



OPEN Aspiration risk prediction models in patients with nasogastric enteral nutrition: a systematic review and meta-analysis

Juan Chen^{1,4}, Yushuang Su^{1,4}, Tianlong Li¹✉, Xiaorong Mao¹, Qinghua Jiang², Qin Yang¹, Qing Wen¹, Zaichun Pu³ & Mengting Liu¹

In recent years, numerous researchers have developed risk prediction models for aspiration in patients with nasogastric enteral nutrition (EN). Nevertheless, comprehensive and systematic comparative studies are lacking. This study systematically review and evaluate the studies on aspiration risk prediction models in patients with nasogastric EN. A computer search was conducted from the database establishment to May 10, 2025. The Prediction Model Risk of Bias Assessment Tool (PROBAST) evaluation tool was used to assess the quality of the included studies, and the meta-analysis was conducted using Stata 17 software to analyze the prediction factors included in the models and the area under the curve (AUC) values of the validated models. Eleven studies were included, with a total of 22 aspiration risk prediction models for patients with nasogastric EN. The AUC ranged from 0.809 to 0.992. The PROBAST evaluation results showed that all 11 included studies had a high risk of bias. The most common predictive factors included the number of diseases, history of aspiration, use of sedative, depth of tube placement, amount of gastric residue, APACHE II score, consciousness disturbance, nutritional risk, age. The pooled AUC value of the four validated models was 0.92 (95% confidence interval: 0.90–0.93), indicating an excellent level of discrimination. The study protocol has been registered with PROSPERO (registration number: CRD42024594672).

Keywords Nasogastric, Enteral nutrition, Respiratory aspiration, Prediction models, Systematic review, Meta-analysis

Enteral nutrition (EN) represents a nutritional support therapy that supplies nutrients via the gastrointestinal tract. The guidelines^{1,2} explicitly indicate that EN is the preferred nutritional intervention approach for those presenting nutritional risks and/or malnutrition, provided that the gastrointestinal tract is functional and can be utilized safely. EN can be delivered either orally or through tubes, such as nasogastric tubes, nasoenteric tubes, gastric or jejunal stomas, directly into the patient's gastrointestinal tract to fulfill their nutritional requirements. Aspiration refers to the entry of food (or non-food) substances, oral secretions, gastroesophageal reflux contents, etc. into the airways below the glottis³. Research has revealed that the incidence of aspiration resulting from nasogastric feeding in patients ranges from 15 to 50%^{4,5}. Aspiration pneumonia is the most severe complication of aspiration⁶, with a mortality rate as high as 58.33%, an average hospital stay extension of 9.43 days, and an average increase in hospital costs of \$9,100.29⁷. Thus, the adverse consequences caused by aspiration have increasingly become a crucial issue influencing the feeding safety of hospitalized patients undergoing nasogastric feeding. Predictive models are founded on mathematical models and by integrating multiple predictive factors, they estimate the probability of an event outcome⁸. In recent years, numerous researchers have developed risk prediction models for aspiration in patients with nasogastric EN. Nevertheless, these models lack comprehensive and systematic comparative studies in aspects such as the construction process, performance evaluation, and data sample bias. It remains unclear whether these predictive models can be accurately and reliably applied in clinical practice. Therefore, this study conducts a systematic review of the risk prediction models for aspiration

¹Sichuan Provincial People's Hospital, School of Medicine, University of Electronic Science and Technology of China, No. 32, West Second Section, 1st Ring Road, Qingyang District, Chengdu 610072, Sichuan Province, China. ²Sichuan Cancer Hospital and Institute, Affiliated Cancer Hospital of University of Electronic Science and Technology of China, Chengdu, China. ³Xindu District People's Hospital of Chengdu, Chengdu, China. ⁴Juan Chen and Yushuang Su equally contributed to this work. ✉email: ajay1989@126.com

in patients with nasogastric EN, with the aim of providing guidance for clinical medical staff in screening aspiration-related risk prediction models, and promoting the extensive application of predictive models in the risk management of aspiration during nasogastric EN, enhancing the scientific and accuracy of clinical decision-making and providing safer and more effective medical services for patients.

Methods

The study protocol has been registered with PROSPERO (registration number: CRD42024594672).

Inclusion and exclusion criteria

Inclusion criteria

(1) The study subjects were patients aged ≥ 18 years and receiving EN support through nasogastric tubes, including nasogastric or nasoenteric tube routes. (2) The study content was the development and/or validation of the aspiration risk prediction model for patients with nasogastric EN. (3) The study design was a prospective or retrospective study. (4) The outcome indicator was respiratory aspiration after EN support through nasogastric tubes. The diagnosis of aspiration referred to the diagnostic criteria for aspiration in the expert consensus of the North American Association⁹. The specific criteria included: ① Coughing, wheezing, accelerated breathing rate, cyanosis of the lips, and discovery of residuals of EN liquid in the oral or nasal cavity occurred in patients during nasogastric feeding; ② Suspected EN liquid was aspirated during sputum suction, and the value measured by a blood glucose meter exceeded 11.1 mmol/L; ③ Gastric pepsin was detected through lung CT examination or determination of bronchial secretions. (5) The language was Chinese or English.

Exclusion criteria

(1) duplicate publications; (2) case reports, review articles, conference abstracts, etc.; (3) articles without access to the original text.

Search strategy

We conducted a comprehensive search of both Chinese and English databases, including the China National Knowledge Infrastructure (CNKI), Full Text Database of Chinese Medical Journals (Yiigle), SinoMed, Wanfang, PubMed, Embase, Medline, Web of Science, from the database's establishment to May 10, 2025. We used a combination of subject terms and free words for the search. The search terms used were "enteral nutrition", "gut nutrition", "intestinal nutrition", "nasal feeding", "nasogastric feeding", "respiratory aspiration", "aspiration", "risk factor", "forecast", "predict", "predict model", "risk prediction model", "forecast model", "nomogram", "alignment diagram", "nomographic chart", and "alignment chart". The specific search strategy is provided in Supplementary Material (Appendix A1). Additionally, we obtained more relevant studies through the references of related studies.

For the systematic review, we adopted the PICOTS system, which is recommended by the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist¹⁰ and is conducive to formulating the purpose, search strategy, and inclusion and exclusion criteria of the review¹¹. The key items of our systematic review are as follows:

P (Population): Hospitalized patients aged ≥ 18 years receiving EN support through nasogastric tube or nasointestinal tube.

I (Intervention model): Aspiration risk prediction model for patients with nasogastric EN that were developed and published.

C (Comparator): No competing model.

O (Outcome): The outcome focused on respiratory aspiration.

T (Timing): The outcome was predicted after evaluating basic information, scoring scale results, and laboratory indicators.

