

Prediction Models for Gastric Cancer Risk in the General Population: A Systematic Review

Jianhua Gu¹, Ru Chen¹, Shao-Ming Wang¹, Minjuan Li¹, Zhiyuan Fan¹, Xinqing Li¹, Jiachen Zhou², Kexin Sun¹, and Wenqiang Wei¹



ABSTRACT

Risk prediction models for gastric cancer could identify high-risk individuals in the general population. The objective of this study was to systematically review the available evidence about the construction and verification of gastric cancer predictive models. We searched PubMed, Embase, and Cochrane Library databases for articles that developed or validated gastric cancer risk prediction models up to November 2021. Data extracted included study characteristics, predictor selection, missing data, and evaluation metrics. Risk of bias (ROB) was assessed using the Prediction model Risk Of Bias Assessment Tool (PROBAST). We identified a total of 12 original risk prediction models that fulfilled the criteria for analysis. The area under the receiver operating characteristic curve (AUC) ranged from 0.73 to 0.93 in derivation sets ($n = 6$), 0.68 to 0.90 in internal validation sets ($n = 5$), 0.71 to 0.92 in external validation sets ($n = 7$). The higher-performing models usually include age, salt preference, *Helicobacter pylori*, smoking, body mass

index, family history, pepsinogen, and sex. According to PROBAST, at least one domain with a high ROB was present in all studies mainly due to methodologic limitations in the analysis domain. In conclusion, although some risk prediction models including similar predictors have displayed sufficient discriminative abilities, many have a high ROB due to methodologic limitations and are not externally validated efficiently. Future prediction models should adhere to well-established standards and guidelines to benefit gastric cancer screening.

Prevention Relevance: Through systematical reviewing available evidence about the construction and verification of gastric cancer predictive models, we found that most models have a high ROB due to methodologic limitations and are not externally validated efficiently. Future prediction models are supposed to adhere to well-established standards and guidelines to benefit gastric cancer screening.

Introduction

According to the latest report of Global Cancer Statistics 2020 (GLOBOCAN 2020), gastric cancer is the fifth most commonly diagnosed and the fourth most life-threatening cancer (1). The endoscopic examination with biopsies is the gold standard for the diagnosis of gastric cancer and has been universally accepted and recommended for routine screening (2, 3). However, the disadvantages of high cost, invasive

examination, and technically demanding extremely restrict the large-scale application, particularly in countries with low incidence rates or limited medical resources. Thus, there is an urgent need to develop efficient methods that enable the early identification of the high-risk groups for subsequent endoscopy.

Risk prediction models, which are designed to estimate future risk based on current and past information (4), can be used for risk stratification in screening programs at the population level. The development of gastric cancer is a protracted process influenced by genetic and environmental factors and accompanied by changes in serologic biomarkers, such as *Helicobacter pylori* (HP) antibody (5), pepsinogen I (PGI), and II (PGII) (6). This provides a basis for the construction of gastric cancer risk prediction models.

An ideal prediction model should have good performance in both discrimination and generalization. Previous studies targeting other diseases such as oropharyngeal cancer (7), esophageal cancer (8), and breast cancer (9) have suggested that the high risk of bias due to inappropriate predictors or methodology could restrict the applicability and predictive power of prediction models. Therefore, assessment of the risk of bias (ROB) and applicability to the intended population and setting for gastric cancer risk prediction models are highly warranted. Although recent years have seen an increasing number of risk

¹Office of National Central Cancer Registry, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P.R. China. ²Department of Epidemiology and Biostatistics, School of Public Health, Xi'an Jiaotong University Health Science Center, Shaanxi, P.R. China.

Note: Supplementary data for this article are available at Cancer Prevention Research Online (<http://cancerprevres.aacrjournals.org/>).

Corresponding Authors: Wenqiang Wei, National Central Cancer Registry, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 17 Panjiayuan, Chaoyang District, Beijing 100021, P.R. China. Phone: 8610-8778-8661; Fax: 8610-6771-3648; E-mail: weiwq@cicams.ac.cn; and Ru Chen, chenru1900@163.com

Cancer Prev Res 2022;15:309-18

doi: 10.1158/1940-6207.CAPR-21-0426

©2022 American Association for Cancer Research

prediction models for gastric cancer, an attempt to comprehensively evaluate model development and validation has not been performed.

In this study, we sought to systematically review the construction and validation of multivariable risk prediction models, that were used to predict the risk of developing gastric cancer in the general population. The study aimed to identify the most common predictors incorporated into prediction models and evaluate study characteristics associated with model performance to guide future model development.

Materials and Methods

The present systematic review was carried out following Preferred Reporting Items for Systematic reviews and Meta-Analyses guidelines (10).

Search strategy

We exhaustively searched PubMed (<https://www.ncbi.nlm.nih.gov/pubmed>), Embase (<https://www.embase.com/>), and Cochrane Library databases (<https://www.cochranelibrary.com/>) for multivariable risk prediction models of gastric cancer. The articles were searched up to November 2021 and only those published in English were included. The search terms were referred to previous studies (11, 12) and detailed retrieval strategies were provided in the Supplementary Appendix (Supplementary Table S1-S3). Two independent reviewers

(Jianhua Gu and Ru Chen) examined the titles and abstracts of the initial screening articles according to the following selection criteria: (i) articles related to gastric cancer prevention or prediction; (ii) a new algorithm or risk prediction model was developed in the general population; (iii) exclude animal studies, reviews and meta-analysis. In addition, we also checked the reference lists of eligible articles to ensure that important articles were not missed.

