Machine learning for lymph node metastasis prediction of in patients with gastric cancer: A systematic review and meta-analysis

Yilin Li, Fengjiao Xie, Qin Xiong, Honglin Lei
and Peimin Feng*

Department of Gastroenterology, Hospital of Chengdu University of Traditional Chinese Medicine,
Chengdu, China

Objective: To evaluate the diagnostic performance of machine learning (ML) in predicting lymph node metastasis (LNM) in patients with gastric cancer (GC) and to identify predictors applicable to the models.

Methods: PubMed, EMBASE, Web of Science, and Cochrane Library were searched from inception to March 16, 2022. The pooled c-index and accuracy were used to assess the diagnostic accuracy. Subgroup analysis was performed based on ML types. Meta-analyses were performed using random-effect models. Risk of bias assessment was conducted using PROBAST tool.

Results: A total of 41 studies (56182 patients) were included, and 33 of the studies divided the participants into a training set and a test set, while the rest of the studies only had a training set. The c-index of ML for LNM prediction in training set and test set was 0.837 [95%CI (0.814, 0.859)] and 0.811 [95%CI (0.785-0.838)], respectively. The pooled accuracy was 0.781 [(95%CI (0.756-0.805)] in training set and 0.753 [95%CI (0.721-0.783)] in test set. Subgroup analysis for different ML algorithms and staging of GC showed no significant difference. In contrast, in the subgroup analysis for predictors, in the training set, the model that included radiomics had better accuracy than the model with only clinical predictors (F = 3.546, p = 0.037). Additionally, cancer size, depth of cancer invasion and histological differentiation were the three most commonly used features in models built for prediction.

Conclusion: ML has shown to be of excellent diagnostic performance in predicting the LNM of GC. One of the models covering radiomics and its ML algorithms showed good accuracy for the risk of LNM in GC. However, the results revealed some methodological limitations in the development process. Future studies should focus on refining and improving existing models to improve the accuracy of LNM prediction.

Background

Gastric cancer (GC) is the fifth most common malignancy and the third leading cause of cancer-associated death worldwide (1–3). Lymph node metastasis (LNM) is one of the most sensitive prognostic factors for patients with GC (4–6). Patients at different lymph node stages may require different degrees of lymphadenectomy or neoadjuvant therapy (7–10), and typically have different outcomes. Therefore, it is of necessity to accurately predict and evaluate LNM before making treatment decisions (11, 12).

Non-invasive imaging modalities, including computed tomography (CT), functional magnetic resonance imaging (fMRI), and B-ultrasonography, have been widely applied for the evaluation of lymph node status in GC patients. However, the performances of these techniques remain controversial due to their sensitivity, specificity, and accuracy (13–19). Sentinel lymph node (SLN) biopsy is an invasive approach that has also been adopted for LNM detection in GC (20), while is still in debates on its effectiveness. Kitagawa et al. (21) and Miyashiro et al. (22) applied two different SLN biopsy methods, but reached different false negative rates (7% and 46.4%, respectively). Endoscopic ultrasonography combined with fine needle aspiration can be used for local lymph node staging and LNM diagnosing, while the former fails to detect distant metastases (23). Several new molecular biomarkers have been found to be useful for predicting LNM of GC, but the application of these agents is limited due to high cost and complex technological requirements (24, 25). There are indeed multiple methods that have potential to diagnose LNM, whereas their performances are tied down by so many limitations and uncertainties, making it an urgent need to find a more applicable and effective method for the identification of LNM status.

Machine learning (ML) algorithm is a newly emerged technique that is capable of accurate raw data-processing, important data connections-analyzing, and accurate decision-making (26, 27). Compared with conventional statistical methods, ML model has higher prediction accuracy (28, 29). It is of critical application value in assisting disease-diagnosing and prognosis-predicting through processing massive and complex medical data (30, 31). Currently, ML has been increasingly applied for LNM prediction in GC patients (32–72). However, different types of ML prediction models have great differences in both included predictors and calculation methods of the models (73, 74), the results produced by different models are far from unanimous (75). More importantly, there is no systematic review and meta-analysis conducted to assess ML for LNM prediction in GC patients. Therefore, we reviewed and synthesized all the related studies published previously to evaluate the accuracy of ML models for LNM prediction in GC patients.

Methods

This systematic review and meta-analysis was performed following the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) guidelines (76). The study was registered on PROSPERO (Registration No. CRD42022320752).

Literature retrieval strategy

PubMed, EMBASE, Web of Science, and Cochrane Library were systematically searched from inception to March 16, 2022 for all the related published articles. Search items mainly contained: "stomach neoplasms," "machine learning," and "lymphatic metastasis." References of included articles were also searched manually for potential eligible studies. Detailed search procedures and strategy are presented in Supplemental File 1.

Inclusion criteria

The studies were selected according to the following criteria:

(1) Patients were diagnosed with GC based on histopathological examination;

(2) Cohort study published in English, with the full-text available;

(3) Reported assessment of the performance of ML algorithm for LNM prediction;

(4) Clearly description for ML models and predictor variables used

(5) Reported the prediction performance indices of ML models and included sufficient data to infer the c-index and/or accuracy.

Studies meeting the following criteria were excluded:

(1) Limited sample size (less than 100);

(2) Letter, editorial, animal study, review, conference summary, consensus, case report, and guidelines;

(3) Research focusing on identifying or analyzing individual predictors, rather than the development and/or verification of models;

(4) Model performance measurements were not reported;

(5) Model building process or method was not described.

Data extraction

Data extraction was processed by two reviewers independently (YL and FX). The list of extraction items was designed based on the modified version of *Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies* (CHARMS) (77). Discrepancies were resolved by a third reviewer (PF). The following data were extracted:

(1) study characteristics (authors, publication date, study-design, and country or region);

(2) cohort characteristics (number of participants, number of patients with positive LNM, and cancer stages);

(3) Feature selection algorithms, number and types of predictor in final model, types of ML prediction model, and model validation and application;

(4) prediction outcomes, including c-index, accuracy, sensitivity, and specificity.

Risk of bias assessment

Risk of bias assessment and applicability of included studies were performed using the *Prediction Model Risk of Bias Assessment Tool* (PROBAST) (78), which includes four domains; participants, predictors, outcomes, and analysis. Risk of bias in each study was assessed based on the four domains, while the applicability was evaluated based on the first three domains. Each study was graded as "high risk", "low risk", or "unclear risk" (78). This process was conducted and cross-

checked by two reviewers (YL and FX) independently. Discrepancies were settled by the third reviewer (PF).

Statistical analysis

Data analysis was performed using R Statistical Software (version 4.1.1) with 'matrix', 'metafor' and 'meta' packages (79, 80). Subgroups were set based on ML algorithms. The c-index and accuracy for LNM prediction in GC patients, which were obtained from each study included, were measured with 95% confidence intervals (95% CIs) in the final analysis. For studies that did not report c-index, we calculated it *via* plotting receiver operating characteristic (ROC) curves based on reported probability distributions. The results from all included studies were pooled, and an overall estimated effect was evaluated using random-effect model which processed heterogeneity among studies (81). We used one-way ANOVA to discuss the differences in c-index and accuracy between the training group and the test group.

Results

Study selection

There were 2582 articles identified through searching the four databases, in which 1210 were excluded after duplicate-checking, 1126 excluded *via* browsing titles and abstracts. Full-texts of the remaining 246 articles were read, and 205 articles were excluded for reasons specified in Figure 1. Finally, a total of 41 studies were included (32–72).

Characteristics of included studies

There were 35 studies (85.4%) that were conducted in China (32, 33, 35–43, 45–49, 51–54, 56–63, 65–69, 71, 72), 5 (12.2%) in Korea (34, 44, 50, 55, 64) and 1 (2.4%) in Germany (70), with the publication date ranged from 2004 to 2022. The number of studies using ML for LNM prediction has gradually increased since 2018 (Figure 2). There were 35 retrospective studies (33, 35–53, 56–59, 61–70, 72) and 6 prospective studies (32, 34, 54, 55, 60, 71), with a total of 56182 participants, in which the number of patients with LNM was 12031. Characteristics of included studies are presented in Table 1.