S (Setting): The intended use of the risk prediction model is to predict the respiratory aspiration in patients with nasogastric EN, facilitating the implementation of preventive measures to prevent adverse events.

Study screening

All the retrieved literature was initially subjected to duplicate removal through EndNote 20 reference management software. Subsequently, an initial screening was conducted by reading the titles and abstracts of the literature in accordance with the inclusion and exclusion criteria. A secondary screening was then carried out by reading the full text and evaluating the methodological section. The entire literature screening process was independently performed by two researchers. In the event of a disagreement between the two, a third researcher would adjudicate. If the information of the included literature was incomplete or questionable, attempts were made to contact the original authors to obtain relevant information before determining whether to include it.

Data extraction

A data collection form was developed based on the prediction model systematic review data extraction toolkit (CHARMS) proposed by Moons et al.¹⁰. The form includes items such as the first author, publication year, country, study design, participants, incidence of aspiration, diagnostic criteria for aspiration, model development method, variable selection, number of models, sample size, missing data handling, continuous variable processing method, predictor variables, model performance, calibration method, validation method, and model presentation. Data extraction was conducted by one researcher, and then checked by another researcher to ensure accuracy and consistency.

Quality evaluation

The Prediction Model Risk of Bias Assessment Tool (PROBAST) was employed to assess the risk of bias and applicability of the literature for studies involving diagnostic or prognostic multivariable prediction models¹². The risk of bias of the model was evaluated in four domains, and the applicability was evaluated in three domains. The assessment method for applicability was similar to that for the risk of bias. The evaluation results of each domain were classified into three grades: "low," "high," and "unclear." The risk of bias and applicability of the included literature were independently evaluated by two researchers. In the event of a disagreement, a third researcher was sought for adjudication. (1) Bias assessment: The bias assessment was carried out by evaluating four domains: the study population, predictor variables, outcomes, and data analysis. It encompassed a total of 20 questions, with 2 to 9 questions in each domain. Each question was answered as "yes," "probably yes," "no," "probably no," or "no information." If any domain was answered with "no" or "probably no," it was considered that the risk of bias in that domain was high. If all the questions in each domain were answered with "yes" or "probably yes," the risk of bias was low. (2) Applicability assessment: The applicability assessment included three domains: the study population, predictor variables, and outcomes. Each domain was evaluated as "good applicability," "poor applicability," or "unclear." If the applicability in any domain was poor, the overall applicability was considered low. Based on the above characteristics, the best model was selected. PROBAST can not only evaluate individual models but also compare multiple models.

Statistical analysis

Meta-analysis was performed using Stata 17 software on the predictive factors included in the model and the area under the curve (AUC) values of the validated models. The Q test and I^2 were employed to evaluate the heterogeneity of the studies. With $P > 0.1$ or $I^2 < 50\%$ regarded that the heterogeneity among the studies was small, and a fixed-effect model was chosen. $P \leq 0.1$ or $I^2 \geq 50\%$, it was considered that the heterogeneity among the studies was considerable, and sensitivity analysis was conducted. If the heterogeneity could not be eliminated, a random-effect model was selected. With $P < 0.05$ was considered statistically significant. Egger's test was used to identify publication bias, with $p > 0.05$ indicating a low likelihood of publication bias.

Results

Study selection

Figure 1 presents the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 flowchart, which depicts the search process and results. We retrieved 559 relevant articles from relevant databases through systematic search, and after deduplication using EndNote 20, there were 503 remaining articles. Based on the inclusion and exclusion criteria for literature, we read the titles and abstracts of the articles and retained 29 of them. We then read the full texts of the articles and excluded 11 that did not meet the research objectives, 7 that did not build a model, and finally included 11 articles^{13–23} for a total of 22 predictive models.

Study characteristics

This study included literature published from 2021 to 2024, all of which were from China (2 articles were published in English), and all had a retrospective study design. The study found that the incidence of aspiration in patients with nasogastric EN was 9.46% to 49.87%. The basic characteristics of the included literature are presented in Table 1.

Models characteristics

Eleven studies^{13–23} constructed a total of 22 aspiration risk prediction models. In model construction, one study¹⁶ used machine learning and generalized linear regression methods to develop separate prediction models, three studies^{17,19,23} used machine learning and Logistic regression methods to establish separate prediction models, and the remaining studies only used Logistic regression methods to establish prediction models. In terms of the handling of continuous variables, five studies^{13,16–19} maintained the continuity of continuous variables, while six studies^{14,15,20–23} converted continuous variables into categorical variables. In terms of variable selection, two studies^{16,23} used machine learning to select variables, while the remaining studies used univariate analysis to select variables; Table 2 shows the detailed information. The AUC of the prediction models included in this study ranged from 0.809 to 0.992, the sensitivity ranged from 64.9 to 95.2%, the specificity ranged from 75.8 to 97.4%, and the predictive performance of all prediction models was good (AUC > 0.8). Nine studies^{13–16,18–22} reported discrimination and calibration; In terms of model validation, five studies^{14,18,19,21,23} conducted external validation, ten studies^{13–21,23} conducted internal validation, and the main method used was Bootstrap. Among them, four studies^{14,17,18,21} did not report the internal validation method, and one study²² only developed the model but did not validate the model. The main presentation forms of the models were the nomogram analysis score ($n = 7$), followed by the classification and regression tree (CART) model ($n = 2$), mathematical formula ($n = 2$), β coefficient drawing risk scoring formula ($n = 1$), and one study¹⁸ did not provide a specific model. The performance and presentation form of the prediction models are shown in Table 3.

Results of quality assessment

Two researchers utilized the PROBAST scale¹² to conduct a quality evaluation of the included literature and meticulously reviewed the evaluation results to guarantee the accuracy of the findings. In the participant domain, all studies were retrospective in design and presented a high risk of bias. In the predictor domain, all studies exhibited a low risk of bias. In the analysis domain, all studies demonstrated a high risk of bias. Among them, three studies^{15,16,18} had insufficient sample sizes, unable to meet the recommended guideline of having more than 20 'events per variable' (EPV), six studies^{14,15,20–23} converted continuous variables into categorical variables, all studies failure to clearly define the approach for handling missing data, nine studies^{13–15,17–22} did not avoid