Data collection

The following information was extracted from the included articles: the first author, publication year, study design, study population (country, ethnicity, age and sex of participants, sample size), number of events, model-related information (statistical methods, model performance, modeling building strategies, validation method, and predictors in final analysis). According to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines (13), we classified the methods of eligible articles published for each risk prediction model. If multiple models were established by age and sex in a given study, all were included separately.

Quality assessment

We used the Prediction model Risk Of Bias ASsessment Tool (PROBAST) to assess the ROB and applicability of the included studies. The PROBAST involves the assessment of four domains: participants, predictors, outcome, and

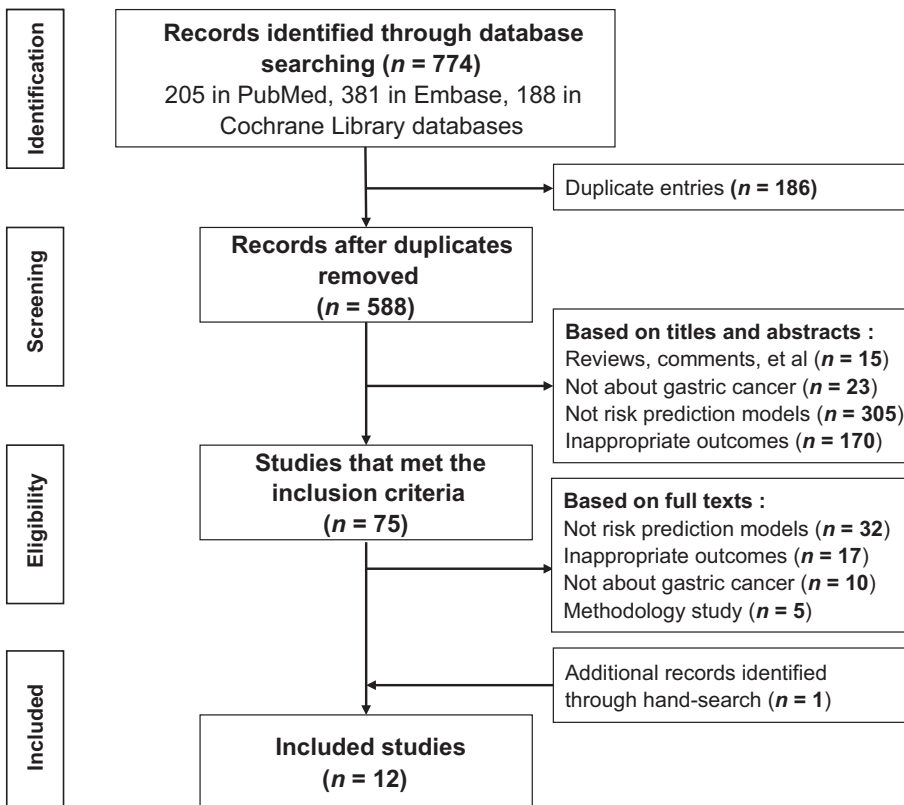


Figure 1. Flowchart of literature-screening process. This figure displays the flow of the literature selection applying the systematic search and selection strategies to identify eligible studies.

Table 1. Summary of included studies.

Authors (Year)	Country	Study design	Study period	Method	Study type ^a	Samples	Cases	Age	Male (%)	c-statistics
Cai Q (2019)	China	Cross-sectional study	2015-2017	Logistic	2a	9,383(D) 5,091(EV)	267(D) 138(EV)	56.2 ± 9.6	46.94	0.76(D) 0.76(IV) 0.73(EV)
Charvat H (2016)	Japan	Cohort study	1993-2009	COX	1b	19,028	412	59.3(51.5-64.9) (cases) 59.3(57.9-67.1) (non-cases)	61.9(cases) 35.7(non-cases)	0.77(IV)
Charvat H (2020)	Japan	Cohort study	1990-1993	COX	4	1,292	33	56.52 ± 5.78	34.1%	0.80(EV)
Eom BW (2015)	Korea	Cohort study	1996-2007(D)	COX (men)	3	1,372,424(D)	19,465(D)	45.1 ± 10.5(men)(D)	77.8(D)	0.78(EV)
					3	484,335(EV)	6,628(EV)	46.8 ± 12.8(men)(EV)	51.0(EV)	
					3	804,077(D)	5,579(D)	48.7 ± 11.0(women)(D)		
					3	466,013(EV)	2,920(EV)	51.1 ± 12.1(women)(EV)		
Iida M (2018)	Japan	Cohort study	1988-2002(D)	COX (women)	3	2,444(D)	90(D)	58 ± 11(D)	41.6(D)	0.79(D)
					3	3,204(EV)	35(EV)	62 ± 13(EV)	42.1(EV)	0.76(EV)
Ikeda F (2016)	Japan	Cohort study	1988-2008	COX	1a	2,446	123	58.3 ± 11.4	41.54	0.77(D)
In H (2020)	America	Case-control	NR	Logistic	1a	140	40	40-79	24(cases)	0.94(D)
Lee DS (2009)	Korea	Case-control	2005.3-8	Logistic	1b	382	183	>35	50(non-cases)	0.90(IV)
Qiu L (2020)	Chinese	Case-control	2009-2011	Logistic	1b	2,287	1,115	≤59: 569 (51%) (cases) ≤59: 593 (51%) (non-cases)	71.1(cases) 70.1(non-cases)	0.684(IV)
So JBY (2021)	Singapore and Korea	Nested case-control study	NR	Cox	3	472(D) 210(EV)	236(D) 94(EV)	61.2 ± 8.4(D)(cases) 68.0 ± 10.9(D)(non-cases)	63.3(D)(cases) 62.7(D)(cases)	0.93(D) 0.92(EV) 0.85(0.81-0.88) (EV)
Taninaga J (2019)	Japan	Case-control study	2006-2017	machine learning	2a	1,144(D) 287(EV)	89	56.7 ± 8.8(cases) 46.2 ± 1.0(non-cases)	78.06	0.87(IV) 0.90(EV)
Tu H(2017)	China	Cross-sectional study	1997-2012	Logistic	1a	9,002	94	61.2 ± 11.4(cases) 50.7 ± 10.1(non-cases)	72.3(cases) 47(non-cases)	0.80(D)

Note: 1a, development only; 1b, development and validation using resampling; 2a, random split-sample development and validation; 2b, nonrandom split-sample development and validation; 3, development and validation using separate data; 4, validation study.

