

Development and Validation of an Explainable Machine Learning Model for Identification of Dysphagia in Patients with COPD

Chiteng Zhou^{1,*}, Shuangwei Hong^{2,*}, Jun Fang³, Ximei Yan³, Min Zang^{1,4}, Ningfei Tang⁵, Lingwei Bao⁴, Huiying Pan¹

¹School of Medicine, Jinhua University of Vocational Technology, Jinhua, Zhejiang, People's Republic of China; ²Department of Acupuncture, Jindong Hospital of Traditional Chinese Medicine, Jinhua, Zhejiang, People's Republic of China; ³Department of Respiratory and Critical Care Medicine, Jinhua Municipal Central Hospital, Jinhua, Zhejiang, People's Republic of China; ⁴Department of General Surgery, The Affiliated Hospital of Jinhua University of Vocational Technology, Jinhua, Zhejiang, People's Republic of China; ⁵Department of Nursing, Wuyi County First People's Hospital, Jinhua, Zhejiang, People's Republic of China

*These authors contributed equally to this work

Correspondence: Huiying Pan, School of Medicine, Jinhua University of Vocational Technology, Jinhua, Zhejiang, People's Republic of China, Tel +86 13586983315, Email panhyhvt@163.com

Purpose: Dysphagia is a common yet often overlooked complication in Chronic Obstructive Pulmonary Disease (COPD) patients, who are prone due to abnormal breathing patterns, impaired airway protection, and generalized frailty. This not only predisposes them to aspiration pneumonia but also serves as a key trigger for acute exacerbations of COPD, substantially increasing the risk of adverse outcomes. Early identification of dysphagia is essential for improving prognosis in COPD. However, research on early detection is limited, particularly regarding machine learning prediction. This study aimed to develop and validate a machine learning based risk assessment model for dysphagia in COPD and deploy it as a user-friendly web-based clinical tool to assist clinicians in risk identification and early intervention.

Patients and Methods: Retrospective medical records from 710 COPD patients admitted between February 2025 and January 2026 were analyzed. Swallowing function was assessed using the Water-Swallowing Test. Univariate and multivariate logistic regression identified independent risk factors, which were used to develop and compare eight machine learning models. Model performance was evaluated using ROC, calibration, and decision curves with Bootstrap internal validation. Key variables were interpreted via Shapley Additive Explanations, and the final model was deployed online.

Results: Dysphagia prevalence was 29.3%. Multivariate regression identified five key risk factors: disease duration, BMI, history of tracheal intubation, muscle strength, and the modified Medical Research Council (mMRC) score for dyspnea severity. Among eight machine learning models, the XGBoost model showed the best performance in the training set (AUC 0.921, 95% CI 0.901–0.940) and demonstrated good calibration and highest clinical net benefit. The model was deployed online (<https://dysphagiamodel.shinyapps.io/COPD-DP/>).

Conclusion: We developed and validated an online machine learning-based dysphagia risk assessment tool for COPD, demonstrating discrimination, calibration, and clinical utility for risk stratification and clinical decision-making.

Keywords: chronic obstructive pulmonary disease, dysphagia, comorbidities, risk prediction, risk calculator tool

Introduction

Chronic obstructive pulmonary disease (COPD) is a respiratory disorder primarily characterized by persistent airflow limitation that typically worsens with time. Several studies have indicated that particulate matter pollution and cigarette smoking are among the major contributors to the global development of COPD.^{1–3} COPD has become a major global public health concern, with consistently high mortality rates. COPD is the third leading cause of death worldwide.⁴ In China, the burden of COPD is particularly heavy, with an estimated 1 million deaths annually; further, approximately



5 million individuals experience disability from the disease, placing significant strain on families and the healthcare system.⁵

Among the complications associated with COPD, dysphagia has gained increasing attention recently. Dysphagia affects ~33% of patients during stable periods of COPD,⁶ and this figure rises to over 50% during acute exacerbations.⁷ The primary mechanism behind dysphagia is impaired coordination between respiration and swallowing,⁸ especially when swallowing is triggered during the inspiratory phase, which significantly increases the risk of aspiration.^{9,10} However, in clinical practice, symptoms of swallowing dysfunction are often overshadowed by the more prominent issue of dyspnea, delaying timely recognition of the associated risks. This failure in recognizing swallowing problems highlights the urgent need for effective identification strategies.¹¹

The presence of dysphagia significantly impacts the disease outcomes for patients with COPD. Dysphagia is one of the key risk factors for aspiration pneumonia in COPD patients,¹² contributing to recurrent acute exacerbations of the disease and increasing the likelihood of unplanned readmissions and in-hospital mortality.^{10,13} Additionally, chronic swallowing difficulties are often associated with reduced food intake, which can worsen malnutrition and lead to skeletal muscle atrophy, weakening respiratory muscle function, and impaired overall respiratory status.^{14,15} From an economic standpoint, patients with COPD who also have dysphagia tend to have prolonged hospital stays, resulting in significantly high healthcare costs.¹⁶ Therefore, for improving clinical outcomes and reducing the healthcare burden on society, it is crucial to accurately identify high-risk patients and provide targeted interventions in areas such as feeding posture, dietary habits, and functional exercises.

Currently, swallowing dysfunction is mainly assessed using bedside functional tests, such as the Water Swallow Test and the Repeated Swallowing of Saliva Test.¹⁷ These tests rely on the clinical experience and operational standards of healthcare professionals. The tests can be conducted only for immediate assessment during hospitalization, and the tests cannot offer early detection or continuous monitoring of swallowing risks. Thus, it is challenging to meet the clinical demand for early identification and intervention of swallowing dysfunction in patients with COPD. As a result, predictive models have been increasingly introduced as a supplementary tool for risk assessment. By integrating multiple clinical features and data, these models can accurately evaluate risk and offer timely guidance for clinical interventions. Currently, risk evaluation for dysphagia is primarily conducted using logistic regression models or nomograms.^{18,19} Although these methods are relatively straightforward and easy to interpret, they still have limitations in their predictive power when addressing common nonlinear relations and interactions among multiple clinical factors.^{20,21}

In recent years, machine learning techniques have been used more frequently in medical prediction models. Unlike traditional statistical methods, machine learning algorithms can identify hidden patterns in multidimensional clinical data without assuming predefined relations between variables. In patients with COPD, the development of dysphagia is often influenced by multiple factors, including respiratory function status, nutritional condition, muscle function, respiratory–swallow coordination, and previous airway interventions, with potential complex interactions among these clinical characteristics. Compared with conventional models, machine learning approaches may offer advantages in integrating complex clinical information and learning high-dimensional features, thereby facilitating more accurate identification and risk stratification of high-risk patients.²² The advantages of machine learning in predicting dysphagia in other high-risk populations, such as persons with stroke or sarcopenia, are well-documented.^{23,24} However, studies investigating the application of machine learning for dysphagia risk identification in patients with COPD remain limited,²⁵ and existing models have mainly focused on predictive performance comparisons rather than their practical application in specific clinical settings.¹⁹ In addition, by integrating readily available bedside clinical information, machine learning models may enable rapid identification of high-risk patients and provide more targeted decision support for swallowing management, feeding care, and aspiration prevention. Therefore, incorporating machine learning techniques to analyze swallowing risk for patients with COPD can not only address the limitations of traditional models but also demonstrate important academic and clinical value in promoting personalized and precision nursing care.

