Original Article

Predictive accuracy of machine learning for radiation-induced temporal lobe injury in nasopharyngeal carcinoma patients: a systematic review and meta-analysis

Yiling Li[1], Fengyuan Gong[2], Yangyang Guo[1], Wai Tong Ng[3,4], Michael Benedict A. Mejia[5], Wen-Long Nei[6], Cuicui Wang[1], Zhanguo Jin[1]

[1]Vertigo Clinic/Research Center of Aerospace Medicine, Air Force Medical Center, PLA, Beijing, China; [2]Graduate School, Hebei North University, Zhangjiakou, China; [3]Clinical Oncology Center and Shenzhen Key Laboratory for Cancer Metastasis and Personalized Therapy, The University of Hong Kong-Shenzhen Hospital, Shenzhen, China; [4]Department of Clinical Oncology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China; [5]Benavides Cancer Institute, UST Hospital, Manila, Philippines; [6]Division of Radiation Oncology, National Cancer Center Singapore, Singapore, Singapore

Contributions: (I) Conception and design: Z Jin; (II) Administrative support: C Wang; (III) Provision of study materials or patients: Y Li, F Gong; (IV) Collection and assembly of data: Y Li, Y Guo; (V) Data analysis and interpretation: Y Li, F Gong; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Zhanguo Jin, MD; Cuicui Wang, MD. Vertigo Clinic/Research Center of Aerospace Medicine, Air Force Medical Center, PLA, No. 30 Fucheng Road, Haidian District, Beijing 100142, China. Email: ccjzg@126.com; cuicuiwang_169@163.com.

Background: Radiotherapy is a common treatment for nasopharyngeal carcinoma (NPC) but can cause radiation-induced temporal lobe injury (RTLI), resulting in irreversible damage. Predicting RTLI at the early stage may help with that issue by personalized adjustment of radiation dose based on the predicted risk. Machine learning (ML) models have recently been used to predict RTLI but their predictive accuracy remains unclear because the reported concordance index (C-index) varied widely from around 0.31 to 0.97. Therefore, a meta-analysis was needed.

Methods: The PubMed, Web of Science, Embase, and Cochrane Library databases were searched from inception to November 2022. Studies that fully develop one or more ML risk models of RTLI after radiotherapy for NPC were included. The Prediction model Risk Of Bias Assessment Tool (PROBAST) was used to assess the risk of bias in the included research. The primary outcome of this review was the C-index, specificity (Spe), and sensitivity (Sen).

Results: The meta-analysis included 14 studies with 15,573 NPC patients reporting a total of 72 prediction models. Overall, 94.44% of models were found to have a high risk of bias. Radiomics was included in 57 models, dosimetric predictors in 28, and clinical data in 27. The pooled C-index for ML models predicting RTLI was 0.77 [95% confidence interval (CI): 0.75–0.79] in the training set and 0.78 (95% CI: 0.75–0.81) in the validation set. The pooled Sen was 0.75 (95% CI: 0.69–0.80) in the training set and 0.70 (95% CI: 0.66–0.73) in the validation set and the pooled Spe was 0.78 (95% CI: 0.73–0.82) in the training set and 0.79 (95% CI: 0.75–0.82) in the validation set. Models with radiomics and clinical data achieved the most excellent discriminative performance, with a pooled C-index of 0.895.

Conclusions: ML models can accurately predict RTLI at an early stage, allowing for timely interventions to prevent further damage. The kind of ML methods and the selection of predictors may influence the predictive accuracy.

Keywords: Nasopharyngeal carcinoma (NPC); temporal lobe injury; machine learning (ML); predictive model; meta-analysis

Submitted May 18, 2023. Accepted for publication Aug 18, 2023. Published online Aug 25, 2023.
doi: 10.21037/tcr-23-859
View this article at: https://dx.doi.org/10.21037/tcr-23-859

Introduction

Nasopharyngeal carcinoma (NPC) is a type of cancer that originates from the epithelium of the nasopharynx. The age-standardized incidence of NPC was 0.4 per 100,000 to 3.0 per 100,000 (1). In 2012, there were over 129,000 new cases, with 71% of them occurring in the east and southeast parts of Asia (1). Radiotherapy is the mainstay treatment of NPC (2). For patients with locoregionally advanced NPC, concurrent chemoradiotherapy was also applied (3). However, 2.3–16% of patients who undergo this treatment may experience radiation-induced temporal lobe injury (RTLI) (4), which can cause irreversible damage to emotional and cognitive functions (5,6). To mitigate this risk, intensity modulation has been used during radiotherapy, and temporal lobe dose tolerance has been established since several dosimetric predictors are significantly associated with RTLI (7). However, dose tolerance varies among individuals, and it is more practical to adjust the dose according to the individual's risk of RTLI, which highlights the importance of predictive models of RTLI. Although previous studies have attempted to create predictive models of RTLI, including normal tissue complication probability (NTCP) models (8-10), most of these models have only included dosimetric predictors and have not provided sufficient predictive accuracy. Recently, additional variables, such as radiomics and clinical data, have been incorporated into predictive models. Radiomics, which is high-dimensional mineable data (11), can contain many details of RTLI and thus improve prediction accuracy. However, previous models were constructed without radiomics due to the limitations of manual data processing in handling large volumes of complex data. An increasing number of promising studies have attempted to employ machine learning (ML) approaches to develop prediction models relying on radiomics (12,13). The biggest strength of this kind of model is the personalized prediction at the early stage of RTLI whereas the biggest weakness is the dependence on imaging data, which may lead to possible additional medical expenses. ML was used to develop predictive models of NPC prognosis and was proven to have great predictive performance (14). Nevertheless, the predictive accuracy of ML models of RTLI is still debated since it varies from around 0.31 to 0.97, which may be caused by different variables and ML approaches (15,16). Also, there is a lack of relevant evidence-based information. Therefore, it is essential to conduct a meta-analysis to evaluate the predictive accuracy of ML for RTLI after radiotherapy. We present this article in accordance with the PRISMA reporting checklist (17) (available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-859/rc).

Highlight box

Key findings

- This study presented evidence-based data for the first time on machine learning (ML) models predicting radiation-induced temporal lobe injury (RTLI) in nasopharyngeal carcinoma (NPC) patients. The best discrimination was achieved by models that utilized both radiomics and clinical features.

What is known and what is new?

- Previous studies attempted to create predictive models of RTLI such as NTCP models. However, most variables in these models only contained dosimetric predictors, and no study provided predictive accuracy.
- Recently, ML models have been employed for early RTLI prediction, utilizing radiomics and clinical data as predictors, but their predictive accuracy remained unclear.

What is the implication, and what should change now?

- ML models have been shown to predict RTLI effectively and can prevent serious RTLI at an early stage. Clinicians can use these models to adjust personalized radiotherapy treatment plans and improve patient outcomes.

Methods

The quality of this meta-analysis was assessed using AMSTAR2 (18). The systematic review was registered on PROSPERO (No. CRD42023380907).

Inclusion and exclusion criteria

Inclusion criteria

A study was included when it was nested case-control studies, cohort studies, case-control studies, or case-cohort studies that investigate the predictive value of ML for RTLI in NPC patients; it fully developed one or more ML risk models of RTLI after radiotherapy for NPC; it was with or without external validation; it used at least one outcome parameter, such as receiver operating characteristic (ROC) curve, C-statistic, concordance index (C-index), accuracy, sensitivity (Sen), specificity (Spe), diagnostic 4 grid table, confusion matrix, F1 score, and calibration curve; it developed different models even if they are based on the same cohort; and it was written in English.

Exclusion criteria

A study was excluded when it was randomized controlled trials (RCTs), review articles, meta-analyses, guidelines, or expert opinions; the sample size of RTLI patients was less than 10 (19); it included only risk factor analysis and an incomplete ML model; any of the outcome parameters listed above were unavailable; it included only validation of a mature predictive model; or it was research on the accuracy of univariate factor prediction.

Search strategy

We searched the Web of Science, PubMed, Embase, and Cochrane Library databases from their inception to November 2022. We used Medical Subject Heading (MeSH) terms and text word terms related to NPC, ML, and radiotherapy were used for the literature search, with no language or region restrictions. The full search strategy is provided in Table S1.

Study selection

We imported the searched articles into Endnote and removed duplicates. After screening the titles and abstracts, we discarded studies that did not meet the criteria and considered the remaining studies as potentially eligible. We downloaded the full texts of these studies and evaluated them. Based on the inclusion/exclusion criteria, we either included or excluded the remaining studies. Two independent reviewers, Li Y and Guo Y carried out the selection procedure. Any disagreements between the reviewers were resolved by consensus or by a third reviewer, if required.

Data extraction

We generated a standardized table to record information from the literature, including the title, author, year of publication, country, type of study, duration of follow-up, treatment, number of total patients, number of RTLI patients, number of test/train patients, type of model, type of predictors, method of predictor selection, statistical outcomes, and outcome parameter. We extracted the data from the included studies according to this table. The process of data extraction was carried out independently by two reviewers, Li Y and Gong F, and any disagreements were resolved through consensus or with the involvement of a third reviewer if necessary.

