



Prediction Models for Prognosis of Cervical Cancer: Systematic Review and Critical Appraisal

Bingjie He^{1†}, Weiye Chen^{1†}, Lili Liu¹, Zheng Hou², Haiyan Zhu³, Haozhe Cheng³, Yixi Zhang³, Siyan Zhan^{1*} and Shengfeng Wang^{1*}

¹ Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, China, ² Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing, China, ³ School of Public Health, Peking University Health Science Center, Beijing, China

OPEN ACCESS

Edited by:

Kevin Lu,
University of South Carolina,
United States

Reviewed by:

Sam Li,
University of Tennessee Health
Science Center (UTHSC),
United States
Xiaomo Xiong,
University of South Carolina,
United States

*Correspondence:

Siyan Zhan
siyan-zhan@bjmu.edu.cn
Shengfeng Wang
shengfeng1984@126.com

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Family Medicine and Primary Care,
a section of the journal
Frontiers in Public Health

Received: 18 January 2021

Accepted: 23 March 2021

Published: 07 May 2021

Citation:

He B, Chen W, Liu L, Hou Z, Zhu H,
Cheng H, Zhang Y, Zhan S and
Wang S (2021) Prediction Models for
Prognosis of Cervical Cancer:
Systematic Review and Critical
Appraisal.
Front. Public Health 9:654454.
doi: 10.3389/fpubh.2021.654454

Objective: This work aims to systematically identify, describe, and appraise all prognostic models for cervical cancer and provide a reference for clinical practice and future research.

Methods: We systematically searched PubMed, EMBASE, and Cochrane library databases up to December 2020 and included studies developing, validating, or updating a prognostic model for cervical cancer. Two reviewers extracted information based on the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modeling Studies checklist and assessed the risk of bias using the Prediction model Risk Of Bias ASsessment Tool.

Results: Fifty-six eligible articles were identified, describing the development of 77 prognostic models and 27 external validation efforts. The 77 prognostic models focused on three types of cervical cancer patients at different stages, i.e., patients with early-stage cervical cancer ($n = 29$; 38%), patients with locally advanced cervical cancer ($n = 27$; 35%), and all-stage cervical cancer patients ($n = 21$; 27%). Among the 77 models, the most frequently used predictors were lymph node status ($n = 57$; 74%), the International Federation of Gynecology and Obstetrics stage ($n = 42$; 55%), histological types ($n = 38$; 49%), and tumor size ($n = 37$; 48%). The number of models that applied internal validation, presented a full equation, and assessed model calibration was 52 (68%), 16 (21%), and 45 (58%), respectively. Twenty-four models were externally validated, among which three were validated twice. None of the models were assessed with an overall low risk of bias. The Prediction Model of Failure in Locally Advanced Cervical Cancer model was externally validated twice, with acceptable performance, and seemed to be the most reliable.

Conclusions: Methodological details including internal validation, sample size, and handling of missing data need to be emphasized on, and external validation is needed to facilitate the application and generalization of models for cervical cancer.

Keywords: cervical cancer, prediction model, predictors, risk of bias, statistical analysis

INTRODUCTION

Cervical cancer is one of the leading causes of cancer-related death among women and ranks fourth among the most common malignancy in women in 2020 worldwide (1, 2). In 2020, there were an estimated 604,127 new cases and more than 341,831 deaths from cervical cancer, representing 7.7% of all female cancer deaths around the world (2). Cervical cancer patients have a poor prognosis in developing countries (3, 4). Improvement in survival remains the ultimate goal of treatment in the clinical setting, and treatment varies according to the stage, metastasis, or recurrence (5, 6). The physician can use known factors to check the personal prognosis and predict the clinical outcomes after a specific therapy such as radical hysterectomy or radiotherapy (7, 8). The prediction of potential personal prognosis facilitates the clinicians to decide the subsequent therapies or follow-up examinations (9, 10). For example, an American study indicated that clinicians could get a net benefit of 0.35 (the benefit from treating true positive patients minus the harm from treating those who do not need treatment) with only delivering intensive therapy to patients with a high risk of 3-year survival (threshold probability: about 0.76) based on the prognostic model, compared with delivering therapy to all patients (11).

A prognostic model, combining a few known factors to predict the individual risk of occurrence, progression, and clinical outcome of cervical cancer, is critical in practice (12). A couple of models have shown great positive net benefits across wide ranges of risk, indicating their favorable clinical utility (11, 13, 14). A prognostic model is a formal combination, usually a statistical equation, of multiple predictors, from which risks of a specific outcome can be calculated for individuals (15–17). As for cervical cancer patients, due to the limited predictive value for the classification of the International Federation of Gynecology and Obstetrics (FIGO) alone, a couple of prognostic models have been proposed to predict and guide treatments based on different tumor and demographic characteristics (13, 18). However, the uneven quality and the diversity of the clinical settings, outcomes, and predictors may limit the practicality of models, and systematic reviews on prognostic models of other diseases also suggested that the methodological features of existing studies varied (19–21). No comprehensive evaluation of prognostic models for cervical cancer has been done. Therefore, in order to help clinicians and the public to select the most appropriate models in practice, there is an urgent need to systematically summarize these models, map their characteristics, and examine their performance.

In this study, we conducted a systematic review to identify and summarize these prognostic models for cervical cancer patients and assess their qualities based on the guideline of the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modeling Studies (CHARMS) and the Prediction model Risk Of Bias ASsessment Tool (PROBAST) (22, 23) in order to provide evidence for determining reliable models and to provide a methodology reference for future research within this field.

MATERIALS AND METHODS

The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines guided this systematic review. **Supplementary Material 1** shows the PRISMA checklist.

Information Sources and Search Strategy

We systematically searched PubMed, EMBASE, and Cochrane library databases from their inception to December 31, 2020 for relevant articles. We combined the following search terms which were used in referring to prediction models (predict*, progn*, risk score, risk calculation, risk assessment, c statistic, discrimination, calibration, AUC, area under the curve, area under the receiver operator characteristic curve) and the disease (cervical, cervix; cancer*, carcinoma*, neoplasm*, and tumor*) based on previous research (19, 24) (* means asterisk wildcard). No other filters were applied. Details of the search strategy are given in **Supplementary Material 2**.

Eligibility Criteria

Studies were included if they reported the development, the update, or the external validation of at least one multivariable prognostic model based on individual characteristics, and the outcome of the prognostic model was any clinical outcome (recurrence, metastasis, death, *etc.*) in patients diagnosed with cervical cancer (**Table 1**). Only published studies were considered.