Characteristics of machine learning in included studies

A total of 61 models were retrieved from included studies (ranged from 1 to 7 models in each study), with various modelling methods applied. The most frequently used ML

FIGURE 1
The PRISMA flow diagram for study selection.
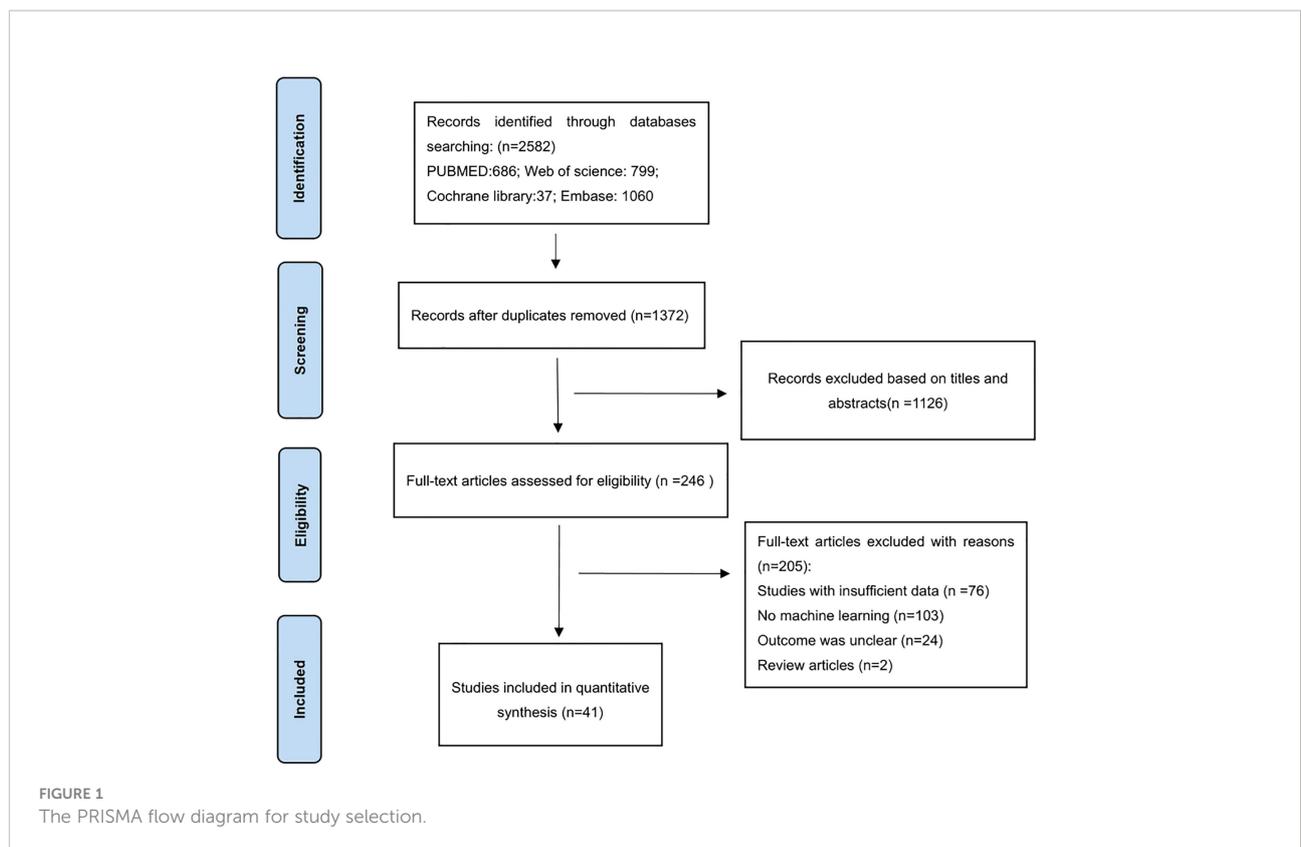
algorithms were logistic regression (LR) (n=30; 49.18%), support vector machine (SVM) (n=5; 8.2%), deep learning (DL) (n=4; 6.56%), and random forest (RF) (n=4; 6.56%) (Table 1). Feature selection is an important step for ML training. The number of features used in the models varied from 2 to 21, and Figure 3 summarizes the 15 most common features. The most commonly used predictors were tumor size (n=35; 14.96%), depth of tumor invasion (n=32; 13.68%), histology differentiation (n=20; 8.55%), imaging techniques (n=17; 7.26%), lymphovascular invasion (n=17; 7.26%), tumor location (n=14; 5.98%), CT-reported LN (n=11; 4.7%), age (n=8; 3.42%), macroscopic features (n=8; 3.42%), and CA199 (n=7; 2.99%).
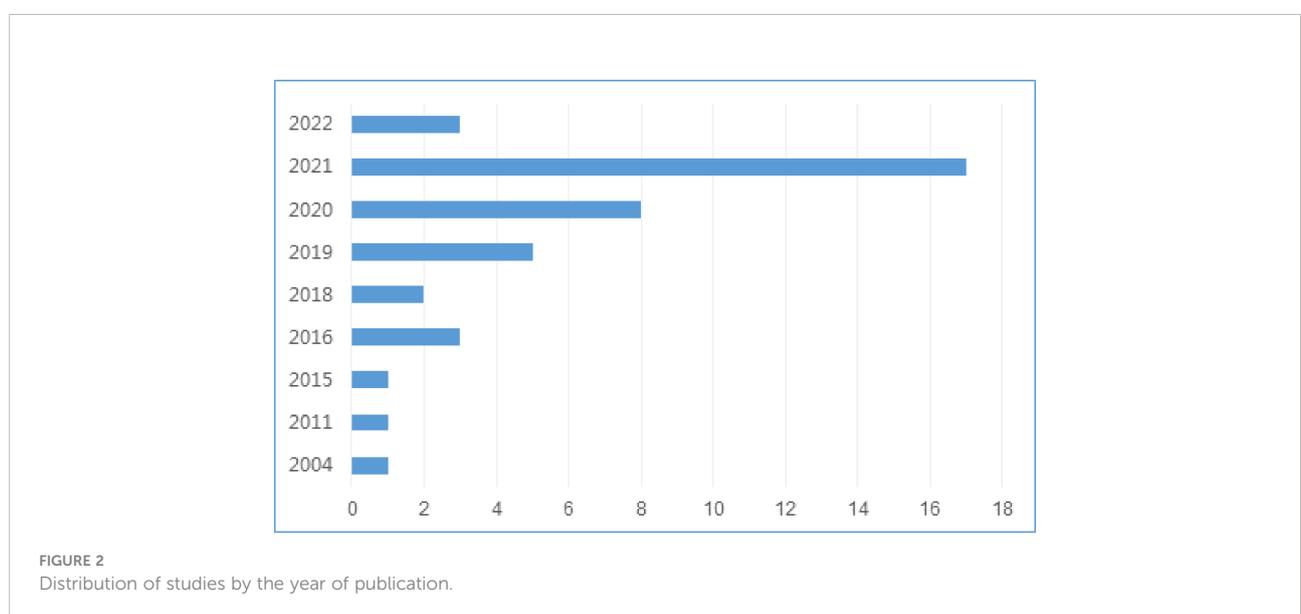

FIGURE 2
Distribution of studies by the year of publication.

TABLE 1 Characteristics of included studies.

Study	Country	Study design	Stage	No. patients in the train set	No. patients in the test set	Technique used for feature selection	Types of machine learning	Data source
Xiao-Peng Zhang (2011)	China	Retro	Early GC Borrmann I-IV	175	NA	LR	SVM	Single institution
Song Liu (2021)	China	Retro	Stages I-IV	122	41	LASSO	SVM, LR	Single institution
C Jin (2021)	China	Retro	Stages I-IV	1172	527	LR, RF	DL	Multiple institution
Xiaoxiao Wang (2021)	China	Retro	T1-2	80	79	LR	LR	Single institution
Xiao-Yi Yin (2020)	China	Retro	T1a, T1b	596	227	LR	LR	Single institution
Bang Wool Eom (2016)	Korea	Retro	T1a, T1b	336	NA	LR	LR	Single institution
Zhixue Zheng (2015)	China	Retro	T1a, T1b	262	NA	LR	LR	Single institution
Jing Li (2020)	China	Retro	Borrmann I-IV	136	68	LR	DL	Single institution
Zhengbing Wang (2021)	China	Retro	T1a, T1b	363	140	LR	LR	Single institution
HuaKai Tian (2022)	China	Retro	T1a, T1b	2294	227	LR	GLM, RPART, RF, GBM, SVM, RDA, ANN	Multiple institution
Zhixue Zheng (2016)	China	Retro	T1a, T1b	597	NA	LR	LR	Single institution
Yu Mei (2021)	China	Pros	T1a, T1b	794	418	LR	LR	Single institution
Jing Li (2018)	China	Retro	Borrmann I-IV	140	70	LR	LR	Single institution
Su Mi Kim (2020)	Korea	Pros	T1a, T1b	10579	2100	LR	LR	Single institution
Miaoquan Zhang (2021)	China	Retro	T1a, T1b	285	NA	LR	LR	Single institution
Yuming Jiang (2019)	China	Retro	Stages I-IV	312	1377	LR	LR	Multiple institution
Qiu-Xia Feng (2019)	China	Retro	Stages I-IV	326	164	SVM	SVM	Single institution
Jianfeng Mu (2019)	China	Retro	T1a, T1b	746	126	LR	LR	Single institution
Shilong Li (2021)	China	Retro	Stages I-IV	144	151	LR	LR	Single institution
Yue Wang (2020)	China	Retro	NA	197	50	LR	RF	Single institution
Chun Guang Guo (2016)	China	Retro	T1a, T1b	256	1273	LR	LR	Multiple institution
Cheng-Mao Zhou (2021)	China	Pros	T1a, T1b	818	351	GBDT	GBDT, XGB, RF, LR, XGB+LR, RF+LR, GBDT+LR	Single institution
Xujie Gao (2021)	China	Retro	T1a, T1b	308	155	LR	LR	Single institution
Xujie Gao (2020)	China	Retro	Stages I-IV	486	240	LR	LR	Single institution
Xu Wang (2021)	China	Retro	T1-4	250	99	LR	LR	Single institution
Siwei Pan (2021)	China	Retro	T1a, T1b	1274	637	LR	LR	Multiple institution