Quality assessment

We used the Prediction model Risk Of Bias Assessment Tool (PROBAST) to assess the quality of the included records. PROBAST has four domains: participants, predictors, outcome, and analysis. The risk of bias in the study was evaluated through these domains, while the applicability was evaluated in the first three domains. Each domain contains 2, 3, 6, and 9 signaling questions, respectively. There were three possible answers to each question: Yes/Probably Yes, No/Probably No, and No information. A domain was judged to have a minimal risk of bias only if all responses to the questions were Yes/Probably Yes. The entire study was considered to have a low risk of bias only when all domains were rated as low risk of bias. If a domain had at least one question with the response No/Probably No, it was considered high risk. The entire study was considered to have a high risk of bias if at least one domain was considered high risk. The quality assessment process was carried out independently by two reviewers, Li Y and Gong F, with any disagreements resolved through consensus or with the involvement of a third reviewer if necessary.

Data synthesis and statistical analysis

The primary outcome of this review was the C-index. We also considered Sen, and Spe as the main outcomes, as the C-index alone may not be sufficient to express the predictive accuracy of ML when there is a difference in the number of individuals between the RTLI group and the non-RTLI group. We performed a meta-analysis based on the outcome parameter above. If both the 95% confidence interval (CI) and the standard error of the C-index were unavailable, we estimated the standard error using the formula (20). We used a random-effects model to conduct the meta-analysis based on the C-index given the differences in predictors used by each ML model. Furthermore, for the Sen and Spe meta-analysis, we utilized a bivariate mixed-effects model. Sensitivity analyses were conducted to evaluate the stability of the meta-analysis. Publication bias was evaluated by Egger's test. We conducted these analyses using Stata 15.1 (StataCorp., College Station, TX, USA).

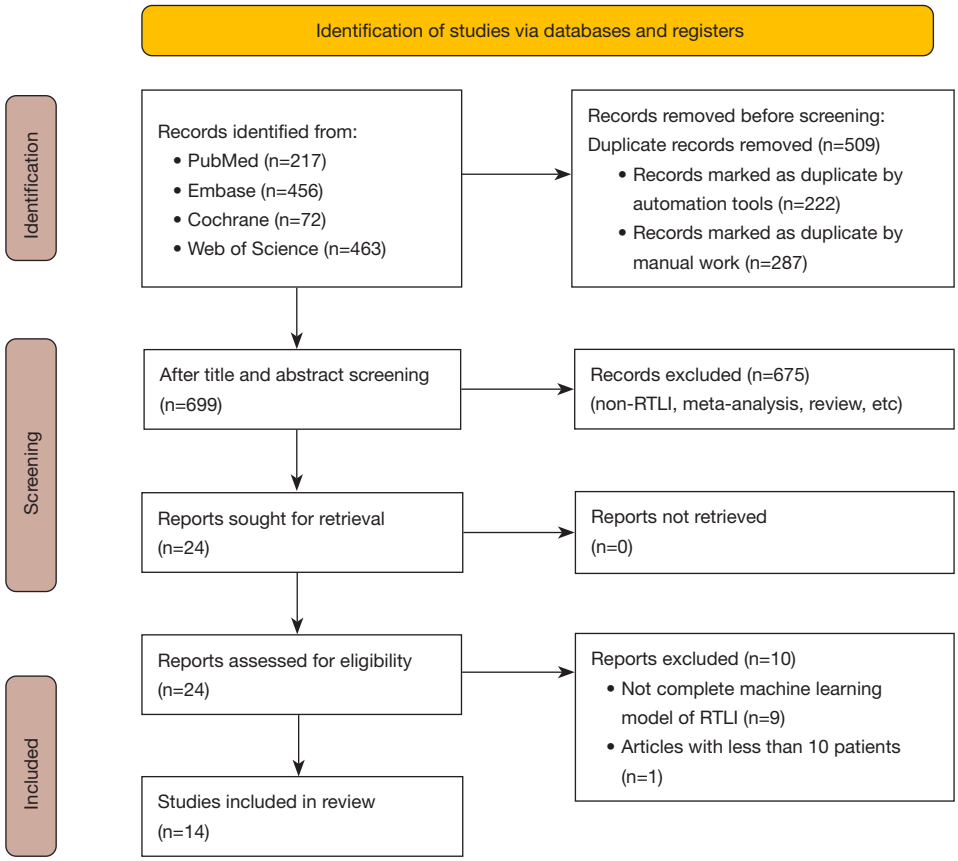


Figure 1 Flow diagram of literature search and selection. RTLI, radiation-induced temporal lobe injury.

Results

Study selection

Figure 1 depicts the study selection process. Initially, we identified 699 unique records, and after assessing the titles and abstracts, we excluded 675 articles, leaving 24 for full-text review. Following a review of the full text of the papers, we excluded 10 studies (Table S2), primarily because they did not provide a complete ML model. Finally, our review finally included 14 studies (4,15,16,21-31).

Characteristics of included studies

Our review included 8 cohort studies, 5 case-control studies, and 1 nested case-control study. All were single-center studies conducted in China. The included studies involved 15,573 NPC patients, with 2,267 (14.56%) of them identified as having RTLI. The authors, last author, publication year, center location, data source, patient inclusion and exclusion criteria, determined criteria of the outcome, time of predictor assessment, and time of outcome assessment of included studies are provided in available online: https://cdn.amegroups.cn/static/public/tcr-23-859-1.xlsx.

Characteristics of included prediction models

The included studies reported a total of 72 prediction models. These models were developed using 10 ML methods, including random forest (RF), Naïve Bayes (NB), k-nearest neighbors (KNN), Adaboost (AB), support vector machines (SVM), generalized linear regression (GLR), logistic regression (LR), Gradient Boosting Trees (GBT), Decision Tree (DT), and Cox proportional hazards (Cox). Radiomics was employed as one of the variables in 57 models, dosimetric predictors in 28 models, and clinical data in 27 models. Age, T/N/overall stage, and gender were the first three major clinical factors in the studies with clinical data models, accounting for 83.33%, 66.67%, and 33.33%, respectively.

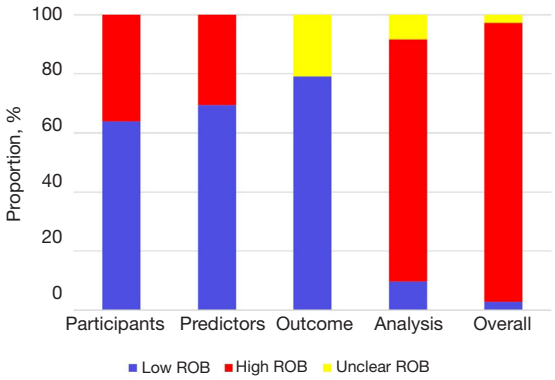


Figure 2 The proportion of models reported by included studies with different risks of bias for each PROBAST domain. ROB, risk of bias; PROBAST, Prediction model Risk of Bias Assessment Tool.

Risk of bias assessment

Table S3 presents the results of the risk of bias assessment for each model in the included articles using the PROBAST tool. *Figure 2* shows the proportion of models with different risks of bias for each domain. Overall, 94.44% of models were found to have a high risk of bias, primarily due to the high risk of bias in the analysis domain (81.94%). This was mainly because of the unreasonable number of participants (70.83%) or not accounting for model overfitting and optimism (9.72%) in model performance. Sensitivity analyses showed that the impact of models with a high risk of bias on this meta-analysis is acceptable (Figure S1). Egger's test showed that there was a publication bias in included models (P<0.01).

Meta-analysis

The included studies reported 48 models and 37 independent validation sets, with most models demonstrating good discrimination.

C-index

As shown in *Table 1*, the pooled C-index of the models was 0.769 (95% CI: 0.749–0.790) in the training set and 0.781 (95% CI: 0.754–0.808) in the validation set. The most common ML methods used to develop models were LR, SVM, and RF, all of which exhibited acceptable discrimination. For example, RF models performed well in predicting RTLI, with pooled C-index of 0.815 (95% CI: 0.781–0.850) in the training set, and 0.764 (95% CI: 0.710–0.818) in the validation set. We also investigated how models with different variables may perform. There were 31 models developed using radiomics data which was the most common variable in the included studies. The pooled C-index of these models was 0.745 (95% CI: 0.716–0.774) in the training set and 0.812 (95% CI: 0.766–0.859) in the validation set. The best-performing models were those established using both radiomics data and clinical data with a pooled C-index of 0.895 (95% CI: 0.860–0.930) in the training set and 0.880 (95% CI: 0.810–0.950) in the validation set (*Table 1*). Figure S2 showed the forest plots of meta-analysis. Sensitivity analyses excluding any one model included in this meta-analysis yielded results that were consistent with the primary analysis (Figure S1).