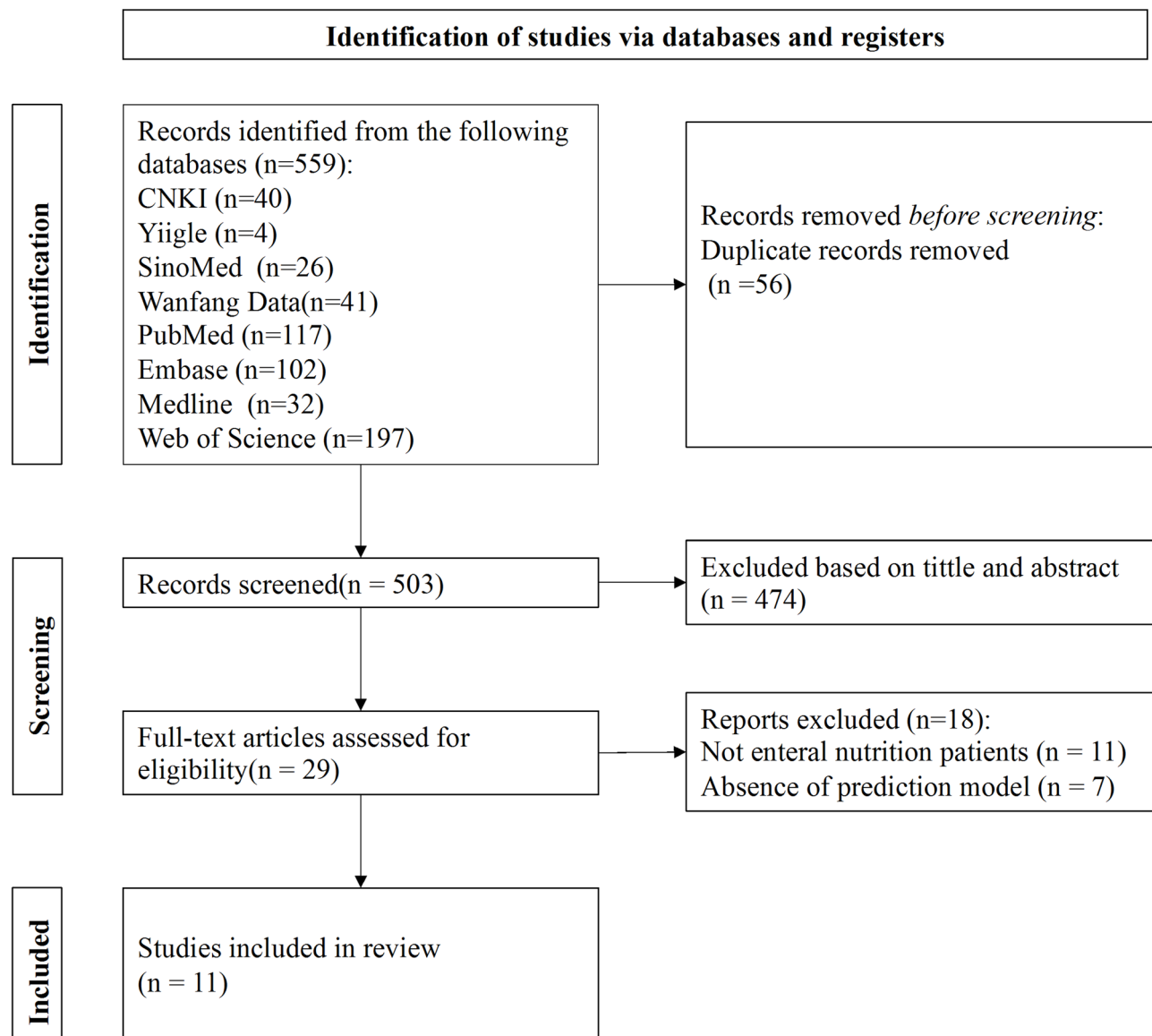


Fig. 1. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) flowchart of literature search and selection.

Author (year)	Country	Study design	Participants	Incidence (%)	Diagnostic criteria
Jing ¹³	China	Retrospective study	patients aged ≥ 18 and received EN	20.0	The expert consensus of the NAA
Peng ¹⁴	China	Retrospective study	Patients with severe cerebral hemorrhage and received EN	39.9	The expert consensus of the NAA
Guan ¹⁵	China	Retrospective study	Patients aged ≥ 18 years with severe acute pancreatitis and received EN	9.46	The expert consensus of the NAA
Guan ¹⁶	China	Retrospective study	Patients aged ≥ 18 years with severe acute pancreatitis and received EN	9.46	The expert consensus of the NAA
Zhang ¹⁷	China	Retrospective study	patients aged > 18 and received EN	20.9	–
Hou ¹⁸	China	Retrospective study	Patients with acute pancreatitis and received EN	15.62	The expert consensus of the NAA
Zhang ¹⁹	China	Retrospective study	Patients aged ≥ 18 with severe acute pancreatitis and received EN	28.31	–
Yu ²⁰	China	Retrospective study	Patients aged ≥ 18 with severe traumatic brain injury and received EN	48.11	The expert consensus of the NAA
Xie ²¹	China	Retrospective study	patients with acute cerebral infarction and received EN	49.87	The expert consensus of the NAA
Shi ²²	China	Retrospective study	Neurosurgical patients aged ≥ 65 and received EN	24.69	the expert consensus of The NAA
Sun ²³	China	Retrospective study	patients aged ≥ 18 and received EN	24.03	the expert consensus of the NAA

Table 1. Basic characteristics of the included studies. NAA: North American Association. “–”, not reported.

Author (year)	Model development method	Variable selection	Number of models	Sample size		Missing data handling	Continuous variable processing method	Predictor variables
				Positive events	Negative events			
Jing ¹³	Logistic regression	Single factor analysis and multiple factor analysis	1	103	412	–	Continuous variable	①②③④
Peng ¹⁴	Logistic regression	Single factor analysis and multiple factor analysis	1	147	221	–	Categorical variables	②③⑥⑦⑧
Guan ¹⁵	Logistic regression	Single factor analysis and multiple factor analysis	1	28	268	–	Categorical variables	④⑨⑩⑪⑬
Guan ¹⁶	Random Forest; Neural Network; CART; Support Vector Machine; Generalized Linear Regression	Feature Importance Selection	5	28	268	–	Continuous variable	④⑨⑩⑪⑬
Zhang ¹⁷	Logistic regression; CART	Single factor analysis and multiple factor analysis	2	46	174	–	Continuous variable	①②③④⑩⑬⑮⑰⑱⑲
Hou ¹⁸	Logistic regression	Single factor analysis and multiple factor analysis	1	25	135	–	Continuous variable	④⑨⑩⑪⑲
Zhang ¹⁹	XGBoost; Logistic regression; Random Forest; Support Vector Machine; KNN; SGD	Single factor analysis and multiple factor analysis	6	77	195	Multiple Imputation	Continuous variable	①②⑩⑪⑭⑲
Yu ²⁰	Logistic regression	Single factor analysis and multiple factor analysis	1	102	110	–	Categorical variables	④⑨⑩⑪⑲⑳
Xie ²¹	Logistic regression	Single factor analysis and multiple factor analysis	1	42	158	–	Categorical variables	②④⑤⑥⑦⑲⑳
Shi ²²	Logistic regression	Single factor analysis and multiple factor analysis	1	79	241	–	Categorical variables	②⑥⑬⑭⑲⑳
Sun ²³	Logistic regression; CART	Feature Importance Selection	2	103	412	–	Categorical variables	①②③④⑮⑲