Studies were not eligible for inclusion if they (1) were conference abstracts, letters, and editorials or non-original studies such as reviews, (2) only examined independent prognostic factors, (3) focused on methodology, (4) developed or validated diagnostic or screening models, (5) did not construct a model to estimate individual risks, and (6) did not have a full text in English or available full text.

Study Selection

Study selection was conducted using EndNote X9. Duplicates were found and removed automatically and manually. Preliminary screening was performed through reviewing of the titles and abstracts by two independent reviewers with

TABLE 1 | Key items for framing the aim, search strategy, and study inclusion and exclusion criteria for systematic review.

Item	Definition
Population	Patients diagnosed as having cervical cancer
Intervention	Any prognostic model to predict clinical outcomes (recurrence, metastasis, death, <i>etc.</i>) in cervical cancer patients
Comparator	Not applicable
Outcomes	Any outcome reported by prognostic models
Timing	No restriction
Setting	No restriction

backgrounds in obstetrics and gynecology and uniform training, using pre-defined eligibility criteria. The full texts of the included articles after preliminary screening were reviewed independently by two reviewers. Any disagreement between reviewers was resolved by consensus.

Data Extraction

For each relevant publication, two reviewers extracted information through a piloted standardized form based on the recommendations in the CHARMS checklist (22). The key items to be extracted from each primary study were grouped within 11 domains, including source of data, participants, outcome(s) to be predicted, candidate predictors, sample size, missing data, model development, model performance, model evaluation, results, interpretation, and discussion. In addition, we extracted the general characteristics of the studies, including title, author, publication year, and specific objective (i.e., to develop or to validate or both). Model performance is typically evaluated using measures of calibration and discrimination. Calibration reflects the disparity between predictions and observed outcomes. Discrimination reflects the ability of a prognostic model to distinguish between individuals who do or do not develop the outcome (25, 26). Discrimination metrics are usually represented by concordance index (C-index) or area under the receiver operator characteristic curve (AUROC). A discrimination metric of 0.5 describes a random prediction, whereas a perfectly predicting model would have a discrimination metric of 1.0. The methods used to assess model calibration and discrimination should be suitable for the corresponding model (25).

The included models can be divided into three categories based on the FIGO stage of the research population: models for patients with early-stage cervical cancer, models for patients with locally advanced cervical cancer, and models for all-stage (including early stage and locally advanced) cervical cancer patients.

Assessment of Risk of Bias

The PROBAST was used to assess the risk of bias (ROB) of each prognostic model identified from the included studies (23, 25). Two investigators assessed ROB for each model independently. Applicability is beyond the scope of this study due to the lack of specific questions in the population, predictors, or outcome. The assessment of ROB in PROBAST involved four domains (participants, predictors, outcome, and analysis), and 20 signaling questions covered the key aspects of prediction model studies. Each question was answered as “yes,” “probably yes,” “probably no,” “no,” or “no information” according to the indications provided by PROBAST, and the ROB of each domain was judged as low, high, or unclear based on the questions in the corresponding domains. The overall assessment of each model was classified as low or high ROB. When all domains were judged as low risk, the overall risk was classified as low ROB; as long as one domain was considered high risk, the overall risk was classified as high ROB.

Data Synthesis

We conducted a descriptive analysis of the characteristics of models and reported mean or median for continuous variables, with differences calculated using *t*-test or Kruskal–Wallis test, or percentages for categorical variables, with differences calculated using χ^2 test or Fisher’s exact test. We compared the number of sample size and events and the proportion of methodological items and predictors used in the models across three categories. We used Stata15.0 for the statistical analysis.

RESULTS

Of the 52,591 screened articles, 56 articles describing the development of 77 prognostic models and the external validation of 24 prognostic models for cervical cancer patients were eligible (Figure 1). The number of prognostic models increased steadily in the past decade, tripling in annual growth in 2019. The prognostic models were mainly developed in the United States ($n = 25$; 32%), China ($n = 23$; 30%), and Korea ($n = 18$; 23%). External validation was performed only in Korea ($n = 10$; 37%), Spain ($n = 8$; 30%), China ($n = 8$; 30%), and the United States ($n = 1$; 4%). For the derivation cohorts, the median sample size and the median number of events were 549 (interquartile range, 203.5–843) and 77 (40.5–187), respectively. For the internal validation cohorts, the median sample size and the median number of events were 484.5 (234–833) and 95 (46–187), respectively.

The research population of eligible models covered all stages of cervical cancer, including patients with early-stage cervical cancer ($n = 29$; 38%), patients with locally advanced cervical cancer ($n = 27$; 35%), and all-stage cervical cancer patients ($n = 21$; 27%) (Supplementary Material 3). These models also focused on a wide range of clinical outcomes. The most common clinical outcomes were overall survival ($n = 34$; 44%), disease-free survival ($n = 12$; 16%), progression-free survival ($n = 9$; 12%), and disease-specific survival ($n = 7$; 9%). The predictors encountered most frequently were lymph nodes status ($n = 57$; 74%), the FIGO stage ($n = 42$; 55%), histological types ($n = 38$; 49%), and tumor size ($n = 37$; 48%) (Figure 2). There were at least six forms of lymph nodes included in the prognostic models. The version of FIGO stage used for patients was presented only in 19 (25%) models. Treatment was considered a predictor in 18 (23%) models.

The modeling method used most frequently in these models was Cox proportional hazard regression ($n = 68$; 88%), and machine learning was also used in five (6%) models. Model presentation was available for most prognostic models ($n = 67$; 87%), but only 16 (21%) models had a regression formula (Table 2).