Abbreviations: D, derivation; EV, external validation; IV, internal validation; NR, Not reported.

^aType of study according to the TRIPOD guidelines.

analysis. These domains further contain 20 signaling questions to facilitate structured judgment of risk of bias (4).

Data availability statement

Data sharing is not applicable to this article as no data were created or analyzed in this study.

Results

Of 774 total citations (205 in PubMed, 381 in Embase, 188 in Cochrane Library databases), 12 full-text articles met the criteria and were included in the final analysis (Fig. 1; refs. 14–25). Among this, one developed two sex-stratified models (17), one was a validation study (15, 16), thus a total of 12 models were identified on risk prediction for gastric cancer.

Description of studies

The primary characteristics of included studies were summarized in Table 1. Eleven studies originated from East Asia [five from Japan (15, 16, 18, 19, 21), three from China (14, 23, 25), two from Korea (17, 22), one from Singapore (24)], while only one study was from America (20). Among this, six were cohort studies (including a nested case-control study; refs. 15–19, 24), four were case-control studies (20–23), and the remaining were cross-sectional studies (14, 25). The sample sizes ranged from 140 to 2,176,501. Among 12 models, six were diagnostic prediction models using Logistic regression (14, 20, 22, 23, 25) and machine learning (21), and six were prognostic prediction models using Cox regression (15–19, 24).

Model characteristics

The number of events per predictor variable (EPV) was defined as the number of gastric cancer events divided by the number of predictor variables. Less than half of the models (41.7%) had EPVs greater than 20 (14, 15, 17, 23), six were 10–20 (18, 19, 21, 22, 24, 25), and only one model had EPVs less than 10 (20). Except for one model that did not state how to handle continuous variables (21), the continuous variables were only retained in two models (16.7%; refs. 23, 24) and converted to a categorical form in nine models (75.0%; refs. 14–20, 22, 25). Six models reported the handling of missing values in the derivation dataset (14, 17, 20, 22, 23), while the missing values were not discussed in the rest. Among these, three models removed records containing missing values (14, 20, 23), and three models imputed missing data (17, 23). Seven models presented complete information, of which four used a complete regression equation (15, 17, 22) and three transformed into a risk scoring system (Table 2; refs. 14, 18, 25).

Model development and validation

All models evaluated the discriminative ability by derivation *c*-statistics, and half of the models assessed calibration ability with the Hosmer–Lemeshow test (14, 15, 17, 18, 22). As shown

in Fig. 2, the AUCs ranged from 0.73 to 0.93 in the derivation sets ($n = 6$), 0.68 to 0.90 in the internal validation sets ($n = 5$), 0.71 to 0.92 in the external validation sets ($n = 7$). Five of the seven external validation models were developed and validated using separate data (15, 17, 18, 24) while the remaining two randomly divided the whole dataset into a training set and test set (14, 21). Four models performed internal validation by bootstrapping (14, 15, 22, 23) and one model performed internal validation by cross-validation (21). Only one model developed in a Chinese cross-sectional study reported the AUC in both the derivation cohort (AUC = 0.76), internal validation cohort (AUC = 0.76) and external validation cohort (AUC = 0.73; ref. 14).

Predictors in the models

Across all models, there were 30 distinct predictors, which could be further divided into four categories: demographic, clinical, lifestyle, and laboratory variables (Fig. 3). Thirteen predictors were presented two or more times, including three demographic variables [age, sex, body mass index (BMI)], five lifestyle variables (salt preference, smoking, alcohol, physical activity, meal regularity), four laboratory variables (HP, PG, hemoglobin A1c level, gastrin-17), and one clinical variable

Table 2. Summary of model characteristics predicting gastric cancer.

Characteristics of analysis	Number of models (%)
EPV in final model	
<10	1 (8.3)
10–20	5 (41.7)
≥21	6 (50)
Handling of continuous variables	
Transforming to multi-categorical variables	9 (75.0)
Keep continuous variables	2 (16.7)
NR	1 (8.3)
Variable selection	
Clinical experience and statistical analysis	9 (75.0)
NR	3 (25.0)
Missing values	
Replaced using imputation	3 (25)
Complete case analysis	2 (16.7)
NR	7 (58.3)
Model performance	
Discrimination	12 (100.0)
Calibration (Hosmer–Lemeshow test)	6 (50.0)
Validation	
Internal	2 (16.7)
External	4 (33.3)
Internal and external	3 (25.0)
NR	3 (25.0)
Internal validation	
Cross-validation	1 (8.3)
Bootstrap	4 (33.3)
Model presentation	
Using complete regression equation	4 (33.3)
Transformed into risk scoring system	3 (25.0)
Complete model not reported	5 (41.7)

Abbreviations: EPV, number of events per predictor variable; NR, not reported.

