. Gray level co-occurrence matrix vi. Gray level run-length matrix iii. Gray level size zone matrix	**Radiomics** Accuracy = 86 % Sensitivity = 91 % Specificity = 75 % AUC = 0.89 **Radiomics-clinical** (Combined Model) Accuracy = 86 % Sensitivity = 91 % Specificity = 75 % AUC = 0.89
Nitsch et al. [33]	2021	Random forest (RF)	90	62	28	–	MRI	i. First Order Statistics ii. Shape feature iii. Texture feature LoG filter (sigma 1–5 mm)	AUC = 0.84
Qiu et al. [47]	2019	Support Vector Machine (SVM) 10-fold cross-validation	106	26	57	–	CT [Arterial enhanced]	i. Tumour intensity ii. Shape feature Texture feature	**With feature Reduction** Sensitivity = 86.6 % Specificity = 84 % AUC = 0.857 **Without feature Reduction** Sensitivity = 88.9 % Specificity = 64 % AUC = 0.721
Wu et al. [31]	2019	i. Decision Tree (DT) ii. Random Forest (RF) iii. K-Nearest Neighbour (KNN) iv. Logistic Regression (LR) The variance threshold, select k best and (LASSO) operator algorithms	369 with 446 lesions	295	74	–	MRI (Pre-contrast)	i. First Order features ii. Shape features iii. 2nd Order features (texture) iv. Higher Order Statistics features	Sensitivity = 0.822 Specificity = 0.714 AUC = 0.89
Nie et al. [41]	2021	Multivariate Logistics Regression (LR) -LASSO	131	93	38	–	CT [Triphasic Contrast CT images]	i. Intensity statistics features ii. Shape features iii. Texture features iv. Filter & Wavelet features	**Radiomics** Accuracy = 68.42 % Sensitivity = 56 % Specificity = 92.31 % AUC = 0.75 **Radiomics Nomogram** Accuracy = 92.11 % Sensitivity = 96 %

(*continued on next page*)

Table 2 (*continued*)

Author	Year	Algorithm	Datasets				Types of Modalities	Types of Radiomic features extracted	Performance Evaluation
			Sample size	Training	Test	Validation			
									Specificity = 84.62 % AUC = 0.94
Peng et al. [48]	2018	Multivariable Logistic Regression (LR) LASSO	304	184	120	–	CT [Contrast-enhanced]	i. Intensity Direct ii. Intensity Histogram iii. Texture feature v. Shape feature	Sensitivity = 75.68 % Specificity = 80.43 % AUC = 0.844
Yang et al. [35]	2021	i. Multiple Logistic Regression (MLR) ii. Support Vector Machine (SVM) iii. Random Forest (RF) iv. Artificial Neural Network (ANN) Multivariate logistic regression [CLINICAL MODEL]	257	143	75	36	MRI	i. Intensity feature ii. Texture feature iii. Shape feature iv. Wavelet feature	A = SLH AUC = 0.726 B = LSCH AUC = 0.79
Yang et al. [36]	2021	Multivariable Logistics Regression mRMR LASSO	201	148	53	–	MRI	i. First Order Feature ii. Shape feature iii. Texture feature Wavelet transformed features	**Radiomics** Accuracy = 66 % Sensitivity = 96.2 % Specificity = 55.6 % AUC = 0.788 **Combined Model** Accuracy = 84.9 % Sensitivity = 88.5 % Specificity = 85.2 % AUC = 0.917
Zhao et al. [37]	2021	Multivariate Logistic Regression LASSO Univariate LR	122	85	37	–	MRI [Contrast-enhanced]	i. Histogram feature ii. Gray-level co-occurrence matrix (GLCM) iii. Gray-level run length matrix (GLRLM) iv. Gray-level size zone matrix (GLSZM) v. Haralick feature vi. Form factors Gaussian transformed feature	Accuracy = 0.73 Sensitivity = 0.833 Specificity = 0.632 AUC = 0.833
Liang et al. [40]	2020	i. Random forests (RF) ii. Artificial Neural Network (ANN) iii. Ridge Regression (RR) iv. Fusion Model (Multivariate LR) -SMOTE	307	205	102	–	CT & MRI [Contrast enhanced]	i. Texture features ii. Wavelet features	**CT** RR, AUC = 0.731 RF, AUC = 0.879 ANN, AUC = 0.763 **MRI** RR, AUC = 0.736 RF, AUC = 0.925 ANN, AUC = 0.769 Fusion Model AUC, CT = 0.966

(*continued on next page*)

Table 2 (*continued*)

Author	Year	Algorithm	Datasets				Types of Modalities	Types of Radiomic features extracted	Performance Evaluation
			Sample size	Training	Test	Validation			
									AUC, MRI = 0.971
Wang et al. [49]	2019	i. Deep Convolutional Neural network (DCNN) ii. Support Vector Machine (SVM) iii. Random Forest (RF) Used Combined Models	167	150	17	–	CT (Multi-phase)	i. High-level temporal & spatial features	AUC = 0.825
Jiang et al. [43]	2021	i. 3D-CNN (Convolutional Neural network) [Deep Learning] ii. XGBoost iii. Models [XGBoost, Radiological model, Radiomics model, Radiological + Radiomics + Clinical Model]	405	324	81	–	CT	**Radiomic features:** i. First order statistics ii. Second order statistics iii. Higher order statistics **Radiological features:** i. Liver morphology ii. Number of hepatic lobes involved iii. Numbers of tumours iv. Peritumoral satellite nodule **Clinical features:** i. Age ii. Sex iii. Background liver disease iv. Diabetes v. Surgery type vi. MELD score	**Radiomic model** Accuracy = 84 % Sensitivity = 90.9 % Specificity = 75.7 % AUC = 0.88 **3D-CNN model** Accuracy = 85.2 % Sensitivity = 93.2 % Specificity = 75.7 % AUC = 0.906
Gao et al. [34]	2021	H-DAR-net (Combination of triplet CNN and simple SE-DenseNet) LASSO	225	168	57	–	MRI [Non-contrast T2 weighted]	i. First Order Statistics (Intensity) ii. Texture feature iii. Wavelet feature	Accuracy = 0.785 Sensitivity = 0.795 Specificity = 0.738 AUC = 0.826
Mao et al. [44]	2020	XGBoost	297	237	60	–	CT [multi-phasic contrast enhanced)	i. First Order Statistics ii. Texture feature iii. Shape feature iv. Wavelet filter	**Radiomics** Accuracy = 68.33 % Sensitivity = 47.83 % Specificity = 81.08 % AUC = 0.7579 **Radiomics + clinical factors** Accuracy = 70 % Sensitivity = 65.22 % Specificity = 72.97 % AUC = 0.8014
Wu et al. [50]	2020	Logistic regression (LR) Sequential forward selection	74	–	–	–	CT	Texture feature	Sensitivity = 0.963 Specificity = 0.75 AUC = 0.836

(*continued on next page*)

Table 2 (*continued*)

Author	Year	Algorithm	Datasets				Types of Modalities	Types of Radiomic features extracted	Performance Evaluation
			Sample size	Training	Test	Validation			
Wang et al. [38]	2021	Support Vector Machine LASSO	235	165	70	–	US [Contrast-enhanced]	i. Gray-level histogram ii. Shape feature iii. Gray-level co-occurrence matrix (GLCM) iv. Gray-level run-length matrix (GLRLM) v. Gray-level size zone matrix (GLSZM) vi. Co-occurrence of local anisotropic gradient orientations (CoLIAGe) vii. Wavelet transformed texture	**Ultrasomics** Accuracy = 67.1 % Sensitivity = 73.7 % Specificity = 64.7 % AUC = 0.72 **Combined Model** Accuracy = 75.7 % Sensitivity = 74.5 % Specificity = 78.9 % AUC = 0.785
Lee et al. [51]	2021	Genetic Algorithm for Predicting Recurrence after Surgery of Liver Cancer (GARSL) = [Support Vector Machine] + Compared with other classifiers (C4.5, RF, Hoeffding tree, LR, Logistic model tree, NB)	517	362	155	–	CT [Contrast-enhanced]	i. Morphology feature ii. Edge feature iii. Intensity feature iv. Haralick feature v. HU-moment invariant feature vi. Discrete wavelet transformed features	Accuracy = 0.729 AUC = 0.739
Yuan et al. [52]	2019	Cox proportional hazards model mRMR LASSO	156	129	55	–	CT [Contrast-enhanced]	i. Shape feature ii. Size feature iii. Intensity feature iv. Gray-level co-occurrence matrix (GLCM) v. Gray-level run length matrix (GLRLM) vi. Gray-level size zone matrix (GLSZM) vii. Neighbouring gray tone difference matrix (NGTDM)	(C-index) AUC = 0.755
Hu et al. [53]	2020	CT-based peritumoral radiomics (PT-RO) prediction model LASSO	203	109	47	47	CT [Contrast-enhanced]	i. Texture feature (Cluster Shade) ii. GLCM (Haralick feature) iii. RLM (Long Run Emphasis & High Gray Level Run Emphasis)	Internal AUC = 0.79 External AUC = 0.63
Liu et al. [39]	2020	i. Deep learning radiomics-based CEUS model (R-DLCEUS) ii. Machine learning radiomics-based time- intensity curve of CEUS model (R-TIC) iii. Machine learning radiomics-based B-Mode images model (R-BMode)	130	89	41	–	US [Contrast-enhanced]	i. Statistics feature ii. Tumour Shape feature iii. Texture feature	Accuracy = 0.9 Sensitivity = 0.893 Specificity = 0.923 AUC = 0.93