Prognostic Models for Patients With Early-Stage Cervical Cancer

Twenty-nine models were developed for patients with early-stage cervical cancer. The sample size ($p = 0.011$), number of events ($p = 0.005$), and events per variable in models that were not based on machine learning ($p = 0.006$) were significantly

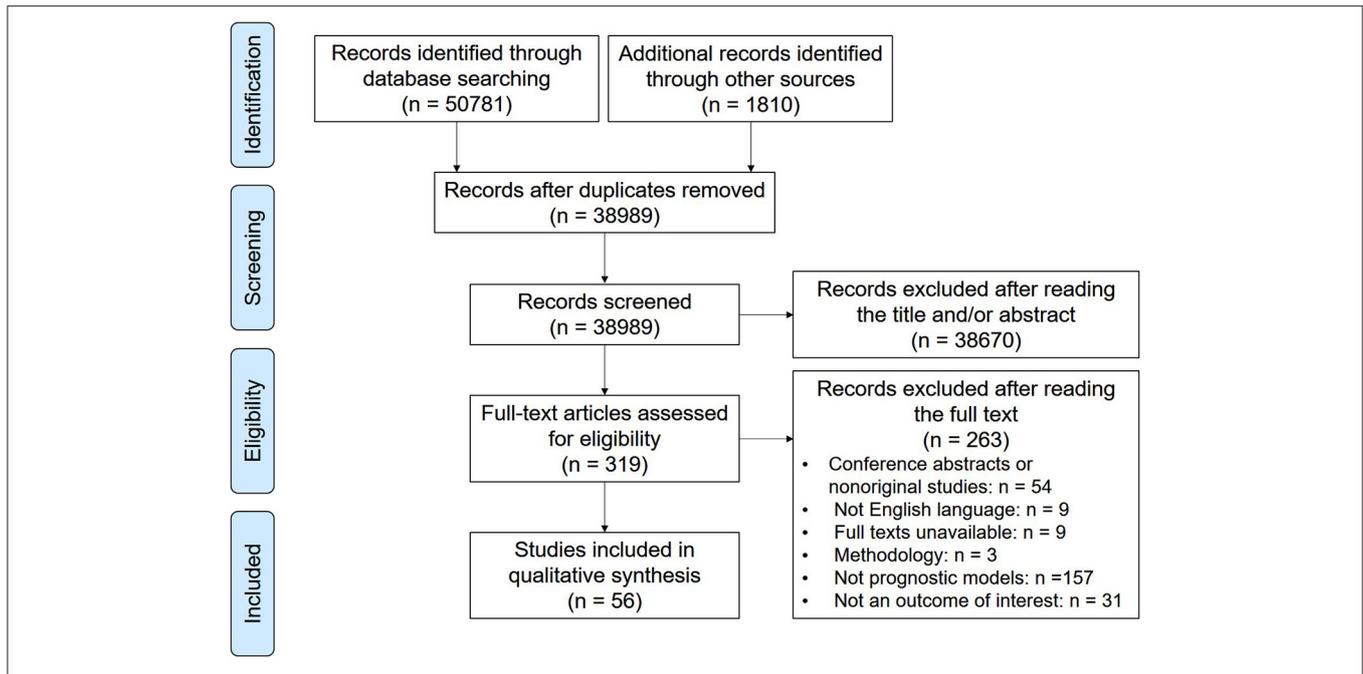


FIGURE 1 | Flow chart of the literature search for prognostic models related to cervical cancer.

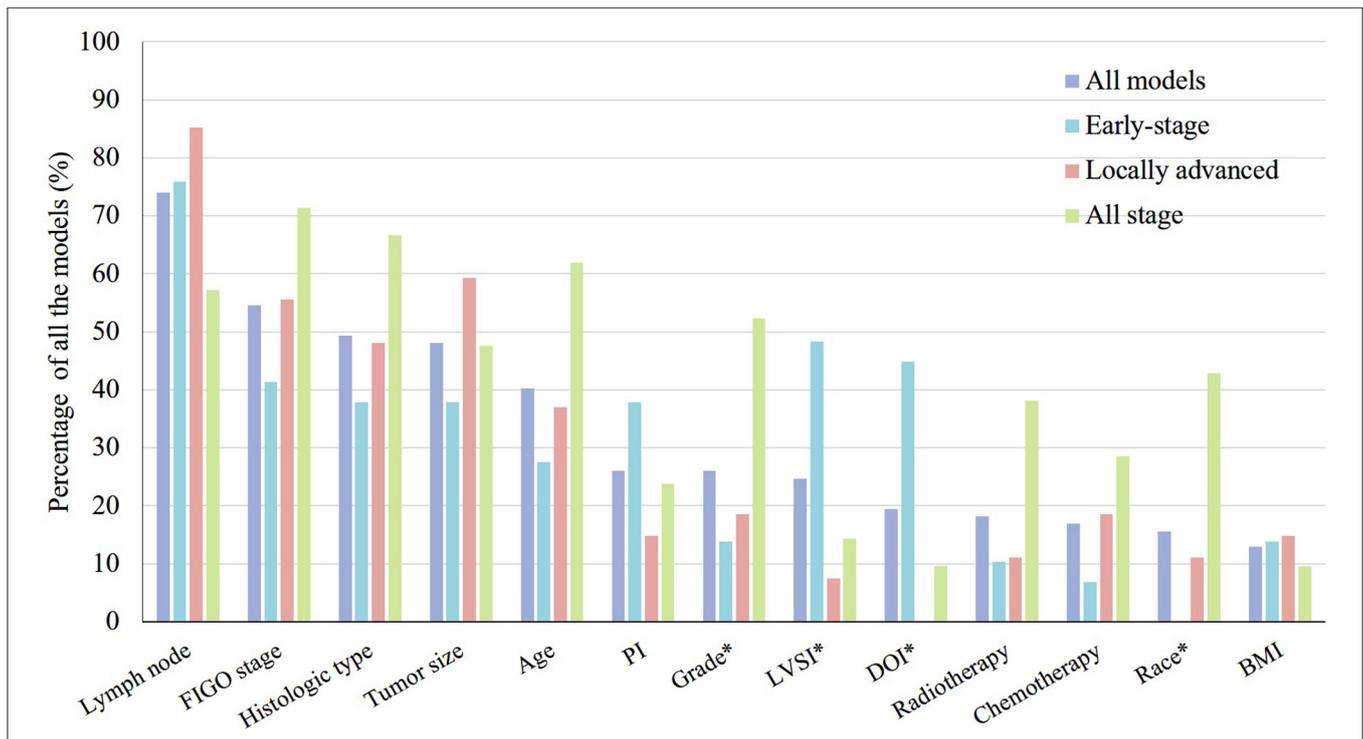


FIGURE 2 | Thirteen most frequently used predictors in 77 prognostic models for the prognosis of cervical cancer patients presented by clinical stage. PI, parametrium invasion; LVSI, lymph vascular space invasion; DOI, depth of stromal invasion; BMI, body mass index. * $p < 0.05$.

smaller in models for patients with early-stage cervical cancer. The predictors used most frequently in these models were lymph nodes status ($n = 22$; 76%), lymph vascular space invasion (LVSI

($n = 14$; 48%), depth of stromal invasion (DOI) ($n = 13$, 45%), FIGO stage ($n = 12$; 41%), histological types ($n = 11$; 38%), tumor size ($n = 11$; 38%), and parametrium invasion ($n =$

11; 38%). LVSI ($p = 0.003$) and DOI ($p < 0.001$) were used significantly more frequently in the models for patients with early-stage cervical cancer (Figure 2).