Table 2. Model establishment. “–”, not reported. CART, Classification and Regression Trees; XGBoost, eXtreme Gradient Boosting; KNN, k-nearest neighbor classifier; SGD, Stochastic Gradient Descent; NIHSS, National Institute of Health stroke scale; APACHE II, Acute Physiology and Chronic Health Evaluation; GCS, Glasgow Coma Scale; ①: Number of comorbidities; ②: History of aspiration; ③: Use of sedatives; ④: Depth of tube insertion; ⑤: Diameter of nasogastric tube; ⑥: Amount of gastric residue; ⑦: NIHSS score; ⑧: Wada drinking test grade; ⑨: APACHE II score; ⑩: Consciousness disturbance; ⑪: Nutritional risk; ⑫: Coexisting diabetes; ⑬: Platelet-lymphocyte ratio; ⑭: Age; ⑮: Days of tube placement; ⑯: Food source; ⑰: Complications; ⑱: Serum C-reactive protein; ⑲: Patient position; ⑳: Mechanical ventilation; ㉑: GCS score; ㉒: Albumin level.

selecting variables based on univariate analysis, two studies^{17,23} failed to report whether model calibration tests were conducted, five studies^{14,17,18,21,22} had non-utilization or non-reporting of internal validation. In the outcome domain, the risk of bias for two studies^{17,19} was unclear, while the remaining studies presented a low risk of bias. Regarding adaptability, in the predictor domain, all studies had a low risk of bias. The overall assessment indicated that although all studies had a high risk of bias, the overall applicability risk of the models was low. Table 4 presents the results of quality assessment.

Results of meta-analysis

A meta-analysis was performed on the predictor factors with a literature quantity of ≥ 3 . The results indicated that the number of comorbidities, history of aspiration, use of sedatives, depth of tube placement, Amount of gastric residue, APACHE II score, consciousness disturbance, nutritional risk, and age were independent risk factors for aspiration in patients with nasogastric EN ($P < 0.05$). In the sensitivity analysis, the effect size was statistically combined after deleting relevant studies one by one, and no significant differences were found. The meta-analysis results are presented in Table 5, and the forest plot is presented in Supplementary Material A2. Due to insufficient reporting of model development details in the included studies, only four studies met the synthesis criteria. Among them, Sun et al. studied two model development methods, both of which were based on the same dataset, so only the CART-developed model was included. The fixed-effects model was used to calculate the combined AUC, which was 0.92 (0.90–0.93) (Fig. 2). The I^2 value was 0.0% ($P > 0.05$), indicating that the heterogeneity among studies was small. The Begg's test showed $p > 0.05$, indicated no significant publication bias (Appendix A3).

According to different disease types, a subgroup analyses was performed on the predictor factors with a literature quantity of ≥ 3 and $I^2 > 50\%$. The disease types including pancreatitis group, neurological disease group, and unrestricted disease group. Subgroup analyses provided an initial indication that disease type was the main source of heterogeneity for the predictor of intubation depth. By further data analysis of the included original literature, we found that patients in the pancreatitis group had nasointestinal tubes and patients in the undefined disease group had gastric tubes, the length of the tube in the two groups was not the same, which was the main factor leading to heterogeneity. Despite the subgroup stratification by disease type, the I^2 of the comorbidities, history of aspiration, and consciousness disturbance were higher than 50%, suggesting that the effects of these three predictors may be affected by unmeasured confounders or have different mechanisms in different disease

Author (year)	Model performance	Calibration method	Validation method		Model presentation
			Internal validation	External validation	
Jing ¹³	C-index = 0.81 [95%CI(0.71,0.89)]	Calibration curve	Bootstrap	–	Nomogram
Peng ¹⁴	A: ACU = 0.955 [95%CI(0.935,0.976)], Sensitivity = 95.2%, Specificity = 84.2% B: ACU = 0.891 [95%CI(0.790,0.991)], Sensitivity = 94.7%, Specificity = 85.3%	Calibration curve	Unclear	temporal validation	Nomogram
Guan ¹⁵	C-index = 0.841 [95%CI(0.316,1.366)]	Hosmer–Lemeshow test	Bootstrap	–	Nomogram
Guan ¹⁶	ACU = 0.976	Decision curve	Bootstrap	–	β coefficient risk scoring formula
Zhang ¹⁷	A:ACU = 0.895 [95%CI(0.822,0.936)], Sensitivity = 82.5%, Specificity = 73.6% A: ACU = 0.902 [95%CI(0.842,0.975)], Sensitivity = 80.6%, Specificity = 75.8%	–	Unclear	–	Nomogram, CART
Hou ¹⁸	A:ACU = 0.926 [95%CI(0.889,0.968)], Sensitivity = 88.4%, Specificity = 85.2% B: ACU = 0.902 [95%CI(0.828,0.986)], Sensitivity = 91.9%, Specificity = 81.8%	Calibration curve	Unclear	temporal validation	unclear
Zhang ¹⁹	B:ACU = 0.904, Sensitivity = 93.8%, Specificity = 77.8% ACU = 0.843, Sensitivity = 66.7%, Specificity = 84.7% ACU = 0.964, Sensitivity = 93.8%, Specificity = 83.1% ACU = 0.984, Sensitivity = 89.6%, Specificity = 96.8% ACU = 0.992, Sensitivity = 89.6%, Specificity = 97.4% ACU = 0.903, Sensitivity = 72.9%, Specificity = 86.8%	Calibration curve	Bootstrap	temporal validation	Nomogram
Yu ²⁰	C-index = 0.857 [95%CI(0.739,0.944)]	Hosmer–Lemeshow test	Bootstrap	–	Nomogram
Xie ²¹	A:ACU = 0.812 [95%CI(0.787,0.837)], Sensitivity = 78.4%, Specificity = 81.3% B: ACU = 0.885 [95%CI(0.816,0.947)], Sensitivity = 85.4%, Specificity = 88.2%	Hosmer–Lemeshow test	Unclear	temporal validation	mathematical formula
Shi ²²	ACU = 0.809 [95%CI(0.757,0.861)]	Hosmer–Lemeshow test	–	–	mathematical formula
Sun ²³	A:ACU = 0.93 [95%CI(0.91,0.96)], Sensitivity = 90.9%, Specificity = 88.6% ACU = 0.96 [95%CI(0.93,0.97)], Sensitivity = 88.3%, Specificity = 96.2% B: ACU = 0.91 [95%CI(0.89,0.98)], Sensitivity = 89.2%, Specificity = 89.7% ACU = 0.920 [95%CI(0.910,0.950)], Sensitivity = 64.9%, Specificity = 96.6%	–	Bootstrap	temporal validation	Nomogra, CART

Table 3. Performance and presentation of models. A, development cohort; B, validation cohort. AUC, area under the curve. We considered AUC = 0.5–0.7 as poor discrimination, 0.7–0.8 as moderate discrimination, 0.8–0.9 as good discrimination, and 0.9–1.0 as excellent discrimination.