(*continued on next page*)

Table 2 (*continued*)

Author	Year	Algorithm	Datasets				Types of Modalities	Types of Radiomic features extracted	Performance Evaluation
			Sample size	Training	Test	Validation			
Xu et al. [54]	2022	Support Vector Machine - LASSO	211	122	–	89	CT	i. Shape statistics ii. First-order statistics iii. Textural features iv. Wavelet-based transformation	**Radiomics** AUC = 0.847 **Radiologists** AUC = 0.659
Li et al. [55]	2023	Deep Learning	262	146	35	81	Dual-energy CT (DECT)	i. Shape statistics ii. Intensity features iii. Textural features iv. Wavelet features v. Deep features	**DL Radiomics Nomogram Model** Accuracy = 0.86 Sensitivity = 0.8 Specificity = 0.9 AUC = 0.89 **Clinical-Radiologic Model** Accuracy = 0.72 Sensitivity = 0.87 Specificity = 0.63 AUC = 0.79
Wang et al. [56]	2023	Support Vector Machine - LASSO	106	72	32	–	CT [Contrast-enhanced]	i. Shape statistics ii. First order statistics iii. Textural features	**Radiomics Model** Accuracy = 0.712 AUC = 0.87 **Clinical Model** Accuracy = 0.792 AUC = 0.816 **Radiomics-Clinical Model** Accuracy = 0.844 AUC = 0.933

2.3. Data extraction and quality assessment

All the articles were evaluated for their appropriateness and relevance to the topic. Those that met the inclusion criteria were considered for further analysis. Data extracted from the chosen articles were collected and executed through Excel spreadsheet with variables listed in Table 1. Six variables were subsequent from the established spreadsheet which were accessible with both qualitative and quantitative data of selected studies. Furthermore, model with the best performance were included in the primary analysis for studies with multiple proposed models. Best performance on specificity and sensitivity of proposed model were also included to perform further analysis.

Methodological quality and reliability of included studies was assessed by authors using risk-of-bias assessment tool which criteria outlined in the Cochrane Handbook for Systematic Reviews of Intervention [30]. This tool analysed six domains related to risk of bias: (i) random sequence generation; (ii) allocation concealment; (iii) blinding of participants and personnel; (iv) incomplete outcome data; (v) selective reporting; and (vi) other bias. Risk of bias figures were generated using Cochrane Revman software (version 5.4, The-Nordic Cochrane Centre, Copenhagen, Denmark) and categorizes the selected studies by either low, unclear or high risk of bias in each domain.

2.4. Statistical analysis

Meta-analysis was conducted for classification proportion on performance metrics such as AUC, accuracy, sensitivity and specificity by using statistical software, Jamovi version 1.1.9.0, a software which utilised R programming language and analysis are operated based on R packages. These metrics of included studies were pooled using a random effects model to assess the predictive performance.

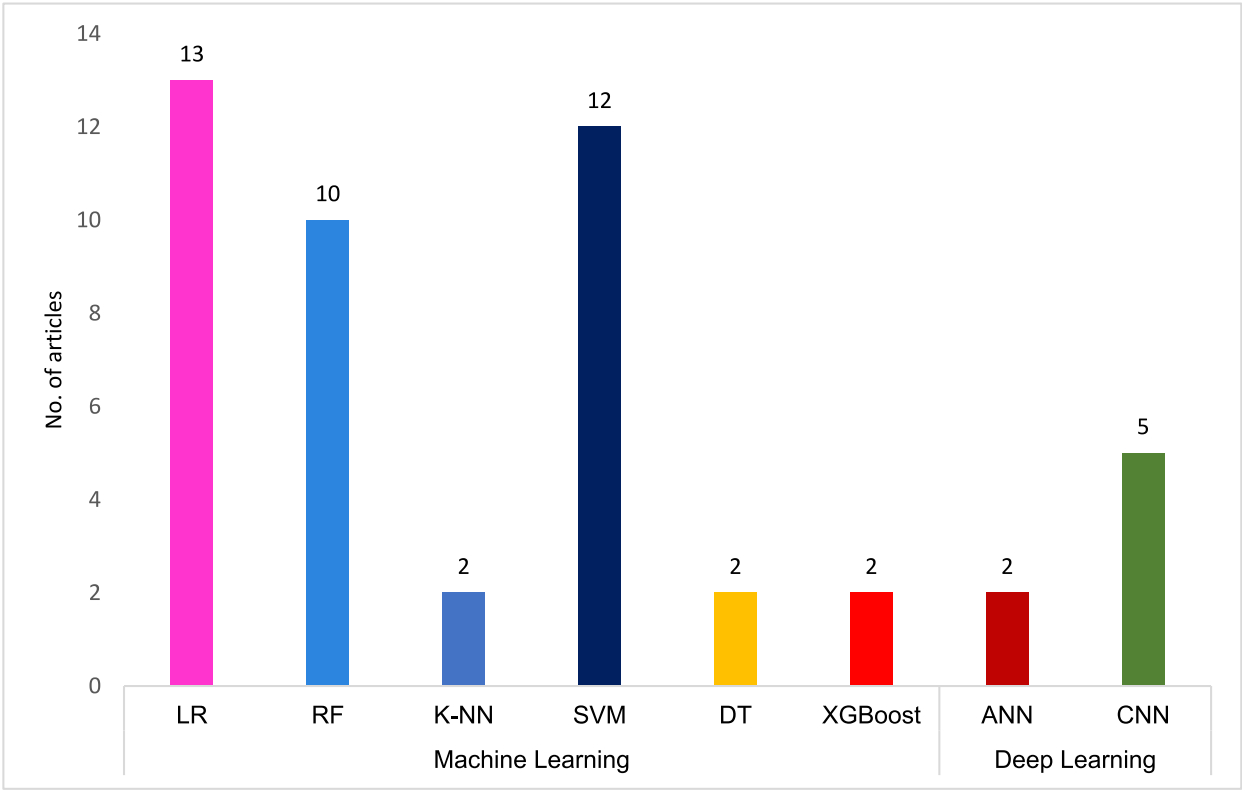


Fig. 2. Types of algorithms used in selected articles as classification model.

The heterogeneity among studies was assessed using chi-square test and Higgins I-squared (I^2) values. A random-effect meta-analysis was performed with 95 % confidence intervals and forest plot were generated for every performance metrics of the selected studies.

3. Results

After conducting a literature search and finding 438 articles, 97 duplicates were removed and 339 were preliminarily screened. Of these, 134 were excluded due to not meeting the selection criteria based on the abstract and title. This round of filtering exclude articles that did not address utilization of radiomics in HCC prediction, but simply uses other features. In addition, articles with lack of performance evaluation data was excluded. 173 articles were ineligible as they were not peer reviewed, not accessible, book chapters, or in foreign languages. This phase was complex and time-consuming due to assessment of full-text articles in order to complete the filtering tasks. Following a full-text analysis, 34 articles were included in the systematic review and meta-analysis, with 12 providing sufficient quantitative data for risk of bias assessment. Fig. 1 presents the flow chart of the review search and data extraction according to PRISMA guidelines.

3.1. Study characteristics and model Methodology: A systematic review

Based on full-text evaluation, 32 studies met our inclusion criteria and were included in systematic review. Table 2 presents study characteristics of each 32 articles. Of the included 32 articles, three types of imaging modalities were applying to identify HCC which were MRI, CT scans and US. Eleven studies select patients undergoes MRI examination with varieties of sequences as their source data [24,26,28,31–37]. Eighteen studies favoured different types of CT images to detect HCC in their research [1,20,23,38–51]. Only three studies used US data to construct and validate model in classifying HCC [4,38,39]. All studies used one specific imaging modality to achieve their objective except two studies which utilised both MRI and CT scans [25,40]. Additional to imaging data, ten studies combined clinical information such as age, gender, types of hepatitis infection, serum alpha-fetoprotein (AFP) level and tumour size with features extracted based on images as input data in model construction [1,24,32,36,41,42,43,44,38,40].

Forty ML and six DL articles were identified and included in the analysis. Twenty-five studies employed conventional ML algorithms such as Logistic Regression (LR) [1,26,32,41,42,48,50], Random Forest (RF) [33,40], K-Nearest Neighbour (K-NN) [20], Support Vector Machine (SVM) [4,57,47,38,56], Decision Tree (DT) [58] and Extreme Gradient Boosting (XGBoost) [44] classifiers. Eight studies proposed DL methods which consists of Artificial Neural Network (ANN) [35,40] and Convolutional Neural Network (CNN) [28,34,49,43]. Fig. 2 shows the types of machine learning and deep learning applied in each study. Amidst imbalanced

Table 3
Summary of algorithm and dataset of CT scans that met the criteria selection.