A C-index or an AUROC was reported for 24 (83%) models, and the remaining five (19%) did not report any discrimination metric. The discrimination estimates ranged from 0.565 to 0.959. Less than half of the models ($n = 14$; 48%) performed internal validation. The calibration plot was used in 13 (45%) models, and the remaining models performed no calibration assessment. The relatively recommended prognostic model for patients with early-stage cervical cancer was the Je et al. model (27).

Prognostic Models for Patients With Locally Advanced Cervical Cancer

In terms of the 27 (35%) prognostic models developed for locally advanced cervical cancer, the predictors encountered most frequently also included lymph node status ($n = 23$; 85%), tumor size ($n = 16$; 59%), FIGO stage ($n = 15$; 56%), histological types ($n = 13$; 48%), and age ($n = 10$; 37%) (Figure 2). The differences of internal validation ($p < 0.001$), model presentation ($p < 0.001$), and discrimination evaluation ($p = 0.044$) were significant across three groups. Bootstrapping and nomogram were used more frequently in the models for patients with locally advanced cervical cancer compared to the others. A C-index or an AUROC was reported for 17 (81%) models, which was significantly higher than the other two types, and the discrimination estimates ranged from 0.62 to 0.86. Calibration was assessed properly for 20 (74%) models with a calibration plot. The Prediction Model of Failure in Locally Advanced Cervical Cancer (PREFACE) was the relatively recommended model (28).

Prognostic Models for All-Stage Cervical Cancer Patients

There were 21 (27%) models developed for the prognosis of all-stage cervical cancer patients. The predictors used commonly were FIGO stage ($n = 15$; 71%), histological type ($n = 14$, 67%), age ($n = 13$; 62%), lymph nodes status ($n = 12$; 57%), grade ($n = 11$, 52%), and tumor size ($n = 10$; 48%). Grade ($p = 0.014$) and race ($p < 0.001$) were used significantly more commonly in the models for all-stage cervical cancer patients (Figure 2). A C-index or an AUROC was reported for 17 (81%) models. The discrimination estimates ranged from 0.616 to 0.897. Two-thirds of the models ($n = 14$; 67%) had internal validation. More than a half of the ($n = 12$; 57%) models were assessed with a calibration plot. The Wang et al. model was relatively recommended (29).

External Validation Studies

Of all models included, 21 (27%) were externally validated once, and three (4%) were validated twice. The median sample size and the median number of events were 211 (101–653) and 39 (18.5–83.5), respectively. Two (7%) external validation studies did not report any discrimination metric, and that of the remaining 25 studies ranged from 0.52 to 0.88. Fifteen (56%) models were assessed with calibration. The prognostic models that were externally validated twice were the SNU/AM (30), KROG 12-08 (8), and PREFACE (28), and the PREFACE model had better performance, with good calibration and a C-index

of 0.67. **Supplementary Material 4** presents detailed information on all the prognostic models.

Risk of Bias of the Included Studies

The PROBAST was used to assess the risk of bias of all studies developing or externally validating prognostic models. None of the models developed or validated were assessed as being at low ROB; the summary by domain is shown in Figure 3. As for the model development studies, 61 (79%), nine (11%), and 33 (47%) models were at low ROB for participants, predictors, and outcome, respectively, but none of the models were at low ROB for analysis. Similarly, among 27 external validation efforts, 15 (56%), four (15%), and five (19%) models were at low ROB for participants, predictors, and outcome, respectively, but none of the models were at low ROB for analysis. Specific signal questions and the result of the assessment are in **Supplementary Material 5**. The full risk of bias assessment table can be found in **Supplementary Material 6**.

The main problems with regards to the analysis domain included the inadequate number of events (68% for development and 93% for validation), the incomplete evaluation of model performance (64% for development and 74% for validation), the improper handling of missing data (86% for development and 89% for validation), and the incomplete procedure of internal validation (97% for development).

DISCUSSION

Our systematic review revealed the detailed characteristics of 77 prognostic models for the prediction of clinical outcomes in cervical cancer patients, but none of the models developed or validated were assessed as being at low ROB with PROBAST mainly because of the limitations in modeling methodology and model presentation. Relatively, the PREFACE model predicting 5-year distant recurrence in locally advanced cervical cancer patients appears to be the most reliable model.

In terms of the geographical distribution of the research population, most of the models were developed and externally validated in countries with high and very high human development index (HDI), especially in the United States, China, and Korea where the incidence of cervical cancer is relatively high rather than in less developed countries like Africa or Melanesia where the incidence and mortality were the highest (2). Prognostic models tailored to less developed countries and external validation are needed before the generalization and application. In addition, from the perspective of predictor choice, the most commonly used predictors in all eligible prognostic models were different types of tumor characteristics, which were also mentioned in the NCCN Clinical Practice Guidelines (31). The LVSI and DOI were more frequently used as predictors in the models for patients with early-stage cervical cancer. The recommended treatment for patients with early-stage cervical cancer is surgery, and these predictive factors can be obtained with more accurate results in postoperative pathological testing (31). Grade and race were also more frequently used as predictors in models for all-stage cervical cancer patients, mainly because of the generalization of these predictors for all cervical cancer