(Continued)

TABLE 1 Continued

Study	Country	Study design	Stage	No. patients in the train set	No. patients in the test set	Technique used for feature selection	Types of machine learning	Data source
Wujie Chen (2019)	China	Retro	T2-4	71	75	LR	LR	Single institution
Bong-Il Song (2020)	Korea	Retro	T1-4	377	189	LR	LR	Single institution
Chao Huang (2020)	China	Retro	NA	466	NA	RF	DT	Single institution
Lili Wang (2021)	China	Retro	T2-4	340	175	LR	LR	Single institution
Seokhwi Kim (2021)	Korea	Retro	T1a	28	108	LR	Bayesian	Multiple institution
Qiufang Liu (2021)	China	Retro	NA	185	NA	RF	DL	Single institution
Wannian Sui (2021)	China	Retro	T1a, T1b	1496	246	LR	LR	Multiple institution
Dexin Chen (2019)	China	Retro	T1a, T1b	232	143	LR	LR	Multiple institution
Lingwei Meng (2021)	China	Retro	T1-4	377	162	LASSO	LR	Multiple institution
D Dong (2020)	China	Retro	T2-4	225	505	Multivariable linear regression analysis, SVM	DL	Multiple institution
Zepang Sun (2021)	China	Pros	T1-4	531	1087	LR	LR	Multiple institution
Ji-Eun Na (2022)	Korea	Pros	T1a, T1b	10332	4428	LR, SVM, RF	LR, SVM, RF	Single institution
Haixing Zhu (2022)	China	Retro	T1a, T1b	1878	470	DT, GBM, LR, ANN, RF, XGBOOST	DT, GBM, LR, ANN, RF, XGBOOST	Multiple institution
Elfriede H Bollschweiler (2004)	Germany	Retro	Stages I-IV	135	NA	ANN	ANN	Single institution
Yinan Zhang (2018)	China	Pros	T1a, T1b	272	81	LR	LR	Single institution

ANN, artificial neural network; DL, deep learning; DT, decision tree; GBM, gradient boosting machine;GC, gastric cancer; GLM, generalized linear model; LASSO, Least Absolute Shrinkage and Selection Operator; LR, logistic regression; NA, not available; No., number; Pros, prospective; RDA, regularized dual averaging; Retro, retrospective; RF, random forest; SVM, support vector machine; XGBOOST, extreme gradient boosting;
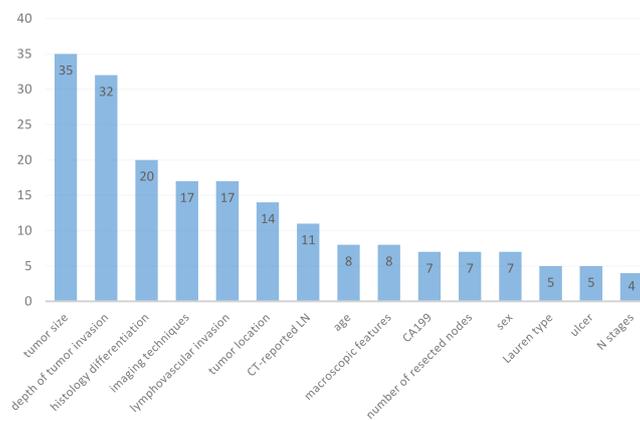

FIGURE 3
15 most frequently used predictors in 61 prediction models for gastric cancer patients.

Risk of bias and applicability assessment

All the included studies were of low risk of bias with respect to the domain of participants and outcome, 19 (46.34%) studies had low risk of bias in predictors (32, 34, 37, 38, 40, 42, 44, 54–56, 60, 62, 63, 66–71), 22 (53.66%) had unclear risk of bias due to that the prediction assessment was performed in the know of outcome data (33, 35, 36, 39, 41, 43, 45–53, 57–59, 61, 64, 65, 72). As for the domain of analysis, the risk of bias in 16 studies was considered high (33, 35, 41, 43, 47, 50, 51, 53, 58, 59, 61, 63, 65, 68, 70, 71), and the reasons were that (1): Insufficient sample size. Eight of the studies did not meet the standard of including at least 100 participants (2). Selection of predictors based on univariable analysis (3). Lack of external validation techniques. Eight studies lacked external validation in model development (35, 50, 51, 53, 59, 63, 65, 70). Concern regarding 'overall applicability' was rated as low in 15 studies (36.59%) (32, 34, 37, 38, 40, 42, 44, 54–56, 60, 62, 66, 67, 69), high in 16 studies (39.02%) (33, 35, 41, 43, 47, 50, 53, 58, 59, 61, 63, 65, 68, 70, 71) and unclear in the remaining 10 (24.39%) (36, 39, 45, 46, 48, 49, 52, 57, 64, 72). Risk of bias and applicability assessment were shown in Table 2.

C-index

There were different numbers for training and test models because there were five studies which only reported the training results (35, 50–53). The overall c-index for ML in training group was 0.837 [(95%CI (0.814, 0.859)] (Table 3; eFigure 1). LR, one of the most commonly used ML methods, resulted in an overall pooled c-index of 0.838 [(95%CI (0.812, 0.865)] (eFigure 2), while non-logistic regression (non-LR) model resulted in an overall pooled c-index of 0.83 [(95%CI (0.786, 0.877)] (eFigure 3).

Furthermore, the pooled c-index in test group was 0.811 [(95%CI (0.785, 0.838)] (Table 3; eFigure 4), which was similar with the result in training group. Subgroup analysis showed that 21 models in LR subgroup had a pooled c-index of 0.824 [95%CI (0.791, 0.858) (eFigure 5), and 13 models that used non-LR model assessment had a pooled c-index of 0.789 [95%CI (0.747, 0.833)] (eFigure 6).

Accuracy

There were different numbers of training and test models because there were ten studies which only reported the results of training group (35, 50–53, 59, 63, 65, 70, 71), whereas two other studies only reported that of test group. The ML models for LNM in training group showed an overall pooled accuracy of 0.781 [95%CI (0.756-0.805)] (Table 4; eFigure 7). Subgroup analysis showed no significant difference in different ML algorithms. LR algorithms had a pooled accuracy of 0.792 [95%CI (0.761-0.82)] (eFigure 8) and non-LR algorithms had that of 0.768 [95%CI (0.725, 0.805)] (eFigure 9).

In test group, the accuracy of the pooled 41 models was 0.753 [95%CI (0.721-0.783)] (Table 4; eFigure 10). Subgroup analysis was conducted based on LR and non-LR algorithms assessment. The pooled accuracy for the 23 models that used LR was 0.787 [95%CI 0.745, 0.824] (eFigure 11). The overall pooled accuracy for non-LR models was 0.707 [95%CI (0.665, 0.746)] (eFigure 12).

Subgroup analysis for early-gastric cancer and advanced gastric cancer

Of the 41 included studies, 21 were early-gastric GC (EGC) (T1) studies (32, 34, 35, 37, 39, 42, 46, 48–53, 55, 56, 60, 64, 65, 69, 71, 72) and 3 were advanced GC (T2-4) studies (43, 45, 67). In EGC, there was a pooled c-index of 0.832 [95%CI (0.804, 0.860)] (Table 5; eFigure 13) and 0.795 [95%CI (0.755, 0.838)] (eFigure 14) for the training and test groups, respectively. As for advanced GC, the pooled c-index for the training and test groups was 0.849 [95%CI (0.801-0.900)](eFigure 15) and 0.804 [95%CI (0.778-0.830)](eFigure 16), respectively.

Thirty-one models evaluated the accuracy of ML for EGC, and their pooled accuracy was 0.765 [95% CI (0.730-0.796)] (eFigure 17) for the training group. In terms of test group, the pooled accuracy for EGC was 0.731 [95%CI (0.686-0.773)] (eFigure 18). As for advanced GC, the training group had a pooled accuracy of 0.821 [95%CI (0.737-0.882)] (eFigure 19) while the test group had a pooled accuracy of 0.844 [95%CI (0.794-0.884)] (eFigure 20).

Subgroup analysis for predictors

Furthermore, we reviewed the predictors in the included original studies and we found three cases: Group A included only clinical predictors, Group B included only radiomic predictors, and Group C included both clinical and radiomic predictors.

In the training group, the c-index of groups A, B, and C was 0.822 ± 0.079 (n = 25), 0.852 ± 0.072 (n = 8), and 0.847 ± 0.063 (n = 8), respectively, with no significant difference between them (F = 0.604, p=0.552) (Table 6). In the test group, the c-index of groups A, B, and C was 0.792 ± 0.092 (n = 20), 0.83 ± 0.07 (n = 8), and 0.817 ± 0.043 (n = 6), respectively, and there was also no significant difference between them (F = 0.664, p = 0.522).

TABLE 2 Risk of bias and applicability assessment by PROBAST criteria.