Author (year)	ROB				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	ROB	Applicability
Jing ¹³	H	L	L	H	L	L	L	H	L
Peng ¹⁴	H	L	L	H	L	L	L	H	L
Guan ¹⁵	H	L	L	H	L	L	L	H	L
Guan ¹⁶	H	L	L	H	L	L	L	H	L
Zhang ¹⁷	H	L	U	H	L	L	L	H	L
Hou ¹⁸	H	L	L	H	L	L	L	H	L
Zhang ¹⁹	H	L	U	H	L	L	L	H	L
Yu ²⁰	H	L	L	H	L	L	L	H	L
Xie ²¹	H	L	L	H	L	L	L	H	L
Shi ²²	H	L	L	H	L	L	L	H	L
Sun ²³	H	L	L	H	L	L	L	H	L

Table 4. PROBAST results of the included studies. PROBAST, Prediction model Risk Of Bias Assessment Tool; ROB, risk of bias. L indicates low ROB/low concern regarding applicability; H indicates high ROB/high concern regarding application; U indicates unclear ROB/unclear concern regarding applicability.

Predictors	Included studies	heterogeneity test		effect model	Meta-analysis		
		I^2 (%)	P		OR (95%CI)	Z	P
Number of comorbidities	4 ^{13,17,19,23}	73.3	<0.05	Random effect	2.09[1.53,2.87]	4.58	<0.001
History of aspiration	7 ^{13,14,17,19,21-23}	94.2	<0.001	Random effect	5.38[1.96,14.75]	3.27	<0.001
Use of sedatives	3 ^{13,17,23}	0	0.547	Fixed effect	2.43[1.91,3.09]	7.20	<0.001
Depth of tube placement	8 ^{13,15-18,20,21,23}	99.1	<0.001	Random effect	1.89[1.05,3.41]	2.11	<0.05
amount of gastric residue	3 ^{14,21,22}	0	0.421	Fixed effect	2.15[1.54,3.01]	4.47	<0.001
APACHE II score	4 ^{15,16,18,20}	0	0.561	Fixed effect	2.72[2.51,2.94]	24.86	<0.001
Consciousness disturbance	6 ¹⁵⁻²⁰	81.3	<0.001	Random effect	2.99[2.36,3.78]	9.13	<0.001
Nutritional risk	5 ^{15,16,18-20}	0	0.850	Fixed effect	4.17[3.79,4.59]	29.11	<0.001
age	4 ^{17,19,20,22}	91.2	<0.001	Random effect	1.55[0.96,2.49]	1.80	0.072

Table 5. Meta-analysis of risk predictors.

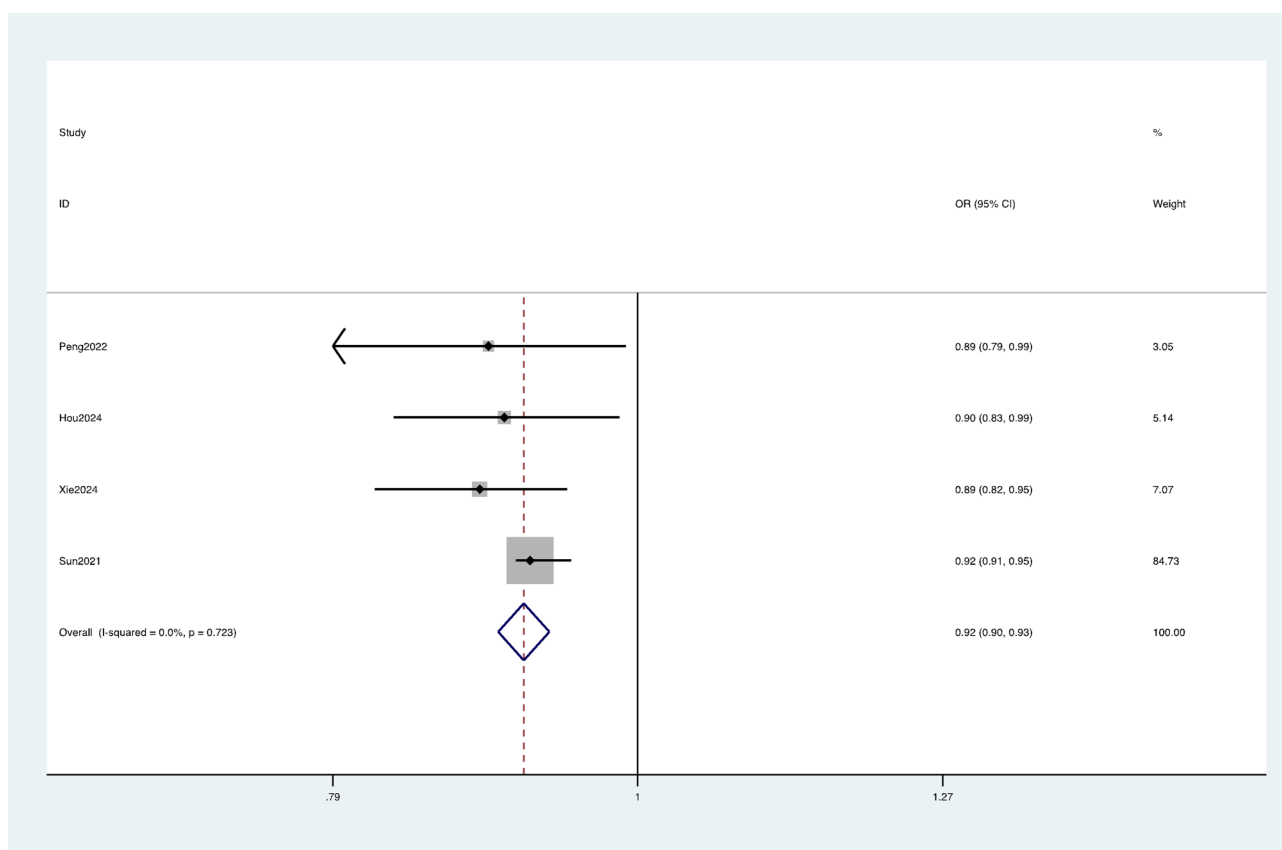


Fig. 2. Forest plot of the random effects meta-analysis of pooled AUC estimates for 4 validation models.

subgroups. Future studies need to be further verified by combining more clinical characteristics or biological markers. Details of the subgroup analyses are provided in the Appendix A4.