Although machine learning models have shown strong predictive performance, in practice, their use by frontline healthcare providers is often constrained by complex model structures and high operational requirements.²⁶ A web-based calculator offers a more amenable approach to translating a model into clinical practice by transforming the risk prediction into an easy-to-use online tool.²⁷ Healthcare professionals can input relevant clinical data and instantly

receive risk assessments, enabling creation of personalized treatment strategies, thereby improving the quality of healthcare.^{28,29} This real-time decision support model not only fits the fast-paced nature of clinical work but also the real-time model aids in developing and implementing personalized treatment plans.³⁰

Against the foregoing backdrop of machine learning tools, there remains a lack of a dysphagia risk assessment tool for patients with COPD that simultaneously offers strong predictive performance, clinical interpretability, and practical usability. In particular, within routine nursing practice, how to achieve early identification and rapid risk stratification of high-risk patients using bedside-accessible clinical information remains an important unresolved issue. Therefore, based on real-world clinical data, we sought to identify from clinical data potential predictive variables of dysphagia for patients with COPD, compare the predictive performance of multiple machine learning models, select the most suitable model for clinical practice, and integrate model interpretation approaches to identify key risk features. On this basis, further develop and validate an online risk prediction tool to aid healthcare professionals in accurate identification and management of dysphagia risk. By constructing a risk assessment model with both predictive capability and clinical applicability, we aimed to provide a novel approach and practical tool for early screening, precision nursing interventions, and clinical decision support for dysphagia in patients with COPD.

Materials and Methods

Population

This investigation was a retrospective cohort study. It included patients with COPD who were hospitalized in the respiratory department of Jinhua Municipal Central Hospital between February 2025 and January 2026. Patients were eligible for inclusion if they met the following criteria: (1) aged ≥ 18 years; (2) diagnosed with COPD according to the Global Initiative for Chronic Obstructive Lung Disease 2026 guidelines,³¹ (3) had adequate reading, comprehension, and communication abilities and were able to cooperate with swallowing assessments; and (4) provided informed consent to participate in the study. Patients were excluded if they (1) were under fasting or fluid restriction; (2) had comorbid conditions known to affect swallowing function, such as laryngeal cancer, esophageal cancer, stroke, or Parkinson's disease; or (3) had unstable clinical conditions that prevented completion of the assessment. The Ethics Committee of Jinhua University of Vocational Technology approved the study protocol (Approval Number: 202502) and performed in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to enrollment. The outcome measure, swallowing function, was assessed during hospitalization, whereas all predictor variables were collected retrospectively from electronic medical records, nursing assessment forms, test reports, and patient interviews. These predictors reflect the patients' historical status and temporally precede the outcome, consistent with the design of a retrospective cohort study.

Data Collection

All predictor data were collected retrospectively from the department's electronic medical records system, nursing assessment forms, test reports, and patient interviews. The selection of potential risk factors was guided by literature and recommendations from clinical professionals, with a focus on the following three key areas: (1) *Demographic characteristics* including sex, age, residence, ethnicity, marital status, occupation, and level of education. (2) *Health-related history and clinical status* including medical history, respiratory conditions such as respiratory failure and asthma, cardiovascular diseases such as hypertension and coronary artery disease, cerebrovascular events, such as transient ischemic attack, stroke (cerebral hemorrhage or infarction), diabetes, gastroesophageal reflux disease, tobacco use, alcohol consumption, body mass index, duration of illness, number of missing teeth, history of tracheal intubation, history of tracheostomy, use of dry powder inhalers, and use of home non-invasive ventilation (HNIV). (3) *Physiological and functional indicators* including respiratory rate, oxygen saturation, cough, dry mouth, muscle strength (assessed by Manual Muscle Testing, MMT), and severity of dyspnea assessed using the modified Medical Research Council scale (mMRC). These measures were collected from clinical records during the early hospital stay, prior to swallowing assessment, to maintain a clear temporal relationship between predictors and outcome.

Sample Size

To ensure an adequate number of participants for developing the prediction model,³² two methods were used to estimate the required sample size, and the larger estimate was selected for this study. (1) The sample size was first calculated using the `pmsampsize` package in R, which is designed for prediction models.^{33,34} On the basis of previous studies, we set the expected R^2 value at 0.5, with 23 candidate predictors. Meta-analysis results indicated a 32.7% dysphagia prevalence for patients with COPD,⁶ resulting in a minimum required sample size of 448. (2) The sample size was also estimated using the Events Per Variable principle, which recommends at least 10 positive events for each predictor.³⁵ With 23 predictors planned for inclusion, at least 230 patients with dysphagia were required. Given a dysphagia prevalence of 32.7%, the sample size was estimated to be 703. Thus, the final sample for this study was 710 patients (Figure 1). Initially, a total of 744 patients were screened for inclusion. Of these, 16 patients were unable to complete the assessments due to severe illness, 6 patients had sequelae of cerebrovascular disease, 4 patients had coexisting Parkinson's disease, 2 patients had esophageal cancer, 1 patient had laryngeal cancer, and 5 patients declined to participate. After applying these exclusion criteria, a total of 710 patients were ultimately included in the study.

Outcome and Definition

During hospitalization, swallowing function was assessed by registered nurses with over 10 years of experience in respiratory nursing and standardized training. The assessment was performed on the day before discharge. In this study, dysphagia was defined as abnormal swallowing function, characterized by difficulty initiating swallowing and impaired transfer of food from the oral cavity to the esophagus, manifesting as choking, uncoordinated swallowing, or the need to swallow in multiple attempts.^{36,37} Swallowing function was assessed using the Water Swallow Test (WST), which is simple, reproducible, and widely used in clinical practice for patients with chronic respiratory diseases.^{17,38} During the assessment, participants were instructed to sit upright and drink 30 mL of warm water, while the time required to swallow and the occurrence of choking or coughing were observed. The WST grading criteria were as follows: Grade 1, swallowing the water smoothly in a single attempt without choking; Grade 2, swallowing the water in two or more

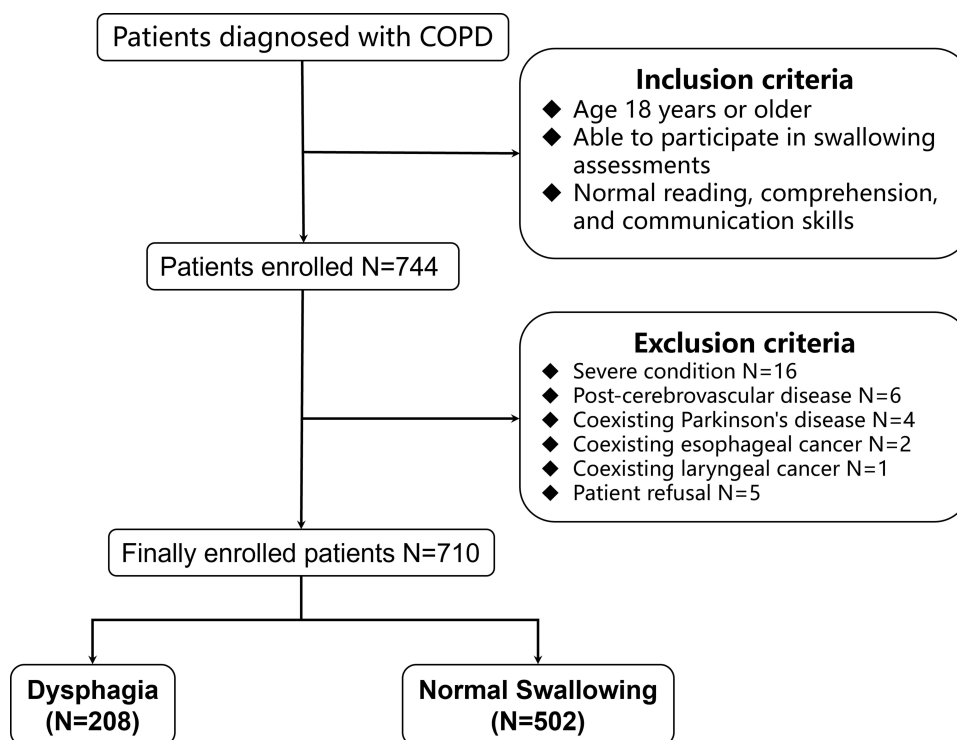


Figure 1 Patient selection flowchart. Flowchart illustrating the patient selection process, including inclusion criteria, exclusion criteria, and final inclusion of patients with COPD for model development and analysis.