Author	Year	Risk of bias				Overall applicability rating
		Participants	Predictors	Outcome	Analysis	
Xiaoxiao Wang	2021	low	unclear	low	high	high
Xiao-Yi Yin	2020	low	unclear	low	low	unclear
Bang Wool Eom	2016	low	unclear	low	high	high
Zhixue Zheng	2015	low	unclear	low	high	high
Zhengbing Wang	2021	low	unclear	low	low	unclear
Zhixue Zheng	2016	low	unclear	low	high	high
Yu Mei	2021	low	low	low	low	low
Jing Li	2018	low	unclear	low	high	high
Su Mi Kim	2020	low	low	low	low	low
Miaoquan Zhang	2021	low	unclear	low	high	high
Yuming Jiang	2019	low	unclear	low	low	unclear
Jianfeng Mu	2019	low	low	low	low	low
Shilong Li	2021	low	low	low	low	low
Chun Guang Guo	2016	low	unclear	low	low	unclear
Xujie Gao	2021	low	unclear	low	low	unclear
Xujie Gao	2020	low	low	low	low	low
Xu Wang	2021	low	unclear	low	high	high
Siwei Pan	2021	low	low	low	low	low
Wujie Chen	2019	low	unclear	low	high	high
Bong-Il Song	2020	low	low	low	low	low
Lili Wang	2021	low	unclear	low	low	unclear
Wannian Sui	2021	low	unclear	low	low	unclear
Dexin Chen	2019	low	unclear	low	low	unclear
Zepang Sun	2021	low	low	low	low	low
Xiao-Peng Zhang	2011	low	unclear	low	high	high
Song Liu	2021	low	unclear	low	high	high
C Jin	2021	low	low	low	low	low
Jing Li	2020	low	low	low	high	high
HuaKai Tian	2022	low	low	low	low	low
Qiu-Xia Feng	2019	low	unclear	low	low	unclear
Yue Wang	2020	low	unclear	low	high	high
Cheng-Mao Zhou	2021	low	low	low	low	low
Chao Huang	2020	low	low	low	high	high
Seokhwi Kim	2021	low	unclear	low	low	unclear
Qiufang Liu	2021	low	unclear	low	high	high
Lingwei Meng	2021	low	low	low	low	low
D Dong	2020	low	low	low	low	low
Ji-Eun Na	2022	low	low	low	low	low
Haixing Zhu	2022	low	low	low	low	low
Elfriede H Bollschweiler	2004	low	low	low	high	high
Yinan Zhang	2018	low	low	low	high	high

In the training group, the accuracy of groups A, B, and C were 0.75 ± 0.087 (n=34), 0.811 ± 0.066 (n = 9), and 0.822 ± 0.073 (n = 7), respectively, and there was a significant difference between them (F = 3.546, p = 0.037) (Table 6), and the model containing radiomics had better accuracy. In the test group, the accuracy of groups A, B, and C were 0.722 ± 0.098 (n = 28), 0.799 ± 0.075 (n = 8), and 0.795 ± 0.04 (n = 5), respectively, although there was no significant difference between them (F = 3.224,

TABLE 3   c-index for prediction models in gastric cancer patients.

Model	Train			Test		
	No. model	c-index	95%CI	No. model	c-index	95%CI
LR	26	0.838	0.812-0.865	21	0.824	0.791-0.858
Non-LR	13	0.83	0.786-0.877	13	0.789	0.747-0.833
DL	3	0.866	0.799-0.938	3	0.835	0.780-0.895
GBDT	1	0.798	0.714-0.892	1	0.788	0.688-0.902
GBDT+LR	1	0.626	0.529-0.740	1	0.65	0.557-0.759
RF	3	0.893	0.817-0.977	3	0.848	0.829-0.868
RF+LR	1	0.691	0.594-0.804	1	0.678	0.584-0.787
SVM	2	0.847	0.804-0.894	2	0.817	0.728-0.917
XGB	1	0.881	0.786-0.987	1	0.762	0.673-0.863
XGB+LR	1	0.739	0.648-0.842	1	0.619	0.521-0.736
Overall	39	0.837	0.814-0.859	34	0.811	0.785-0.838

No. model indicates the number of prediction models. DL, deep learning; LR, logistic regression; No., number; Non-LR, non logistic regression; RF, random forest; SVM, support vector machine.

p=0.051), the significance probability p-value was close to the critical value of 0.05. The mean value of accuracy was higher for models containing radiomics in the test cohort than for models containing only clinical predictors.

In summary, the model covering radiomics and its machine learning algorithms has better accuracy for the risk of lymph node metastasis in gastric cancer.

Discussion

The number of studies that apply ML to LNM prediction has been gradually increasing since 2018, making it important to systematically review the published studies so as to provide guidance for future research. To our knowledge, this is the first systematic review and meta-analysis that evaluated ML

TABLE 4   Results of meta-analyses of accuracy for prediction models for gastric cancer patients.

Model	Train			Test		
	No. model	accuracy	95%CI	No. model	accuracy	95%CI
LR	28	0.792	0.761-0.820	23	0.787	0.745-0.824
Non-LR	22	0.768	0.725-0.805	18	0.707	0.665-0.746
ANN	2	0.707	0.574-0.812	1	0.634	0.589-0.678
DL	2	0.818	0.755-0.868	1	0.765	0.646-0.859
DT	1	0.794	0.754-0.830	1	0.632	0.587-0.676
GBDT	1	0.835	0.808-0.860	1	0.815	0.770-0.854
GBDT+LR	1	0.903	0.881-0.923	1	0.573	0.519-0.625
GBM	1	0.618	0.597-0.638	1	0.687	0.643-0.729
GLM	1	0.667	0.647-0.686	NA	NA	NA
RDA	1	0.668	0.649-0.688	1	0.700	0.636-0.759
RF	4	0.793	0.710-0.858	4	0.723	0.678-0.764
RF+LR	1	0.644	0.610-0.677	1	0.578	0.525-0.631
RPART	1	0.625	0.604-0.645	NA	NA	NA
SVM	4	0.765	0.678-0.835	2	0.789	0.693-0.861
XGB	1	0.863	0.838-0.886	1	0.678	0.626-0.727
XGB+LR	1	0.806	0.777-0.832	1	0.581	0.528-0.633
Bayesian	NA	NA	NA	1	0.824	0.739-0.891
XGBOOST	NA	NA	NA	1	0.691	0.648-0.733
Overall	50	0.781	0.756-0.805	41	0.753	0.721-0.783

No. model indicates the number of prediction models. ANN, artificial neural network; DL, deep learning; DT, decision tree; GBM, gradient boosting machine; GLM, generalized linear model; LR, logistic regression; NA, not available; No., number; Non-LR, non logistic regression; RDA, regularized dual averaging; RF, random forest; SVM, support vector machine; XGBOOST, extreme gradient boosting.

TABLE 5 Subgroup analysis for early-gastric cancer and advanced gastric cancer.

Stage	Train		Test		Train		Test	
	No. model	c-index(95% CI)	No. model	c-index(95% CI)	No. model	Accuracy(95% CI)	No. model	Accuracy(95% CI)
EGC	24	0.832 (0.804-0.860)	19	0.795 (0.755-0.838)	31	0.765 (0.730-0.796)	26	0.731 (0.686-0.773)
Advanced GC	3	0.849 (0.801-0.900)	3	0.804 (0.778-0.830)	2	0.821 (0.737-0.882)	2	0.844 (0.794-0.884)

No. model indicates the number of prediction models. EGC, early-gastric cancer; GC, gastric cancer.

performance in the assessment of LNM in GC patients. ML-related studies can be methodologically categorized into LR and non-LR study. Several included studies were assessed to be of high or unclear risk of bias in the domains of prediction, analysis and overall applicability, which highlighted the current state of technology, as well as the need for methodological quality improvement.

This study demonstrated that ML had an excellent diagnostic performance in predicting LNM with great repeatability, which was in consistence with other studies. The pooled c-index and accuracy were 0.837 [95%CI (0.814,0.859)] and 0.781 [95%CI (0.756–0.805)], respectively. Significant heterogeneity existed between the studies, which could be caused by multiple factors. EGC is defined as a tumor limited to the mucosa and submucosa, regardless of the LNM (82). A subgroup analysis was performed since the difference in the order of magnitude characteristics of LNM between EGC and advanced gastric cancer may have a certain impact on the results of machine learning. It showed no significant difference in c-index or accuracy between EGC and advanced gastric cancer. In addition, since the included studies used LR or non-LR, subgroup analysis based on this variable was conducted to observe the changes in heterogeneity between the two groups. There was also no significant difference in c-index or accuracy among different ML algorithms. The type of ML algorithm had no effect on LNM prediction. Most importantly, this study was not designed to identify one superior algorithm from the other ones.

Feature selection was also critical to the performance and interpretation of the model. The most commonly used variables in the model development were tumor size, depth of tumor

invasion, histology differentiation, imaging techniques, lymphovascular invasion, tumor location, CT-reported LN, age, macroscopic features, and CA199. These variables are either anthropometric characteristics serving as markers of disease severity, or important factors contributing to the natural disease progression. These predictive indicators are easy to measure. Another merit of these predictors is the low risk of bias in measurement, resulting in a minimal possibility of exposure misclassification. Previous studies have revealed that the size of tumor is closely related to the incidence of LNM in patients with GC (83–85). Larger tumor typically indicates a higher risk of LNM (86–88), which might be attributed to easiness of invasion for larger tumor to the surrounding tissues. Depth of tumor invasion was found to be a strong predictor in 32 models (32, 34–37, 39, 41, 42, 46, 49, 50, 52, 54–56, 62, 69–71), which is in consistence with substantial evidence supporting its use as a predictor of LNM (36, 63, 89–91). There were 20 models considered histology differentiation as a vital factor for predicting LNM (34–37, 41, 42, 48–55, 61–63). Deeply infiltrated and poorly differentiated tumors might have sufficient nutritional support to facilitate its invasion to tissues, capillaries and lymphatic vessels, and thus to have the potential to grow and metastasize faster (69).