Author	Year	Datasets				Algorithm	Feature Selector	Types Features Extracted
		Sample size	No. of training data	No. of test data	No. of validation data			
Nie et al. [1]	2020	156	119	37	–	Multiple Logistic Regression	LASSO	i. Intensity statistics features ii. Shape features iii. Texture features iv. Filter and wavelet features
Zhang et al. [46]	2020	637	451	111	75	Multivariable Logistic Regression (MLR)	LASSO	i. Intensity feature ii. Texture feature iii. Wavelet feature
Li et al. [32]	2021	301	131	113	57	Logistic Regression (LR)	LASSO	i. Gray-level histogram Form factor parameters
Shen et al. [42]	2021	114	80	34	–	Random forest (RF)	–	Difference-in-difference (DD) features
Xu et al. [57]	2019	619	350	145	495	Support Vector Machine (SVM)	Recursive feature selection Support Vector Machine (ref-SVM)	i. Shape feature ii. Texture feature iii. First order statistics iv. Higher order statistic v. Filter and wavelet feature
Mao et al. [44]	2020	297	237	60	–	XGBoost	–	i. First order statistics ii. Texture feature iii. Shape feature iv. Wavelet filter

classification issue, two studies used Synthetic Minority Oversampling Technique (SMOTE) to overcome adverse impact of imbalance dataset on model's performance by adjusting dataset class distribution [20,40]. Fifteen studies applied Least Absolute Shrinkage and Selection Operator (LASSO) regression to pinpoint critical features. Six studies opted different types of algorithm to act as feature selector such as backward step-wise elimination [45], sparse representation [4], recursive feature selection [57], random forest [42], sequential forward selection [50] and particle component analysis (PCA) [25]. Logistic regression (LR) was the most commonly used classifier, and LASSO was the most frequently employed feature selector.

Table 2 shows the amount of dataset used for each study after dividing into test and training data. The range amount of sample size is between 42 and 637, training data from 26 to 451, where test data range from 17 to 155 while validation data from 26 to 495. Furthermore, most studies have excess data to be used as validation set. However, it is unnecessarily needed for construction of classification model. Concerning total amount of data will affect performance of classification models. Most studies extracted shape, texture, first and higher order statistical, filter, and wavelet features, with the exception of eight studies as they used additional features such as difference-in-difference (DD) features, form factor parameters, Haralick feature and delayed features [4,20,28,32,49,42,43,51].

Tables 3 and 4 summarize the parameters used to construct classification models using CT scans and MRI scans, respectively, including sample sizes, number of training and test data, classifiers and feature selectors employed, and types of features extracted. The sample size range for CT scans is between 114 and 637 patients, while for MRI it is between 122 and 369. From the tables, it can be seen that most of the studies did not use a validation set for their model, with the exception of the studies from Ji et al. (2019) and Zhang et al. (2020) for CT scans and Li et al. (2021) for MRI scans [32,45,46].

3.2. Evaluation performance of machine learning: meta-analysis Quantification

Twenty-nine studies were included to estimate the AUC of classification of HCC methods. Only twelve studies provide accuracy of their classification model while seventeen studies provide both sensitivity and specificity. There are eight studies clearly specify all three data sets; training, test and validation [20,28,32,35,45,57,46,53].

Random-effect model meta-analysis were performed to demonstrate summary proportions using sample size and AUC across studies. The classification AUC was 0.86 (95 % CI: 0.83–0.89). The I^2 was 66.50 % of the total variance between studies which was significantly high. The graphical representation of meta-analysis summary is illustrated in Fig. 3. We made further analysis on three different performance metrics: accuracy, sensitivity and specificity. Summary proportions of accuracy and sample size is presented in Fig. 4. This classification accuracy was 0.83 (95 % CI: 0.78–0.88). True heterogeneity across studies for accuracy is high as I^2 was

Table 4
Summary of algorithm and dataset of MRI procedure that met the criteria selection.

Author	Year	Datasets				Algorithm	Feature Selector	Types Features Extracted
		Sample size	No. of training data	No. of test data	No. of validation data			
Wu et al. [31]	2019	369	295	74	–	i. Decision Tree (DT) ii. Random Forest (RF) iii. K-Nearest Neighbour (KNN) iv. Logistic Regression (LR)	–	i. First Order features ii. Shape features iii. 2nd Order features (texture) iv. Higher Order Statistics features
Yang et al. [35]	2021	257	143	111	–	i. Multiple Logistic Regression (MLR) ii. Support Vector Machine (SVM) iii. Random Forest (RF) iv. Artificial Neural Network (ANN)	–	i. Intensity feature ii. Texture feature iii. Shape feature iv. Wavelet feature
Zhao et al. [37]	2021	122	85	37	–	Multivariate Logistic Regression (MLR)	i. LASSO ii. Univariate LR	i. Histogram feature ii. Gray-level co-occurrence matrix (GLCM) iii. Gray-level run length matrix (GLRLM) iv. Gray-level size zone matrix (GLSZM) v. Haralick feature vi. Form factors vii. Gaussian transformed feature
Gao et al. [34]	2021	225	168	57	–	H-DAR-net (Combination of triplet CNN and simple SE-DenseNet)	LASSO	i. First Order Statistics (Intensity) ii. Texture feature iii. Wavelet feature

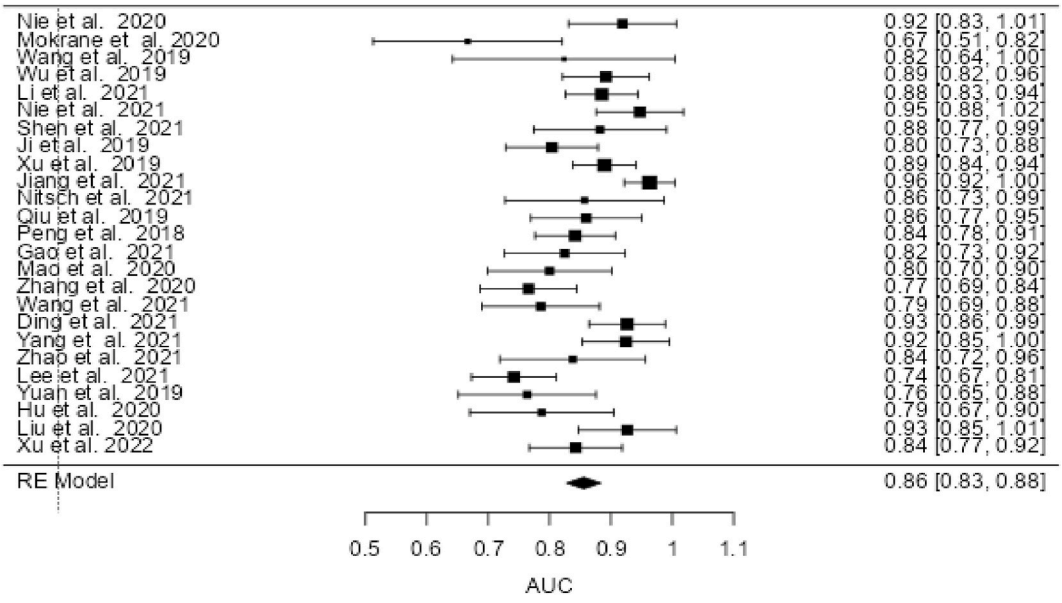


Fig. 3. Forest plots showing the proportion of classification AUC ML models for HCC.

69.52 % Figs. 5 and 6 demonstrates summary proportions of sensitivity and specificity. The classification sensitivity and specificity were 0.8 (95 % CI: 0.75–0.84) and 0.84 (95 % CI: 0.80–0.88), respectively. The analysis revealed I^2 was 77.45 % for sensitivity and I^2 for specificity was 67.01 % All four-performance metrics indicate heterogeneity supported by I^2 values of each metric were more than 50 % (p-value <0.001) due to variability of studies methods and other design aspects.

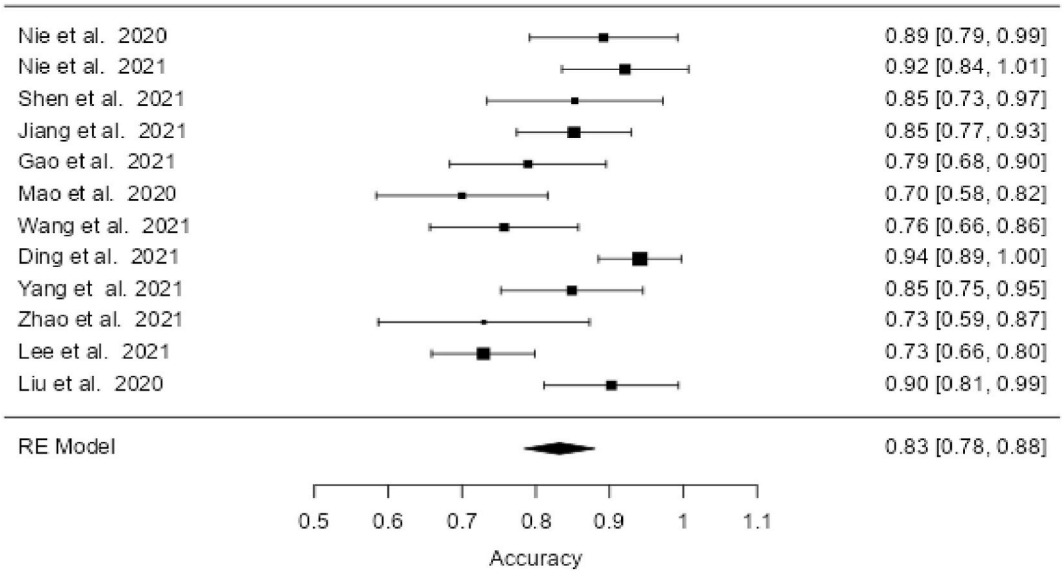


Fig. 4. Forest plots showing the proportion of classification accuracy ML models for HCC.