Discussion

This study found that the advantages of the existing models were excellent prediction performance ($AUC > 0.8$), diverse model forms, which could meet the needs of different application scenarios, with relatively low applicability risk and practical application potential. Of course, the limitations of this model are also obvious, including a high risk of methodological bias, mainly manifested in the retrospective study design, insufficient methodological standardization, and insufficient validation.

In this study, most of the models demonstrated high discrimination; however, there was no in-depth optimization in aspects such as model development, validation, and result reporting. The entire process of model construction encompasses clarifying the research question, determining the data source, selecting predictive variables, processing predictive variables, fitting the predictive model, evaluating the predictive

model, presenting the predictive model, and reporting the research results²⁴. Regarding the data source, all 22 predictive models included in this study were based on retrospective data. Considering that cohort studies have good representativeness, it is recommended that in future model construction or optimization, prospective data or registered data be selected as the modeling data²⁴ to reduce the risk of bias^{12,25}. In the selection of predictive variables, most studies relied on univariate logistic regression analysis, which might increase the risk of inappropriate predictor selection²⁶. Some studies²⁷ have proposed new variable selection methods, such as LASSO regression, Ridge regression, and ElasticNet regression. These methods can reduce the risk of overfitting. It is suggested that in future research, new selection methods can be adopted based on the actual clinical situation to enhance the accuracy of variable selection. Additionally, reporting and handling of missing data can prevent overfitting of the model²⁸. It is recommended that future studies handle the missing data appropriately. In the variable processing stage, converting continuous data into categorical variables for modeling might lead to the loss of model efficacy; However, in the clinical promotion stage, such data conversion can be carried out for the convenience of researchers' application²⁹. For the performance evaluation of the model, the core indicators include discrimination and calibration. The discrimination is commonly measured by AUC or the C-index. When the value is greater than 0.7, it indicates that the model has good discrimination. The calibration is evaluated through methods such as the Hosmer- Lemeshow test or calibration curves. If the model lacks performance evaluation, overfitting might occur, which to a certain extent limits the model's applicability. The average sample size of the included studies was 307, falling within the scope of small sample studies. In such studies, the lack of internal validation might lead to an overestimation of the model's performance. Therefore, internal validation is essential. Furthermore, external validation is an important approach to enhance the generalizability of the model³⁰, but when conducting external validation, it is necessary to compare the baseline characteristics of the modeling and validation data. Overall, current research on the aspiration prediction model for patients with nasogastric EN has made certain progress, however, there are still some issues that require further in-depth investigation in aspects such as model construction, validation, and reporting.

The existing prediction models reported in this review also have clinical implications, and the high-frequency predictors have implications for nursing practice and future research. The most commonly used predictors were depth of tube placement and history of aspiration, which appeared in 8 and 7 models, respectively. Other commonly used predictors were consciousness disturbance and nutritional risk, which were used in 6 and 5 models, respectively. The number of comorbidities, APACHE II score, and age were used in four models, and use of sedatives, Amount of gastric residue were used in three models. The above nine predictors can be prioritized when developing new models. Most of these factors are objective and easy to collect.

The EN via nasal feeding include nasogastric tube feeding and nasoenteric tube feeding. For patients with a risk of reflux and aspiration or feeding intolerance, post-pyloric feeding with a nasoenteric tube should be adopted. For the majority of patients, it is sufficient for the tip of the nasoenteric tube to pass the pylorus. However, for patients with a high risk of reflux and aspiration, the tube should be advanced to 95–105 cm, that is, through the duodenum, which can effectively reduce the incidence of aspiration risk³¹. For patients receiving EN support via nasogastric tube, if the gastric tube is inserted too deeply, it may stimulate the patient's gastric peristalsis, causing food reflux and thereby increasing the risk of aspiration¹⁵. If the insertion is too shallow, with the top of the gastric tube only reaching the cardia, when the patient is exposed to external stimuli such as coughing, it is prone to reflux and aspiration. Qian et al.³² demonstrated that when the insertion depth of the nasogastric tube reaches the prepyloric part of the antrum, the occurrence of complications such as aspiration and diarrhea in enteral nutrition can be reduced. The recommended insertion length of the gastric tube is Nose to Ear to Xiphisternum (NEX), approximately 45–55 cm for adults^{33,34}. However, some studies have questioned this standard, suggesting that it poses a relatively high risk of aspiration^{35,36}. Taylor et al.³⁷ proposed that when the catheterization depth is less than 48 cm, it is prone to ectopic placement of the nasal feeding tube, leading to aspiration. Therefore, they suggested “NEX + 10” as the reference standard for the catheterization depth in nasal feeding patients, that is, extending the catheterization depth by 10 cm based on the NEX standard. Therefore, to prevent the occurrence of aspiration in EN, it is recommended that nurses fully consider the individual circumstances of the patient when performing gastric tube insertion, select an appropriate catheterization depth, and when necessary, use visual ultrasound technology equipment to determine the location where the gastric tube reaches.

In our study, we found that there was a high heterogeneity in age, which was caused by the inconsistent age stratification criteria of the included studies. Advanced age has always been considered as a risk factor for aspiration, and the risk of aspiration is significantly increased in elderly patients due to the deterioration of swallowing function, weakened cough reflex and a variety of chronic diseases^{40–42}. Therefore, nurses should be alert to the possibility of aspiration in elderly patients with nasogastric enteral nutrition. In addition, some studies have shown that intra-abdominal hypertension (IAH) is an independent risk factor for aspiration in patients with gastric tube feeding (OR = 0.225–0.329)⁴³, and that mechanical ventilation and IAH increase the risk of aspiration⁴³. However, IAH was not included in the models included in this study, and only one study included mechanical ventilation. Further studies are needed to confirm the predictive value of these risk factors.

To enhance the effectiveness of clinical decision-making and practice guidance, developing precise prediction models for the risk of aspiration in tube-feeding EN patients is particularly important. For the future research on such prediction models, the following dimensions can be deepened and optimized: (1) Model development: When developing new models, the nine predictive factors in previous meta-analyses can be referenced to enhance the model's predictive ability. (2) Data collection method: To ensure the objectivity and accuracy of the data, it is recommended to collect data using a blind method. (3) Sample size determination: The sample size required for the study can be scientifically and rationally determined using EPV value or the sample size calculation method proposed by Riley et al.³⁹. (4) Variable handling: Before converting continuous variables into categorical variables, the specific basis for grouping should be clarified to ensure the rationality and consistency

of the conversion. (5) Variable selection: It is advisable to avoid selecting variables solely through univariate analysis and to conduct a more comprehensive and in-depth variable selection by combining machine learning algorithms and relevant professional background knowledge. (6) Research scope and scale: It is recommended to conduct large-sample, multi-center studies to cover a broader patient population and disease types, thereby improving the model's applicability and generalization ability. (7) Follow the TRIPOD statement for reporting of individual prognosis or diagnosis-specific multivariable prediction models to enhance transparency.