attempts without choking; Grade 3, swallowing the water in a single attempt with choking or coughing; Grade 4, swallowing the water in two or more attempts with choking or coughing; and Grade 5, frequent choking or coughing with inability to swallow all the water. In this study, Grade 1 was classified as normal swallowing function, whereas Grades 2–5 were classified as dysphagia, which is consistent with previous validation studies of the WST in patients with COPD and other hospitalized populations.⁷

Data Preprocessing

Prior to model development, the dataset was carefully cleaned. Patients with more than 30% missing values in their medical records were excluded to ensure data quality. Remaining missing values were imputed using the k-nearest neighbors algorithm with $K = 5$. Because the study included 710 patients, and to maximize data use for accurate model training, no random train-test split was performed. Instead, all 710 sets of data were used as the training set specifically for model development. Internal validation was conducted using Bootstrap resampling with 1000 iterations to assess model stability and generalizability.

Statistical Analysis

Baseline characteristics were assessed for normality prior to analysis. Continuous variables with a normal distribution were presented as mean \pm standard deviation and compared using independent samples t tests. Continuous variables with non-normal distributions were reported as medians with interquartile ranges [M (P25, P75)] and compared using the Mann–Whitney U -test. Categorical variables were expressed as frequencies and percent and compared using the chi-square test or Fisher's exact test, as appropriate.

Variables Selection

All analyses were conducted using R software (version 4.4.0). A total of 23 candidate variables were initially included in the analysis. Univariate logistic regression was first performed for each candidate variable to identify potential risk factors associated with dysphagia. Variables with a P value <0.1 in the univariate analysis were selected for subsequent multivariate logistic regression to avoid prematurely excluding variables that might influence the model and to ensure sufficient statistical power for further analysis. In the multivariate analysis, forward stepwise regression based on the Akaike Information Criterion (AIC) was used for variable selection, with a significance threshold of P value <0.05 . The Spearman correlation analysis was used to assess independence among predictors, and the results were visualized using a heatmap. To address class imbalance, if the proportion of positive events in the training set is below 20%, oversampling techniques will be applied to balance the classes and improve the model's predictive performance for the minority class.

Model Development

Based on the predictors identified in the multivariate analysis, we developed and compared eight machine learning algorithms: Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forests (RF), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Support Vector Machines (SVM) and Neural Network (NN). All models used the five variables retained from the multivariate regression as input features to ensure comparability across algorithms. Continuous variables were standardized prior to modeling in KNN, SVM, and neural network. All models were built using preset hyperparameters (Table 1).

Model Evaluation

Model performance was evaluated using the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive and negative predictive value, and accuracy. Calibration curves were generated to assess agreement between predicted and observed probabilities, and decision curve analysis was performed to evaluate clinical net benefit. Internal validation was performed using the Bootstrap method with 1000 repetitions. In each iteration, a Bootstrap sample of the same size as the original dataset was drawn with replacement for model training, and the out-of-bag samples not selected in that iteration were used for model validation. The model was refitted in each repetition, and performance metrics including AUC, sensitivity, specificity, positive predictive value, negative predictive value, and

Table 1 Preset Hyperparameter Configurations for the Eight Machine Learning Models

Model	Main Parameter Settings
Logistic Regression	Binary logistic regression with logit link function
K-Nearest Neighbors	k = 5; continuous predictors standardized
Decision Tree	rpart algorithm; complexity parameter cp = 0.01
Random Forest	500 trees; mtry set to the square root of the number of predictors
XGBoost	Binary: logistic objective; evaluated by AUC; nrounds = 100; max_depth = 6; eta = 0.3
LightGBM	Binary objective; evaluated by AUC; num_iterations = 100; max_depth = 6; learning_rate = 0.1
Support Vector Machine	Radial basis function kernel; outputting predicted probabilities
Neural Network	Single hidden layer; size = 10; decay = 0.01; maxit = 200

Note: All models were implemented using R software (version 4.4.0).

Abbreviations: cp, complexity parameter; mtry, number of variables randomly sampled at each split; AUC, area under the receiver operating characteristic curve; eta, learning rate; maxit, maximum number of iterations.

accuracy were calculated. The stability of the models was then assessed based on the mean, standard deviation, and 95% percentile confidence interval of the AUC across all iterations.

Final Model Selection Criteria

The final model was primarily selected based on the AUC to ensure good discriminative performance, and models with better performance in both the training set and higher AUC in the validation set were preferentially considered. When the differences in validation AUC among models were small (<0.03), calibration performance was further compared. When calibration performance was comparable, models with higher net benefit in decision curve analysis were selected.

Model Interpretation

To explain the prediction mechanism of the top-performing model, we used SHapley Additive exPlanations (SHAP).³⁹ This approach, grounded in cooperative game theory, breaks down a model's predictions by calculating Shapley values, showing how each feature contributes to the final output and ensuring interpretability within a unified framework. For global feature importance, we calculated the mean of the absolute SHAP values for each feature across all samples and visualized these values using a bar plot to identify the most influential features of a model's predictions. Additionally, we used a bee swarm plot to assess each feature's specific impact on the prediction for individual instances, both in terms of direction and magnitude.

Development of a Web Application

To facilitate clinical implementation, the final prediction model was deployed as a Shiny-based web application. By entering patient-specific characteristics, clinicians can obtain an individualized probability of dysphagia, providing timely and quantitative support for clinical decision-making.

TRIPOD AI Statement

This study is reported in accordance with the TRIPOD + AI guidelines.⁴⁰ The relevant items from the TRIPOD AI checklist were referenced and followed throughout the model development, validation, performance evaluation, and reporting processes to ensure transparency and completeness.

Results

Patient Characteristics

From February 2025 to January 2026, we selected 744 inpatients with COPD, from whom 710 valid questionnaires were collected. No patients were excluded in this study. Remaining minor missing values were imputed using the K-Nearest Neighbor method (K=5). Missing values were identified in the following variables: tracheostomy status (n = 8), missing teeth (n = 7), HNIV (n = 6), dry mouth (n = 2), and muscle strength (n = 1), with missing rates for all variables below

2%. In the final analysis, 710 samples were included. Among them, 208 patients (29.3%) had dysphagia, while 502 patients (70.7%) did not. The proportion of positive cases was 29.3%, which did not fall below the conventional threshold for severe class imbalance (<20%). Therefore, no additional class balancing techniques were applied. Table 2 presents the general characteristics of the patients.