Sen and Spe

As shown in *Table 2*, the summary of Sen was 0.75 (95% CI: 0.69–0.80) in the training set and 0.70 (95% CI: 0.66–0.73) in the validation set. Additionally, the summary of Spe was 0.78 (95% CI: 0.73–0.82) in the training set and 0.79 (95% CI: 0.75–0.82) in the validation set. Among models developed using different ML methods, those developed using SVM showed the highest Sen (pooled Sen in the training set 0.80, 95% CI: 0.70–0.87; pooled Sen in the validation set 0.72, 95% CI: 0.57–0.83), while those developed using RF showed the highest Spe (pooled Spe in the training set 0.83, 95% CI: 0.63–0.94; pooled Spe in the validation set 0.76, 95% CI: 0.68–0.82). Among models developed using different variables, those established using radiomics data, clinical data, and dosimetric predictors showed the highest Sen (pooled Sen in the training set 0.78, 95% CI: 0.71–0.83; pooled Sen in the validation set 0.71, 95% CI: 0.62–0.79), while those established using both radiomics data and clinical data showed the highest Spe (pooled Spe in the training set 0.88, 95% CI: 0.78–0.98; pooled Spe in the validation set 0.90, 95% CI: 0.84–0.96).

Discussion

The meta-analysis results indicate that ML is a robust technique for predicting RTLI, with radiomics playing a crucial role as a predictor. LR, SVM, and RF were the most common ML methods used to develop models and all performed well in predicting RTLI, which provided a solution for NPC patients that the radiation dose could be reduced individually at the early stage in the patients with high predicted risk of RTLI. Models developed with both radiomics and clinical features demonstrated the best

Table 1 Meta-analysis for C-index of models, in the training and validation sets separately

Subgroup	Training set		Validation set	
	Number	C-index	Number	C-index
ML method type				
RF	7	0.815 (0.781–0.850)	7	0.764 (0.710–0.818)
SVM	8	0.756 (0.668–0.844)	7	0.898 (0.843–0.953)
Cox	6	0.793 (0.763–0.823)	2	0.753 (0.730–0.775)
LR	14	0.810 (0.773–0.847)	17	0.780 (0.725–0.835)
AB	3	0.687 (0.580–0.794)		
KNN	4	0.649 (0.509–0.789)	4	0.608 (0.411–0.805)
GLR	3	0.562 (0.510–0.613)		
NB	1	0.83		
DT	1	0.72		
GBT	1	0.88		
Variable type				
R	31	0.745 (0.716–0.774)	7	0.812 (0.766–0.859)
D	5	0.781 (0.741–0.821)	3	0.728 (0.707–0.748)
C+D	5	0.791 (0.759–0.822)	4	0.776 (0.749–0.802)
C	1	0.74	2	0.705 (0.640–0.770)
R+C	2	0.895 (0.860–0.930)	2	0.880 (0.810–0.950)
R+C+D	4	0.827 (0.805–0.849)	19	0.768 (0.721–0.815)
Overall	48	0.769 (0.749–0.790)	37	0.781 (0.754–0.808)

C-index, concordance index; ML, machine learning; RF, random forest; SVM, support vector machines; Cox, Cox proportional hazards; LR, logistic regression; AB, AdaBoost; KNN, k-nearest neighbors; GLR, generalized linear regression; NB, Naïve Bayes; DT, Decision Tree; GBT, Gradient Boosting Trees; R, radiomics; D, dosimetric predictors; C, clinical data.

discrimination, with age and T/N/overall stage being the most commonly used clinical features. These findings can serve as a reference for future research in this area.

The radical treatment for NPC always involves the delivery of radiotherapy but RTLI is a severe late complication that can cause irreversible emotional and cognitive impairment. Medications such as bevacizumab and corticosteroids can be administered for RTLI to alleviate symptoms and prevent further deterioration (32,33). More importantly, researchers have focused on adjusting radiation doses to find the most appropriate dose to prevent RTLI, using the "as low as reasonably practicable" principle (34). However, inconsistent results have been reported, with some identifying D1cc as the most important predictor and a tolerance dose of 62.8 Gy (8,9), while others found D2cc (10) and D0.5cc (35) to be the best predictor. Given the inconsistent findings and differentiated actual tolerance dose of each person, it is challenging to recommend a tolerance dose. Additionally, the proximity of targets to organs at risk involves a delicate balancing act that involves tumor coverage and normal tissue. Therefore, more studies are focusing on predicting RTLI, enabling clinicians to estimate the risk before radiotherapy and make individualized radiation plans or adjust the dose in a timely manner.

Radiomics was viewed as a valuable feature that should be explored for use in developing RTLI prediction models. It has been frequently employed in radiation-induced toxicity prediction models and has outperformed clinical and dosimetric parameters in predicting toxicities such as xerostomia in NPC patients, cardiac toxicity in breast cancer patients, and lung damage in esophageal cancer

Table 2 Meta-analysis for sensitivity and specificity of models, in the training and validation sets separately

Subgroup	Training set			Validation set		
	Number	Sen	Spe	Number	Sen	Spe
ML method type						
RF	4	0.63 (0.30–0.87)	0.83 (0.63–0.94)	7	0.67 (0.58–0.74)	0.76 (0.68–0.82)
SVM	5	0.80 (0.70–0.87)	0.80 (0.65–0.89)	7	0.72 (0.57–0.83)	0.81 (0.72–0.87)
LR	14	0.73 (0.67–0.78)	0.78 (0.72–0.83)	16	0.70 (0.65–0.75)	0.78 (0.73–0.83)
KNN	1	0.950	0.690	3	0.59 (0.42–0.75)	0.84 (0.75–0.90)
NB	1	0.800	0.690			
DT	1	0.770	0.650			
GBT	1	0.840	0.780			
Variable type						
R	15	0.75 (0.64–0.83)	0.79 (0.72–0.85)	7	0.70 (0.63–0.76)	0.78 (0.65–0.87)
D	3	0.78 (0.67–0.86)	0.66 (0.62–0.71)	2	0.57 (0.46–0.66)	0.74 (0.72–0.77)
C+D	3	0.72 (0.62–0.80)	0.72 (0.69–0.75)	3	0.68 (0.59–0.76)	0.76 (0.70–0.80)
C	1	0.590	0.870	1	0.59	0.71
R+C	2	0.75 (0.72–0.79)	0.88 (0.78–0.98)	2	0.76 (0.71–0.81)	0.90 (0.84–0.96)
R+C+D	3	0.78 (0.71–0.83)	0.75 (0.69–0.80)	18	0.71 (0.62–0.79)	0.79 (0.73–0.83)
Overall	27	0.75 (0.69–0.80)	0.78 (0.73–0.82)	33	0.70 (0.66–0.73)	0.79 (0.75–0.82)

Sen, sensitivity; Spe, specificity; ML, machine learning; RF, random forest; SVM, support vector machines; LR, logistic regression; KNN, k-nearest neighbors; NB, Naïve Bayes; DT, Decision Tree; GBT, Gradient Boosting Trees; R, radiomics; D, dosimetric predictors; C, clinical data.

patients after radiotherapy (36). As expected, radiomics performed well in predicting RTLI as it can detect microstructural changes in the temporal lobe early (11), enhancing prediction accuracy, as seen in the meta-analysis for each subgroup above. ML methods are required to process high-dimensional mineable data like radiomics, and these methods have been shown to play a crucial part in predictive models based on radiomics and standard clinical characteristics, such as predicting rapidly deteriorating mild cognitive impairment in Alzheimer's disease (13) and the prognosis of acute ischemic stroke (12). LR is the most often used ML approach for building models due to its ease of use and consistently good performance. However, to determine the best-performing model, comparisons of various ML methods based on the same dataset should be made (15,23,24). As indicated in the meta-analysis above, SVM appears to have the highest performance, particularly in the validation set, although further comparison studies are needed to confirm this finding.

In addition, age and T/N/overall stage were shown to be major clinical predictors in several studies (4,24), emphasizing the need for paying attention to the RTLI risk of elderly or have advanced-stage NPC patients and developing specific radiation strategies for them. Elderly patients may have a lower tolerance dose due to their susceptibility to cerebrovascular injury, which significantly affects RTLI (37). Moreover, patients with advanced-stage NPC are treated with higher doses in a wider area to control tumors, resulting in more damage to normal brain tissue (5,34).

The greatest strength of our study is its originality. This is the first study to present evidence-based data for ML models predicting RTLI in NPC patients after radiotherapy. Besides, the quality of this meta-analysis was high based on AMSTAR2 (Appendix 1). Nevertheless, this research has several limitations. First, although we conducted a comprehensive search, the number of included studies was relatively limited and all included studies were conducted in China, due to the geographical distribution of the disease. Second, since only a few models were included,

further research is required to design and identify the most optimal ML model for predicting RTLI. Third, there was a publication bias in included models, which may cause the overestimation of prediction accuracy of them. Fourth, it would be better to verify if the SVM is the ML method with highest predictive performance in more clinical experimental research.

Conclusions

Our study demonstrated that ML methods can effectively predict RTLI in NPC patients after radiotherapy. While the primary goal of radiotherapy is to maximize survival with the appropriate radiation dose, stratifying the risk of RTLI through ML can help adjust personalized radiation doses for NPC patients. Therefore, more multicenter and multi-ethnic studies are necessary to develop tools for predicting RTLI, particularly ML models with radiomics, to prevent or mitigate it by timely adjusting the treatment.

Acknowledgments

Funding: This study was supported by the Key Project of Equipment Development of the Central Military Commission (No. KJ2019IA050325, to ZJ), Shenzhen Key Laboratory for cancer metastasis and personalized therapy (No. ZDSYS20210623091811035, to WTN), and the Shenzhen Fundamental Research Program, China (No. CYJ20210324114404013, to WTN).