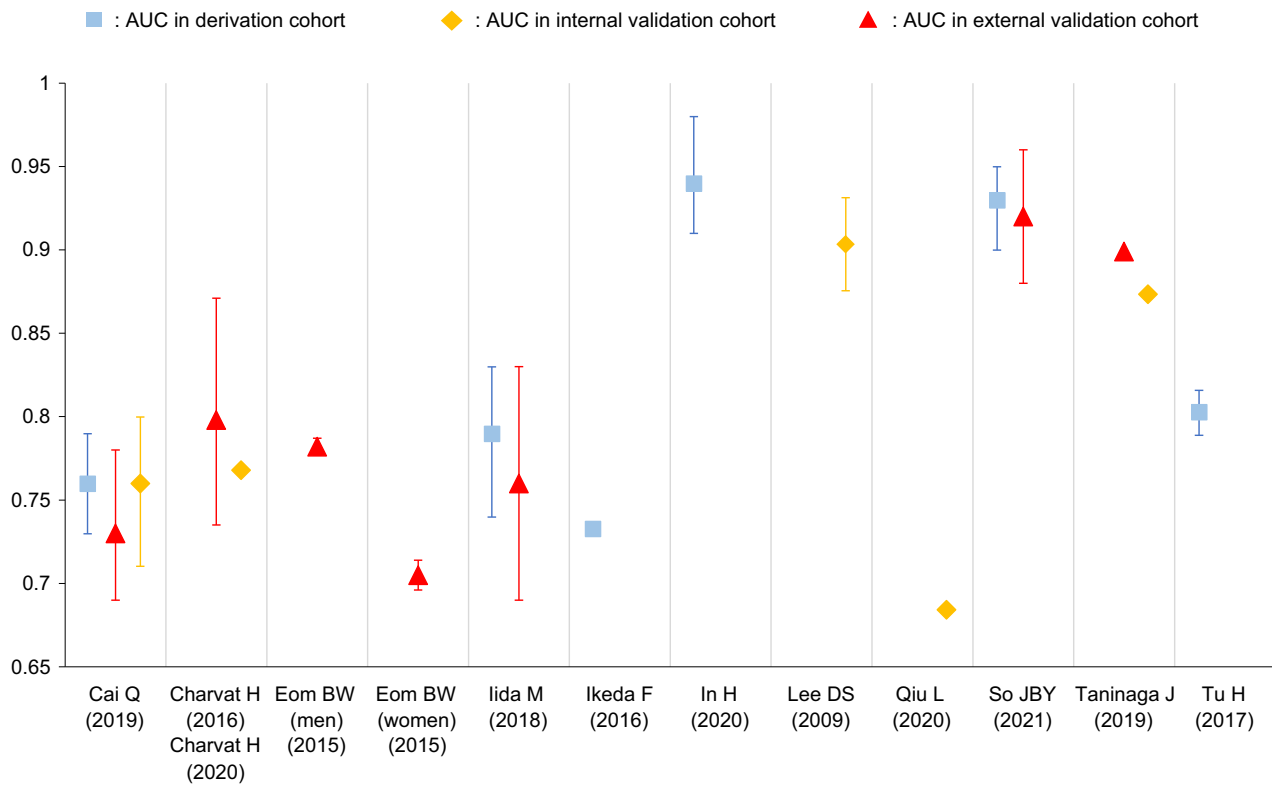


Figure 2.

The AUC of risk prediction models in the derivation cohort, internal validation cohort, and external validation cohort. Note: Charvat H (2020) was a validation study that can be seen as complementary to Charvat H (2016). Hence, we merged the two studies to analyze them. Eom BW (2015) developed two sex-stratified models, so we analyzed two models separately. AUC, area under the receiver operating characteristic curve.

(family history). The details of predictors for each study across the above categories are presented in Supplementary Table S4. In terms of individual variables, age was the most frequently selected variable which appeared nine times. The second-ranking variables, a high-salt diet and HP, appeared six times. Aside from this, the usage frequency did not exceed 50% for any other predictors. The type of variables was strongly dependent on the purpose and design of the study.

Study quality and ROB

Table 3 and Fig. 4 summarize the ROB for all included studies. Overall, there were two modes at a high ROB for participants due to the unreasonable inclusion criteria (21, 22), one at a high ROB for predictors (22), and one at a high ROB for outcomes (22). In the domain of analysis concerns, all responses were “high risk” because the processing of missing values or overfitting was unsatisfactory in most studies. The details of the ROB for each study across the above domains are presented in Supplementary Fig. S1. Two models were assessed as high concern regarding applicability because one included hospital-based cases in a case-control study (22) and the other did not have a clear pathologic diagnosis for some a small selection of patients who underwent endoscopic examination (21).

Discussion

In this systematic review, we synthesized and evaluated the prediction performance, predictors, and ROB in 12 risk prediction models of gastric cancer developed at the population level. Although most risk prediction models presented a high discriminative ability, less than half of the models performed external validation in diverse datasets. The risk prediction models with high discriminative ability tended to include age, salt preference, HP, smoking, BMI, family history, pepsinogen, and sex. At least one domain with a high ROB was present in all studies mainly due to methodologic limitations in the analysis domain, which resulted in an overall high ROB for each model assessed by PROBAST.