The novel PROBAST was applied for assessment of risk of bias and applicability of included prediction model studies, which allowed more details of the model, such as data source, processing, number of events per variable, feature selection, model development, and model validation, to be checked intensively (92–95). The PROBAST quality assessment revealed some other issues that could be avoided in future studies. First, external validation is rarely performed, which

TABLE 6 Subgroup analysis for predictors.

Model	Indicator	CP		RP		CP+RP		F	P
		n	mean(sd)	n	mean(sd)	n	mean(sd)		
Train	c-index	25	0.822(0.079)	8	0.852(0.072)	6	0.847(0.063)	0.604	0.552
	accuracy	34	0.75(0.087)	9	0.811(0.066)	7	0.822(0.073)	3.546	0.037
Test	c-index	20	0.792(0.092)	8	0.830(0.07)	6	0.817(0.043)	0.664	0.522
	accuracy	28	0.722(0.098)	8	0.799(0.075)	5	0.795(0.04)	3.224	0.051

n indicates the number of prediction models. CP, Clinical Predictors; RP, Radiomics Predictors.

might be a primary limitation in studies of this field. Simple determination of samples for modeling would lead to an overestimation for the model performance (96), and further accuracy verification for these models would be advisable. Guidelines that include external validation should be followed when reporting ML models (97). On the other hand, most of the included studies were retrospective design, leading to confounding and selection bias. More prospective studies are needed to produce evidence of high quality. The included studies also demonstrated another roadblock to the clinical implementation of ML. The data used in most of the included studies were from single institution, which resulted in limited datasets for training and failed to exert the advantage of ML that it is effective in processing large samples on multiple dimensions (98). Also, limited number of studies were less likely to be of broad public health significance. Future studies should take into accounts the expansion of datasets from multiple centers to increase the sample size and to improve classifier performance.

We also note the importance of preoperative assessment of peritoneal metastasis of GC for prognosis. Currently, the assessment of peritoneal metastases is mainly in the form of radiomics, but the data obtained by radiomics is obtained from a variety of sources, usually by CT, which may affect the results obtained (99–106). The heterogeneity of the results can be brought about by the different parameters and bits of CT and the artificial partitioning by different investigators through their own experience, so that the prediction of preoperative peritoneal metastases based on radiomics can be highly heterogeneous. At the same time, the application of radiomics generates a large amount of high-dimensional data, and the screening of these high-dimensional data is a great challenge in clinical practice. Therefore, although the prediction of preoperative peritoneal metastasis based on radiomics has been favored by a large number of researchers in recent years, these studies have not reached a clear consensus, thus resulting in a great variation in C-index (C-index ranged from 0.712 to 0.981) (99–106). We also expect subsequent studies based on radiomics to guide preoperative peritoneal metastases.

There were several limitations in this study. The first limitation was the significant heterogeneity. The sample sizes and distributions varied in different studies, as well as heterogenous variety of feature selection methods and ML algorithms, which compromised the performance and applicability of each model. However, such heterogeneity could be deemed as a key finding that should be addressed by future studies. Second, the estimation for prediction performance was based on limited data due to the incomplete reports of the results in several studies. Third, most of the reviewed studies included GC patients in different cancer stages, which might represent a possible confounding factor that disrupted ML performance in differential diagnosis. Fourth,

our findings should be interpreted prudently considering the potential significant publication bias. It is not suggested for investigators to report a test of unsatisfactory prediction values. It is probable that there might be instances in which ML might not have optimal prediction accuracies and so that has not been published yet (107). Last but not least, eight of the included studies that did not report the test set also affected the robustness of this study by causing a false high-performance result. It would be even better if all the studies provided external test results.

Conclusion

ML has shown excellent diagnostic performance for LNM prediction in GC patients, and ML models based on radiomics and clinical features could be a better potential prediction method. However, there were some methodological limitations in their development, and there is still room for improvement in predictive value. Future studies are needed to explore efficient, minimally invasive, and easily collected predictors for LNM so as to build more effective ML models and improve the accuracy of LNM prediction.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

Concept and design, YL, FX, QX, HL, and PF. Acquisition of data, YL, FX, and PF. Statistical analysis, YL, QX, and PF. Interpretation of data, YL, HL, and PF. Writing original draft, YL and PF. Writing review and editing, all authors. All authors have made substantial contributions to this work and have approved the final version of the manuscript.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.946038/full#supplementary-material

SUPPLEMENTAL FILE 1
Search strategy

SUPPLEMENTARY FIGURE 1
The overall pooled c-index of machine learning for lymph node metastasis prediction in train set

SUPPLEMENTARY FIGURE 2
The overall pooled c-index of logistic regression models for lymph node metastasis prediction in train set

SUPPLEMENTARY FIGURE 3
The overall pooled c-index of non-logistic regression models for lymph node metastasis prediction in train set

SUPPLEMENTARY FIGURE 4
The overall pooled c-index of machine learning for lymph node metastasis prediction in test set

SUPPLEMENTARY FIGURE 5
The overall pooled c-index of logistic regression models for lymph node metastasis prediction in test set

SUPPLEMENTARY FIGURE 6
The overall pooled c-index of non-logistic regression models for lymph node metastasis prediction in test set

SUPPLEMENTARY FIGURE 7
The overall pooled accuracy of machine learning for lymph node metastasis prediction in train set

SUPPLEMENTARY FIGURE 8
The overall pooled accuracy of logistic regression models for lymph node metastasis prediction in train set

SUPPLEMENTARY FIGURE 9
The overall pooled accuracy of non-logistic regression models for lymph node metastasis prediction in train set

SUPPLEMENTARY FIGURE 10
The overall pooled accuracy of machine learning for lymph node metastasis prediction in test set

SUPPLEMENTARY FIGURE 11
The overall pooled accuracy of logistic regression models for lymph node metastasis prediction in test set

SUPPLEMENTARY FIGURE 12
The overall pooled accuracy of non-logistic regression models for lymph node metastasis prediction in test set

SUPPLEMENTARY FIGURE 13
The overall pooled c-index for predicting lymph node metastasis of early-gastric cancer in train set

SUPPLEMENTARY FIGURE 14
The overall pooled c-index for predicting lymph node metastasis of early-gastric cancer in test set

SUPPLEMENTARY FIGURE 15
The overall pooled c-index for predicting lymph node metastasis of advanced gastric cancer in train set

SUPPLEMENTARY FIGURE 16
The overall pooled c-index for predicting lymph node metastasis of advanced gastric cancer in test set

SUPPLEMENTARY FIGURE 17
The overall pooled accuracy for predicting lymph node metastasis of early-gastric cancer in train set

SUPPLEMENTARY FIGURE 18
The overall pooled accuracy for predicting lymph node metastasis of early-gastric cancer in test set

SUPPLEMENTARY FIGURE 19
The overall pooled accuracy for predicting lymph node metastasis of advanced gastric cancer in train set

SUPPLEMENTARY FIGURE 20
The overall pooled accuracy for predicting lymph node metastasis of advanced gastric cancer in test set

References

1. Joshi SS, Badgwell BD. Current treatment and recent progress in gastric cancer. *CA: Cancer J Clin* (2021) 71(3):264–79. doi: 10.3322/caac.21657

2. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. *Lancet (London England)* (2020) 396(10251):635–48. doi: 10.1016/S0140-6736(20)31288-5

3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492

4. Isomoto H, Shikuwa S, Yamaguchi N, Fukuda E, Ikeda K, Nishiyama H, et al. Endoscopic submucosal dissection for early gastric cancer: a large-scale feasibility study. *Gut* (2009) 58(3):331–6. doi: 10.1136/gut.2008.165381

5. Ke B, Liang H. Current status of lymph node dissection in gastric cancer. *Chin J Cancer Res = Chung-kuo yen cheng yen chiu* (2021) 33(2):193–202. doi: 10.21147/j.issn.1000-9604.2021.02.07

6. Rawicz-Pruszyński K, Ciseł B, Mlak R, Mielko J, Skórzewska M, Kwietniewska M, et al. The role of the lymph node ratio in advanced gastric cancer after neoadjuvant chemotherapy. *Cancers* (2019) 11(12):1914. doi: 10.3390/cancers11121914

7. Coccolini F, Montori G, Ceresoli M, Cima S, Valli MC, Nita GE, et al. Advanced gastric cancer: What we know and what we still have to learn. *World J Gastroenterol* (2016) 22(3):1139–59. doi: 10.3748/wjg.v22.i3.1139

8. Degiuli M, De Manzoni G, Di Leo A, D'Ugo D, Galasso E, Marrelli D, et al. Gastric cancer: Current status of lymph node dissection. *World J Gastroenterol* (2016) 22(10):2875–93. doi: 10.3748/wjg.v22.i10.2875