Table 2 Baseline Characteristics of Patients

Variable	Overall (N=710)	Normal Swallowing (N=502)	Dysphagia (N=208)	P
N	710	502	208	
Sex				0.748
Male	523 (73.7%)	372 (74.1%)	151 (72.6%)	
Female	187 (26.3%)	130 (25.9%)	57 (27.4%)	
Age (years)	75 (69, 80)	75 (68, 80)	76 (70, 81)	0.103
Residence				0.124
Urban	233 (32.8%)	174 (34.7%)	59 (28.4%)	
Rural	477 (67.2%)	328 (65.3%)	149 (71.6%)	
Ethnicity				1.000
Han	702 (98.9%)	496 (98.8%)	206 (99%)	
Non-Han	8 (1.1%)	6 (1.2%)	2 (1%)	
Past history				0.540
Respiratory conditions				0.857
No	622 (87.6%)	441 (87.8%)	181 (87%)	
Yes	88 (12.4%)	61 (12.2%)	27 (13%)	
Cardiovascular diseases				0.887
No	405 (57%)	285 (56.8%)	120 (57.7%)	
Yes	305 (43%)	217 (43.2%)	88 (42.3%)	
Cerebrovascular events				0.608
No	662 (93.2%)	466 (92.8%)	196 (94.2%)	
Yes	48 (6.8%)	36 (7.2%)	12 (5.8%)	
Diabetes				0.377
No	632 (89%)	443 (88.2%)	189 (90.9%)	
Yes	78 (11%)	59 (11.8%)	19 (9.1%)	
Gastroesophageal reflux disease				0.599
No	685 (96.5%)	486 (96.8%)	199 (95.7%)	
Yes	25 (3.5%)	16 (3.2%)	9 (4.3%)	
Others				0.111
No	624 (87.9%)	448 (89.2%)	176 (84.6%)	
Yes	86 (12.1%)	54 (10.8%)	32 (15.4%)	
Marital status				0.511
Unmarried	3 (0.4%)	3 (0.6%)	0 (0%)	
Married/Cohabiting	672 (94.6%)	474 (94.4%)	198 (95.2%)	
Separated	3 (0.4%)	3 (0.6%)	0 (0%)	
Divorced	8 (1.1%)	4 (0.8%)	4 (1.9%)	
Widowed	24 (3.4%)	18 (3.6%)	6 (2.9%)	
Occupation				0.830
Farmer	506 (71.3%)	354 (70.5%)	152 (73.1%)	
Worker	77 (10.8%)	54 (10.8%)	23 (11.1%)	
Staff	64 (9%)	48 (9.6%)	16 (7.7%)	
Other	63 (8.9%)	46 (9.2%)	17 (8.2%)	

(Continued)

Table 2 (Continued).

Variable	Overall (N=710)	Normal Swallowing (N=502)	Dysphagia (N=208)	P
Education				0.595
Primary and below	513 (72.3%)	356 (70.9%)	157 (75.5%)	
Junior high school	109 (15.4%)	83 (16.5%)	26 (12.5%)	
High school	59 (8.3%)	41 (8.2%)	18 (8.7%)	
Junior College	26 (3.7%)	20 (4%)	6 (2.9%)	
College or above	3 (0.4%)	2 (0.4%)	1 (0.5%)	
Smoking				0.665
No	478 (67.3%)	335 (66.7%)	143 (68.8%)	
Yes	232 (32.7%)	167 (33.3%)	65 (31.2%)	
Alcohol consumption				0.182
No	563 (79.3%)	391 (77.9%)	172 (82.7%)	
Yes	147 (20.7%)	111 (22.1%)	36 (17.3%)	
BMI (kg/m²)	20.8 (18.7, 23.5)	21.3 (19.1, 23.7)	19.7 (17.8, 22.1)	0.000
Duration	10 (5, 10)	10 (5, 10)	10 (7, 15.25)	0.000
Respiratory rate	20 (20, 21)	20 (20, 21)	20 (19, 21)	0.385
SpO₂	96 (94, 97)	96 (94, 97)	96 (94, 97)	0.142
Missing teeth	1 (0, 5)	1 (0, 4)	2 (1, 7)	0.000
Tracheal Intubation				0.000
No	610 (85.9%)	448 (89.2%)	162 (77.9%)	
Yes	100 (14.1%)	54 (10.8%)	46 (22.1%)	
Tracheostomy				0.242
No	702 (98.9%)	498 (99.2%)	204 (98.1%)	
Yes	8 (1.1%)	4 (0.8%)	4 (1.9%)	
Dry powder inhalers				0.271
No	280 (39.4%)	205 (40.8%)	75 (36.1%)	
Yes	430 (60.6%)	297 (59.2%)	133 (63.9%)	
HNIV				0.093
No	445 (62.7%)	325 (64.7%)	120 (57.7%)	
Yes	265 (37.3%)	177 (35.3%)	88 (42.3%)	
Cough				0.137
No	118 (16.6%)	94 (18.7%)	24 (11.5%)	
Grade 1	306 (43.1%)	219 (43.6%)	87 (41.8%)	
Grade 2	128 (18%)	86 (17.1%)	42 (20.2%)	
Grade 3	101 (14.2%)	64 (12.7%)	37 (17.8%)	
Grade 4	40 (5.6%)	28 (5.6%)	12 (5.8%)	
Grade 5	17 (2.4%)	11 (2.2%)	6 (2.9%)	
Dry mouth				0.031
No	285 (40.1%)	218 (43.4%)	67 (32.2%)	
Grade 1	332 (46.8%)	226 (45%)	106 (51%)	
Grade 2	59 (8.3%)	37 (7.4%)	22 (10.6%)	
Grade 3	34 (4.8%)	21 (4.2%)	13 (6.2%)	
Muscle Strength				0.000
Grade 2	5 (0.7%)	2 (0.4%)	3 (1.4%)	
Grade 3	5 (0.7%)	2 (0.4%)	3 (1.4%)	
Grade 4	154 (21.7%)	79 (15.7%)	75 (36.1%)	
Grade 5	546 (76.9%)	419 (83.5%)	127 (61.1%)	

(Continued)

Table 2 (Continued).

Variable	Overall (N=710)	Normal Swallowing (N=502)	Dysphagia (N=208)	P
mMRC				0.000
Grade 0	87 (12.3%)	72 (14.3%)	15 (7.2%)	
Grade 1	212 (29.9%)	170 (33.9%)	42 (20.2%)	
Grade 2	169 (23.8%)	122 (24.3%)	47 (22.6%)	
Grade 3	146 (20.6%)	83 (16.5%)	63 (30.3%)	
Grade 4	96 (13.5%)	55 (11%)	41 (19.7%)	

Notes: Values are presented as n (%) for categorical variables and median (interquartile range) for continuous variables. P-values: Chi-square or Fisher's exact test for categorical variables, Mann-Whitney U-test for continuous variables. Swallowing function was assessed by the Water-Swallowing Test (WST). **Cough:** No: no cough present; Grade 1: audible airflow without cough; Grade 2: faint or barely audible cough; Grade 3: clearly audible cough; Grade 4: strong cough; Grade 5: continuous strong cough. **Dry Mouth:** No: no dry mouth; Grade 1: mild dry mouth, occasional feeling of dryness; Grade 2: moderate dry mouth, noticeable dryness during daily activities; Grade 3: severe dry mouth, significantly affects daily activities. **Muscle strength** (assessed by Manual Muscle Testing, MMT): Grade 2, movement with gravity eliminated but limited strength; Grade 3, movement against gravity but not against resistance; Grade 4, movement against resistance with reduced strength; Grade 5, normal muscle strength with full resistance. **mMRC:** Grade 0, no breathlessness except with strenuous exercise; Grade 1, dyspnea when hurrying or walking up a slight hill; Grade 2, walks slower than peers on level ground or stops due to dyspnea; Grade 3, stops for breath after walking about 100 meters or a few minutes on level ground; Grade 4, too breathless to leave the house or breathless when dressing or undressing.

Abbreviations: BMI, body mass index; HNIV, home non-invasive ventilation; mMRC, modified Medical Research Council.