The models warranting promotion in clinical practice or primary care should have good discriminative and generalization capability. When the same dataset is used for both development and validation, performance is usually overestimated due to overfitting issues. Therefore, the generalization capability should be evaluated with a testing dataset independent of model construction (26). In addition, internal validation, which provides a more objective discriminative capacity and stability is also the crucial step for model evaluation (4). In our review, only five of the seven external validation models were developed and validated using separate data while the

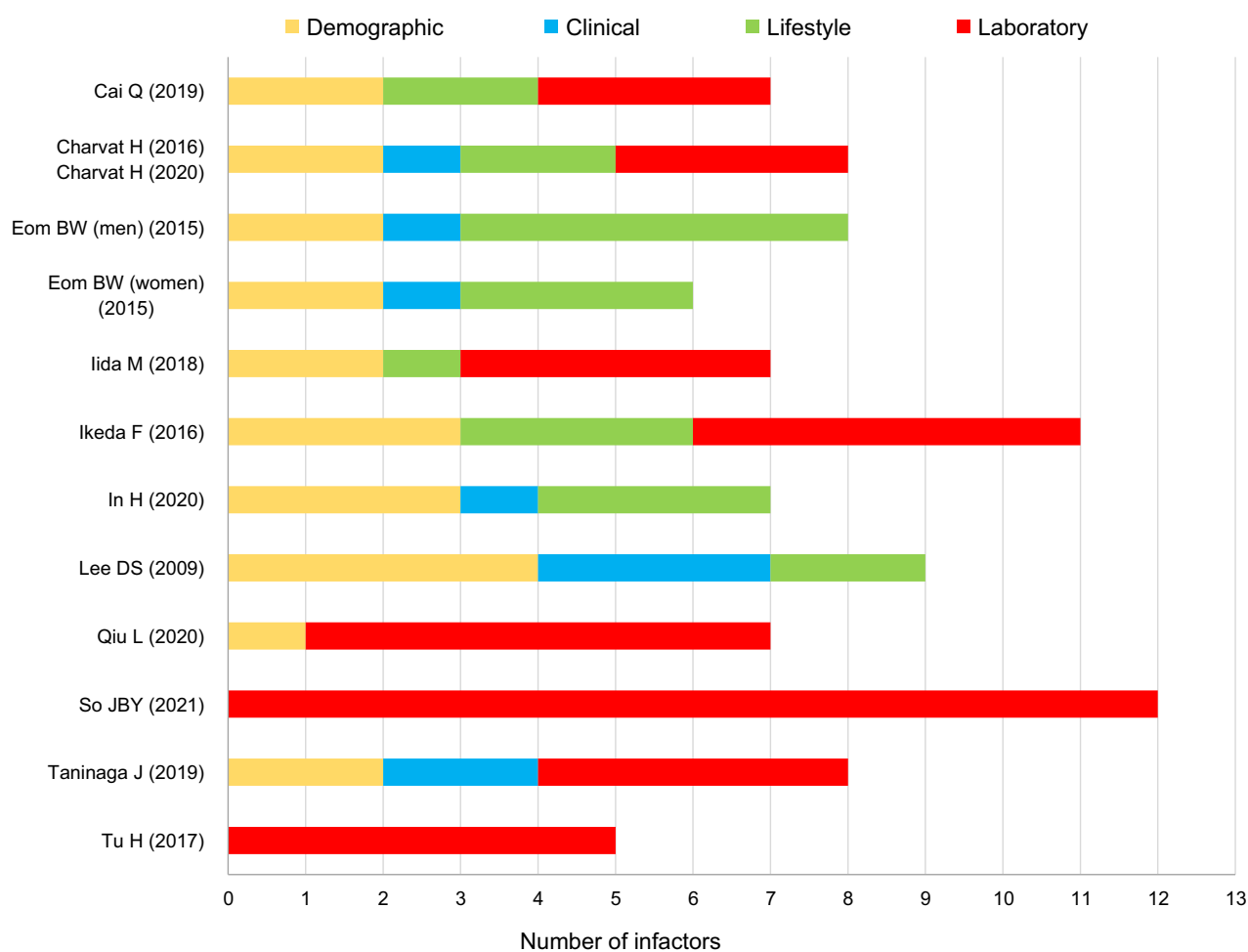


Figure 3. Classification of predictors identified in risk prediction models for gastric cancer. Note: Charvat H (2020) was a validation study that can be seen as complementary to Charvat H (2016). Hence, we merged the two studies to analyze them. Eom BW (2015) developed two sex-stratified models, so we analyzed two models separately.

Table 3. Risk of bias and applicability of included studies using PROBAST.

	ROB				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of bias	Applicability
Cai Q (2019)	+	+	+	-	+	+	+	-	+
Charvat H (2016)	+	+	+	-	+	+	+	-	+
Charvat H (2020) ^a									
Eom BW (2015)	+	+	+	-	+	+	+	-	+
Iida M (2018)	+	+	+	-	+	+	+	-	-
Ikeda F (2016)	+	+	+	-	+	+	+	-	+
In H (2020)	+	+	+	-	+	+	+	-	+
Lee DS (2009)	-	-	-	-	-	+	+	-	-
Qiu L (2020)	+	+	+	-	+	+	+	-	+
So JBY (2021)	+	+	+	-	+	+	+	-	+
Taninaga J (2019)	-	+	+	-	-	+	+	-	-
Tu H (2017)	+	+	+	-	+	+	+	-	+

Abbreviation: ROB, risk of bias.

^aCharvat H (2020) was a validation study that can be seen as complementary to Charvat H (2016). Hence, we merged the two studies to analyze them. + indicates low ROB/low concern regarding applicability; - indicates high ROB/high concern regarding applicability.