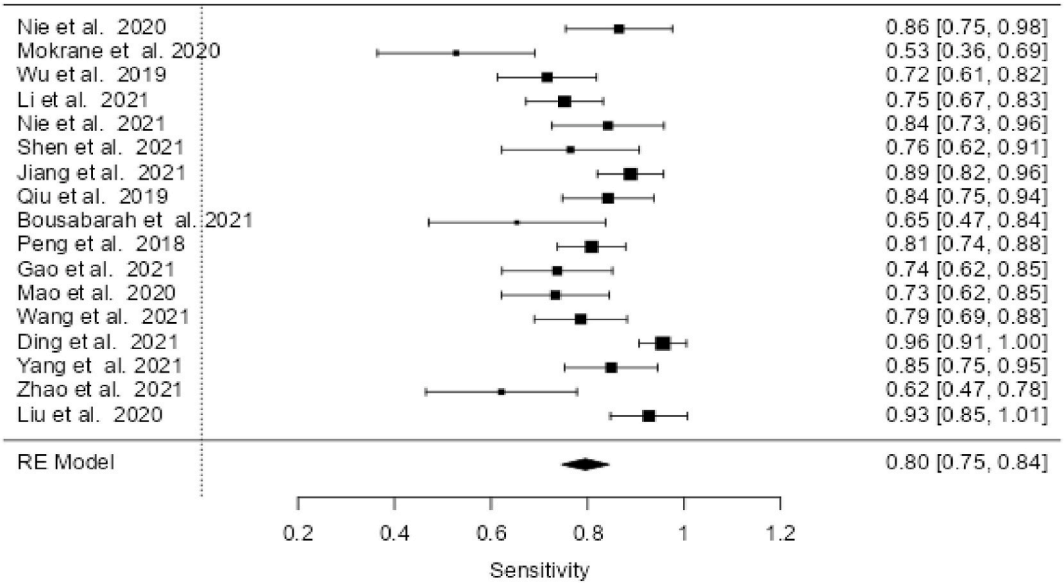


Fig. 5. Forest plots showing the proportion of classification sensitivity ML models for HCC.

3.3. Risk of bias

Twelve studies that met the inclusion criteria were included in risk of bias assessment. Fig. 7 (a) showcases the risk of bias of each study, with a summarized version of the risk of bias presented in Fig. 7 (b). Table 5 shows the five risk of bias questions and the score key. These questions relate to: (i) generation of sample sequence, (ii) concealment of knowledge on the allocation sequence, (iii) exclusion of the outcome, (iv) outcome reporting and (iv) other source of bias. For (i), a study was deemed to high risk of bias when it describe a non-random component in the sequence generation process, unclear risk when information was not provided or low risk when it describes a random sequence generation process on the sample. For (ii), studies at high risk introduced to selection bias due to investigators could possibly foresee assignment given, posed an unclear risk when there is insufficient information provided. The studies become a low risk when the assignments are adequately concealed. For (iii), studies at high risk when there is an attempt blinding key study participants and lack of blinding influence the outcome measurements, unclear when the studies did not address this outcome, or studies at low risk of bias when blinding of key study is ensured. For (iv), studies deemed to be high risk when outcome in the methods were not reported in the results, unclear risk when there was incomplete report or low risk when there is no missing

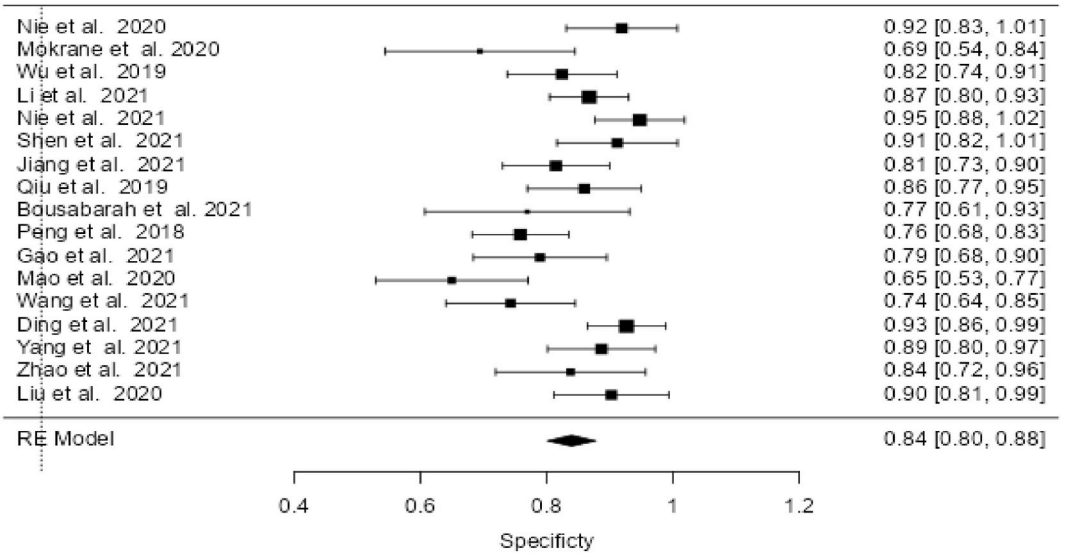


Fig. 6. Forest plots showing the proportion of classification specificity ML models for HCC.

outcome data. For (v), studies with a stated potential of other source of bias were high risk, unclear risk where there is insufficient information to assess the existence of other bias or low risk when studies is free from other bias. Table 6 describes in details the judgement of risk for each studies.

Specifically, among the 12 studies, 61.11 % were answered as "Yes", 29.17 % as "No" and the remaining 9.72 % as "Unclear". All studies were assessed as low risk of attrition and reporting bias as there was an absence of incomplete data and all pre-specified outcomes were reported. Only one study had a low risk of bias in random sequence generation as relevant information was not provided [47]. The remaining studies were judged as high risk of bias because patients included in each study were selected according on specific years and criteria (91.67 %). Allocation concealment and blinding of radiologists and pathologists were judged as low risk of bias (66.67–75 %). Three studies had an unclear risk of bias for performance bias as there was an absence of information to justify blinding of participants (25 %) [24,33,54]. The potential source of other bias was evaluated as high risk of bias (75 %).

4. Discussions

The diagnosis of cancer has always been a difficult task for clinicians, as it is a complex and heterogeneous disease. Recently, the development of precision medicine has been facilitated by advances in technology, such as the use of AI for cancer stage classification. ML has been used to make the most of medical images in order to provide personalized medicine. Research has shown that ML-based approaches have been successful in predicting and classifying HCC, however, there is still no proper implementation in clinical practice. This systematic review and meta-analysis study aimed to evaluate the relationship between ML-based approaches or factors such as the type of classifier, feature selector, and amount of input features extracted, and the performance of constructed classification models.

Computers are used to classify HCC, with the extracted data split into two or three sets, such as the training, test and validation sets. Most studies employ a simple method of dividing the data by using the split data function. However, two studies utilize Leave-One-Out Cross-Validation (LOOCV) which only leaves one subject for testing and the remainder for training [4,26]. In general, the amount of training set is higher than test set and the train-test split ratio in this study are 2:1, 4:1, 7:3 and 9:1. There was no fixed separation ratio for training and test set but larger patient cohort would be necessary in order for the classifier to distinguish between two or more classes [33]. This is in line with the reported from several researchers where they overcome the problems with small datasets using either 10-fold cross validation or LOOCV [4,26,47].

In most studies, features extracted from ROIs in medical images and act as inputs for machine learning and deep learning models are texture features, shape features, first order statistics, higher order statistics, filter, and wavelet features. In addition, there are several features are added such as Difference in difference (DD) features, arterial features, portal venous features, delayed features dual phase features to obtain more relevant information [20,28,42]. However, only certain features hold valuable information that describes selected area of regions of interest (ROIs) in the medical images. Thus, feature selector is needed to remove insignificant features. Feature selection is essential method in constructing computer-learning model especially in classification of cancer. Having abundant of radiomic features extracted from ROIs can lead to overfitting in the classification model. This finding was in agreement with a study which claimed that model with features reduction was more efficient in classifying healthy liver and HCC compared to model with original features and also reduces complexity of the model [47]. In this study, most studies involving machine learning uses feature selectors to select important features while removing redundant features. This process is convenient for those with small sample size but with large radiomics features to avoid overfitting. In addition, it was found that 14 studies used LASSO regression