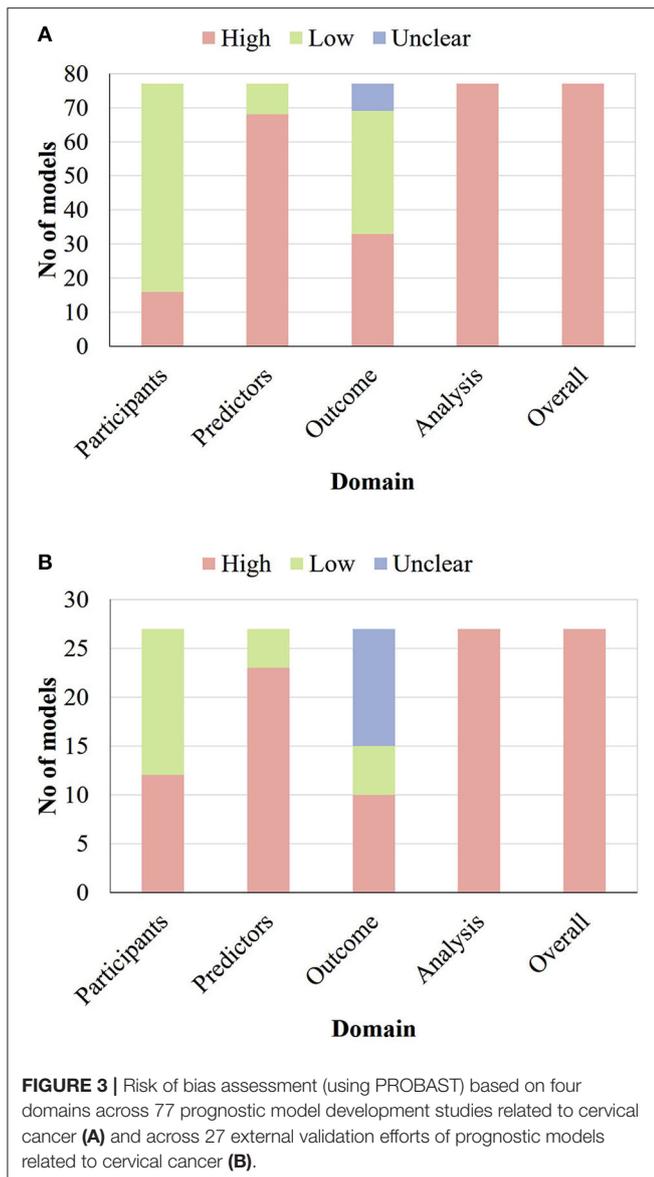
TABLE 2 | Methodological characteristics of the development of prognostic models for patients with cervical cancer by clinical stage.

Methodological items	All models (77 models)	Early stage (29 models)	Locally advanced (27 models)	All stage (21 models)	P-value
Sample size, median (IQR)	549 (203.5–843)	330 (119–788)	314 (234–833)	1,501 (371–4,220)	0.011
Number of events, median (IQR)	77 (40.5–187)	47 (19.5–96.5)	106 (52.5–246.75)	166 (45.25–696.75)	0.005
Events per variable					
Not machine learning					0.006
EPV <10	35 (49)	19 (70)	11 (41)	5 (28)	
EPV 10–20	3 (4)	1 (4)	0 (0)	2 (11)	
EPV >20	22 (31)	2 (7)	12 (44)	8 (44)	
No information	12 (17)	5 (19)	4 (15)	3 (17)	
Machine learning					1.000
EPV <100	5 (100)	2 (100)	0 (0)	3 (100)	
EPV ≥100	0 (0)	0 (0)	0 (0)	0 (0)	
Internal validation					<0.001
Bootstrapping	35 (45)	10 (34)	20 (74)	5 (24)	
Cross-validation	9 (12)	2 (7)	3 (11)	4 (19)	
Random split	6 (8)	0 (0)	1 (4)	5 (24)	
Resampling	1 (1)	1 (3)	0 (0)	0 (0)	
Not reported	1 (1)	1 (3)	0 (0)	0 (0)	
No internal validation	25 (32)	15(52)	3 (11)	7 (33)	
Modeling method					0.113
Cox hazard model	68 (88)	26 (90)	26 (96)	16 (76)	
Logistic regression	2 (3)	1 (3)	1 (4)	0 (0)	
Machine learning	5 (6)	2 (7)	0 (0)	3 (14)	
Discriminant analysis	1 (1)	0 (0)	0 (0)	1 (5)	
Not reported	1 (1)	0 (0)	0 (0)	1 (5)	
Missing data handling					0.092
Multiple imputation	4 (5)	1 (3)	3 (11)	0 (0)	
Complete case analysis	31 (40)	9 (31)	9 (33)	13 (62)	
No information	42 (55)	19 (66)	15 (66)	8 (38)	
Model presentation					<0.001
Full equation	9 (12)	8 (28)	0 (0)	1 (5)	
Nomogram	46 (60)	10 (34)	25 (93)	11 (52)	
Sum score	4 (5)	2 (7)	0 (0)	2 (10)	
CART	1 (1)	1 (3)	0 (0)	0 (0)	
More than one method	7 (9)	3 (10)	2 (7)	2 (10)	
None	10 (13)	5 (17)	0 (0)	5 (24)	
Discrimination					0.044
C-index or AUROC	68 (88)	24 (83)	27 (100)	17 (81)	
None	9 (12)	5 (17)	0 (0)	4 (19)	
Calibration					0.087
Calibration plot	45 (58)	13 (45)	20 (74)	12 (57)	
None	32 (42)	16 (55)	7 (26)	9 (43)	

Values are numbers (percentages^a) unless stated otherwise.

^aSome percentages do not add up to 100%, owing to rounding off.

IQR, interquartile range; EPV, events per variable; CART, Classification and Regression Tree; C-index, concordance index; AUROC, area under the receiver operating characteristic.



patients. Our research suggested that some issues related to predictors should be noted when developing, validating, or using prognostic models. First, the forms of predictors included in prognostic models are not consistent. For example, at least six forms of lymph node status were involved in existing models (29). A practical problem is that the effect of each form may be different, and which one is superior requires research to examine, which can help subsequent studies choose the appropriate one. Second, the FIGO stage is also crucial to the progression of cervical cancer, but the staging system has been revised at least three times during the past decades (32, 33). Researchers should confirm the version of FIGO stage they used ahead of model development or application. Third, the accuracy of measurement and the availability of predictors should be also considered. The tumor characteristics mentioned above are

usually measured by different imaging evaluation and pathologic examination, which is subjective and may have inadequate accuracy (34). Therefore, parallel assessment and blindness are necessary during measurement procedures to ensure quality. Additionally, despite the increasing number of research on the predictive value of omics (35–37), the inclusion of omics into a conventional prediction model needs to be discussed considering the accessibility of these predictors and cost-effectiveness issues, especially in countries with middle or low income. Besides this, treatment is very important for the prognosis of cervical cancer patients, and different surgical methods may lead to different survival outcomes (38), but treatments were considered a predictor only in 18 models, which indicates that prognostic models tailored to different treatment modalities are needed.