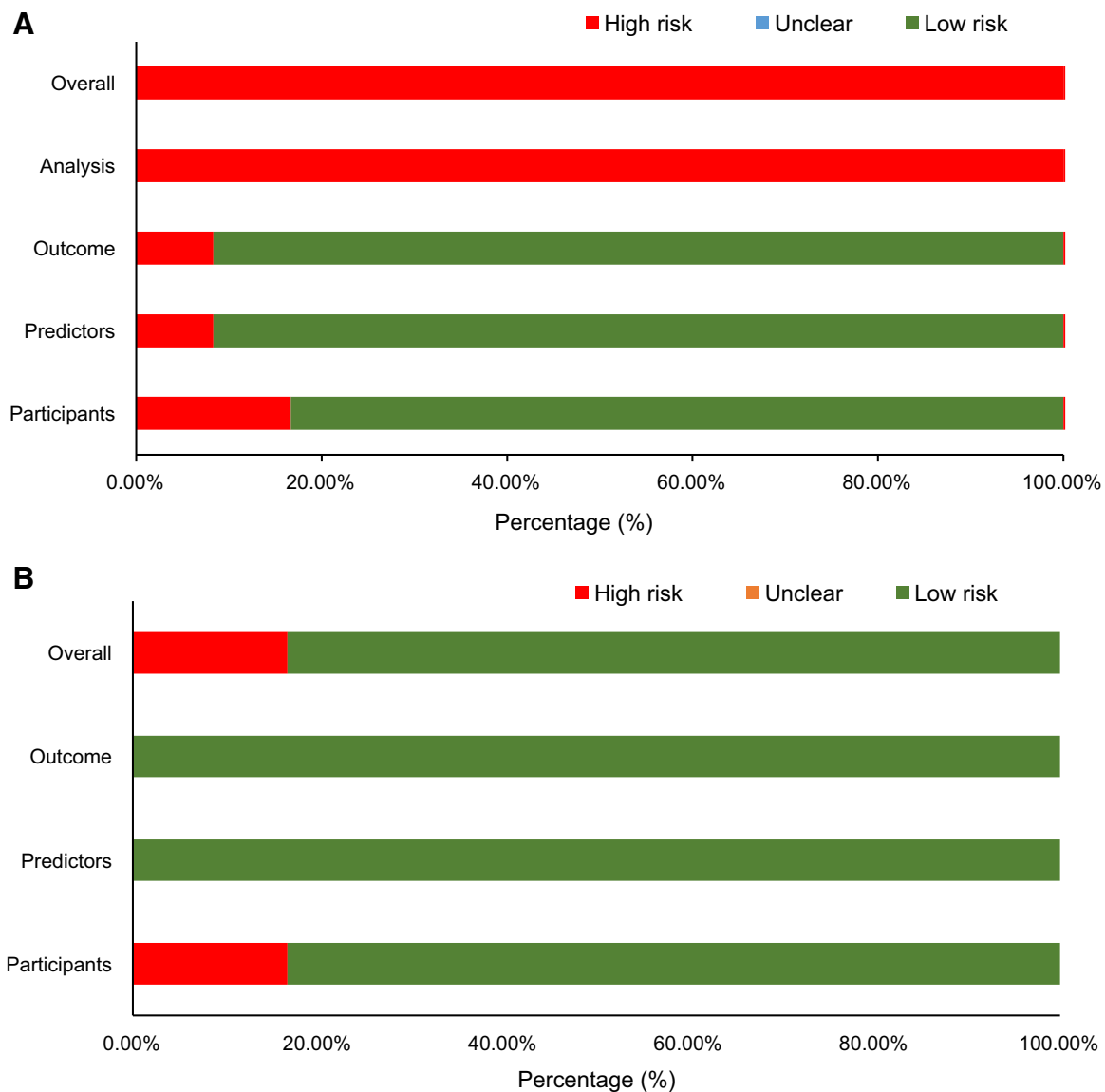


Figure 4.

A, Risk of bias assessment according to the PROBAST. **B,** Applicability assessment according to the PROBAST.

remaining two randomly divided the whole dataset into a training set and testing set. Although splitting the dataset is a reliable approach for external validation, a better approach would be to use one or more independent datasets. In addition, only few studies performed internal validation, and fewer studies considered both internal and external validation. Similar concerns were identified from other modeling studies performed on breast cancer (9) or lung cancer (27). In addition, only a limited number of studies provided complete risk equations, whereas the remaining studies were unlikely to be independently externally validated since complete risk equations are required for risk assessment. Therefore, further external validations of existing models using independent datasets are required to guarantee their generalizability.

A suitable prediction model should consist of plausible variables that have been epidemiologically and clinically associated with gastric cancer. There is substantial evidence that the incidence of gastric cancer in the general population increases with age and is higher in males, especially for those over 40 years old (28, 29). HP infection is the strongest known risk factor for the development of gastric cancer (30), and a high-salt diet (31), familial history of gastric cancer (32), and smoking (28) are also associated with an increased risk of gastric cancer. The above factors, which provide a theoretical basis for model construction, are the most common variables in gastric cancer models. In addition, the type of variables was strongly dependent on the purpose and design of the study. Except for three exploratory

studies that used single-type variables such as genotype variables or miRNA, the other models strive to incorporate multivariate factors to explore high-risk groups. The models developed by In and colleagues (20) and Lee and colleagues (22) had the highest prediction performance in derivation and internal validation, respectively. However, their studies were not subjected to external validation, and the predictors were dominated by lifestyle and demographic factors only assessed by questionnaires, and a lack of specific diagnostic indicators such as HP infection or serum pepsinogen. Atrophic gastritis caused by HP could affect the secretion of PGs and gastrin, which serve as promising biomarkers for predicting the status of the gastric mucosa (33). Asian countries with a high incidence of gastric cancer, such as China and Japan, have implemented screening programs using a combination of the presence of HP antibody and serum PG concentration (34, 35). Although most of these serological markers are relatively difficult to obtain in primary and community care settings, they account for higher specificity and result in improved generalization performances (25).