Footnote

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-859/rc

Peer Review File: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-859/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-859/coif). WTN reports funding support from the Shenzhen Key Laboratory for cancer metastasis and personalized therapy (No. ZDSYS20210623091811035) and the Shenzhen Fundamental Research Program, China (No. CYJ20210324114404013). The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
2. Chen YP, Chan ATC, Le QT, et al. Nasopharyngeal carcinoma. Lancet 2019;394:64-80.
3. Ou D, Blanchard P, El Khoury C, et al. Induction chemotherapy with docetaxel, cisplatin and fluorouracil followed by concurrent chemoradiotherapy or chemoradiotherapy alone in locally advanced non-endemic nasopharyngeal carcinoma. Oral Oncol 2016;62:114-21.
4. Wen DW, Lin L, Mao YP, et al. Normal tissue complication probability (NTCP) models for predicting temporal lobe injury after intensity-modulated radiotherapy in nasopharyngeal carcinoma: A large registry-based retrospective study from China. Radiother Oncol 2021;157:99-105.
5. Su SF, Huang Y, Xiao WW, et al. Clinical and dosimetric characteristics of temporal lobe injury following intensity modulated radiotherapy of nasopharyngeal carcinoma. Radiother Oncol 2012;104:312-6.
6. Chen W, Qiu S, Li J, et al. Diffusion tensor imaging study on radiation-induced brain injury in nasopharyngeal carcinoma during and after radiotherapy. Tumori 2015;101:487-90.
7. Su SF, Huang SM, Han F, et al. Analysis of dosimetric factors associated with temporal lobe necrosis (TLN) in patients with nasopharyngeal carcinoma (NPC) after intensity modulated radiotherapy. Radiat Oncol 2013;8:17.
8. Kong C, Zhu XZ, Lee TF, et al. LASSO-based NTCP model for radiation-induced temporal lobe injury

developing after intensity-modulated radiotherapy of nasopharyngeal carcinoma. Sci Rep 2016;6:26378.
9. Zeng L, Huang SM, Tian YM, et al. Normal Tissue Complication Probability Model for Radiation-induced Temporal Lobe Injury after Intensity-modulated Radiation Therapy for Nasopharyngeal Carcinoma. Radiology 2015;276:243-9.
10. Feng M, Huang Y, Fan X, et al. Prognostic variables for temporal lobe injury after intensity modulated-radiotherapy of nasopharyngeal carcinoma. Cancer Med 2018;7:557-64.
11. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278:563-77.
12. Wang X, Lyu J, Meng Z, et al. Small vessel disease burden predicts functional outcomes in patients with acute ischemic stroke using machine learning. CNS Neurosci Ther 2023;29:1024-33.
13. Zhao X, Sui H, Yan C, et al. Machine-Based Learning Shifting to Prediction Model of Deteriorative MCI Due to Alzheimer's Disease - A Two-Year Follow-Up Investigation. Curr Alzheimer Res 2022;19:708-15.
14. Wang Y, He Y, Duan X, et al. Construction of diagnostic and prognostic models based on gene signatures of nasopharyngeal carcinoma by machine learning methods. Transl Cancer Res 2023;12:1254-69.
15. Lin X, Li Z, Chen S, et al. Divergent white matter changes in patients with nasopharyngeal carcinoma post-radiotherapy with different outcomes: a potential biomarker for prediction of radiation necrosis. Eur Radiol 2022;32:7036-47.
16. Zhao LM, Kang YF, Gao JM, et al. Functional Connectivity Density for Radiation Encephalopathy Prediction in Nasopharyngeal Carcinoma. Front Oncol 2021;11:687127.
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. PLoS Med 2021;18:e1003583.
18. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:j4008.
19. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020;368:m441.
20. Debray TP, Damen JA, Riley RD, et al. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. Stat Methods Med Res 2019;28:2768-86.
21. Zhang B, Lian Z, Zhong L, et al. Machine-learning based MRI radiomics models for early detection of radiation-induced brain injury in nasopharyngeal carcinoma. BMC Cancer 2020;20:502.
22. Bao D, Zhao Y, Li L, et al. A MRI-based radiomics model predicting radiation-induced temporal lobe injury in nasopharyngeal carcinoma. Eur Radiol 2022;32:6910-21.
23. Zhong L, Zhang X, Xi Y, et al. Deep Longitudinal Feature Representations for Detection of Postradiotherapy Brain Injury at Presymptomatic Stage. IEEE Access; 2020;8:184710-21.
24. Bin X, Zhu C, Tang Y, et al. Nomogram Based on Clinical and Radiomics Data for Predicting Radiation-induced Temporal Lobe Injury in Patients with Non-metastatic Stage T4 Nasopharyngeal Carcinoma. Clin Oncol (R Coll Radiol) 2022;34:e482-92.
25. Guan W, Xie K, Fan Y, et al. Development and Validation of a Nomogram for Predicting Radiation-Induced Temporal Lobe Injury in Nasopharyngeal Carcinoma. Front Oncol 2020;10:594494.
26. Hou J, Li H, Zeng B, et al. MRI-based radiomics nomogram for predicting temporal lobe injury after radiotherapy in nasopharyngeal carcinoma. Eur Radiol 2022;32:1106-14.
27. Fang J, Li A, Ouyang PY, et al. Weighted Concordance Index Loss-Based Multimodal Survival Modeling for Radiation Encephalopathy Assessment in Nasopharyngeal Carcinoma Radiotherapy. Springer, Cham; 2022.
28. Wang J, Miao Y, Ou X, et al. Development and validation of a model for temporal lobe necrosis for nasopharyngeal carcinoma patients with intensity modulated radiation therapy. Radiat Oncol 2019;14:42.
29. Zhang YM, Kang YF, Zeng JJ, et al. Surface-Based Falff: A Potential Novel Biomarker for Prediction of Radiation Encephalopathy in Patients With Nasopharyngeal Carcinoma. Front Neurosci 2021;15:692575.
30. Kang YF, Chen RT, Ding H, et al. Structure-Function Decoupling: A Novel Perspective for Understanding the Radiation-Induced Brain Injury in Patients With Nasopharyngeal Carcinoma. Front Neurosci 2022;16:915164.
31. Du QH, Gan YX, Wang RS, et al. Half-Brain Delineation for Prediction of Radiation-Induced Temporal Lobe Injury in Nasopharyngeal Carcinoma Receiving Intensity-Modulated Radiotherapy. Front Oncol 2021;11:599942.
32. Leavitt RJ, Limoli CL, Baulch JE. miRNA-based therapeutic potential of stem cell-derived extracellular

vesicles: a safe cell-free treatment to ameliorate radiation-induced brain injury. Int J Radiat Biol 2019;95:427-35.
33. Xu Y, Rong X, Hu W, et al. Bevacizumab Monotherapy Reduces Radiation-induced Brain Necrosis in Nasopharyngeal Carcinoma Patients: A Randomized Controlled Trial. Int J Radiat Oncol Biol Phys 2018;101:1087-95.
34. Lee AW, Ng WT, Pan JJ, et al. International Guideline on Dose Prioritization and Acceptance Criteria in Radiation Therapy Planning for Nasopharyngeal Carcinoma. Int J Radiat Oncol Biol Phys 2019;105:567-80.
35. Sun Y, Zhou GQ, Qi ZY, et al. Radiation-induced temporal lobe injury after intensity modulated radiotherapy in nasopharyngeal carcinoma patients: a dose-volume-outcome analysis. BMC Cancer 2013;13:397.
36. Desideri I, Loi M, Francolini G, et al. Application of Radiomics for the Prediction of Radiation-Induced Toxicity in the IMRT Era: Current State-of-the-Art. Front Oncol 2020;10:1708.
37. Smart D. Radiation Toxicity in the Central Nervous System: Mechanisms and Strategies for Injury Reduction. Semin Radiat Oncol 2017;27:332-9.