Our systematic review showed several methodological pitfalls in the development or validation of the models, especially in the analysis domain which was the main cause of high ROB. Similar to reviews assessing the quality of prognostic models for other diseases including oropharyngeal cancer, chronic lymphocytic leukemia, and chronic obstructive pulmonary disease (19, 20, 39, 40), inadequate sample size, improper handling of missing data, and incomplete evaluation of model performance were the main problems in analysis domain and needed to be emphasized. Besides this, several other methodological details should also be noted. First, some common problems of internal validation exist, and the proportion of external validation was low. Only a half of the studies applied internal validation properly by using bootstrapping or cross-validation techniques to overcome overfitting; however, almost none of the bootstrapping or cross-validation techniques used in these studies were applied in all model development procedures such as variable selection, which was also consistent with other studies (41). It is one of the most overlooked technical aspects at present, but it violates the guidance and recommendations of PROBAST on internal validation (25, 42), indicating that the recommendation needs to be further strengthened. Additionally, only one third of the models have been externally validated, although the demonstration of the performance of a model in an independent population is a necessary step before recommending its widespread use (43). Researchers generally do not have access to multiple data, and external validation might be limited because of that. Data sharing can offer the possibility of making the most of all available data, which should be promoted in the future (44). Second, besides the most common Cox proportional hazard regression that was suitable for survival data, machine learning was also adopted for prognostic prediction in cervical cancer from 2019 (45). The discrimination parameter of the machine learning model was better than the traditional one (0.795 vs. 0.784) (45). The model performance could therefore be partially improved, but researchers should also attend to the difficulty of interpretation of algorithm and requirement of large sample size (46). In addition, it should be noted that different predicted outcomes have corresponding suitable modeling methods, and the same type of outcome also has alternative statistical models; therefore, researchers should consider the characteristics of different modeling methods when choosing one (47). Third, the model presentation was also inappropriate, with almost

one-seventh of the models having no presentation and only one-fifth providing the full model equation. Researchers should provide not only concise and proper presentation formats but also the full model equation to enable independent external validation, update, and recalibration (48). Compared with the above-mentioned two issues, this one is much easier to realize. Both authors and journals should pay attention to it to directly facilitate the validation and generalization of prognostic models.

In terms of the differences of methodological details across the three groups, the problem of adequate sample size was most severe in models for patients with early-stage cervical cancer. Early-stage patients have a better prognosis, leading to a fewer number of events. The methodological details (including internal validation, model presentation, and discrimination evaluation) of models for patients with locally advanced cervical cancer were generally better than the other two categories. We think that the reason for those differences might not be relevant to the features of patients in different stages. Indeed the all-stage cervical cancer patients included both patients with early-stage cancer and patients with locally advanced cancer. We think that the choice of method should be mainly related to the researchers. We added new analysis to compare the year of publication and the involvement of methodological experts across the three groups. We found that the publication years of the models were similar across the three groups, and the proportion of models involving methodological experts was not high in the locally advanced group (**Supplementary Material 3**). Although the current clues cannot easily explain the above-mentioned phenomenon, this result indicated that a couple of research teams did better at methodological details, and it also takes lead to improve the qualities of models in this field.

Although all of the existing models have high ROB, current research still inevitably needs the guidance of prognostic models on clinical intervention, therapeutic strategy, clinical trial, and so on. Clinicians need to choose the most appropriate and corresponding models to predict the prognosis of patients at different stages. Appropriate models should possess high quality, good model performance, strict validation, and available predictors. Before the development of new models with better quality, the Je et al. model (27), PREFACE model (28), and Wang et al. model (29) were relatively recommended for early stage, locally advanced, and all-stage cervical cancer patients, respectively. The Je et al. (27) model was recommended based on the relatively standardized measurement of predictors and outcome, large sample size, complete evaluation and good result of model performance, and internal validation. The PREFACE model was recommended based on the barely enough sample size and external validation in both Korea and Spain with acceptable performance. The Wang et al. (29) model was based on a large database with good data quality and enough sample size and was externally validated in an independent population with good performance. Of course, researchers should notice the high ROB and existing problems of these models when using them. The serum squamous cell carcinoma antigen, a predictor included in the PREFACE model, can only predict the prognosis of cervical squamous cell carcinoma (49, 50). The Je et al. (27) model still needs to be externally validated, and some

analysis details including missing data were handled improperly in the Wang et al. (27) model. Therefore, external validation and more high-quality prognostic models tailored to different cervical cancer patients are needed before models can be implemented in clinical practice.

To get prognostic models with good predictive capability for clinical practice, researchers should pay attention to each detail in the modeling process. The larger value of discrimination and calibration does not always mean a better model; instead it is also important to assess the risk of bias of the model from multiple domains. Although these domains are very detailed, it is still recommended to pay attention and value to avoid misuse. On the basis of the aforementioned problems, besides some common issues including recruiting enough participants to obtain enough outcome events, applying imputation techniques to handle missing data, and providing complete model performance measures, the following recommendations could be stated to improve the research on prognostic models for cervical cancer. First, model development studies should apply internal validation to adjust for overfitting and include all model development procedures. Second, several modeling guidelines of methodology and reporting specific to machine learning have been developed recently (51, 52), and researchers should follow these guidelines to correctly apply machine learning for prediction in the future. Third, researchers should carefully consider the intended users, settings, and timing when choosing the format to present the model, and a final prognostic model equation should always be presented. Finally, research is mostly concentrated in countries with high and very high HDI now; similar studies in other populations are needed.

To our knowledge, this is the first systematic review to provide a comprehensive overview of available studies on prognostic models for different outcomes of cervical cancer according to the CHARMS and PROBAST tools. Nevertheless, it has some limitations. First, we only included studies published in English and did not search gray literature, but the missing models due to this are limited in usage and usually of relatively low quality. Second, a quantitative analysis was not conducted due to the heterogeneity of participants and outcomes.

CONCLUSION

The number of prognostic models for cervical cancer has increased steadily in the past decade, tripling in annual growth in 2019, and the performance of these models varies. However, all of them are at high risk of bias mainly because of unsatisfactory statistical analysis. High-quality prognostic models tailored to different stages of cervical cancer, external validation by an independent population, and head-to-head comparisons of existing models are needed to inform clinical practice better.

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/s.