Our risk of bias assessment showed that all 12 gastric cancer models scored a high ROB in the analysis domain of the PROBAST tool because poor methods were often employed and many were inadequately reported. In addition to the limitations of the absence of effective external verification, another issue worthy of mention is the handling of missing values. Six models reported the handling of missing values in the derivation dataset. Of these, three used multiple imputations (17, 23) and three removed missing values (14, 20, 23). Deletion of outliers decreases the discriminatory power of multivariable models compared with values replaced with multiple imputations also leads to a biased effect estimation, making the sample less representative (36). Furthermore, the EPV has been studied in relation to model performance, and a model with an EPV higher than 20 is less likely to have overfitting (37). Our review showed that only six of 12 models met the criteria, implying that the remaining models were likely to have been overfitted. For predictors with low prevalence presented in a model, such as occupational exposures and lymphocyte ratio, a higher EPV is needed to eliminate bias in regression coefficients and improve predictive accuracy.

Strengths and limitations

The main strength of this review was that it provided a comprehensive mapping of the available studies on risk prediction models for gastric cancer in the general population. We used robust methodology to synthesize the evidence by following the TRIPOD guidelines and thus provided a detailed description of the features of each gastric cancer model. In addition, the ROB of the included studies was assessed according to PROBAST, which is a new tool for assessing the methodologic quality of risk prediction models, to provide a

better understanding of the overall quality of the current gastric cancer models.

Several limitations exist in the present systematic review. First, we were unable to perform subgroup analysis by anatomic site since almost all of the included studies did not provide typing information. Given that the etiology is different between gastric cardia and noncardia cancer (38), it would be inappropriate to predict a high-risk population of the two subgroups by the same model. Second, due to the lack of evaluation metrics, such as 95% confidence intervals, we did not perform a meta-analysis for prediction performance in the model development and validation. Third, we only included full-text articles published in English, which may have led to a potential publication bias.

Conclusion

In conclusion, although some risk prediction models including similar predictors have displayed sufficient discriminative abilities, many have a high ROB due to methodologic limitations and are not externally validated efficiently. Up-to-date guidelines, such as PROBAST and TRIPOD, have proposed a set of practical guidelines for the normalization process of model construction and validation. For a more reliable risk prediction model for gastric cancer based on the general population, future studies should take inspiration from previous studies and adhere to well-established methodologic standards and guidelines specifically developed for the prediction model.

Authors' Disclosures

No disclosures were reported.

Authors' Contributions

J. Gu: Writing—original draft, writing—review and editing. **R. Chen:** Funding acquisition, writing—original draft, writing—review and editing. **S.-M. Wang:** Writing—review and editing. **M. Li:** Data curation, investigation. **Z. Fan:** Investigation. **X. Li:** Supervision. **J. Zhou:** Investigation. **K. Sun:** Supervision, funding acquisition. **W. Wei:** Conceptualization, supervision.

Acknowledgments

We thank all authors for their contributions to this study. This study is supported by the Beijing Natural Science Foundation <https://doi.org/10.13039/501100005089> (7204294, to R. Chen), National Natural Science Foundation of China <https://doi.org/10.13039/501100001809> (81903403, to R. Chen, 81974493, to W. Wei), and the National Key R&D Program of China (2016YFC0901404, to W. Wei, 2018YFC1311704, to K. Sun).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 21, 2021; revised November 15, 2021; accepted January 7, 2022; published first January 10, 2022.