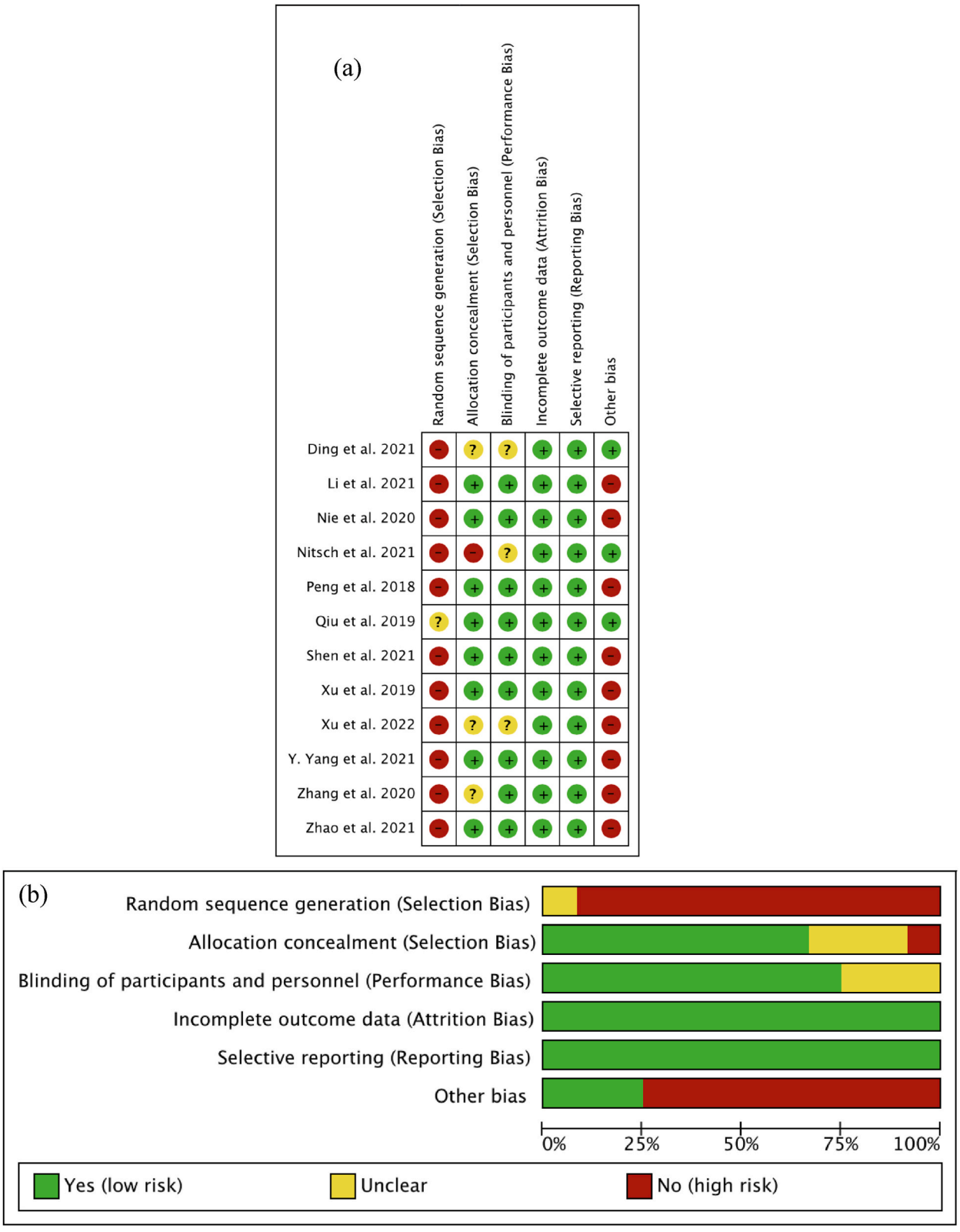


Fig. 7. (a) Risk of bias of selected study and (b) overview of the risk of bias.

Table 5
Risk of bias tool. Yes = Low risk of bias, Unclear = incomplete information or not reported, No = High risk of bias.

	Score: Yes, Unclear, No
1. Did the authors generate allocation sequence adequately ?	
2. Was the authors adequately concealed the assignments given?	
3.Was knowledge of the allocated intervention adequately prevented?	
4. Were all the outcome measured in the methods addressed in the results?	
5. Was the study free of other problems that could be considered as a high risk of bias?	

algorithm to select valuable features in the datasets. LASSO was chosen as it said to be suitable to analyse large radiomics features with a small sample size [26]. This method is applicable for high dimensional data such as radiomic features to select most significant features and obtain subset features [59]. Although majority of included studies utilised LASSO as feature selection, it has a high tendency to be influenced by the data correlations that could lead to lower performance.

A classifier is an algorithm that assigns numerical features to discrete categories. In this study, supervised machine learning was used to construct classifiers for the classification of HCC with normal liver, the stages of HCC, and the detection of recurrence of HCC. This approach involves training a model using a set of samples with known output categories, in order to build a classifier that can accurately categorize new inputs [31]. Four studies use deep learning to distinguish the HCC. According to Bousabarah et al. (2021), their approach can automatically segment liver and HCC which facillitate an efficient workflow in clinical practices [28]. Fig. 2 shows number of research that uses unique algorithm for classification, and as observed the logistic regression (LR) was preferred. This is because logistic regression is one of the most basic classification algorithms. Its often use to solve binary classification problems. This model describes sigmoid relationships between continuous independent variable and binary outcomes to discover the line of separation. As LR is one of the linear classifiers, it assumes the linearity between independent variables and log-odds function which is used to model the binary outcomes [60]. Although majority of studies preferred LR as the classification algorithm, the performance of LR is limited by the data linearity of radiomics features. In addtion, there were previous studies employ multi-class classification problems with the multivariate logistic regression (MLR) [1,35–37,41,42,48,46]. This algorithm predicts multiple outcomes using multiple independent variables and it has ability to correlate complex relationships of the variables.

The results of pooled predictive performance of machine learning algorithms for the classification of hepatocellular carcinoma (HCC) showed high values, demonstrating the ability of radiomic features to capture distinct characteristics of each HCC phenotype. The meta-analysis yielded an overall area under the curve (AUC) of 0.86 % (95 % CI: 0.83–0.88), followed by 0.83 (95 % CI: 0.78–0.88) for accuracy, 0.80 (95 % CI: 0.75–0.84) for sensitivity, and 0.84 (95 % CI: 0.80–0.88) for specificity. These results indicate that proposed ML approaches have promising performance for HCC prediction from image-based diagnosis. However, one study had the lowest proportions for AUC, sensitivity, and specificity, which can be attributed to the overlap between different pathologies obtained from cirrhotic liver [20].

Performance of classification model can be analysed using performance metrics such as accuracy, sensitivity, specificity, and area under curve (AUC). Information was extracted from medical images via radiomics method and can be further used as input for classification model. Optimized radiomics features have been reported to be helpful as biomarkers in detecting HCC. In 2019, Wu et al. (2019) prove that radiomics features can be used to distinguish between HCC and hepatic haemangioma (HH) [31]. Although radiomic features describe the phenotypes of HCC, it can be seen that additional information from clinical and radiological features does increase the performance of predictive models as this method provide additional information that could be useful to predict HCC. Comparisons of performance were studied by Ding et al. (2021) and Nie et al. (2021) and found that combined model more efficient and high performance in predicting outcomes compared with solely radiomic model or clinical model [41]. Hence, by combining and integrating other data in the radiomic signature could improve the models' performance.

Imaging modalities such as ultrasounds, computed tomography (CT) and magnetic resonance imaging (MRI) are widely used in HCC diagnosis. These imaging modalities play important role in prediction of HCC as quality of medical images produced by the imaging modalities could affect the process extraction process of radiomic features. An images with high spatial resolutions from CT and MRI increase the potential of generating informative features that can describe further the HCC phenotypes. Thus, lead to better performance for the constructed predictive models. There is also a study that analyses performance of combined models based on multimodal imaging data such as MRI and CT images. According to Liang et al. (2020) imply that the MRI has slightly higher efficiency rather than CT images [40]. In addition, the sample size used in the study does affect the performance of the model as a small sample size could lead to overfitting [61]. In 2021, Liu et al. (2021) demonstrates that having a small sample size lead to degradation of predictive model [25]. This systematic review was limited to articles written in English language, which could have caused the authors to overlook relevant articles available in the chosen database. Furthermore, the differences in input features, feature selectors and ML approaches caused the heterogeneity across studies to be high.

5. Conclusion

The incorporation of AI in the medical sector has been a significant breakthrough in recent years, particularly with regards to the analysis of big data in healthcare. In this study, various methods were evaluated for their performance in detecting HCC and associated factors, including feature selectors, classifiers, and input features. It was found that radiomics features extracted from medical images provided sufficient information for AI to accurately detect HCC. Radiomics, which involves the extraction of a large number of features

Table 6
Detailed judgement for risk of bias assessments.

Author	Bias	Author's Judgement	Support for Judgement
Ding et al., 2021 [24]	Random sequence generation (Selection Bias)	No	" (…) medical records were viewed to identify all consecutive cases seen between May 2015 and May 2019."
	Allocation concealment (Selection Bias)	Unclear	" (…) by drawing the outline of tumor tissue layer-after-layer and avoiding the bile duct and vessels by Radiologists 1 and 2." (No specific information provided).
	Blinding of participants and personnel (Performance Bias)	Unclear	No specific information provided.
	Incomplete outcome data (Attrition Bias)	Yes	No losses to follow-up.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	Yes	" (…) number of samples was still limited compared to the large number of features."
Author	Bias	Author's Judgement	Support for Judgement
Li et al., 2021 [32]	Random sequence generation (Selection Bias)	No	"The dataset for the entire cohort was obtained from the January 2015 to December 2019."
	Allocation concealment (Selection Bias)	Yes	" (…) who were blinded to the clinical data of the patients." "liver imaging who were blinded to the patients' clinical data, and who assessed the imaging features randomly and independently."
	Blinding of participants and personnel (Performance Bias)	Yes	Blinding of two pathologists and three radiologists ensured and unlikely that the blinding could have been broken.
	Incomplete outcome data (Attrition Bias)	Yes	No losses to follow-up.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	No	" (…) tumour area segmentation has to be per- formed manually by radiologists."
Author	Bias	Author's Judgement	Support for Judgement
Nie et al., 2020 [1]	Random sequence generation (Selection Bias)	No	"A total of 156 patients with FNH (n = 55, 32 men and 23 women; mean age, 31.82 ± 12.55 years) and HCC (n = 101, 85 men and 16 women; mean age, 57.10 ± 9.89 years) were enrolled in this study ac- cording to several inclusion criterias".
	Allocation concealment (Selection Bias)	Yes	"Blinded to the clinic-pathologic data, (…)"
	Blinding of participants and personnel (Performance Bias)	Yes	Blinding of two radiologists ensured and unlikely that the blinding could have been broken.
	Incomplete outcome data (Attrition Bias)	Yes	No losses to follow-up.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	No	" (…) potential selection bias may hamper the reproducibility …)"
Author	Bias	Author's Judgement	Support for Judgement
Nitsch et al., 2021 [33]	Random sequence generation (Selection Bias)	No	"This was a retrospective study using MRI scans of patients with cirrhosis who were undergoing hepatocellular carcinoma (HCC) screening at Brigham and Women's Hospital (BWH) from June 1, 2015, to June 1, 2018".
	Allocation concealment (Selection Bias)	No	" (.) was performed by two hepatologists with a combined experience of 15 years to confirm the diagnosis of cirrhosis (using clinical history, liver biopsy, elastography) and to classify the presence of any liver-related decompensation (…)"
	Blinding of participants and personnel (Performance Bias)	Unclear	No specific information provided.
	Incomplete outcome data (Attrition Bias)	Yes	No losses to follow-up.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	Yes	No suggestion of other likely bias.
Author	Bias	Author's Judgement	Support for Judgement
Peng et al., 2018 [48]	Random sequence generation (Selection Bias)	No	" (…) 304 patients were finally selected for this study."
	Allocation concealment (Selection Bias)	Yes	" (…) who were blinded to information on clinical, laboratory, pathologic, and MVI status."
	Blinding of participants and personnel (Performance Bias)	Yes	Blinding of radiologists ensured, and unlikely that the blinding could have been broken.