Feature Selection

We treated swallowing dysfunction as the dependent variable, and univariate logistic regression was conducted for each predictor variable. Variables with a *P* value <0.1 were selected, and 17 variables were included in the multivariate analysis (Table 3). We conducted a multivariate logistic regression analysis using the forward stepwise regression method based on the Akaike Information Criterion. The final model retained five predictor variables: disease duration, BMI, tracheal intubation, muscle strength, and mMRC score (Table 4). Specifically, longer disease duration was associated with an increased risk of dysphagia (OR = 1.032, 95% CI: 1.013–1.051, *P* = 0.001). Higher BMI was identified as

Table 3 Univariate Logistic Regression Analysis

Variable	β	SE	OR (95% CI)	P
BMI	-0.114	0.025	0.892 (0.849–0.937)	0.000
Duration	0.037	0.009	1.038 (1.020–1.056)	0.000
mMRC_Grade3	1.293	0.329	3.643 (1.911–6.948)	0.000
Tracheal Intubation	0.857	0.221	2.356 (1.529–3.630)	0.000
mMRC_Grade4	1.275	0.351	3.578 (1.799–7.118)	0.000
Cough_Grade3	0.817	0.308	2.264 (1.238–4.142)	0.008
Missing teeth	0.027	0.011	1.027 (1.006–1.050)	0.013
Dry mouth_Grade1	0.423	0.183	1.526 (1.067–2.183)	0.021
Cough_Grade2	0.649	0.296	1.913 (1.070–3.418)	0.029
Dry mouth_Grade2	0.660	0.303	1.935 (1.068–3.506)	0.030
mMRC_Grade2	0.615	0.332	1.849 (0.965–3.543)	0.064
Dry mouth_Grade3	0.700	0.380	2.014 (0.957–4.238)	0.065
Age	0.018	0.010	1.018 (0.998–1.038)	0.072
HNIV	0.298	0.169	1.347 (0.968–1.874)	0.078
MMT_Grade5	-1.599	0.918	0.202 (0.033–1.223)	0.082
SpO₂	-0.046	0.027	0.955 (0.907–1.007)	0.086
Cough_Grade1	0.442	0.261	1.556 (0.932–2.598)	0.091

Notes: Univariate logistic regression analysis for dysphagia. Variables with *P* < 0.10 are shown.

Abbreviations: MMT, manual muscle testing; β , regression coefficient; SE, standard error; OR, odds ratio; CI, confidence interval.

Table 4 Multivariate Logistic Regression Analysis

Variable	β	SE	OR (95% CI)	P
Duration	0.031	0.010	1.032 (1.013–1.051)	0.001
BMI	−0.084	0.027	0.919 (0.872–0.968)	0.002
Tracheal Intubation	0.772	0.246	2.164 (1.336–3.506)	0.002
mMRC_Grade3	0.937	0.353	2.553 (1.277–5.105)	0.008
mMRC_Grade4	0.682	0.387	1.978 (0.927–4.221)	0.078
mMRC_Grade2	0.467	0.351	1.595 (0.802–3.173)	0.183
MMT_Grade5	−1.240	0.952	0.289 (0.045–1.869)	0.193
mMRC_Grade1	0.384	0.351	1.469 (0.738–2.921)	0.273
MMT_Grade4	−0.572	0.955	0.564 (0.087–3.668)	0.549
MMT_Grade3	−0.026	1.318	0.974 (0.074–12.895)	0.984

Notes: Multivariate logistic regression analysis was performed to identify independent predictors of dysphagia in COPD patients. Variable selection was based on stepwise regression using the Akaike Information Criterion ($P < 0.05$). mMRC and MMT grades were entered as categorical variables, with the lowest grade used as reference.

a protective factor (OR = 0.919, 95% CI: 0.872–0.968, $P = 0.002$). A history of tracheal intubation significantly increased the risk of dysphagia (OR = 2.164, 95% CI: 1.336–3.506, $P = 0.002$). Regarding COPD severity, compared with lower mMRC grades, mMRC grade 3 showed the strongest association with dysphagia (OR = 2.553, 95% CI: 1.277–5.105, $P = 0.008$). Although mMRC grade 2 and grade 4 were not statistically significant ($P = 0.183$ and $P = 0.078$, respectively), and muscle strength grading did not reach statistical significance (all $P > 0.05$), these variables were retained in the final model due to their contribution to overall model fit based on AIC-driven stepwise regression. Finally, these five variables were used as input features for all machine learning models.

To evaluate potential collinearity issues between variables, we calculated the Spearman correlation coefficient matrix and visualized the results using a heatmap (Figure 2). The correlation coefficients between variables were all below 0.7, suggesting that there were no significant collinearity issues and that the variables could be included in the subsequent modeling analysis.

Model Performance

On the basis of the multivariate regression analysis, we used eight machine learning models to predict the occurrence of swallowing dysfunction: LR, XGBoost, LightGBM, DT, RF, SVM, NN, and KNN. All models were evaluated using the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value, negative predictive value, and accuracy (Table 5). In the training set, all models demonstrated strong discriminative power. The XGBoost model achieved the best overall performance, with an AUC of 0.921 (95% CI: 0.901–0.940), sensitivity of 0.851, positive predictive value of 0.639, negative predictive value of 0.928, and accuracy of 0.815, all of which were the highest among the models; its specificity was 0.801. RF (AUC = 0.888) and LightGBM (AUC = 0.861) also showed strong discriminative performance. KNN (AUC = 0.811) and NN (AUC = 0.785) demonstrated moderate performance. LR (AUC = 0.703) and SVM (AUC = 0.698) showed relatively lower discriminative ability. The DT model had the lowest AUC (0.637) and the poorest sensitivity (0.375), indicating limited ability to identify positive cases. Figure 3 presents the ROC curves of all eight models. Consistent with the AUC values in Table 5, the curve of the XGBoost model is located in the upper-left region of the plot, while RF and LightGBM also lie close to this region; in contrast, the DT and SVM curves are closer to the diagonal line. Overall, ensemble tree-based methods (XGBoost, RF, and LightGBM) outperformed linear models and simpler algorithms.

We performed internal validation using the Bootstrap method with 1000 resampling iterations. Table 6 presents the performance metrics of all eight models after bootstrap validation. The mean AUC values for all eight models ranged from 0.6 to 0.7, suggesting reasonable predictive stability. Among all models, LR achieved the highest validation AUC (0.684), followed by XGBoost (0.664), LightGBM (0.646), NN (0.646), RF (0.644), SVM (0.617), DT (0.608), and KNN (0.598). Compared with the training performance (Table 5), all models showed a decline in AUC. In particular, the