9. Japanese Gastric cancer treatment guidelines 2018 (5th edition). *Gastric Cancer* (2021) 24(1):1–21. doi: 10.1007/s10120-020-01042-y

10. Wang FH, Zhang XT, Li YF, Tang L, Qu XJ, Ying JE, et al. The Chinese society of clinical oncology (CSCO): Clinical guidelines for the diagnosis and treatment of gastric cancer, 2021. *Cancer Commun (London England)* (2021) 41(8):747–95. doi: 10.1002/cac2.12193

11. Abdelfatah MM, Barakat M, Lee H, Kim JJ, Uedo N, Grimm I, et al. The incidence of lymph node metastasis in early gastric cancer according to the expanded criteria in comparison with the absolute criteria of the Japanese gastric cancer association: a systematic review of the literature and meta-analysis. *Gastrointestinal Endoscopy* (2018) 87(2):338–47. doi: 10.1016/j.gie.2017.09.025

12. Yang J, Wu Q, Xu L, Wang Z, Su K, Liu R, et al. Integrating tumor and nodal radiomics to predict lymph node metastasis in gastric cancer. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2020) 150:89–96. doi: 10.1016/j.radonc.2020.06.004

13. Cardoso R, Coburn N, Seevaratnam R, Sutradhar R, Lourenco LG, Mahar A, et al. A systematic review and meta-analysis of the utility of EUS for preoperative staging for gastric cancer. *Gastric Cancer* (2012) 15 Suppl 1:S19–26. doi: 10.1007/s10120-011-0115-4

14. Kwee RM, Kwee TC. Imaging in assessing lymph node status in gastric cancer. *Gastric Cancer* (2009) 12(1):6–22. doi: 10.1007/s10120-008-0492-5

15. Zhong J, Zhao W, Ren F, Qi S, Wang X, Lv T, et al. Lymph node metastasis in patients with gastric cancer: a multi-modality, morphologic and functional imaging study. *Am J Trans Res* (2016) 8(12):5601–9.

16. Pang L, Wang J, Fan Y, Xu R, Bai Y, Bai L. Correlations of TNM staging and lymph node metastasis of gastric cancer with MRI features and VEGF expression. *Cancer Biomarkers Section A Dis Markers* (2018) 23(1):53–9. doi: 10.3233/CBM-181287

17. Kim HJ, Kim AY, Oh ST, Kim JS, Kim KW, Kim PN, et al. Gastric cancer staging at multi-detector row CT gastrography: comparison of transverse and volumetric CT scanning. *Radiology* (2005) 236(3):879–85. doi: 10.1148/radiol.2363041101

18. Chen J, Cheong JH, Yun MJ, Kim J, Lim JS, Hyung WJ, et al. Improvement in preoperative staging of gastric adenocarcinoma with positron emission tomography. *Cancer* (2005) 103(11):2383–90. doi: 10.1002/cncr.21074

19. Seevaratnam R, Cardoso R, McGregor C, Lourenco L, Mahar A, Sutradhar R, et al. How useful is preoperative imaging for tumor, node, metastasis (TNM) staging of gastric cancer? a meta-analysis. *Gastric Cancer* (2012) 15 Suppl 1:S3–18. doi: 10.1007/s10120-011-0069-6

20. Hiratsuka M, Miyashiro I, Ishikawa O, Furukawa H, Motomura K, Ohigashi H, et al. Application of sentinel node biopsy to gastric cancer surgery. *Surgery* (2001) 129(3):335–40. doi: 10.1067/msy.2001.111699

21. Kitagawa Y, Takeuchi H, Takagi Y, Natsugoe S, Terashima M, Murakami N, et al. Sentinel node mapping for gastric cancer: a prospective multicenter trial in Japan. *J Clin Oncol* (2013) 31(29):3704–10. doi: 10.1200/JCO.2013.50.3789

22. Miyashiro I, Hiratsuka M, Sasako M, Sano T, Mizusawa J, Nakamura K, et al. High false-negative proportion of intraoperative histological examination as a serious problem for clinical application of sentinel node biopsy for early gastric cancer: final results of the Japan clinical oncology group multicenter trial JCOG0302. *Gastric Cancer* (2014) 17(2):316–23. doi: 10.1007/s10120-013-0285-3

23. Li B, Zheng P, Zhu Q, Lin J. Accurate preoperative staging of gastric cancer with combined endoscopic ultrasonography and PET-CT. *Tohoku J Exp Med* (2012) 228(1):9–16. doi: 10.1620/tjem.228.9

24. Huang KH, Lan YT, Fang WL, Chen JH, Lo SS, Li AF, et al. The correlation between miRNA and lymph node metastasis in gastric cancer. *BioMed Res Int* (2015) 2015:543163. doi: 10.1155/2015/543163

25. Li W, Ye F, Wang D, Sun X, Tong W, Lian G, et al. Protein predictive signatures for lymph node metastasis of gastric cancer. *Int J Cancer* (2013) 132(8):1851–9. doi: 10.1002/ijc.27864

26. Oliveira AL. Biotechnology, big data and artificial intelligence. *Biotechnol J* (2019) 14(8):e1800613. doi: 10.1002/biot.201800613

27. Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine learning and integrative analysis of biomedical big data. *Genes* (2019) 10(2):87. doi: 10.3390/genes10020087

28. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning? a primer for the epidemiologist. *Am J Epidemiol* (2019) 188(12):2222–39. doi: 10.1093/aje/kwz189

29. Wang Z, Li H, Carpenter C, Guan Y. Challenge-enabled machine learning to drug-response prediction. *AAPS J* (2020) 22(5):106. doi: 10.1208/s12248-020-00494-5

30. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Internal Med* (2018) 284(6):603–19. doi: 10.1111/joim.12822

31. Fleuren LM, Klausch TLT, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med* (2020) 46(3):383–400. doi: 10.1007/s00134-019-05872-y

32. Mei Y, Wang S, Feng T, Yan M, Yuan F, Zhu Z, et al. Nomograms involving HER2 for predicting lymph node metastasis in early gastric cancer. *Front Cell Dev Biol* (2021) 9:781824. doi: 10.3389/fcell.2021.781824

33. Li J, Fang M, Wang R, Dong D, Tian J, Liang P, et al. Diagnostic accuracy of dual-energy CT-based nomograms to predict lymph node metastasis in gastric cancer. *Eur Radiol* (2018) 28(12):5241–9. doi: 10.1007/s00330-018-5483-2

34. Kim SM, Min BH, Ahn JH, Jung SH, An JY, Choi MG, et al. Nomogram to predict lymph node metastasis in patients with early gastric cancer: a useful clinical tool to reduce gastrectomy after endoscopic resection. *Endoscopy* (2020) 52(6):435–43. doi: 10.1055/a-1117-3059

35. Zhang M, Ding C, Xu L, Feng S, Ling Y, Guo J, et al. A nomogram to predict risk of lymph node metastasis in early gastric cancer. *Sci Rep* (2021) 11(1):22873. doi: 10.1038/s41598-021-02305-z

36. Jiang Y, Wang W, Chen C, Zhang X, Zha X, Lv W, et al. Radiomics signature on computed tomography imaging: Association with lymph node metastasis in patients with gastric cancer. *Front Oncol* (2019) 9:340. doi: 10.3389/fonc.2019.00340

37. Mu J, Jia Z, Yao W, Song J, Cao X, Jiang J, et al. Predicting lymph node metastasis in early gastric cancer patients: development and validation of a model. *Future Oncol (London England)* (2019) 15(31):3609–17. doi: 10.2217/fon-2019-0377

38. Li S, Zhao Z, Yang H, Wang D, Sun W, Li S, et al. Construction and validation of a nomogram for the preoperative prediction of lymph node metastasis in gastric cancer. *Cancer Control J Moffitt Cancer Center* (2021) 28:10732748211027160. doi: 10.1177/10732748211027160

39. Guo CG, Chen YJ, Ren H, Zhou H, Shi JF, Yuan XH, et al. A nomogram for predicting the likelihood of lymph node metastasis in early gastric signet ring cell carcinoma: A single center retrospective analysis with external validation. *Medicine* (2016) 95(46):e5393. doi: 10.1097/MD.0000000000005393

40. Gao X, Ma T, Cui J, Zhang Y, Wang L, Li H, et al. A radiomics-based model for prediction of lymph node metastasis in gastric cancer. *Eur J Radiol* (2020) 129:109069. doi: 10.1016/j.ejrad.2020.109069

41. Wang X, He Q, Liang H, Liu J, Xu X, Jiang K, et al. A novel robust nomogram based on preoperative hemoglobin and albumin levels and lymphocyte and platelet counts (HALP) for predicting lymph node metastasis of gastric cancer. *J Gastrointestinal Oncol* (2021) 12(6):2706–18. doi: 10.21037/jgo-21-507

42. Pan S, An W, Tan Y, Chen Q, Liu P, Xu H. Prediction model of lymph node metastasis risk in elderly patients with early gastric cancer before endoscopic resection: A retrospective analysis based on international multicenter data. *J Cancer* (2021) 12(18):5583–92. doi: 10.7150/jca.56702