(*continued on next page*)

Table 6 (*continued*)

Author	Bias	Author's Judgement	Support for Judgement
	Incomplete outcome data (Attrition Bias)	Yes	No losses to follow-up.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	No	" (…) larger sample size to obtain more convincing evidence in favor of clinical application of the radiomics nomogram."

Author	Bias	Author's Judgement	Support for Judgement
Qiu et al., 2019 [47]	Random sequence generation (Selection Bias)	Unclear	"A total of 106 patients at Shandong Cancer Hospital Affiliated to Shandong University between December 2015 and October 2017 were randomly enrolled in this research." (No specific information provided).
	Allocation concealment (Selection Bias)	Yes	"None of the radiation oncologists had access to clinical patient information other than the CT scans."
	Blinding of participants and personnel (Performance Bias)	Yes	Blinding of oncologists ensured, and unlikely that the blinding could have been broken.
	Incomplete outcome data (Attrition Bias)	Yes	No losses to follow-up.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	Yes	No suggestion of other likely bias.

Author	Bias	Author's Judgement	Support for Judgement
Shen et al., 2021 [42]	Random sequence generation (Selection Bias)	No	"We reviewed all patients with HCC who underwent surgical resection or ablation in a single center from January 2009 to April 2018."
	Allocation concealment (Selection Bias)	Yes	"The radiologists were blinded when evaluating the images."
	Blinding of participants and personnel (Performance Bias)	Yes	Blinding of radiologists ensured, and unlikely that the blinding could have been broken.
	Incomplete outcome data (Attrition Bias)	Yes	No missing outcome data.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	No	" (…) retrospective nature with selective bias and a single-center study with a limited sample size."

Author	Bias	Author's Judgement	Support for Judgement
Xu et al., 2019 [57]	Random sequence generation (Selection Bias)	No	" (…) surgically confirmed cases of HCC between January 2009 and August 2017."
	Allocation concealment (Selection Bias)	Yes	" (…) who were blinded to the clinical and pathological data." " (…) who were not involved in the LI-score assignment were involved in radiomics analysis."
	Blinding of participants and personnel (Performance Bias)	Yes	Blinding of radiologists ensured, and unlikely that the blinding could have been broken.
	Incomplete outcome data (Attrition Bias)	Yes	No missing outcome data.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	No	"Potential selection bias may hamper the reproducibility and comparability of the results (…)"

Author	Bias	Author's Judgement	Support for Judgement
Xu et al., 2022 [54]	Random sequence generation (Selection Bias)	No	"All patients with path- ologic results of liver cancer underwent noncontrast CT at our institution between August 2018 and November 2019."
	Allocation concealment (Selection Bias)	Unclear	"The radiologists were aware of the diagnostic criteria and blinded to the clinical radiological details." (Insufficient information to decide the risks).
	Blinding of participants and personnel (Performance Bias)	Unclear	Insufficient information to decide the risks.
	Incomplete outcome data (Attrition Bias)	Yes	No missing outcome data.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	No	" (…) retrospective study with some considerable risk of bias (…)"

Author	Bias	Author's Judgement	Support for Judgement
Yang et al., 2021 [36]	Random sequence generation (Selection Bias)	No	"Between May 2015 and October 2020, patients who were pathologically diagnosed with primary HCCs and underwent Gd-EOB-DTPA-enhanced MRI examinations were consecutively included in this study."
	Allocation concealment (Selection Bias)	Yes	"They were blinded to MVI status and other clinical information."

(*continued on next page*)

Table 6 (*continued*)

Author	Bias	Author's Judgement	Support for Judgement
	Blinding of participants and personnel (Performance Bias)	Yes	Blinding of radiologists ensured, and unlikely that the blinding could have been broken.
	Incomplete outcome data (Attrition Bias)	Yes	No losses to follow-up.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	No	" (…) the ROI were semiautomatically drawn."
Author	Bias	Author's Judgement	Support for Judgement
Zhang et al., 2020 [46]	Random sequence generation (Selection Bias)	No	"Patients in the two institutions who met the inclusion criteria were collected from March 2015 to March 2018."
	Allocation concealment (Selection Bias)	Unclear	"A senior radiologist checked all of the tumor segmentation results." (No specific information provided).
	Blinding of participants and personnel (Performance Bias)	Yes	No blinding and outcome measurement are not influenced by lack of blinding.
	Incomplete outcome data (Attrition Bias)	Yes	No losses to follow-up.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	No	" (…), the morphologic features of HCC were not evaluated (…)"
Author	Bias	Author's Judgement	Support for Judgement
Zhao et al., 2021 [37]	Random sequence generation (Selection Bias)	No	"Between February 2008 and November 2019, 328 consecutive patients with HCC (…)"
	Allocation concealment (Selection Bias)	Yes	" (…) who were aware that the patients had HCC but were blinded to clinical data and imaging report."
	Blinding of participants and personnel (Performance Bias)	Yes	Blinding of radiologists ensured, and unlikely that the blinding could have been broken.
	Incomplete outcome data (Attrition Bias)	Yes	No losses to follow-up.
	Selective reporting (Reporting Bias)	Yes	Pre-specified outcome was reported.
	Other bias	No	" (…) small population as well as the long duration of the inclusion period, may affect the robustness (…)"

from medical images using data-characterization algorithms, has shown great promise in enhancing the diagnostic accuracy of AI models. The high-dimensional data derived from radiomics can capture intricate details and patterns that are often imperceptible to the human eye, thereby improving the sensitivity and specificity of HCC detection. Moreover, advanced machine learning algorithms, such as deep learning and ensemble methods, have demonstrated superior performance in processing radiomics data. These algorithms can handle the complexity and heterogeneity of medical imaging data, enabling more precise and reliable predictions. The integration of radiomics with AI algorithms not only aids in the early detection of HCC but also in the assessment of tumour characteristics and prognosis, which are crucial for personalized treatment planning.

Funding

This work was supported by Fundamental Research Grant Scheme from Ministry of Higher Education of Malaysia and Universiti Putra Malaysia under grant number FRGS/1/2020/STG07/UPM/02/3.

Data availability statement

Data will be made available on request.