(English Language Editor: J. Jones)

Cite this article as: Li Y, Gong F, Guo Y, Ng WT, Mejia MBA, Nei WL, Wang C, Jin Z. Predictive accuracy of machine learning for radiation-induced temporal lobe injury in nasopharyngeal carcinoma patients: a systematic review and meta-analysis. Transl Cancer Res 2023;12(9):2361-2370. doi: 10.21037/tcr-23-859

Supplementary

Table S1 Literature search strategy

1. PubMed

Search number	Query	Results
#1	"Nasopharyngeal Carcinoma"[Mesh]	5,718
#2	((((((((((((((((((((((Nasopharyngeal Carcinoma[Title/Abstract]) OR (Nasopharyngeal Carcinomas[Title/Abstract])) OR (nasopharynx tumor[Title/Abstract])) OR (epipharynx tumor[Title/Abstract])) OR (epipharynx tumour[Title/Abstract])) OR (nasopharyngeal neoplasms[Title/Abstract])) OR (nasopharyngeal tumor[Title/Abstract])) OR (nasopharyngeal tumour[Title/Abstract])) OR (nasopharynx tumour[Title/Abstract])) OR (rhinopharyngeal tumor[Title/Abstract])) OR (rhinopharynx tumor[Title/Abstract])) OR (rhinopharynx tumour[Title/Abstract])) OR (nasopharynx cancer[Title/Abstract])) OR (epipharynx cancer[Title/Abstract])) OR (nasopharyngeal cancer[Title/Abstract])) OR (rhinopharyngioma[Title/Abstract])) OR (rhinopharynx cancer[Title/Abstract])) OR (epipharyngeal carcinoma[Title/Abstract])) OR (epipharynx carcinoma[Title/Abstract])) OR (naso-pharyngeal carcinoma[Title/Abstract])) OR (nasopharyngeal carcinoma[Title/Abstract])) OR (postnasal space carcinoma[Title/Abstract])) OR (rhino-pharyngeal carcinoma[Title/Abstract])) OR (rhinopharyngeal carcinoma[Title/Abstract])) OR (rhinopharynx carcinoma[Title/Abstract])	20,075
#3	((((((((((((((((((((((Nasopharyngeal Carcinoma[Title/Abstract]) OR (Nasopharyngeal Carcinomas[Title/Abstract])) OR (nasopharynx tumor[Title/Abstract])) OR (epipharynx tumor[Title/Abstract])) OR (epipharynx tumour[Title/Abstract])) OR (nasopharyngeal neoplasms[Title/Abstract])) OR (nasopharyngeal tumor[Title/Abstract])) OR (nasopharyngeal tumour[Title/Abstract])) OR (nasopharynx tumour[Title/Abstract])) OR (rhinopharyngeal tumor[Title/Abstract])) OR (rhinopharynx tumor[Title/Abstract])) OR (rhinopharynx tumour[Title/Abstract])) OR (nasopharynx cancer[Title/Abstract])) OR (epipharynx cancer[Title/Abstract])) OR (nasopharyngeal cancer[Title/Abstract])) OR (rhinopharyngioma[Title/Abstract])) OR (rhinopharynx cancer[Title/Abstract])) OR (epipharyngeal carcinoma[Title/Abstract])) OR (epipharynx carcinoma[Title/Abstract])) OR (naso-pharyngeal carcinoma[Title/Abstract])) OR (nasopharyngeal carcinoma[Title/Abstract])) OR (postnasal space carcinoma[Title/Abstract])) OR (rhino-pharyngeal carcinoma[Title/Abstract])) OR (rhinopharyngeal carcinoma[Title/Abstract])) OR (rhinopharynx carcinoma[Title/Abstract])) OR ("Nasopharyngeal Carcinoma"[Mesh])	20,192
#4	"Machine Learning"[Mesh]	51,433
#5	((((((((((((((((((((((machine learning[Title/Abstract]) OR (Transfer Learning[Title/Abstract])) OR (Deep learning[Title/Abstract])) OR (Hierarchical Learning[Title/Abstract])) OR (Ensemble Learning[Title/Abstract])) OR (artificial intelligence[Title/Abstract])) OR (Prediction model[Title/Abstract])) OR (random forest[Title/Abstract])) OR (neural network[Title/Abstract])) OR (ANN[Title/Abstract])) OR (Support vector machine[Title/Abstract])) OR (SVM[Title/Abstract])) OR (Gradient Boosting Machine[Title/Abstract])) OR (GBM[Title/Abstract])) OR (Nomogram[Title/Abstract])) OR (XGboost[Title/Abstract])) OR (Adaboose[Title/Abstract])) OR (Decision tree[Title/Abstract])) OR (External validation[Title/Abstract])) OR (Risk Prediction[Title/Abstract])) OR (Risk-Prediction[Title/Abstract])) OR (Radiomics[Title/Abstract])) OR (Radiomic[Title/Abstract])) OR (statistical learning[Title/Abstract])) OR (predictive analytics[Title/Abstract])	265,748
#6	((((((((((((((((((((((machine learning[Title/Abstract]) OR (Transfer Learning[Title/Abstract])) OR (Deep learning[Title/Abstract])) OR (Hierarchical Learning[Title/Abstract])) OR (Ensemble Learning[Title/Abstract])) OR (artificial intelligence[Title/Abstract])) OR (Prediction model[Title/Abstract])) OR (random forest[Title/Abstract])) OR (neural network[Title/Abstract])) OR (ANN[Title/Abstract])) OR (Support vector machine[Title/Abstract])) OR (SVM[Title/Abstract])) OR (Gradient Boosting Machine[Title/Abstract])) OR (GBM[Title/Abstract])) OR (Nomogram[Title/Abstract])) OR (XGboost[Title/Abstract])) OR (Adaboose[Title/Abstract])) OR (Decision tree[Title/Abstract])) OR (External validation[Title/Abstract])) OR (Risk Prediction[Title/Abstract])) OR (Risk-Prediction[Title/Abstract])) OR (Radiomics[Title/Abstract])) OR (Radiomic[Title/Abstract])) OR (statistical learning[Title/Abstract])) OR (predictive analytics[Title/Abstract])) OR ("Machine Learning"[Mesh])	271,051
#7	"Radiotherapy"[Mesh]	204,166
#8	((((((((((((Radiotherapies[Title/Abstract]) OR (Radiotherapy[Title/Abstract])) OR (Radiation Therapy[Title/Abstract])) OR (Radiation Therapies[Title/Abstract])) OR (Radiation Treatment[Title/Abstract])) OR (Radiation Treatments[Title/Abstract])) OR (Targeted Radiotherapies[Title/Abstract])) OR (Targeted Radiotherapy[Title/Abstract])) OR (Targeted Radiation Therapy[Title/Abstract])) OR (Targeted Radiation Therapies[Title/Abstract])) OR (bioradiant therapy[Title/Abstract])) OR (x ray therapy[Title/Abstract])) OR (x ray treatment[Title/Abstract])) OR (x-ray therapy[Title/Abstract])	282,256
#9	((((((((((((Radiotherapies[Title/Abstract]) OR (Radiotherapy[Title/Abstract])) OR (Radiation Therapy[Title/Abstract])) OR (Radiation Therapies[Title/Abstract])) OR (Radiation Treatment[Title/Abstract])) OR (Radiation Treatments[Title/Abstract])) OR (Targeted Radiotherapies[Title/Abstract])) OR (Targeted Radiotherapy[Title/Abstract])) OR (Targeted Radiation Therapy[Title/Abstract])) OR (Targeted Radiation Therapies[Title/Abstract])) OR (bioradiant therapy[Title/Abstract])) OR (x ray therapy[Title/Abstract])) OR (x ray treatment[Title/Abstract])) OR (x-ray therapy[Title/Abstract])) OR ("Radiotherapy"[Mesh])	376,249
#10	((((((((((((((((Radiotherapies[Title/Abstract]) OR (Radiotherapy[Title/Abstract])) OR (Radiation Therapy[Title/Abstract])) OR (Radiation Therapies[Title/Abstract])) OR (Radiation Treatment[Title/Abstract])) OR (Radiation Treatments[Title/Abstract])) OR (Targeted Radiotherapies[Title/Abstract])) OR (Targeted Radiotherapy[Title/Abstract])) OR (Targeted Radiation Therapy[Title/Abstract])) OR (Targeted Radiation Therapies[Title/Abstract])) OR (bioradiant therapy[Title/Abstract])) OR (x ray therapy[Title/Abstract])) OR (x ray treatment[Title/Abstract])) OR (x-ray therapy[Title/Abstract])) OR ("Radiotherapy"[Mesh])) AND ((((((((((((((((((((((((machine learning[Title/Abstract]) OR (Transfer Learning[Title/Abstract])) OR (Deep learning[Title/Abstract])) OR (Hierarchical Learning[Title/Abstract])) OR (Ensemble Learning[Title/Abstract])) OR (artificial intelligence[Title/Abstract])) OR (Prediction model[Title/Abstract])) OR (random forest[Title/Abstract])) OR (neural network[Title/Abstract])) OR (ANN[Title/Abstract])) OR (Support vector machine[Title/Abstract])) OR (SVM[Title/Abstract])) OR (Gradient Boosting Machine[Title/Abstract])) OR (GBM[Title/Abstract])) OR (Nomogram[Title/Abstract])) OR (XGboost[Title/Abstract])) OR (Adaboose[Title/Abstract])) OR (Decision tree[Title/Abstract])) OR (External validation[Title/Abstract])) OR (Risk Prediction[Title/Abstract])) OR (Risk-Prediction[Title/Abstract])) OR (Radiomics[Title/Abstract])) OR (Radiomic[Title/Abstract])) OR (statistical learning[Title/Abstract])) OR (predictive analytics[Title/Abstract])) OR ("Machine Learning"[Mesh]))) AND ((((((((((((((((((((((((Nasopharyngeal Carcinoma[Title/Abstract]) OR (Nasopharyngeal Carcinomas[Title/Abstract])) OR (nasopharynx tumor[Title/Abstract])) OR (epipharynx tumor[Title/Abstract])) OR (epipharynx tumour[Title/Abstract])) OR (nasopharyngeal neoplasms[Title/Abstract])) OR (nasopharyngeal tumor[Title/Abstract])) OR (nasopharyngeal tumour[Title/Abstract])) OR (nasopharynx tumour[Title/Abstract])) OR (rhinopharyngeal tumor[Title/Abstract])) OR (rhinopharynx tumor[Title/Abstract])) OR (rhinopharynx tumour[Title/Abstract])) OR (nasopharynx cancer[Title/Abstract])) OR (epipharynx cancer[Title/Abstract])) OR (nasopharyngeal cancer[Title/Abstract])) OR (rhinopharyngioma[Title/Abstract])) OR (rhinopharynx cancer[Title/Abstract])) OR (epipharyngeal carcinoma[Title/Abstract])) OR (epipharynx carcinoma[Title/Abstract])) OR (naso-pharyngeal carcinoma[Title/Abstract])) OR (nasopharyngeal carcinoma[Title/Abstract])) OR (postnasal space carcinoma[Title/Abstract])) OR (rhino-pharyngeal carcinoma[Title/Abstract])) OR (rhinopharyngeal carcinoma[Title/Abstract])) OR (rhinopharynx carcinoma[Title/Abstract])) OR ("Nasopharyngeal Carcinoma"[Mesh]))	217