43. Chen W, Wang S, Dong D, Gao X, Zhou K, Li J, et al. Evaluation of lymph node metastasis in advanced gastric cancer using magnetic resonance imaging-based radiomics. *Front Oncol* (2019) 9:1265. doi: 10.3389/fonc.2019.01265

44. Song BI. Nomogram using f-18 fluorodeoxyglucose positron emission tomography/computed tomography for preoperative prediction of lymph node metastasis in gastric cancer. *World J Gastrointestinal Oncol* (2020) 12(4):447–56. doi: 10.4251/wjgo.v12.i4.447

45. Wang L, Gong J, Huang X, Lin G, Zheng B, Chen J, et al. CT-based radiomics nomogram for preoperative prediction of No.10 lymph nodes metastasis in advanced proximal gastric cancer. *Eur J Surg Oncol J Eur Soc Surg Oncol Br Assoc Surg Oncol* (2021) 47(6):1458–65. doi: 10.1016/j.ejso.2020.11.132

46. Sui W, Chen Z, Li C, Chen P, Song K, Wei Z, et al. Nomograms for predicting the lymph node metastasis in early gastric cancer by gender: A retrospective multicentric study. *Front Oncol* (2021) 11:616951. doi: 10.3389/fonc.2021.616951

47. Wang X, Li C, Fang M, Zhang L, Zhong L, Dong D, et al. Integrating No.3 lymph nodes and primary tumor radiomics to predict lymph node metastasis in T1-2 gastric cancer. *BMC Med Imaging* (2021) 21(1):58. doi: 10.1186/s12880-021-00587-3

48. Yin XY, Pang T, Liu Y, Cui HT, Luo TH, Lu ZM, et al. Development and validation of a nomogram for preoperative prediction of lymph node metastasis in early gastric cancer. *World J Surg Oncol* (2020) 18(1):2. doi: 10.1186/s12957-019-1778-2

49. Chen D, Chen G, Jiang W, Fu M, Liu W, Sui J, et al. Association of the collagen signature in the tumor microenvironment with lymph node metastasis in early gastric cancer. *JAMA Surgery* (2019) 154(3):e185249. doi: 10.1001/jamasurg.2018.5249

50. Eom BW, Joo J, Park B, Jo MJ, Choi SH, Cho SJ, et al. Nomogram incorporating CD44v6 and clinicopathological factors to predict lymph node metastasis for early gastric cancer. *PloS One* (2016) 11(8):e0159424. doi: 10.1371/journal.pone.0159424

51. Zheng Z, Zhang Y, Zhang L, Li Z, Wu A, Wu X, et al. Nomogram for predicting lymph node metastasis rate of submucosal gastric cancer by analyzing

clinicopathological characteristics associated with lymph node metastasis. *Chin J Cancer Res = Chung-kuo yen cheng yen chiu* (2015) 27(6):572–9. doi: 10.3978/j.issn.1000-9604.2015.12.06

52. Wang Z, Liu J, Luo Y, Xu Y, Liu X, Wei L, et al. Establishment and verification of a nomogram for predicting the risk of lymph node metastasis in early gastric cancer. *Rev Espanola Enfermedades Digestivas Organo Oficial la Sociedad Espanola Patologia Digestiva* (2021) 113(6):411–7. doi: 10.17235/reed.2020.7102/2020

53. Zheng Z, Zhang Y, Zhang L, Li Z, Wu X, Liu Y, et al. A nomogram for predicting the likelihood of lymph node metastasis in early gastric patients. *BMC Cancer* (2016) 16:92. doi: 10.1186/s12885-016-2132-5

54. Sun Z, Jiang Y, Chen C, Zheng H, Huang W, Xu B, et al. Radiomics signature based on computed tomography images for the preoperative prediction of lymph node metastasis at individual stations in gastric cancer: A multicenter study. *Radiother Oncol J Eur Soc Ther Radiol Oncol* (2021) 165:179–90. doi: 10.1016/j.radonc.2021.11.003

55. Na JE, Lee YC, Kim TJ, Lee H, Won HH, Min YW, et al. Machine learning model to stratify the risk of lymph node metastasis for early gastric cancer: A single-center cohort study. *Cancers* (2022) 14(5):1121. doi: 10.3390/cancers14051121

56. Zhu H, Wang G, Zheng J, Zhu H, Huang J, Luo E, et al. Preoperative prediction for lymph node metastasis in early gastric cancer by interpretable machine learning models: A multicenter study. *Surgery* (2022) 171(6):1543–51. doi: 10.1016/j.surg.2021.12.015

57. Feng QX, Liu C, Qi L, Sun SW, Song Y, Yang G, et al. An intelligent clinical decision support system for preoperative prediction of lymph node metastasis in gastric cancer. *J Am Coll Radiol JACR.* (2019) 16(7):952–60. doi: 10.1016/j.jacr.2018.12.017

58. Wang Y, Liu W, Yu Y, Liu JJ, Xue HD, Qi YF, et al. CT radiomics nomogram for the preoperative prediction of lymph node metastasis in gastric cancer. *Eur Radiol* (2020) 30(2):976–86. doi: 10.1007/s00330-019-06398-z

59. Zhang XP, Wang ZL, Tang L, Sun YS, Cao K, Gao Y. Support vector machine model for diagnosis of lymph node metastasis in gastric cancer with multidetector computed tomography: a preliminary study. *BMC Cancer* (2011) 11:10. doi: 10.1186/1471-2407-11-10

60. Zhou CM, Wang Y, Ye HT, Yan S, Ji M, Liu P, et al. Machine learning predicts lymph node metastasis of poorly differentiated-type intramucosal gastric cancer. *Sci Rep* (2021) 11(1):1300. doi: 10.1038/s41598-020-80582-w

61. Liu S, Qiao X, Xu M, Ji C, Li L, Zhou Z. Development and validation of multivariate models integrating preoperative clinicopathological parameters and radiographic findings based on late arterial phase CT images for predicting lymph node metastasis in gastric cancer. *Acad Radiol* (2021) 28 Suppl 1:S167–s78. doi: 10.1016/j.acra.2021.01.011

62. Jin C, Jiang Y, Yu H, Wang W, Li B, Chen C, et al. Deep learning analysis of the primary tumour and the prediction of lymph node metastases in gastric cancer. *Br J Surgery* (2021) 108(5):542–9. doi: 10.1002/bjs.11928

63. Huang C, Hu C, Zhu J, Zhang W, Huang J, Zhu Z. Establishment of decision rules and risk assessment model for preoperative prediction of lymph node metastasis in gastric cancer. *Front Oncol* (2020) 10:1638. doi: 10.3389/fonc.2020.01638

64. Kim S, Bae WJ, Ahn JM, Heo JH, Kim KM, Choi KW, et al. MicroRNA signatures associated with lymph node metastasis in intramucosal gastric cancer. *Modern Pathol* (2021) 34(3):672–83. doi: 10.1038/s41379-020-00681-x

65. Liu Q, Li J, Xin B, Sun Y, Feng D, Fulham MJ, et al. (18)F-FDG PET/CT radiomics for preoperative prediction of lymph node metastases and nodal staging in gastric cancer. *Front Oncol* (2021) 11:723345. doi: 10.3389/fonc.2021.723345

66. Meng L, Dong D, Chen X, Fang M, Wang R, Li J, et al. 2D and 3D CT radiomic features performance comparison in characterization of gastric cancer: A multi-center study. *IEEE J Biomed Health Informatics* (2021) 25(3):755–63. doi: 10.1109/JBHI.2020.3002805

67. Dong D, Fang MJ, Tang L, Shan XH, Gao JB, Giganti F, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. *Ann Oncol* (2020) 31(7):912–20. doi: 10.1016/j.annonc.2020.04.003

68. Li J, Dong D, Fang M, Wang R, Tian J, Li H, et al. Dual-energy CT-based deep learning radiomics can improve lymph node metastasis risk prediction for gastric cancer. *Eur Radiol* (2020) 30(4):2324–33. doi: 10.1007/s00330-019-06621-x

69. Tian H, Ning Z, Zong Z, Liu J, Hu C, Ying H, et al. Application of machine learning algorithms to predict lymph node metastasis in early gastric cancer. *Front Med* (2021) 8:759013. doi: 10.3389/fmed.2021.759013

70. Bollschweiler EH, Mönig SP, Hensler K, Baldus SE, Maruyama K, Hölscher AH. Artificial neural network for prediction of lymph node metastases in gastric cancer: a phase II diagnostic study. *Ann Surg Oncol* (2004) 11(5):506–11. doi: 10.1245/ASO.2004.04.018

71. Zhang Y, Liu Y, Zhang J, Wu X, Ji X, Fu T, et al. Construction and external validation of a nomogram that predicts lymph node metastasis in early gastric cancer patients using preoperative parameters. *Chin J Cancer Res = Chung-kuo yen cheng yen chiu* (2018) 30(6):623–32. doi: 10.21147/j.issn.1000-9604.2018.06.07

72. Gao X, Ma T, Cui J, Zhang Y, Wang L, Li H, et al. A CT-based radiomics model for prediction of lymph node metastasis in early stage gastric cancer. *Acad Radiol* (2021) 28(6):e155–e64. doi: 10.1016/j.acra.2020.03.045