Limitation

This study has certain limitations. Firstly, this study only retrieved Chinese studies on the prediction models of nasogastric EN aspiration, and there is currently a lack of relevant studies in other countries, which may limit the applicability of the research results in the population of other countries and require adjustments when applied in different regions. Secondly, there is inconsistency in the disease types included in the studies, and some predictive factors have been studied in fewer studies, without being subjected to a meta-analysis, which may have certain impacts on the predictive results. In summary, all 22 prediction models included in this study showed good discriminatory power. However, there is still considerable room for improvement in the quality of the models. Therefore, in the future, rigorous methods and transparent reporting are needed to further refine related studies. Finally, since this review only includes studies published in English and Chinese, there may be problems related to language barriers, and the results of studies published in other major languages have not been included in this review.

Conclusion

This systematic review encompassed 11 studies and 22 models. The results indicated that the combined AUC of the 4 validated models was 0.92 (0.90–0.93), suggesting an excellent discriminatory capacity of the models. Nevertheless, based on PROBAST, all the included studies were evaluated as having a high risk of bias. Currently, the prediction models for the aspiration risk in patients with nasogastric EN do not comply with the PROBAST criteria. Hence, researchers need to be acquainted with the PROBAST checklist and adhere to the reporting guidelines in the TRIPOD statement to enhance the quality of future studies. Future research should give priority to the development of new models with larger sample sizes, rigorous study designs, and multi-center external validation.

Data availability

Data available on request from the first corresponding author.

Received: 24 October 2024; Accepted: 16 July 2025

Published online: 19 July 2025

References

- Zhang, Z.G., Yan, M.Y., Tan, S.J. & Wu, H.G. Interpretation of ESPEN practical and partially revised guideline: Clinical nutrition in the intensive care unit. *Chinese Journal of Practical Surgery* **43**, 1335–1343 (2023). <https://doi.org/10.19538/j.cjps.issn1005-2208.2023.12.03>
- Chinese Society of Parenteral and Enteral Nutrition (CSPEN), 2023. Guideline for clinical application of parenteral and enteral nutrition in adults patients in China (2023 edition). *Natl Med J China*, 946–974 (2023).
- R Leonard 2019 Predicting aspiration risk in patients with dysphagia: Evidence from fluoroscopy Laryngoscope Investig Otolaryngol **4** 83–88 <https://doi.org/10.1002/lio2.226>
- Chen, S. *et al.* Risk of regurgitation and aspiration in patients infused with different volumes of enteral nutrition. *Asia Pac J Clin Nutr* **24**, 212–218 (2015). <https://doi.org/10.6133/apjcn.2015.24.2.12>
- JL Wang XJ Dai Q Shi SS Xu L Guo 2018 Research of aspiration after stroke in China: a bibliometric analysis *Chin. Nurs. Res.* **32** 1475–1477
- Anderson, S., Wobeser, B., Duke-Novakowski, T. & Singh, B. Pathology in Practice. Aspiration pneumonia. *J Am Vet Med Assoc* **251**, 409–411 (2017). <https://doi.org/10.2460/javma.251.4.409>
- Schwarz, M., Cocchetti, A., Murdoch, A. & Cardell, E. The impact of aspiration pneumonia and nasogastric feeding on clinical outcomes in stroke patients: A retrospective cohort study. *J Clin Nurs* **27**, e235–e241 (2018). <https://doi.org/10.1111/jocn.13922>
- Paul, G., Bohle, W. & Zoller, W. Risk Factors for the Development of Esophagorespiratory Fistula in Esophageal Cancer. *J Gastrointest Liver Dis* **28**, 265–270 (2019). <https://doi.org/10.15403/jgld-271>
- McClave, S. A. *et al.* North American Summit on Aspiration in the Critically Ill Patient: consensus statement. *JPEN J Parenter Enteral Nutr* **26**, S80–85 (2002). <https://doi.org/10.1177/014860710202600613>
- Moons, K. G. *et al.* Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* **11**, e1001744 (2014). <https://doi.org/10.1371/journal.pmed.1001744>
- Debray, T. P. *et al.* A guide to systematic review and meta-analysis of prediction model performance. *Bmj* **356**, i6460 (2017). <https://doi.org/10.1136/bmj.i6460>
- Moons, K. G. M. *et al.* PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Ann Intern Med* **170**, W1–w33 (2019). <https://doi.org/10.7326/m18-1377>
- Jing, Y.X. & Wu, Y.L. Development and validation of a prediction model for aspiration risk in patients with nasal feeding. *Clinical Education of General Practice* **21**, 1131–1134 (2023). <https://doi.org/10.13558/j.cnki.issn1672-3686.2023.012.022>
- Y Peng LY Sha ZL Liu Y Liu J Yi 2022 Construction and validation of risk prediction model for aspiration in patients with severe intracerebral hemorrhage receiving enteral nutrition support *Chinese Nursing Management* **22** 1391–1397
- Y Guan X Wang L Hong W Wei 2022 Construction of early enteral nutrition aspiration risk prediction model in patients with severe acute pancreatitis *J. Nurs. Sci.* **37** 94–96
- Y Guan GJ Zhang Y Luo 2022 Establishment and validation of early enteral nutrition aspiration risk prediction model for patients with severe acute pancreatitis based on machine learning algorithm *J Clin Surg* **30** 634–638
- JH Zhang YP Ji HX Wei 2022 Research on the prediction model of aspiration risk of nasal feeding patients based on decision tree and Nomogram *Chinese Evidencebased Nursing* **8** 971–975
- Hou, P. *et al.* Prediction model establishment and validation for enteral nutrition aspiration during hospitalization in patients with acute pancreatitis. *World J Gastrointest Surg* **16**, 2583–2591 (2024). <https://doi.org/10.4240/wjgs.v16.i8.2583>