CRediT authorship contribution statement

Nurin Syazwina Mohd Haniff: Writing – original draft, Investigation, Formal analysis. **Kwan Hoong Ng:** Visualization, Validation, Data curation. **Izdihar Kamal:** Writing – review & editing, Visualization, Project administration. **Norhayati Mohd Zain:** Visualization, Validation, Resources. **Muhammad Khalis Abdul Karim:** Writing – review & editing, Supervision, Methodology, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] P. Nie, G. Yang, J. Guo, J. Chen, X. Li, Q. Ji, J. Wu, J. Cui, W. Xu, A CT-based radiomics nomogram for differentiation of focal nodular hyperplasia from hepatocellular carcinoma in the non-cirrhotic liver, Cancer Imag. 20 (2020).
[2] C. Schraml, S. Kaufmann, H. Rempp, R. Syha, D. Ketelsen, M. Notohamiprodjo, K. Nikolaou, Imaging of HCC-current state of the art, Diagnostics 5 (2015) 513–545.
[3] H. Liao, T. Xiong, J. Peng, L. Xu, M. Liao, Z. Zhang, Z. Wu, K. Yuan, Y. Zeng, Classification and prognosis prediction from histopathological images of hepatocellular carcinoma by a fully automated pipeline based on machine learning, Ann. Surg Oncol. 27 (2020) 2359–2369.
[4] Y. Dong, Q.M. Wang, Q. Li, L.Y. Li, Q. Zhang, Z. Yao, M. Dai, J. Yu, W.P. Wang, Preoperative prediction of microvascular invasion of hepatocellular carcinoma: radiomics algorithm based on ultrasound original radio frequency signals, Front. Oncol. 9 (2019).
[5] Z. Ramli, M.K.A. Karim, N. Effendy, M.A. Abd Rahman, M.M.A. Kechik, M.J. Ibahim, N.S.M. Haniff, Stability and reproducibility of radiomic features based on various segmentation techniques on cervical cancer DWI-MRI, Diagnostics 12 (2022) 1–14.
[6] U.R. Acharya, Y. Hagiwara, V.K. Sudarshan, W.Y. Chan, K.H. Ng, Towards precision medicine: from quantitative imaging to radiomics, J. Zhejiang Univ. - Sci. B 19 (2018) 6–24.
[7] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R.G.P.M. Van Stiphout, P. Granton, C.M.L. Zegers, R. Gillies, R. Boellard, A. Dekker, H.J.W.L. Aerts, Radiomics: extracting more information from medical images using advanced feature analysis, Eur. J. Cancer 48 (2012) 441–446.
[8] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data, Radiology 278 (2016) 563–577.
[9] C. Parmar, P. Grossmann, J. Bussink, P. Lambin, H.J.W.L. Aerts, Machine learning methods for quantitative radiomic biomarkers, Sci. Rep. 5 (2015) 1–11.
[10] S.F. Mat Radzi, M.K. Abdul Karim, M.I. Saripan, M.A. Abd Rahman, N.H. Osman, E.Z. Dalah, N. Mohd Noor, Impact of image contrast enhancement on stability of radiomics feature quantification on a 2D mammogram radiograph, IEEE Access 8 (2020) 127720–127731.
[11] L. Vaugier, L. Ferrer, L. Mengue, E. Jouglar, Radiomics for radiation oncologists: are we ready to go? BJR|Open 2 (2020) 20190046.
[12] S.F.M. Radzi, M.K.A. Karim, M.I. Saripan, M.A.A. Rahman, I.N.C. Isa, M.J. Ibahim, Hyperparameter tuning and pipeline optimization via grid search method and tree-based autoML in breast cancer prediction, J. Pers. Med. 11 (2021).
[13] M.M. Yunus, A. Sabarudin, M.K.A. Karim, P.N.E. Nohuddin, I.A. Zainal, M.S.M. Shamsul, A.K.M. Yusof, Reproducibility and repeatability of coronary computed tomography angiography (ccta) image segmentation in detecting atherosclerosis: a radiomics study, Diagnostics 12 (2022) 2007.
[14] M.E. Mayerhoefer, A. Materka, G. Langs, I. Häggström, P. Szczypiński, P. Gibbs, G. Cook, Introduction to radiomics, J. Nucl. Med. 61 (2020) 488–495.
[15] N.S. Mohd Haniff, M.K. Abdul Karim, N.H. Osman, M.I. Saripan, I.N. Che Isa, M.J. Ibahim, Stability and reproducibility of radiomic features based various segmentation technique on MR images of hepatocellular carcinoma (HCC), Diagnostics 11 (2021).
[16] N.S.M. Haniff, M.K.B.A. Karim, N.S. Ali, M.A.A. Rahman, N.H. Osman, M.I. Saripan, Magnetic resonance imaging radiomics analysis for predicting hepatocellular carcinoma, in: 2021 Int. Congr. Adv. Technol. Eng. ICOTEN, 2021, 2021.
[17] A.K. Jha, S. Mithun, V. Jaiswar, U.B. Sherkhane, N.C. Purandare, K. Prabhash, V. Rangarajan, A. Dekker, L. Wee, A. Traverso, Repeatability and reproducibility study of radiomic features on a phantom and human cohort, Sci. Rep. 11 (2021) 1–12.
[18] Q. Qiu, J. Duan, G. Gong, Y. Lu, D. Li, J. Lu, Y. Yin, Reproducibility of radiomic features with GrowCut and GraphCut semiautomatic tumor segmentation in hepatocellular carcinoma, Transl, Cancer Res. 6 (2017) 940–948.
[19] C. Haarburger, G. Müller-Franzes, L. Weninger, C. Kuhl, D. Truhn, D. Merhof, Radiomics feature reproducibility under inter-rater variability in segmentations of CT images, Sci. Rep. 10 (2020) 1–10.
[20] F.Z. Mokrane, L. Lu, A. Vavasseur, P. Otal, J.M. Peron, L. Luk, H. Yang, S. Ammari, Y. Saenger, H. Rousseau, B. Zhao, L.H. Schwartz, L. Dercle, Radiomics machine-learning signature for diagnosis of hepatocellular carcinoma in cirrhotic patients with indeterminate liver nodules, Eur. Radiol. 30 (2020) 558–570.
[21] K. Izdihar, M.K.A. Karim, N.N. Aresli, S.F.M. Radzi, A. Sabarudin, M.M. Yunus, M.A. Rahman, S. Shamsul, Detection of novel coronavirus from chest X-ray radiograph images via automated machine learning and CAD4COVID, IEEE (2021) 1–4.
[22] M.M. Yunus, A. Sabarudin, N.I. Hamid, A.K.M. Yusof, P.N.E. Nohuddin, M.K.A. Karim, Automated classification of atherosclerosis in coronary computed tomography angiography images based on radiomics study using automatic machine learning, in: Proc. Int. Conf. Electron. Renew. Syst. ICEARS 2022, IEEE, 2022, pp. 1895–1903.
[23] Q.Y. Shan, H.T. Hu, S.T. Feng, Z.P. Peng, S.L. Chen, Q. Zhou, X. Li, X.Y. Xie, M. De Lu, W. Wang, M. Kuang, CT-based peritumoral radiomics signatures to predict early recurrence in hepatocellular carcinoma after curative tumor resection or ablation, Cancer Imag. 19 (2019) 1–11.
[24] Z. Ding, K. Lin, J. Fu, Q. Huang, G. Fang, Y. Tang, W. You, Z. Lin, Z. Lin, X. Pan, Y. Zeng, An MR-based radiomics model for differentiation between hepatocellular carcinoma and focal nodular hyperplasia in non-cirrhotic liver, World J. Surg. Oncol. 19 (2021).
[25] X. Liu, F. Khalvati, K. Namdar, S. Fischer, S. Lewis, B. Taouli, M.A. Haider, K.S. Jhaveri, Can machine learning radiomics provide pre-operative differentiation of combined hepatocellular cholangiocarcinoma from hepatocellular carcinoma and cholangiocarcinoma to inform optimal treatment planning? Eur. Radiol. 31 (2021) 244–255.
[26] H. Dai, M. Lu, B. Huang, M. Tang, T. Pang, B. Liao, H. Cai, M. Huang, Y. Zhou, X. Chen, H. Ding, S.T. Feng, Considerable effects of imaging sequences, feature extraction, feature selection, and classifiers on radiomics-based prediction of microvascular invasion in hepatocellular carcinoma using magnetic resonance imaging, Quant. Imag. Med. Surg. 11 (2021) 1836–1853.
[27] A. Oyama, Y. Hiraoka, I. Obayashi, Y. Saikawa, S. Furui, K. Shiraishi, S. Kumagai, T. Hayashi, J. Kotoku, Hepatic tumor classification using texture and topology analysis of non-contrast-enhanced three-dimensional T1-weighted MR images with a radiomics approach, Sci. Rep. 9 (2019) 2–11.
[28] K. Bousabarah, B. Letzen, J. Tefera, L. Savic, I. Schobert, T. Schlachter, L.H. Staib, M. Kocher, J. Chapiro, M. De Lin, Automated detection and delineation of hepatocellular carcinoma on multiphasic contrast-enhanced MRI using deep learning, Abdom. Radiol. 46 (2021) 216–225.
[29] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, D. Altman, G. Antes, D. Atkins, V. Barbour, N. Barrowman, J.A. Berlin, J. Clark, M. Clarke, D. Cook, R. D'Amico, J.J. Deeks, P.J. Devereaux, K. Dickersin, M. Egger, E. Ernst, P.C. Gøtzsche, J. Grimshaw, G. Guyatt, J. Higgins, J.P.A. Ioannidis, J. Kleijnen, T. Lang, N. Magrini, D. McNamee, L. Moja, C. Mulrow, M. Napoli, A. Oxman, B. Pham, D. Rennie, M. Sampson, K.F. Schulz, P.G. Shekelle, D. Tovey, P. Tugwell, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, PLoS Med. 6 (2009) e1000097.
[30] J.P. Higgins, D.G. Altman, Assessing risk of bias in included studies, Cochrane Handb. Syst. Rev. Interv. Cochrane B. Ser. (2008) 187–241.
[31] J. Wu, A. Liu, J. Cui, A. Chen, Q. Song, L. Xie, Radiomics-based classification of hepatocellular carcinoma and hepatic haemangioma on precontrast magnetic resonance images, BMC Med. Imaging 19 (2019).
[32] L. Li, Q. Su, H. Yang, Preoperative prediction of microvascular invasion in hepatocellular carcinoma: a radiomic nomogram based on MRI, Clin. Radiol. 77 (2021) e269–e279.
[33] J. Nitsch, J. Sack, M.W. Halle, J.H. Moltz, A. Wall, A.E. Rutherford, R. Kikinis, H. Meine, MRI-based radiomic feature analysis of end-stage liver disease for severity stratification, Int. J. Comput. Assist. Radiol. Surg. 16 (2021) 457–466.
[34] F. Gao, K. Qiao, B. Yan, M. Wu, L. Wang, J. Chen, D. Shi, Hybrid network with difference degree and attention mechanism combined with radiomics (H-DARnet) for MVI prediction in HCC, Magn. Reson. Imaging 83 (2021) 27–40.
[35] F. Yang, Y. Wan, L. Xu, Y. Wu, X. Shen, J. Wang, D. Lu, C. Shao, S. Zheng, T. Niu, X. Xu, MRI-radiomics prediction for cytokeratin 19-positive hepatocellular carcinoma: a multicenter study, Front. Oncol. 11 (2021) 1.
[36] Y. Yang, W.J. Fan, T. Gu, L. Yu, H.L. Chen, Y.F. Lv, H. Liu, G.X. Wang, D. Zhang, Radiomic features of multi-ROI and multi-phase MRI for the prediction of microvascular invasion in solitary hepatocellular carcinoma, Front. Oncol. 11 (2021).
[37] Y. Zhao, N. Wang, J. Wu, Q. Zhang, T. Lin, Y. Yao, Z. Chen, M. Wang, L. Sheng, J. Liu, Q. Song, F. Wang, X. An, Y. Guo, X. Li, T. Wu, A.L. Liu, Radiomics analysis based on contrast-enhanced MRI for prediction of therapeutic response to transarterial chemoembolization in hepatocellular carcinoma, Front. Oncol. 11 (2021).
[38] W. Wang, S.S. Wu, J.C. Zhang, M.F. Xian, H. Huang, W. Li, Z.M. Zhou, C.Q. Zhang, T.F. Wu, X. Li, M. Xu, X.Y. Xie, M. Kuang, M. De Lu, H.T. Hu, Preoperative pathological grading of hepatocellular carcinoma using ultrasomics of contrast-enhanced ultrasound, Acad. Radiol. 28 (2021) 1094–1101.