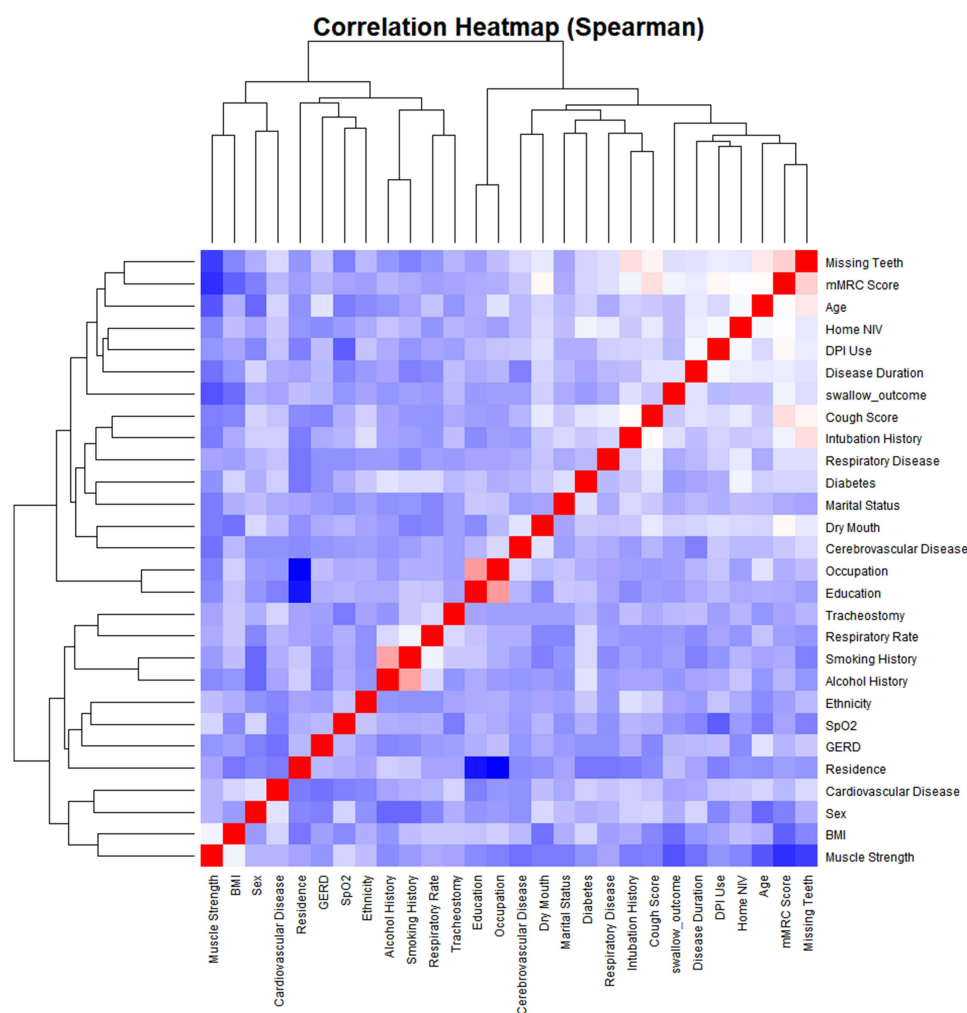


Figure 2 Correlation Analysis Heatmap. The heatmap displays pairwise correlation coefficients between all input features. Red indicates positive correlation, blue indicates negative correlation, and the intensity of the color represents the strength of the correlation. Darker colors correspond to higher absolute correlation values.

XGBoost model decreased from 0.921 in the training set to 0.664 in the validation set ($\Delta = 0.257$), indicating potential overfitting. RF ($\Delta = 0.244$) and LightGBM ($\Delta = 0.215$) also exhibited substantial performance degradation. In contrast, LR demonstrated relatively stable performance in the validation set ($\Delta = 0.019$), outperforming most ensemble and

Table 5 Machine Learning Modeling of Dysphagia for Patients with COPD

Model	AUC	Sensitivity	Specificity	PPV	NPV	Accuracy
XGBoost	0.921	0.851	0.801	0.639	0.928	0.815
RF	0.888	0.803	0.811	0.637	0.908	0.808
LightGBM	0.861	0.712	0.819	0.619	0.873	0.787
KNN	0.811	0.774	0.705	0.521	0.883	0.725
NN	0.785	0.721	0.707	0.505	0.860	0.711
LR	0.703	0.697	0.635	0.442	0.835	0.654
SVM	0.698	0.591	0.779	0.526	0.821	0.724
DT	0.637	0.375	0.873	0.549	0.771	0.727

Notes: Performance of eight machine learning models for dysphagia prediction in COPD was evaluated in the training set. All models were developed using the five selected features from Table 3. **Abbreviations:** XGBoost, Extreme Gradient Boosting; RF, random forest; LightGBM, Light Gradient Boosting Machine; KNN, k-nearest neighbors; NN, neural network; LR, logistic regression; SVM, support vector machine; DT, decision tree; AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value.

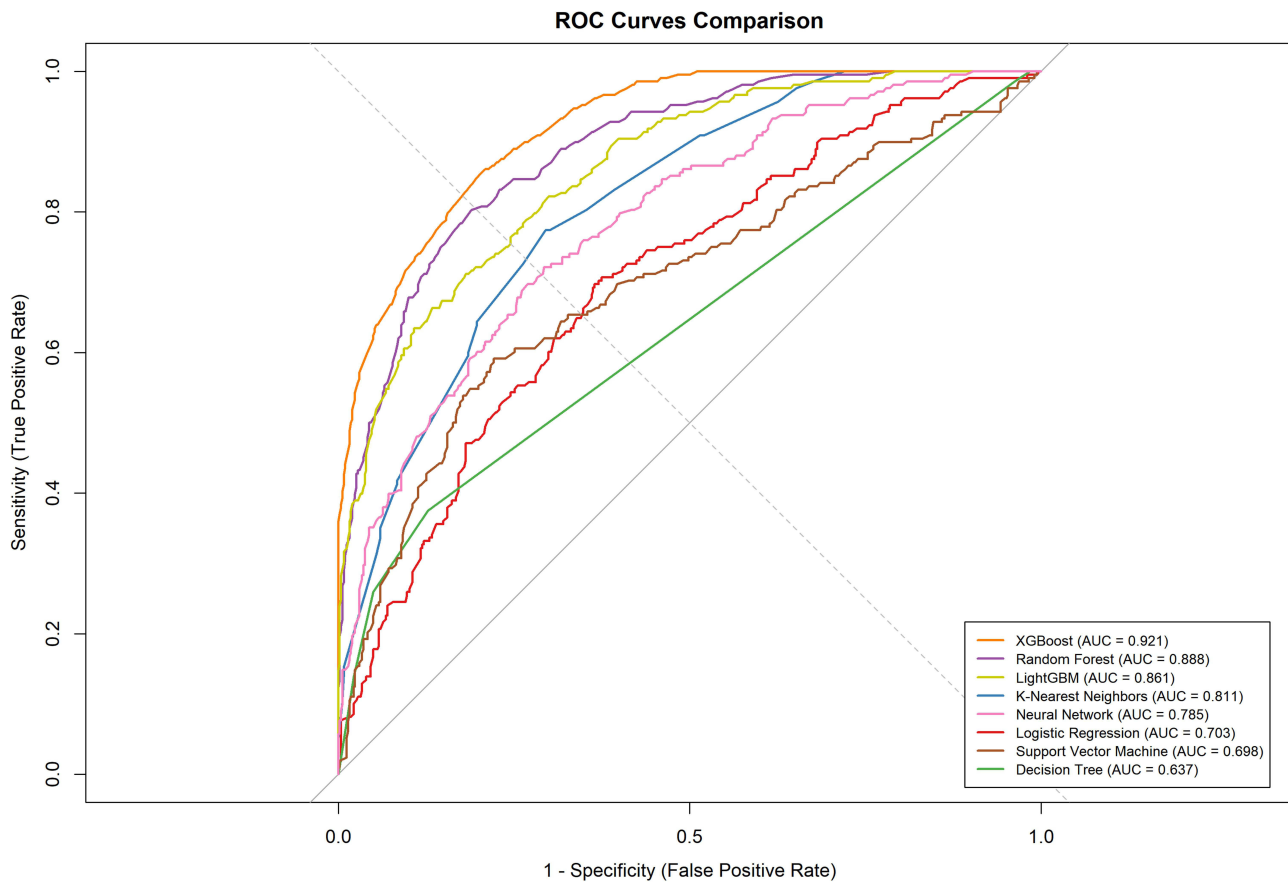


Figure 3 ROC curves of the eight machine learning models in the training set. The curves compare the discriminative performance of each model for predicting swallowing dysfunction in patients with COPD.

nonlinear models in terms of stability. Figure 4 illustrates the distribution of AUC values during bootstrap validation for each model. The AUC distribution of LR was the narrowest and most concentrated, whereas those of XGBoost, RF, and LightGBM were wider and more dispersed, reflecting greater variability across bootstrap samples, which is consistent with their observed overfitting behavior.