https://dx.doi.org/10.21037/tcr-23-859

2. Cochrane

Search number	Query	Results
#1	MeSH descriptor: [Nasopharyngeal Carcinoma] explode all trees	256
#2	(Nasopharyngeal Carcinoma):ti,ab,kw OR (Nasopharyngeal Carcinomas):ti,ab,kw OR (nasopharynx tumor):ti,ab,kw OR (epipharynx tumor):ti,ab,kw OR (epipharynx tumour):ti,ab,kw (Word variations have been searched)	1767
#3	(nasopharyngeal neoplasms):ti,ab,kw OR (nasopharyngeal tumor):ti,ab,kw OR (nasopharyngeal tumour):ti,ab,kw OR (nasopharynx tumour):ti,ab,kw OR (rhinopharyngeal tumor):ti,ab,kw (Word variations have been searched)	1308
#4	(rhinopharynx tumor):ti,ab,kw OR (rhinopharynx tumour):ti,ab,kw OR (nasopharynx cancer):ti,ab,kw OR (epipharynx cancer):ti,ab,kw OR (nasopharyngeal cancer):ti,ab,kw (Word variations have been searched)	1144
#5	(rhinopharyngioma):ti,ab,kw OR (rhinopharynx cancer):ti,ab,kw OR (epipharyngeal carcinoma):ti,ab,kw OR (epipharynx carcinoma):ti,ab,kw OR (naso-pharyngeal carcinoma):ti,ab,kw (Word variations have been searched)	6
#6	(nasopharyngeal carcinoma):ti,ab,kw OR (postnasal space carcinoma):ti,ab,kw OR (rhino-pharyngeal carcinoma):ti,ab,kw OR (rhinopharyngeal carcinoma):ti,ab,kw OR (rhinopharynx carcinoma):ti,ab,kw (Word variations have been searched)	1749
#7	#1 OR #2 OR #3 OR #4 OR #5 OR #6	2608
#8	MeSH descriptor: [Machine Learning] explode all trees	274
#9	(machine learning):ti,ab,kw OR (Transfer Learning):ti,ab,kw OR (Deep learning):ti,ab,kw OR (Hierarchical Learning):ti,ab,kw OR (Ensemble Learning):ti,ab,kw (Word variations have been searched)	5292
#10	(artificial intelligence):ti,ab,kw OR (Prediction model):ti,ab,kw OR (random forest):ti,ab,kw OR (neural network):ti,ab,kw OR (ANN):ti,ab,kw (Word variations have been searched)	30863
#11	(Support vector machine):ti,ab,kw OR (SVM):ti,ab,kw OR (Gradient Boosting Machine):ti,ab,kw OR (GBM):ti,ab,kw OR (Nomogram):ti,ab,kw (Word variations have been searched)	2866
#12	(XGboost):ti,ab,kw OR (Adaboose):ti,ab,kw OR (Decision tree):ti,ab,kw OR (External validation):ti,ab,kw OR (Risk Prediction):ti,ab,kw (Word variations have been searched)	29984
#13	(Risk-Prediction):ti,ab,kw OR (Radiomics):ti,ab,kw OR (Radiomic):ti,ab,kw OR (statistical learning):ti,ab,kw OR (predictive analytics):ti,ab,kw (Word variations have been searched)	8734
#14	#8 OR #9 OR #10 OR #11 OR #12 OR #13	60698
#15	MeSH descriptor: [Radiotherapy] explode all trees	6696
#16	(Radiotherapy):ti,ab,kw OR (Radiotherapies):ti,ab,kw OR (Radiation Therapy):ti,ab,kw OR (Radiation Therapies):ti,ab,kw OR (Radiation Treatment):ti,ab,kw (Word variations have been searched)	48473
#17	(Radiation Treatments):ti,ab,kw OR (Targeted Radiotherapies):ti,ab,kw OR (Targeted Radiotherapy):ti,ab,kw OR (Targeted Radiation Therapy):ti,ab,kw OR (Targeted Radiation Therapies):ti,ab,kw (Word variations have been searched)	22991
#18	(bioradiant therapy):ti,ab,kw OR (x ray therapy):ti,ab,kw OR (x ray treatment):ti,ab,kw OR (x-ray therapy):ti,ab,kw (Word variations have been searched)	14410
#19	#15 OR #16 OR #17 OR #18	61883
#20	#7 AND #14 AND #19	72

https://dx.doi.org/10.21037/tcr-23-859

3. Embase

Search number	Query	Results
#1	'nasopharynx tumor'/exp	32970
#2	'nasopharyngeal carcinomas':ab,ti OR 'nasopharynx tumor':ab,ti OR 'epipharynx tumor':ab,ti OR 'epipharynx tumour':ab,ti OR 'nasopharyngeal neoplasms':ab,ti OR 'nasopharyngeal tumor':ab,ti OR 'nasopharyngeal tumour':ab,ti OR 'nasopharynx tumour':ab,ti OR 'rhinopharyngeal tumor':ab,ti OR 'rhinopharynx tumor':ab,ti OR 'rhinopharynx tumour':ab,ti OR 'nasopharynx cancer':ab,ti OR 'epipharynx cancer':ab,ti OR 'nasopharyngeal cancer':ab,ti OR rhinopharyngioma:ab,ti OR 'rhinopharynx cancer':ab,ti OR 'epipharyngeal carcinoma':ab,ti OR 'epipharynx carcinoma':ab,ti OR 'naso-pharyngeal carcinoma':ab,ti OR 'nasopharyngeal carcinoma':ab,ti OR 'postnasal space carcinoma':ab,ti OR 'rhino-pharyngeal carcinoma':ab,ti OR 'rhinopharyngeal carcinoma':ab,ti OR 'rhinopharynx carcinoma':ab,ti	21961
#3	#1 OR #2	34881
#4	'machine learning'/exp	348615
#5	'machine learning':ab,ti OR 'transfer learning':ab,ti OR 'deep learning':ab,ti OR 'hierarchical learning':ab,ti OR 'ensemble learning':ab,ti OR 'artificial intelligence':ab,ti OR 'prediction model':ab,ti OR 'random forest':ab,ti OR 'neural network':ab,ti OR ann:ab,ti OR 'support vector machine':ab,ti OR svm:ab,ti OR 'gradient boosting machine':ab,ti OR gbm:ab,ti OR nomogram:ab,ti OR xgboost:ab,ti OR adaboose:ab,ti OR 'decision tree':ab,ti OR 'external validation':ab,ti OR 'risk prediction':ab,ti OR radiomics:ab,ti OR radiomic:ab,ti OR 'statistical learning':ab,ti OR 'predictive analytics':ab,ti	374079
#6	#4 OR #5	568645
#7	'radiotherapy'/exp	661590
#8	radiotherapy:ab,ti OR radiotherapies:ab,ti OR 'radiation therapy':ab,ti OR 'radiation therapies':ab,ti OR 'radiation treatment':ab,ti OR 'radiation treatments':ab,ti OR 'targeted radiotherapies':ab,ti OR 'targeted radiotherapy':ab,ti OR 'targeted radiation therapy':ab,ti OR 'targeted radiation therapies':ab,ti OR 'bioradiant therapy':ab,ti OR 'x ray therapy':ab,ti OR 'x ray treatment':ab,ti OR 'x-ray therapy':ab,ti	419142
#9	#8 OR #9	768587
#10	#3 AND #6 AND #9	456