AUTHOR CONTRIBUTIONS

BH, WC, LL, and SW designed the study. BH, WC, HZ, HC, and YZ did the literature search and the data extraction. BH and WC performed the initial analysis of data and wrote the first draft of the manuscript. ZH, SW, and SZ contributing to the interpretation of data. All the authors contributed to the critical revision of the manuscript for important intellectual content and approved the final version.

FUNDING

This research was supported by the National Key Research and Development Program of China (2018YFC1311704)

REFERENCES

- Small W Jr, Bacon MA, Bajaj A, Chuang LT, Fisher BJ, Harkenrider MM, et al. Cervical cancer: a global health crisis. *Cancer*. (2017) 123:2404–12. doi: 10.1002/cncr.30667
- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. (2021) 1–41. doi: 10.3322/caac.21660
- Sankaranarayanan R, Swaminathan R, Brenner H, Chen K, Chia KS, Chen JG, et al. Cancer survival in Africa, Asia, and Central America: a population-based study. *Lancet Oncol*. (2010) 11:165–73. doi: 10.1016/S1470-2045(09)70335-3
- Sengayi-Muchengeti M, Joko-Fru WY, Miranda-Filho A, Egue M, Akele-Akpo MT, N'da G, et al. Cervical cancer survival in sub-Saharan Africa by age, stage at diagnosis and Human Development Index: A population-based registry study. *Int J Cancer*. (2020) 147:3037–48. doi: 10.1002/ijc.33120
- WHO. *WHO Guidelines for Screening and Treatment of Precancerous Lesions for Cervical Cancer Prevention*. Geneva, IL (2013).
- ESGO. *Cervical Cancer Guidelines*. Geneva, IL (2017).
- Sevin BU, Nadji M, Lampe B, Lu Y, Hilsenbeck S, Koechli OR, et al. Prognostic factors of early stage cervical cancer treated by radical hysterectomy. *Cancer*. (1995) 76(10 Suppl):1978–86. doi: 10.1002/1097-0142(19951115)76:10+<1978::AID-CNCR2820761313>3.0.CO;2-K
- Je Hu, Han S, Kim YS, Nam JH, Kim HJ, Kim JW, et al. A nomogram predicting the risks of distant metastasis following postoperative radiotherapy for uterine cervical carcinoma: a Korean radiation oncology group study (KROG 12-08). *Radiother Oncol*. (2014) 111:437–41. doi: 10.1016/j.radonc.2014.03.025
- Yang J, Tian G, Pan Z, Zhao F, Feng X, Liu Q, et al. Nomograms for predicting the survival rate for cervical cancer patients who undergo radiation therapy: a SEER analysis. *Future Oncol*. (2019) 15:3033–45. doi: 10.2217/fon-2019-0029
- Wang W, Liu X, Meng Q, Zhang F, Hu K. Nomograms predicting survival and patterns of failure in patients with cervical cancer treated with concurrent chemoradiotherapy: a special focus on lymph nodes metastases. *PLoS ONE*. (2019) 14:e0214498. doi: 10.1371/journal.pone.0214498
- Zhang S, Wang X, Li Z, Wang W, Wang L. Score for the overall survival probability of patients with first-diagnosed distantly metastatic cervical cancer: a novel nomogram-based risk assessment system. *Front Oncol*. (2019) 9:1106. doi: 10.3389/fonc.2019.01106
- Paik ES, Lim MC, Kim MH, Kim YH, Song ES, Seong SJ, et al. Prognostic model for survival and recurrence in patients with early stage cervical cancer: a Korean Gynecologic Oncology Group study (KGOG 1028). *Cancer Res Treat*. (2019) 52:320–33. doi: 10.4143/crt.2019.124
- Lora D, Gómez de la Cámara A, Fernández SP, Enríquez de Salamanca R, Gómez JFPR. Prognostic models for locally advanced cervical cancer: external validation of the published models. *J Gynecol Oncol*. (2017) 28:e58. doi: 10.3802/jgo.2017.28.e58
- Xie G, Wang R, Shang L, Qi C, Yang L, Huang L, et al. Calculating the overall survival probability in patients with cervical cancer: a nomogram and decision curve analysis-based study. *BMC Cancer*. (2020) 20:833. doi: 10.1186/s12885-020-07349-4
- Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. (2009) 338:b375. doi: 10.1136/bmj.b375
- Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ*. (2009) 338:b604. doi: 10.1136/bmj.b604
- Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. (2015) 162:55–63. doi: 10.7326/M14-0697
- Yoon WS, Yang DS, Lee JA, Lee NK, Park YJ, Kim CY, et al. Validation of nomograms for survival and metastases after hysterectomy and adjuvant therapy in uterine cervical cancer with risk factors. *Biomed Res Int*. (2017) 2017:2917925. doi: 10.1155/2017/2917925
- 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*. (2019) 367:l5358. doi: 10.1136/bmj.l5358
- Palazon-Bru A, Mares-Garcia E, Lopez-Bru D, Mares-Arambul E, Folgado-de la Rosa DM, Carbonell-Torregrosa MLA, 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) 42:763–73. doi: 10.1002/hed.26025
- Kreuzberger N, Damen JA, Trivella M, Estcourt LJ, Aldin A, Umlauff L, et al. Prognostic models for newly-diagnosed chronic lymphocytic leukaemia in adults: a systematic review and meta-analysis. *Cochrane Database Syst Rev*. (2020) 7:Cd012022. doi: 10.1002/14651858.CD012022.pub2
- 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:e1001744. doi: 10.1371/journal.pmed.1001744
- Wolff RE, 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 Intern Med*. (2019) 170:51–8. doi: 10.7326/M18-1376
- Chen S, Tao M, Zhao L, Zhang X. The association between diabetes/hyperglycemia and the prognosis of cervical cancer patients: a systematic review and meta-analysis. *Medicine*. (2017) 96:e7981. doi: 10.1097/MD.0000000000007981
- Moons KGM, Wolff RE, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med*. (2019) 170:W1–33. doi: 10.7326/M18-1377
- Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, et al. Discrimination and calibration of clinical prediction

and the National Research Program for Key Issues in Air Pollution Control (DQGG0404). The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2021.654454/full#supplementary-material>