73. Wang ZK, Lin JX, Li P, Xie JW, Wang JB, Lu J, et al. Higher risk of lymph node metastasis in young patients with early gastric cancer. *J Cancer* (2019) 10(18):4389–96. doi: 10.7150/jca.30260

74. Chu YN, Yu YN, Jing X, Mao T, Chen YQ, Zhou XB, et al. Feasibility of endoscopic treatment and predictors of lymph node metastasis in early gastric cancer. *World J Gastroenterol* (2019) 25(35):5344–55. doi: 10.3748/wjg.v25.i35.5344

75. Chen C, Qin Y, Chen H, Zhu D, Gao F, Zhou X. A meta-analysis of the diagnostic performance of machine learning-based MRI in the prediction of axillary lymph node metastasis in breast cancer patients. *Insights into Imaging* (2021) 12(1):156. doi: 10.1186/s13244-021-01034-1

76. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ (Clinical Res ed).* (2009) 339:b2700. doi: 10.1136/bmj.b2700

77. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PloS Med* (2014) 11(10):e1001744. doi: 10.1371/journal.pmed.1001744

78. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Ann Internal Med* (2019) 170(1):51–8. doi: 10.7326/M18-1376

79. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with r: a practical tutorial. *Evidence-Based Ment Health* (2019) 22(4):153–60. doi: 10.1136/ebmental-2019-300117

80. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* (2010) 36(3):1–48. doi: 10.18637/jss.v036.i03

81. Dersimonian R, Nan L. Meta-analysis in clinical trials. *Controlled Clinical Trials* (1986) 7(3):177. doi: 10.1016/0197-2456(86)90046-2

82. Ono H, Yao K, Fujishiro M, Oda I, Uedo N, Nimura S, et al. Guidelines for endoscopic submucosal dissection and endoscopic mucosal resection for early gastric cancer (second edition). *Digestive Endoscopy* (2021) 33(1):4–20. doi: 10.1111/den.13883

83. Wang X, Wan F, Pan J, Yu GZ, Chen Y, Wang JJ. Tumor size: a non-neglectable independent prognostic factor for gastric cancer. *J Surg Oncol* (2008) 97(3):236–40. doi: 10.1002/jso.20951

84. Adachi Y, Oshiro T, Mori M, Maehara Y, Sugimachi K. Tumor size as a simple prognostic indicator for gastric carcinoma. *Ann Surg Oncol* (1997) 4(2):137–40. doi: 10.1007/BF02303796

85. Liu X, Xu Y, Long Z, Zhu H, Wang Y. Prognostic significance of tumor size in T3 gastric cancer. *Ann Surg Oncol* (2009) 16(7):1875–82. doi: 10.1245/s10434-009-0449-x

86. Novotny AR, Schuhmacher C. Predicting lymph node metastases in early gastric cancer: radical resection or organ-sparing therapy? *Gastric Cancer* (2008) 11(3):131–3. doi: 10.1007/s10120-008-0479-2

87. Ohashi S, Okamura S, Urano F, Maeda M. Clinicopathological variables associated with lymph node metastasis in submucosal invasive gastric cancer. *Gastric Cancer* (2007) 10(4):241–50. doi: 10.1007/s10120-007-0442-7

88. Ma X, Zhang Q, Zhu S, Zhang S, Sun X. Risk factors and prediction model for non-curative resection of early gastric cancer with endoscopic resection and the evaluation. *Front Med* (2021) 8:637875. doi: 10.3389/fmed.2021.637875

89. Abe N, Watanabe T, Suzuki K, Machida H, Toda H, Nakaya Y, et al. Risk factors predictive of lymph node metastasis in depressed early gastric cancer. *Am J Surgery* (2002) 183(2):168–72. doi: 10.1016/S0002-9610(01)00860-1

90. Chen S, Nie RC, OuYang LY, Li YF, Xiang J, Zhou ZW, et al. Nomogram analysis and external validation to predict the risk of lymph node metastasis in gastric cancer. *Oncotarget* (2017) 8(7):11380–8. doi: 10.18632/oncotarget.14535

91. Nakagawa M, Choi YY, An JY, Chung H, Seo SH, Shin HB, et al. Difficulty of predicting the presence of lymph node metastases in patients with clinical early stage gastric cancer: a case control study. *BMC Cancer* (2015) 15:943. doi: 10.1186/s12885-015-1940-3

92. Palazón-Bru A, Mares-García E, López-Bru D, Mares-Arambul E, Folgado-de la Rosa DM, Carbonell-Torregrosa M, et al. A critical appraisal of the clinical

applicability and risk of bias of the predictive models for mortality and recurrence in patients with oropharyngeal cancer: Systematic review. *Head neck* (2020) 42 (4):763–73. doi: 10.1002/hed.26025

93. Di Tanna GL, Wirtz H, Burrows KL, Globe G. Evaluating risk prediction models for adults with heart failure: A systematic literature review. *PloS One* (2020) 15(1):e0224135. doi: 10.1371/journal.pone.0224135

94. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ (Clinical Res ed).* (2020) 368:m689. doi: 10.1136/bmj.m689

95. Dretzke J, Chuchu N, Agarwal R, Herd C, Chua W, Fabritz L, et al. Predicting recurrent atrial fibrillation after catheter ablation: a systematic review of prognostic models. *Europace Eur Pacing Arrhythmias Cardiac Electrophysiology J Working Groups Cardiac Pacing Arrhythmias Cardiac Cell Electrophysiology Eur Soc Cardiol* (2020) 22(5):748–60. doi: 10.1093/europace/euaa041

96. Bellou V, Belbasis L, Konstantinidis AK, Tzoulaki I, Evangelou E. Prognostic models for outcome prediction in patients with chronic obstructive pulmonary disease: systematic review and critical appraisal. *BMJ (Clinical Res ed).* (2019) 367: l5358. doi: 10.1136/bmj.l5358

97. Bedrikovetski S, Dudi-Venkata NN, Kroon HM, Seow W, Vather R, Carneiro G, et al. Artificial intelligence for pre-operative lymph node staging in colorectal cancer: a systematic review and meta-analysis. *BMC Cancer* (2021) 21 (1):1058. doi: 10.1186/s12885-021-08773-w

98. Al-Jarrah OY, Yoo PD, Muhaidat S, Karagiannidis GK, Taha K. Efficient machine learning for big data: A review. *Big Data Res* (2015) 2(3):87–93. doi: 10.1016/j.bdr.2015.04.001

99. Dong D, Tang L, Li ZY, Fang MJ, Gao JB, Shan XH, et al. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer. *Ann Oncol* (2019) 30(3):431–8. doi: 10.1093/annonc/mdz001

100. Xue B, Jiang J, Chen L, Wu S, Zheng X, Zheng X, et al. Development and validation of a radiomics model based on (18)F-FDG PET of primary gastric cancer for predicting peritoneal metastasis. *Front Oncol* (2021) 11:740111. doi: 10.3389/fonc.2021.740111

101. Huang W, Zhou K, Jiang Y, Chen C, Yuan Q, Han Z, et al. Radiomics nomogram for prediction of peritoneal metastasis in patients with gastric cancer. *Front Oncol* (2020) 10:1416. doi: 10.3389/fonc.2020.01416

102. Chen Y, Xi W, Yao W, Wang L, Xu Z, Wels M, et al. Dual-energy computed tomography-based radiomics to predict peritoneal metastasis in gastric cancer. *Front Oncol* (2021) 11:659981. doi: 10.3389/fonc.2021.659981

103. Liu D, Zhang W, Hu F, Yu P, Zhang X, Yin H, et al. A bounding box-based radiomics model for detecting occult peritoneal metastasis in advanced gastric cancer: A multicenter study. *Front Oncol* (2021) 11:777760. doi: 10.3389/fonc.2021.777760

104. Liu S, He J, Liu S, Ji C, Guan W, Chen L, et al. Radiomics analysis using contrast-enhanced CT for preoperative prediction of occult peritoneal metastasis in advanced gastric cancer. *Eur Radiol* (2020) 30(1):239–46. doi: 10.1007/s00330-019-06368-5

105. Mirniaharikandehei S, Heidari M, Danala G, Lakshmivarahan S, Zheng B. Applying a random projection algorithm to optimize machine learning model for predicting peritoneal metastasis in gastric cancer patients using CT images. *Comput Methods Programs Biomed* (2021) 200:105937. doi: 10.1016/j.cmpb.2021.105937

106. Wang L, Lv P, Xue Z, Chen L, Zheng B, Lin G, et al. Novel CT based clinical nomogram comparable to radiomics model for identification of occult peritoneal metastasis in advanced gastric cancer. *Eur J Surg Oncol J Eur Soc Surg Oncol Br Assoc Surg Oncol* (2022) 30:S0748-7983(22)00539-X. doi: 10.1016/j.ejso.2022.06.034

107. Iannattone PA, Zhao X, VanHouten J, Garg A, Huynh T. Artificial intelligence for diagnosis of acute coronary syndromes: A meta-analysis of machine learning approaches. *Can J Cardiol* (2020) 36(4):577–83. doi: 10.1016/j.cjca.2019.09.013