https://dx.doi.org/10.21037/tcr-23-859

4. Web of Science

Search number	Query	Results
#1	"Nasopharyngeal Carcinoma (Topic) OR Nasopharyngeal Carcinomas (Topic) OR nasopharynx tumor (Topic) OR epipharynx tumor (Topic) OR epipharynx tumour (Topic) OR nasopharyngeal neoplasms (Topic) OR nasopharyngeal tumor (Topic) OR nasopharyngeal tumour (Topic) OR nasopharynx tumour (Topic) OR rhinopharyngeal tumor (Topic) OR rhinopharynx tumor (Topic) OR rhinopharynx tumour (Topic) OR nasopharynx cancer (Topic) OR epipharynx cancer (Topic) OR nasopharyngeal cancer (Topic) OR rhinopharyngioma (Topic) OR epipharyngeal carcinoma (Topic) OR rhinopharynx cancer (Topic) OR epipharynx carcinoma (Topic) OR naso-pharyngeal carcinoma (Topic) OR nasopharyngeal carcinoma (Topic) OR postnasal space carcinoma (Topic) OR rhino-pharyngeal carcinoma (Topic) OR rhinopharyngeal carcinoma (Topic) OR rhinopharynx carcinoma (Topic)"	25554
#2	"machine learning (Topic) OR Transfer Learning (Topic) OR Deep learning (Topic) OR Hierarchical Learning (Topic) OR Ensemble Learning (Topic) OR artificial intelligence (Topic) OR Prediction model (Topic) OR random forest (Topic) OR neural network (Topic) OR ANN (Topic) OR Support vector machine (Topic) OR SVM (Topic) OR Gradient Boosting Machine (Topic) OR GBM (Topic) OR Nomogram (Topic) OR XGboost (Topic) OR Adaboose (Topic) OR Decision tree (Topic) OR External validation (Topic) OR Risk Prediction (Topic) OR Risk-Prediction (Topic) OR Radiomics (Topic) OR Radiomic (Topic) OR statistical learning (Topic) OR predictive analytics (Topic)"	2088885
#3	"Radiotherapy (Topic) OR Radiotherapies (Topic) OR Radiation Therapy (Topic) OR Radiation Therapies (Topic) OR Radiation Treatment (Topic) OR Radiation Treatments (Topic) OR Targeted Radiotherapies (Topic) OR Targeted Radiotherapy (Topic) OR Targeted Radiation Therapy (Topic) OR Targeted Radiation Therapies (Topic) OR bioradiant therapy (Topic) OR x ray therapy (Topic) OR x ray treatment (Topic) OR x-ray therapy (Topic)"	565709
#4	"((#3) AND #2) AND #1"	463

Table S2 List of excluded studies

DOI	Title
10.1016/j.ijrobp.2022.01.047	NTCP Modeling for High-Grade Temporal Radionecroses in a Large Cohort of Patients Receiving Pencil Beam Scanning Proton Therapy for Skull Base and Head and Neck Tumors
10.1016/j.radonc.2022.06.008	Longitudinal study of irradiation-induced brain functional network alterations in patients with nasopharyngeal carcinoma
10.1186/s40644-019-0203-y	Application of a machine learning method to whole brain white matter injury after radiotherapy for nasopharyngeal carcinoma
10.1002/hbm.23852	Radiation-induced brain structural and functional abnormalities in presymptomatic phase and outcome prediction
10.5599/admet.5.4.484	Identify the radiotherapy-induced abnormal changes in the patients with nasopharyngeal carcinoma
10.1016/j.ijrobp.2016.06.2111	Multimodal testing of DNA damage response markers for prediction of normal tissue toxicities following head and neck intensity modulated radiation therapy
10.1200/JCO.2022.40.16_suppl.e18063	Voxel-based radiomics outlines spatial heterogeneity of cerebral radiation necrosis (RN) associated with bevacizumab (Bev) response in head and neck radiotherapy (RT) patients
10.1016/j.ijrobp.2022.03.027	Efficacy and Safety of Apatinib for Radiation- induced Brain Injury Among Patients With Head and Neck Cancer: An Open-Label, Single-Arm, Phase 2 Study
10.3389/fonc.2021.720417	Blood-Brain Barrier Repair of Bevacizumab and Corticosteroid as Prediction of Clinical Improvement and Relapse Risk in Radiation-Induced Brain Necrosis: A Retrospective Observational Study
10.1158/1078-0432.CCR-20-1264	A radiomics model for predicting the response to bevacizumab in brain necrosis after radiotherapy

https://dx.doi.org/10.21037/tcr-23-859

Table S3 Risk of bias assessment

No.	Author	Year	Risk of bias for all the four domains	Participants			Predictors				Outcome							Analysis									
				question1	question2	ALL	question1	question2	question3	ALL	question1	question2	question3	question4	question5	question6	ALL	question1	question2	question3	question4	question5	question6	question7	question8	question9	ALL
1	Bin Zhang	2020	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
2	Bin Zhang	2020	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
3	Bin Zhang	2020	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
4	Dan-Wan Wen	2021	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	High
5	Dan-Wan Wen	2021	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	High
6	Dan-Wan Wen	2021	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	High	Low	High
7	Dan Bao	2022	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
8	Dan Bao	2022	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
9	Dan Bao	2022	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	High	High
10	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
11	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
12	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
13	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
14	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
15	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
16	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
17	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
18	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
19	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
20	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
21	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
22	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
23	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
24	Liming Zhong	2020	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
25	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
26	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
27	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
28	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
29	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
30	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
31	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
32	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
33	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
34	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
35	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
36	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
37	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
38	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
39	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
40	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
41	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
42	X. Bin	2022	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
43	Wenqiang Guan	2020	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	High
44	Jing Hou	2022	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	High
45	Jing Hou	2022	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	High
46	Jing Hou	2022	High	High	Low	High	Low	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	High	High
47	Jiansheng Fang	2022	High	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear
48	Jiansheng Fang	2022	High	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear
49	Jiansheng Fang	2022	High	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear
50	Jiansheng Fang	2022	High	High	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear
51	Jiazhou Wang	2019	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
52	Jiazhou Wang	2019	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
53	Jiazhou Wang	2019	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	High
54	Jiazhou Wang	2019	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
55	You-ming Zhang	2021	Unclear	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low	Low	Low	Unclear
56	Lin-Mei Zhao	2021	Unclear	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low	Low	Low	Unclear
57	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
58	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
59	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
60	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
61	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
62	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
63	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
64	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
65	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
66	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
67	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
68	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
69	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
70	Xiaoshan Lin	2022	High	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Unclear	Low	Low	High	
71	Ya-fei Kang	2022	High	High	Low	High	Low	High	Low	High	Low	Unclear	Low	Unclear	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Low	Low	Low	High
72	Qing-Hua Du	2021	High	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low	High	Low	High

https://dx.doi.org/10.21037/tcr-23-859

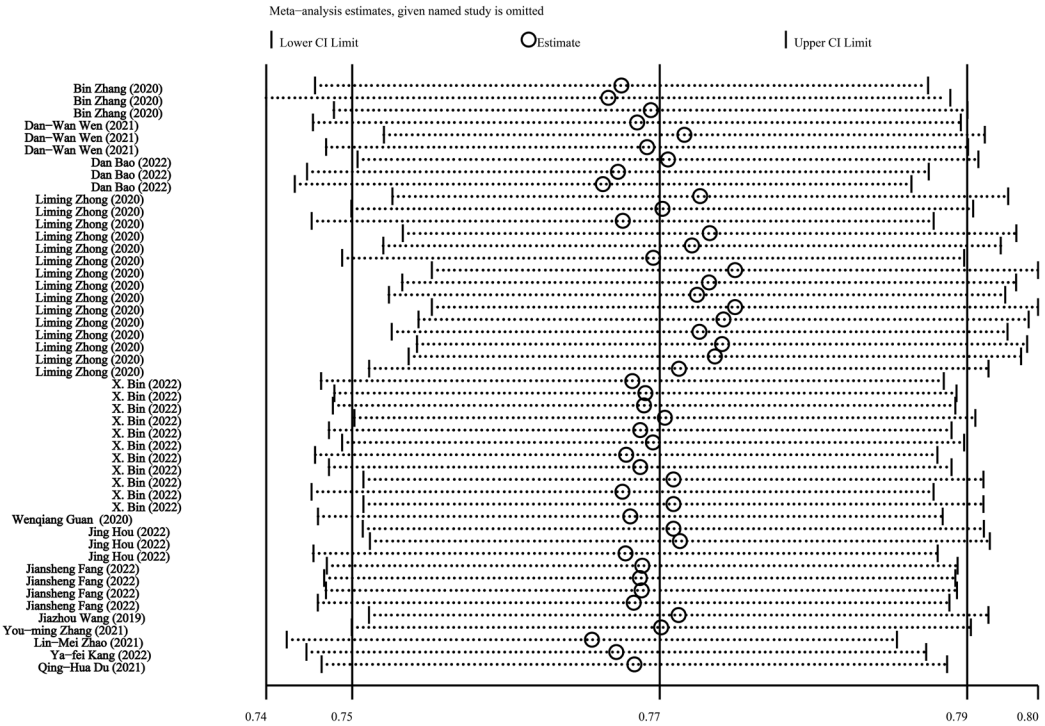
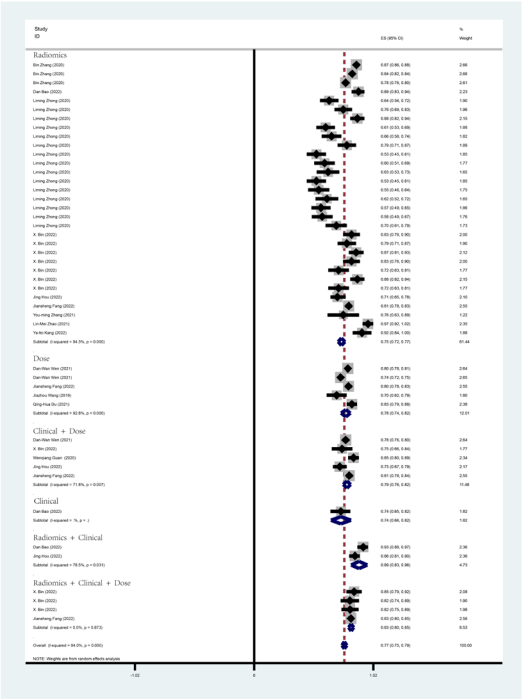