- models: users' guides to the medical literature. *JAMA*. (2017) 318:1377–84. doi: 10.1001/jama.2017.12126
27. Je HU, Han S, Kim YS, Nam JH, Park W, Song S, et al. Risk prediction model for disease-free survival in women with early-stage cervical cancers following postoperative (Chemo)radiotherapy. *Tumori*. (2018) 104:105–10. doi: 10.5301/tj.5000697
 28. Kang S, Nam BH, Park JY, Seo SS, Ryu SY, Kim JW, et al. Risk assessment tool for distant recurrence after platinum-based concurrent chemoradiation in patients with locally advanced cervical cancer: a Korean gynecologic oncology group study. *J Clin Oncol*. (2012) 30:2369–74. doi: 10.1200/JCO.2011.37.5923
 29. Wang C, Yang C, Wang W, Xia B, Li K, Sun F, et al. A prognostic nomogram for cervical cancer after surgery from SEER database. *J Cancer*. (2018) 9:3923–8. doi: 10.7150/jca.26220
 30. Kim MK, Jo H, Kong HJ, Kim HC, Kim JW, Kim YM, et al. Postoperative nomogram predicting risk of recurrence after radical hysterectomy for early-stage cervical cancer. *Int J Gynecol Cancer*. (2010) 20:1581–6. doi: 10.1111/IGC.0b013e3181f7b353
 31. Koh WJ, Abu-Rustum NR, Bean S, Bradley K, Campos SM, Cho KR, et al. Cervical Cancer, Version 3.2019, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. (2019) 17:64–84. doi: 10.6004/jcncc.2019.0001
 32. Pecorelli S. Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. *Int J Gynaecol Obstet*. Cham (2009) 105:103–4. doi: 10.1016/j.ijgo.2009.02.012
 33. Bhatla N, Aoki D, Sharma DN, Sankaranarayanan R. Cancer of the cervix uteri. *Int J Gynaecol Obstet*. (2018) 143(Suppl. 2):22–36. doi: 10.1002/ijgo.12611
 34. Du R, Li L, Ma S, Tan X, Zhong S, Wu M. Lymph nodes metastasis in cervical cancer: incidences, risk factors, consequences and imaging evaluations. *Asia Pac J Clin Oncol*. (2018) 14:e380–e5. doi: 10.1111/ajco.12997
 35. Xie F, Dong D, Du N, Guo L, Ni W, Yuan H, et al. An 8-gene signature predicts the prognosis of cervical cancer following radiotherapy. *Mol Med Rep*. (2019) 20:2990–3002. doi: 10.3892/mmr.2019.10535
 36. Uleberg KE, Øvestad IT, Munk AC, Brede C, Diermen BV, Gudlaugsson E, et al. Prediction of spontaneous regression of cervical intraepithelial neoplasia lesions grades 2 and 3 by proteomic analysis. *Int J Proteomics*. (2014) 2014:129064. doi: 10.1155/2014/129064
 37. Takada A, Yokota H, Watanabe M, Horikoshi T, Uno T. MRI radiomics analysis for predicting prognosis of cervical cancer after definitive radiotherapy. *Radiother Oncol*. (2019) 133:S264–5. doi: 10.1016/S0167-8140(19)30929-6
 38. Ramirez PT, Frumovitz M, Pareja R, Lopez A, Vieira M, Ribeiro R, et al. Minimally invasive versus abdominal radical hysterectomy for cervical cancer. *N Engl J Med*. (2018) 379:1895–904. doi: 10.1056/NEJMoa1806395
 39. Carrillo-Larco RM, Altez-Fernandez C, Pacheco-Barríos N, Bambs C, Irazola V, Miranda JJ, et al. Cardiovascular disease prognostic models in Latin America and the Caribbean: a systematic review. *Glob Heart*. (2019) 14:81–93. doi: 10.1016/j.gheart.2019.03.001
 40. Carrillo-Larco RM, Aparcana-Granda DJ, Mejia JR, Barengo NC, Bernabe-Ortiz A. Risk scores for type 2 diabetes mellitus in Latin America: a systematic review of population-based studies. *Diabet Med*. (2019) 36:1573–84. doi: 10.1111/dme.14114
 41. Heestermans T, Payne B, Kayode GA, Amoakoh-Coleman M, Schuit E, Rijken MJ, et al. Prognostic models for adverse pregnancy outcomes in low-income and middle-income countries: a systematic review. *BMJ Glob Health*. (2019) 4:e001759. doi: 10.1136/bmjgh-2019-001759
 42. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. (2001) 54:774–81. doi: 10.1016/S0895-4356(01)00341-9
 43. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. external validation, model updating, and impact assessment. *Heart*. (2012) 98:691–8. doi: 10.1136/heartjnl-2011-301247
 44. Choi YI, Park SJ, Chung JW, Kim KO, Cho JH, Kim YJ, et al. Development of machine learning model to predict the 5-year risk of starting biologic agents in patients with Inflammatory Bowel Disease (IBD): K-CDM network study. *J Clin Med*. (2020) 9:E3427. doi: 10.3390/jcm9113427
 45. Matsuo K, Purushotham S, Jiang B, Mandelbaum RS, Takiuchi T, Liu Y, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ*. (2020) 368:m689. doi: 10.1136/bmj.m689
 46. 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*. (2020) 368:m689. doi: 10.1136/bmj.m689
 47. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. 2nd edn. Springer Nature Switzerland. (2019) doi: 10.1007/978-3-030-16399-0
 48. Bonnett LJ, Snell KIE, Collins GS, Riley RD. Guide to presenting clinical prediction models for use in clinical settings. *BMJ*. (2019) 365:l737. doi: 10.1136/bmj.l737
 49. Duk JM, Groenier KH, de Bruijn HW, Hollema H, ten Hoor KA, van der Zee AG, et al. Pretreatment serum squamous cell carcinoma antigen: a newly identified prognostic factor in early-stage cervical carcinoma. *J Clin Oncol*. (1996) 14:111–118. doi: 10.1200/JCO.1996.14.1.111
 50. Chen P, Jiao L, Ren F, Wang DB. Clinical value of serum squamous cell carcinoma antigen levels in predicting chemosensitivity, lymph node metastasis, and prognosis in patients with cervical squamous cell carcinoma. *BMC Cancer*. (2020) 20:423. doi: 10.1186/s12885-020-06934-x
 51. Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. *Lancet*. (2019) 393:1577–9. doi: 10.1016/S0140-6736(19)30037-6
 52. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. *J Med Internet Res*. (2016) 18:e323. doi: 10.2196/jmir.5870

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.

Copyright © 2021 He, Chen, Liu, Hou, Zhu, Cheng, Zhang, Zhan and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.