19. Bo Zhang, Huanqing Xu, Qigui Xiao et al. Machine learning predictive model for aspiration risk in early enteral nutrition patients with severe acute pancreatitis, 05 March 2024, PREPRINT (Version 1) available at Research Square [<https://doi.org/10.21203/rs.3.rs-3969940/v1>].
20. Yu, L.Z., Lin, X., Zhu, X.M., Qin, Y. & Zhou, S.J. Development of patients with severe craniocerebral injury based on nomogram study on risk prediction model of enteral aspiration. *Trauma and Critical Care Medicine* **10**, 467–470 (2022). <https://doi.org/10.16048/j.issn.2095-5561.2022.06.22>
21. Xie, X.L. Develop and validate a risk prediction model for enteral nutrition aspiration in emergency inpatients with cerebral infarction, 2536–2541 (2024).
22. J Shi R Liu 2023 Investigation of nasal feeding reflux aspiration in elderly coma patients in neurosurgery department and analysis of its risk factors *Geriatr Health Care* **29**, 1253–1257
23. Sun, W.J., Xie, L.L., Huang, L.X., You, H.L., Qu, F.R. & Zuo, Y. External validation and comparison of 2 aspiration risk prediction models in patients receiving nasogastric feeding. *Chin J Nurs* **56**, 1352–1356 (2021). <https://doi.org/10.3761/j.issn.0254-1769.2021.09.012>
24. Gu, H.Q., Wang, J.F., Zhang, Z.H. & Zhou, Z.R. Clinical Prediction Models: Model Development. *Chin J Evid Based Cardiovasc Med* **11**, 14–16+23 (2019).
25. Chen, X.P., Zhang, Y., Zhuang, Y.Y. & Zhang, Z.H. PROBAST: PROBAST: a tool for assessing risk of bias in the study of diagnostic or prognostic multi-factorial predictive models. *Chin J. EVIDENCE-BASED Med* **20**, 737–744 (2020).
26. A Ziegler 2020 Clinical prediction models: A practical approach to development, validation, and updating [J] *Biom. J.* **62**, 4–1122–1123
27. Ranalli, M. G., Salvati, N., Petrella, L. & Pantalone, F. M-quantile regression shrinkage and selection via the Lasso and Elastic Net to assess the effect of meteorology and traffic on air quality. *Biom J* **65**, e2100355 (2023). <https://doi.org/10.1002/bimj.202100355>
28. Zhang, Q., Yuan, K. H. & Wang, L. Asymptotic bias of normal-distribution-based maximum likelihood estimates of moderation effects with data missing at random. *Br J Math Stat Psychol* **72**, 334–354 (2019). <https://doi.org/10.1111/bmsp.12151>
29. Steyerberg, E. W. & Vergouwe, Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* **35**, 1925–1931 (2014). <https://doi.org/10.1093/eurheartj/ehu207>
30. Steyerberg, E. W. & Harrell, F. E., Jr. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol* **69**, 245–247 (2016). <https://doi.org/10.1016/j.jclinepi.2015.04.005>
31. YW Deng 2022 Expert consensus on ultrasound-guided nasoenteric intubation in adults *Modern Clinical Nursing* **21**, 1–6
32. D Qian Y Cao H Pan 2020 Safety and efficacy of under bedside ultrasound guidance for critical patients enteral nutrition support *Chin J Prac Nurs* **36**, 1479–1481
33. Best, C. How to insert a nasogastric tube and check gastric position at the bedside. *Nurs Stand* **30**, 36–40 (2016). <https://doi.org/10.7748/ns.30.38.36.s43>
34. Best, C. How to set up and administer an enteral feed via a nasogastric tube. *Nurs Stand* **31**, 42–47 (2017). <https://doi.org/10.7748/ns.2017.e10509>
35. Taylor, S. J. Methods of Estimating Nasogastric Tube Length: All, Including "NEX," Are Unsafe. *Nutr Clin Pract* **35**, 864–870 (2020). <https://doi.org/10.1002/ncp.10497>
36. Chen, Y. C. et al. Potential risk of malposition of nasogastric tube using nose-ear-xiphoid measurement. *PLoS One* **9**, e88046 (2014). <https://doi.org/10.1371/journal.pone.0088046>
37. Taylor, S. J., Allan, K., McWilliam, H. & Toher, D. Nasogastric tube depth: the 'NEX' guideline is incorrect. *Br J Nurs* **23**, 641–644 (2014). <https://doi.org/10.12968/bjon.2014.23.12.641>
38. Friedrich, J. O., Adhikari, N. K. & Beyene, J. Ratio of means for analyzing continuous outcomes in meta-analysis performed as well as mean difference methods. *J Clin Epidemiol* **64**, 556–564 (2011). <https://doi.org/10.1016/j.jclinepi.2010.09.016>
39. Riley, R. D. et al. Calculating the sample size required for developing a clinical prediction model. *Bmj* **368**, m441 (2020). <https://doi.org/10.1136/bmj.m441>
40. Yanagita, Y. et al. The severity of nutrition and pneumonia predicts survival in patients with aspiration pneumonia: A retrospective observational study. *Clin Respir J* **16**, 522–532 (2022). <https://doi.org/10.1111/crj.13521>
41. Kurita, D. et al. Handgrip strength predicts early postoperative dysphagia after thoracoscopic-laparoscopic esophagectomy in male patients with esophageal cancer. *Esophagus* **19**, 586–595 (2022). <https://doi.org/10.1007/s10388-022-00925-6>
42. Mittal, A., Patel, M., Wang, D., Khrais, A. & Chyn, E. T. Y. Does Dysphagia Predict Inpatient Morbidity and Mortality in Geriatric Patients Admitted for Aspiration Pneumonia? *Cureus* **15**, e39223 (2023). <https://doi.org/10.7759/cureus.39223>
43. Miao, P., Zhang, Y. & Zhong, A. Risk factors of aspiration occurrence with different feeding patterns in elderly intensive care unit patients: a cross-sectional study. *J Thorac Dis* **15**, 2585–2600 (2023). <https://doi.org/10.21037/jtd-23-430>

Acknowledgements

The authors would like to thank all the researchers and health professionals involved in this study.

What is already known

The aspiration risk prediction models in patients with nasogastric EN can predict the occurrence of aspiration risk and provide early intervention guidance for preventing aspiration. Recent studies have shown an increasing number of research on nasogastric EN aspiration risk prediction models.

What this paper adds

This review evaluated 22 published risk prediction models for aspiration in patients with nasogastric EN. A meta-analysis was conducted on the predictive factors and efficacy of the models, and the methodology and predictive factors of the models were discussed and summarized. This study recommends that future researchers employ the Prediction Model Risk of Bias Assessment Tool (PROBAST) for model development to mitigate bias risks. Additionally, researchers are encouraged to adhere to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement to enhance the transparency of reporting.

Author contributions

Juan Chen: Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization. Yushuang Su: Software, Formal analysis, Data curation. Qinghua Jiang: Writing – review & editing. Qin Yang: Validation, Methodology, Formal analysis. Qing Wen: Validation, Data curation. Zaichun Pu: Visualization. Mengting Liu: Resources. Xiaorong Mao: Supervision, Project administration, Methodology. TianLong Li: Visualization, Resources, Investigation.

Funding

This research received no specific grant from any funding agency in the public, commercial, and not-for-profit sectors.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-12252-8>.

Correspondence and requests for materials should be addressed to T.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025