The calibration curves demonstrated that most machine learning models showed good agreement between predicted and observed probabilities of dysphagia in COPD patients (Figure 5). Specifically, the calibration curves of the DT, XGBoost, LightGBM, and NN models were generally close to the ideal diagonal reference line, indicating good

Table 6 Internal Validation Results of Machine Learning Models for Dysphagia for Patients with COPD

Model	AUC	Sensitivity	Specificity	PPV	NPV	Accuracy
LR	0.684	0.669	0.653	0.451	0.829	0.658
XGBoost	0.664	0.610	0.683	0.454	0.811	0.662
LightGBM	0.646	0.608	0.647	0.434	0.805	0.635
NN	0.646	0.595	0.673	0.439	0.803	0.650
RF	0.644	0.620	0.635	0.425	0.806	0.631
SVM	0.617	0.539	0.713	0.447	0.790	0.662
DT	0.608	0.463	0.753	0.452	0.774	0.668
KNN	0.598	0.570	0.616	0.399	0.781	0.602

Notes: Internal validation was conducted using bootstrap resampling with 1000 iterations to assess the performance of eight machine learning models.

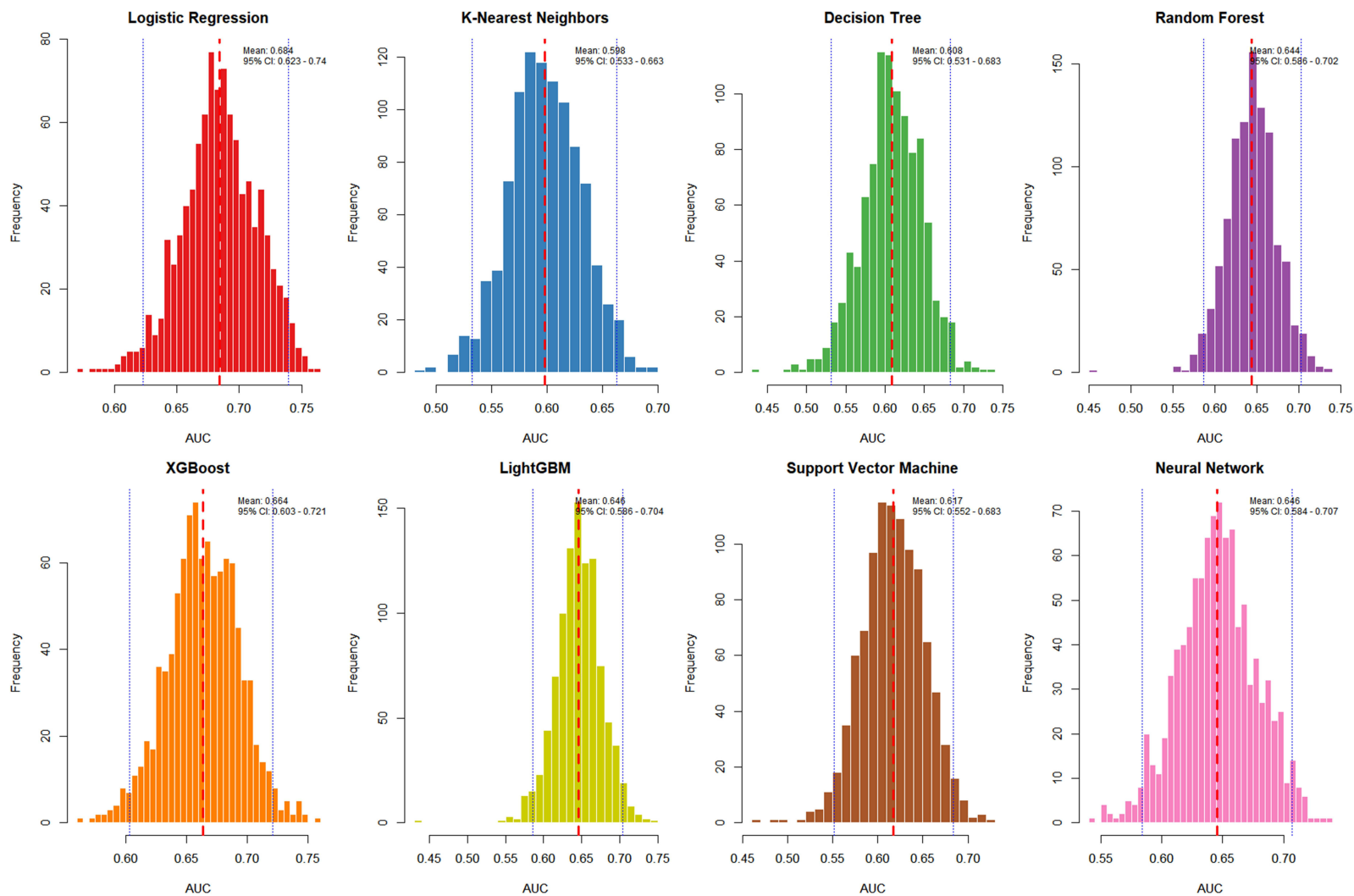


Figure 4 Distribution of the AUC values obtained from 1000 Bootstrap internal validation repetitions for the eight machine learning models. The red dashed line represents the mean AUC, while the blue dotted lines indicate the corresponding 95% confidence interval.