References

1. 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;71:209–49.
2. Zhang X, Li M, Chen S, Hu J, Guo Q, Liu R, et al. Endoscopic screening in asian countries is associated with reduced gastric cancer mortality: a meta-analysis and systematic review. *Gastroenterology* 2018;155:347–54.
3. Chen R, Liu Y, Song G, Li B, Zhao D, Hua Z, et al. Effectiveness of one-time endoscopic screening programme in prevention of upper gastrointestinal cancer in China: a multicentre population-based cohort study. *Gut* 2021;70:251–60.
4. Moons KGM, Wolff RF, 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–W33.
5. Weck MN, Gao L, Brenner H. Helicobacter pylori infection and chronic atrophic gastritis: associations according to severity of disease. *Epidemiology* 2009;20:569–74.
6. Samloff IM, Varis K, Ihamaki T, Siurala M, Rotter JI. Relationships among serum pepsinogen I, serum pepsinogen II, and gastric mucosal histology. A study in relatives of patients with pernicious anemia. *Gastroenterology* 1982;83:204–9.
7. Palazón-Bru A, Mares-García E, López-Bru D, Mares-Arambul E, Folgado-de la Rosa DM, Carbonell-Torregrosa MLÁ, 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.
8. Li H, Sun D, Cao M, He S, Zheng Y, Yu X, et al. Risk prediction models for esophageal cancer: a systematic review and critical appraisal. *Cancer Med* 2021;10:7265–76.
9. Anothaisintawee T, Teerawattananon Y, Wiratkapun C, Kasamesup V, Thakkinstian A. Risk prediction models of breast cancer: a systematic review of model performances. *Breast Cancer Res Treat* 2012; 133:1–10.
10. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
11. Tian P, Liu Y, Bian S, Li M, Zhang M, Liu J, et al. Laparoscopic proximal gastrectomy versus laparoscopic total gastrectomy for proximal gastric cancer: a systematic review and meta-analysis. *Front Oncol* 2021;10:607922.
12. Sahle BW, Owen AJ, Chin KL, Reid CM. Risk prediction models for incident heart failure: a systematic review of methodology and model performance. *J Card Fail* 2017;23:680–7.
13. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.
14. Cai Q, Zhu C, Yuan Y, Feng Q, Feng Y, Hao Y, et al. Development and validation of a prediction rule for estimating gastric cancer risk in the Chinese high-risk population: a nationwide multicentre study. *Gut* 2019;68:1576–87.
15. Charvat H, Sasazuki S, Inoue M, Iwasaki M, Sawada N, Shimazu T, et al. Prediction of the 10-year probability of gastric cancer occurrence in the Japanese population: the JPHC study cohort II. *Int J Cancer* 2016;138:320–31.
16. Charvat H, Shimazu T, Inoue M, Iwasaki M, Sawada N, Yamaji T, et al. Estimation of the performance of a risk prediction model for gastric cancer occurrence in Japan: Evidence from a small external population. *Cancer Epidemiol* 2020;67:101766.
17. Eom BW, Joo J, Kim S, Shin A, Yang HR, Park J, et al. Prediction model for gastric cancer incidence in korean population. *PLoS One* 2015;10: e0132613.
18. Iida M, Ikeda F, Hata J, Hirakawa Y, Ohara T, Mukai N, et al. Development and validation of a risk assessment tool for gastric cancer in a general Japanese population. *Gastric Cancer* 2018;21: 383–90.
19. Ikeda F, Shikata K, Hata J, Fukuhara M, Hirakawa Y, Ohara T, et al. Combination of helicobacter pylori antibody and serum pepsinogen as a good predictive tool of gastric cancer incidence: 20-year prospective data from the hisayama study. *J Epidemiol* 2016;26: 629–36.
20. In H, Solsky I, Castle PE, Schechter CB, Parides M, Friedmann P, et al. Utilizing cultural and ethnic variables in screening models to identify individuals at high risk for gastric cancer: a pilot study. *Cancer Prev Res* 2020;13:687–98.
21. Taninaga J, Nishiyama Y, Fujibayashi K, Gunji T, Sasabe N, Iijima K, et al. Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: a case-control study. *Sci Rep* 2019;9:12384.
22. Lee DS, Yang HK, Kim JW, Yook JW, Jeon SH, Kang SH, et al. Identifying the risk factors through the development of a predictive model for gastric cancer in South Korea. *Cancer Nurs* 2009;32:135–42.
23. Qiu L, Qu X, He J, Cheng L, Zhang R, Sun M, et al. Predictive model for risk of gastric cancer using genetic variants from genome-wide association studies and high-evidence meta-analysis. *Cancer Med* 2020;9:7310–6.
24. So JBY, Kapoor R, Zhu F, Koh C, Zhou L, Zou R, et al. Development and validation of a serum microRNA biomarker panel for detecting gastric cancer in a high-risk population. *Gut* 2021;70:829–37.
25. Tu H, Sun L, Dong X, Gong Y, Xu Q, Jing J, et al. A Serological biopsy using five stomach-specific circulating biomarkers for gastric cancer risk assessment: a multi-phase study. *Am J Gastroenterol* 2017;112: 704–15.
26. 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.
27. Gray EP, Teare MD, Stevens J, Archer R. Risk Prediction models for lung cancer: a systematic review. *Clin Lung Cancer* 2016;17:95–106.
28. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kaman-gar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. *Cancer Epidemiol Biomarkers Prev* 2014;23:700–13.
29. Park JY, von Karsa L, Herrero R. Prevention strategies for gastric cancer: a global perspective. *Clin Endosc* 2014;47:478–89.
30. Hooi JKY, Lai WY, Ng WK, Suen MMY, Underwood FE, Tanyingoh D, et al. Global prevalence of helicobacter pylori infection: systematic review and meta-analysis. *Gastroenterology* 2017;153:420–9.
31. Fang X, Wei J, He X, An P, Wang H, Jiang L, et al. Landscape of dietary factors associated with risk of gastric cancer: a systematic review and dose-response meta-analysis of prospective cohort studies. *Eur J Cancer* 2015;51:2820–32.
32. Choi IJ, Kim CG, Lee JY, Kim YI, Kook MC, Park B, et al. Family history of gastric cancer and *Helicobacter pylori* treatment. *N Engl J Med* 2020;382:427–36.
33. Yamaguchi Y, Nagata Y, Hiratsuka R, Kawase Y, Tominaga T, Takeuchi S, et al. Gastric cancer screening by combined assay for

- serum anti-helicobacter pylori IgG antibody and serum pepsinogen levels—the ABC method. *Digestion* 2016;93:13–8.
34. Chinese Society of Digestive Endoscopy. Consensus on screening and endoscopic diagnosis and treatment of early gastric cancer in China (Changsha, 2014). *Chinese Journal of Digestive Endoscopy*. 2014;31: 361–77.
 35. Hamashima C; Systematic Review Group and Guideline Development Group for Gastric Cancer Screening Guidelines. Update version of the Japanese Guidelines for Gastric Cancer Screening. *Jpn J Clin Oncol* 2018;48:673–83.
 36. Janssen KJ, Donders AR, Harrell FE Jr, Vergouwe Y, Chen Q, Grobbee DE, et al. Missing covariate data in medical research: to impute is better than to ignore. *J Clin Epidemiol* 2010;63: 721–7.
 37. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med* 2004;66:411–21.
 38. Balakrishnan M, George R, Sharma A, Graham DY. Changing trends in stomach cancer throughout the world. *Curr Gastroenterol Rep* 2017;19:36.