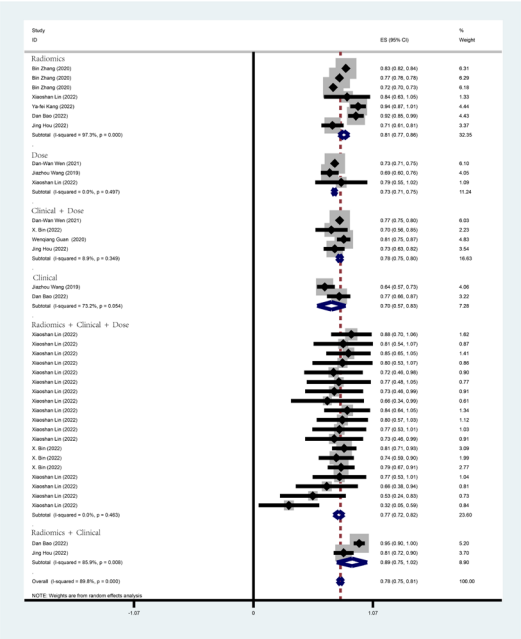
Figure S1 Sensitivity analyses of included models about C-index.

https://dx.doi.org/10.21037/tcr-23-859

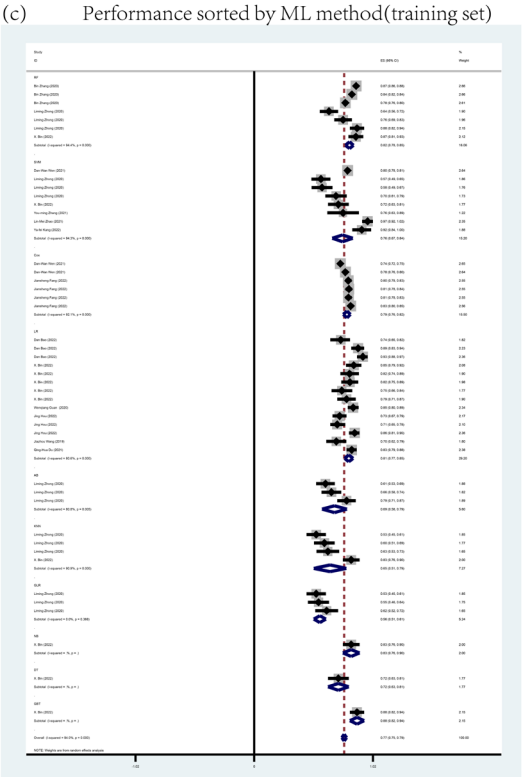
(a) Performance sorted by variable(training set)



(b) Performance sorted by variable(validation set)



(c) Performance sorted by ML method(training set)



(d) Performance sorted by ML method(validation set)

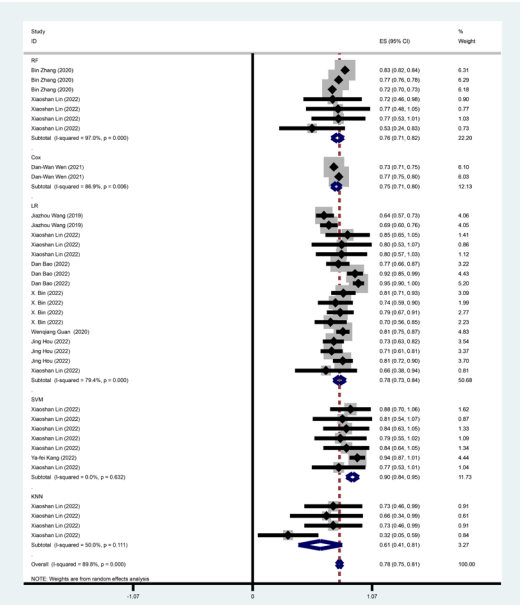


Figure S2 Forest plots of meta-analysis. (A) Performance sorted by variable (training set). (B) Performance sorted by variable (validation set). (C) Performance sorted by ML method (training set). (D) Performance sorted by ML method (validation set). ES, effect size; ML, machine learning; RF, random forest; KNN, k-nearest neighbors; AB, AdaBoost; GLR, generalized linear regression; SVM, support vector machines; LR, logistic regression; NB, Naive Bayes; DT, Decision Tree; GBT, Gradient Boosting Trees; Cox, cox proportional hazards.

https://dx.doi.org/10.21037/tcr-23-859

Appendix 1

AMSTAR 2

1. Did the research questions and inclusion criteria for the review include the components of PICO?

For Yes:	Optional (recommended)	
√ Population	Timeframe for follow-up	√ Yes
√ Intervention		No
√ Comparator group		
√ Outcome		

2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

For Partial Yes:	For Yes:	
The authors state that they had a written protocol or guide that included ALL the following:	As for partial yes, plus the protocol should be registered and should also have specified:	√ Yes
√ review question(s)	√ a meta-analysis/synthesis plan, if appropriate, *and*	Partial Yes
√ a search strategy	√ a plan for investigating causes of heterogeneity	No
√ inclusion/exclusion criteria	√ justification for any deviations from the protocol	
√ a risk of bias assessment		

3. Did the review authors explain their selection of the study designs for inclusion in the review?

For Yes, the review should satisfy ONE of the following:	
Explanation for including only RCTs	√ Yes
√ OR *Explanation for* including only NRSI	No
OR *Explanation for* including both RCTs and NRSI	

4. Did the review authors use a comprehensive literature search strategy?

For Partial Yes (all the following):	For Yes, should also have (all the following):	
√ searched at least 2 databases (relevant to research question)	searched the reference lists/bibliographies of included studies	Yes
√ provided key word and/or search strategy	√ searched trial/study registries	√ Partial Yes
√ justified publication restrictions (eg, language)	included/consulted content experts in the field	No
	where relevant, searched for grey literature	
	conducted search within 24 months of completion of the review	

5. Did the review authors perform study selection in duplicate?

For Yes, either ONE of the following:	
√ at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include	√ Yes
OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 per cent), with the remainder selected by one reviewer	No

6. Did the review authors perform data extraction in duplicate?

For Yes, either ONE of the following:	
√ at least two reviewers achieved consensus on which data to extract	√ Yes

https://dx.doi.org/10.21037/tcr-23-859

	from included studies OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 per cent), with the remainder extracted by one reviewer	No

7. Did the review authors provide a list of excluded studies and justify the exclusions?

For Partial Yes:	For Yes, must also have:	
√ provided a list of all potentially relevant studies that were read in full text form but excluded from the review	√ Justified the exclusion from the review of each potentially relevant study	√ Yes Partial Yes No

8. Did the review authors describe the included studies in adequate detail?

For Partial Yes (ALL the following):	For Yes, should also have ALL the following:	
√ described populations √ described interventions √ described comparators √ described outcomes √ described research designs	√ described population in detail √ described intervention and comparator in detail (including doses where relevant) √ described study's setting √ timeframe for follow-up	√ Yes Partial Yes No

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

RCTs		
For Partial Yes, must have assessed RoB from unconcealed allocation, *and* lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all cause mortality)	For Yes, must also have assessed RoB from: allocation sequence that was not truly random, *and* selection of the reported result from among multiple measurements or analyses of a specified outcome	Yes Partial Yes No Includes only NRSI
NRSI		
For Partial Yes, must have assessed RoB: √ from confounding, *and* √ from selection bias	For Yes, must also have assessed RoB: √ methods used to ascertain exposures and outcomes, *and* √ selection of the reported result from among multiple measurements or analyses of a specified outcome	√ Yes Partial Yes No Includes only RCTs

10. Did the review authors report on the sources of funding for the studies included in the review?

For Yes	
√ Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies	√ Yes No

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

RCTs	
For Yes: The authors justified combining the data in a meta-analysis AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present	Yes No No meta-analysis

https://dx.doi.org/10.21037/tcr-23-859

	AND investigated the causes of any heterogeneity	conducted
For NRSI For Yes: √ The authors justified combining the data in a meta-analysis √ AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present √ AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review	√	Yes No No meta-analysis conducted Not Applicable

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

For Yes: included only low risk of bias RCTs √ OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect	√	Yes No No meta-analysis conducted

13. Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review?

For Yes: included only low risk of bias RCTs √ OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results	√	Yes No

14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

For Yes: There was no significant heterogeneity in the results √ OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review	√	Yes No

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

For Yes: √ performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias	√	Yes No No meta-analysis conducted

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

For Yes: √ The authors reported no competing interests OR The authors described their funding sources and how they managed potential conflicts of interest	√	Yes No

https://dx.doi.org/10.21037/tcr-23-859