[39] D. Liu, F. Liu, X. Xie, L. Su, M. Liu, X. Xie, M. Kuang, G. Huang, Y. Wang, H. Zhou, K. Wang, M. Lin, J. Tian, Accurate prediction of responses to transarterial chemoembolization for patients with hepatocellular carcinoma by using artificial intelligence in contrast-enhanced ultrasound, Eur. Radiol. 30 (2020) 2365–2376.
[40] W. Liang, J. Shao, W. Liu, S. Ruan, W. Tian, X. Zhang, D. Wan, Q. Huang, Y. Ding, W. Xiao, Differentiating hepatic epithelioid angiomyolipoma from hepatocellular carcinoma and focal nodular hyperplasia via radiomics models, Front. Oncol. 10 (2020).
[41] P. Nie, N. Wang, J. Pang, G. Yang, S. Duan, J. Chen, W. Xu, CT-based radiomics nomogram: a potential tool for differentiating hepatocellular adenoma from hepatocellular carcinoma in the noncirrhotic liver, Acad. Radiol. 28 (2021) 799–807.
[42] J. xian Shen, Q. Zhou, Z. hang Chen, Q. feng Chen, S. ling Chen, S. ting Feng, X. Li, T. fan Wu, S. Peng, M. Kuang, Longitudinal radiomics algorithm of posttreatment computed tomography images for early detecting recurrence of hepatocellular carcinoma after resection or ablation, Transl, Oncol. 14 (2021) 100866.
[43] Y.Q. Jiang, S.E. Cao, S. Cao, J.N. Chen, G.Y. Wang, W.Q. Shi, Y.N. Deng, N. Cheng, K. Ma, K.N. Zeng, X.J. Yan, H.Z. Yang, W.J. Huan, W.M. Tang, Y. Zheng, C. K. Shao, J. Wang, Y. Yang, G.H. Chen, Preoperative identification of microvascular invasion in hepatocellular carcinoma by XGBoost and deep learning, J. Cancer Res. Clin. Oncol. 147 (2021) 821–833.
[44] B. Mao, L. Zhang, P. Ning, F. Ding, F. Wu, G. Lu, Y. Geng, J. Ma, Preoperative prediction for pathological grade of hepatocellular carcinoma via machine learning–based radiomics, Eur. Radiol. 30 (2020) 6924–6932.
[45] G.W. Ji, F.P. Zhu, Q. Xu, K. Wang, M.Y. Wu, W.W. Tang, X.C. Li, X.H. Wang, Machine-learning analysis of contrast-enhanced CT radiomics predicts recurrence of hepatocellular carcinoma after resection: a multi-institutional study, EBioMedicine 50 (2019) 156–165.
[46] X. Zhang, S. Ruan, W. Xiao, J. Shao, W. Tian, W. Liu, Z. Zhang, D. Wan, J. Huang, Q. Huang, Y. Yang, H. Yang, Y. Ding, W. Liang, X. Bai, T. Liang, Contrast-enhanced CT radiomics for preoperative evaluation of microvascular invasion in hepatocellular carcinoma: a two-center study, Clin. Transl. Med. 10 (2020).
[47] Q. Qiu, J. Duan, Z. Duan, X. Meng, C. Ma, J. Zhu, J. Lu, T. Liu, Y. Yin, Reproducibility and non-redundancy of radiomic features extracted from arterial phase CT scans in hepatocellular carcinoma patients: impact of tumor segmentation variability, Quant. Imag. Med. Surg. 9 (2019) 453–464.
[48] J. Peng, J. Zhang, Q. Zhang, Y. Xu, J. Zhou, L. Liu, A radiomics nomogram for preoperative prediction of microvascular invasion risk in hepatitis b virus-related hepatocellular carcinoma, Diagnostic Interv, Radiol. 24 (2018) 121–127.
[49] W. Wang, Q. Chen, Y. Iwamoto, X. Han, Q. Zhang, H. Hu, L. Lin, Y.W. Chen, Deep learning-based radiomics models for early recurrence prediction of hepatocellular carcinoma with multi-phase CT images and clinical data, in: Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, Annu Int Conf IEEE Eng Med Biol Soc, 2019, pp. 4881–4884.
[50] H. Wu, X. Han, Z. Wang, L. Mo, W. Liu, Y. Guo, X. Wei, X. Jiang, Prediction of the Ki-67 marker index in hepatocellular carcinoma based on CT radiomics features, Phys. Med. Biol. 65 (2020) 235048.
[51] I.C. Lee, J.Y. Huang, T.C. Chen, C.H. Yen, N.C. Chiu, H.E. Hwang, J.G. Huang, C.A. Liu, G.Y. Chau, R.C. Lee, Y.P. Hung, Y. Chao, S.Y. Ho, Y.H. Huang, Evolutionary learning-derived clinical-radiomic models for predicting early recurrence of hepatocellular carcinoma after resection, Liver Cancer 10 (2021) 572–582.
[52] C. Yuan, Z. Wang, D. Gu, J. Tian, P. Zhao, J. Wei, X. Yang, X. Hao, D. Dong, N. He, Y. Sun, W. Gao, J. Feng, Prediction early recurrence of hepatocellular carcinoma eligible for curative ablation using a Radiomics nomogram, Cancer Imag. 19 (2019).
[53] H. tong Hu, Q. yuan Shan, S. ling Chen, B. Li, S. ting Feng, E. jiao Xu, X. Li, J. yan Long, X. yan Xie, M. de Lu, M. Kuang, J. xian Shen, W. Wang, CT-based radiomics for preoperative prediction of early recurrent hepatocellular carcinoma: technical reproducibility of acquisition and scanners, Radiol, Méd. 125 (2020) 697–705.
[54] X. Xu, Y. Mao, Y. Tang, Y. Liu, C. Xue, Q. Yue, Q. Liu, J. Wang, Y. Yin, Classification of hepatocellular carcinoma and intrahepatic cholangiocarcinoma based on radiomic analysis, Comput. Math. Methods Med. 2022 (2022).
[55] M. Li, Y. Fan, H. You, C. Li, M. Luo, J. Zhou, A. Li, L. Zhang, X. Yu, W. Deng, J. Zhou, D. Zhang, Z. Zhang, H. Chen, Y. Xiao, B. Huang, J. Wang, Dual-energy CT deep learning radiomics to predict macrotrabecular-massive hepatocellular carcinoma, Radiology 308 (2023) e230255.
[56] S.Y. Wang, K. Sun, S. Jin, K.Y. Wang, N. Jiang, S.Q. Shan, Q. Lu, G.Y. Lv, J.H. Dong, Predicting the outcomes of hepatocellular carcinoma downstaging with the use of clinical and radiomics features, BMC Cancer 23 (2023).
[57] X. Xu, H.L. Zhang, Q.P. Liu, S.W. Sun, J. Zhang, F.P. Zhu, G. Yang, X. Yan, Y.D. Zhang, X.S. Liu, Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma, J. Hepatol. 70 (2019) 1133–1144.
[58] M. Wu, H. Tan, F. Gao, J. Hai, P. Ning, J. Chen, S. Zhu, M. Wang, S. Dou, D. Shi, Predicting the grade of hepatocellular carcinoma based on non-contrast-enhanced MRI radiomics signature, Eur. Radiol. 29 (2019) 2802–2811.
[59] S. Wu, J. Zheng, Y. Li, Z. Wu, S. Shi, M. Huang, H. Yu, W. Dong, J. Huang, T. Lin, Development and validation of an MRI-based radiomics signature for the preoperative prediction of lymph node metastasis in bladder cancer, EBioMedicine 34 (2018) 76–84.
[60] P. Schober, T.R. Vetter, Logistic regression in medical research, Anesth. Analg. 132 (2021) 365–366.
[61] X. Su, N. Chen, H. Sun, Y. Liu, X. Yang, W. Wang, S. Zhang, Q. Tan, J. Su, Q. Gong, Q. Yue, Automated machine learning based on radiomics features predicts H3 K27M mutation in midline gliomas of the brain, Neuro, Oncol. 22 (2020) 393–401.