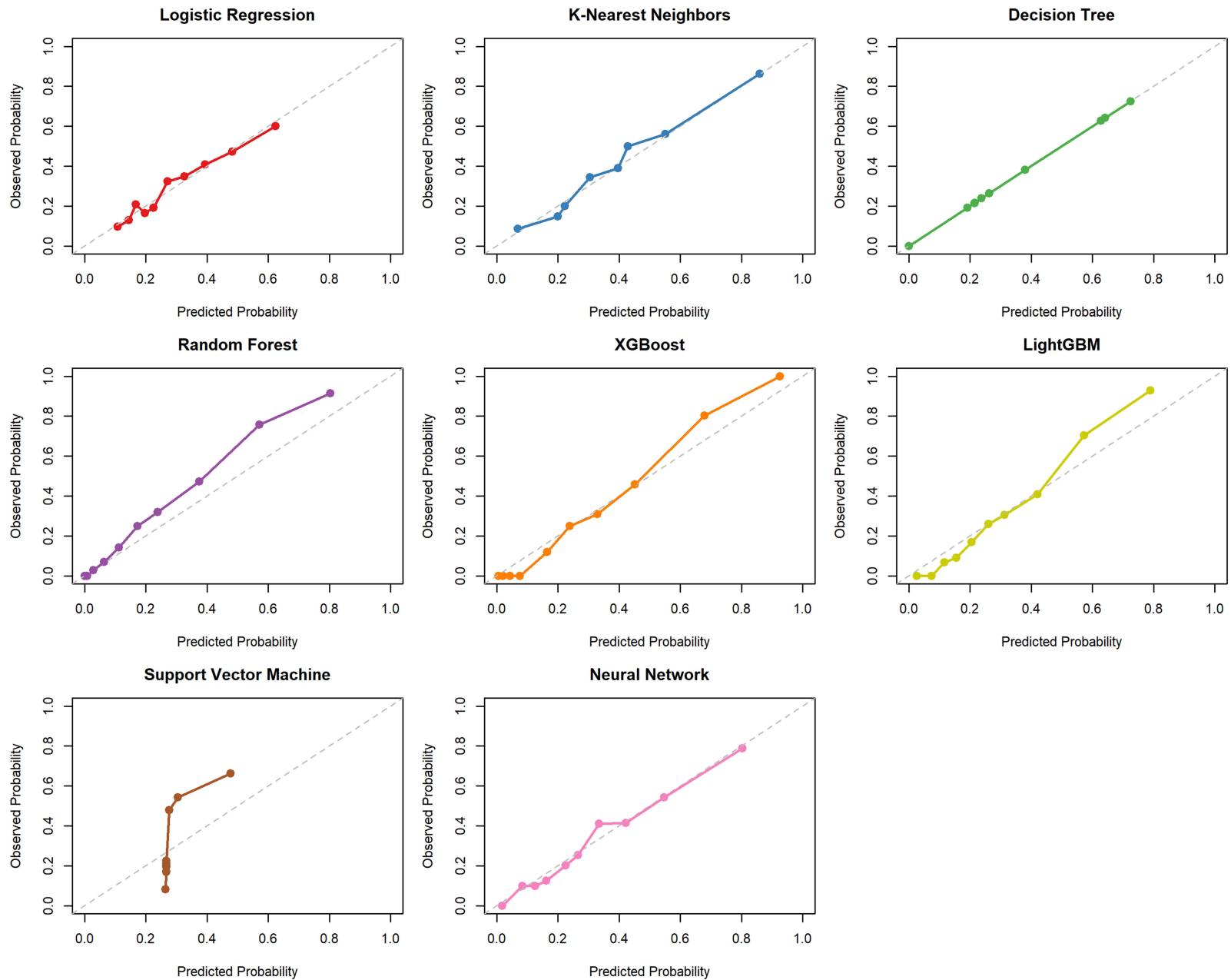


Figure 5 Calibration curves of the eight machine learning models showing the agreement between predicted probabilities and observed outcomes of swallowing dysfunction in patients with COPD. The grey dashed line indicates perfect calibration, where predicted probabilities exactly match observed outcomes. The coloured lines depict the calibration curves of the individual machine learning models.

calibration performance and reliable probability estimation. LR and KNN models also showed relatively good calibration, although slight deviations from the reference line were observed in some probability ranges. RF model tended to overestimate the risk at higher predicted probabilities, whereas the SVM model showed the poorest calibration performance, with substantial deviation from the diagonal line, suggesting limited accuracy in probability estimation.

Decision curve analysis demonstrated that most machine learning models provided greater net benefit than the “treat-all” and “treat-none” strategies across a wide range of threshold probabilities (Figure 6). Among all models, XGBoost consistently achieved the highest net benefit over most threshold ranges, indicating superior clinical utility in identifying dysphagia risk in COPD patients. RF and LightGBM also showed relatively favorable decision performance, with net benefit curves remaining above those of several other models across moderate to high threshold probabilities. LR model showed moderate clinical utility at lower threshold probabilities; however, its net benefit declined progressively as the threshold increased, and it was outperformed by most machine learning models in the mid to high threshold range. In contrast, the SVM model showed comparatively limited net benefit, particularly at higher threshold probabilities. These findings suggest that XGBoost may provide the greatest clinical decision support value for dysphagia risk assessment.

Based on overall performance, the XGBoost model achieved the highest AUC in the training set (0.921), while in the validation set its AUC was 0.664, ranking second to LR (0.684), with a minimal difference ($\Delta = 0.02$). Given the small difference in AUC between models, further model comparison was guided by calibration curves and decision curve analysis. The XGBoost model demonstrated good calibration, with the curve closely approximating the diagonal line, and showed the highest net clinical benefit in DCA. Therefore, considering its overall superior performance across multiple metrics, the XGBoost model was selected as the final predictive model.

Model Interpretability

We ranked the importance of the prediction features in XGBoost model. On the basis of the average absolute SHAP values, we found that the top five most important features, ie, BMI, disease duration, mMRC, tracheal intubation history, and muscle strength, were the most influential variables (Figure 7). Figure 8 illustrates the detailed relation between each feature and the occurrence of swallowing dysfunction. The features positively correlated with swallowing dysfunction

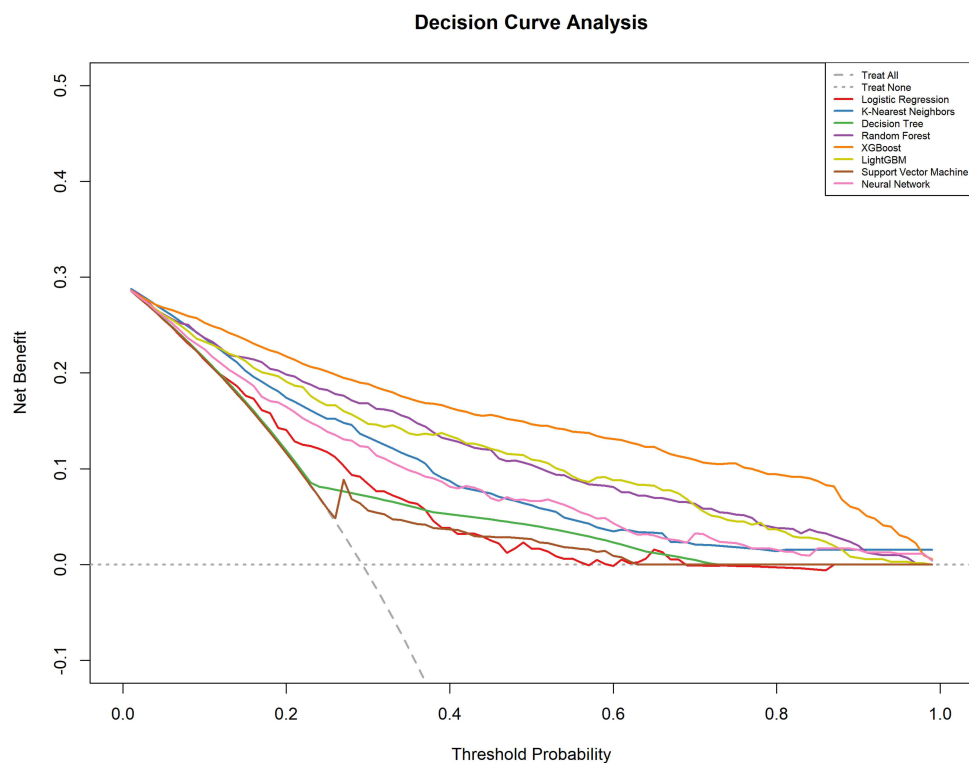


Figure 6 Decision curve analysis of the eight machine learning models. The curves demonstrate the net clinical benefit of each model across different threshold probabilities compared with the “treat-all” and “treat-none” strategies.

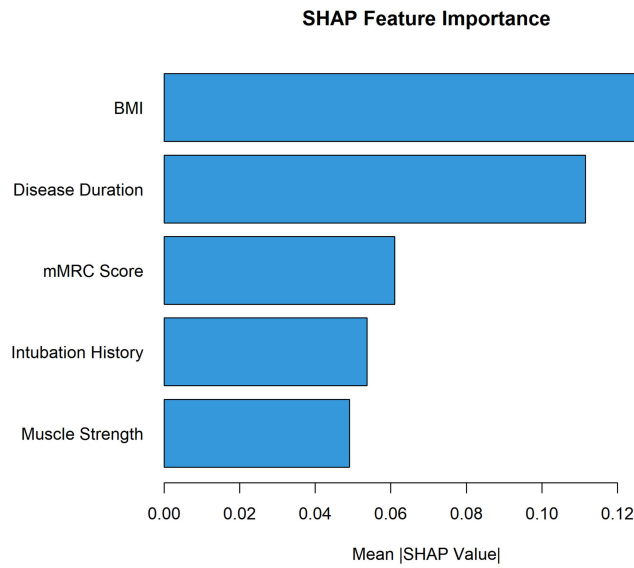


Figure 7 Ranking of feature importance in the optimal machine learning model, demonstrating the relative contribution of each predictor variable to the prediction of swallowing dysfunction risk.

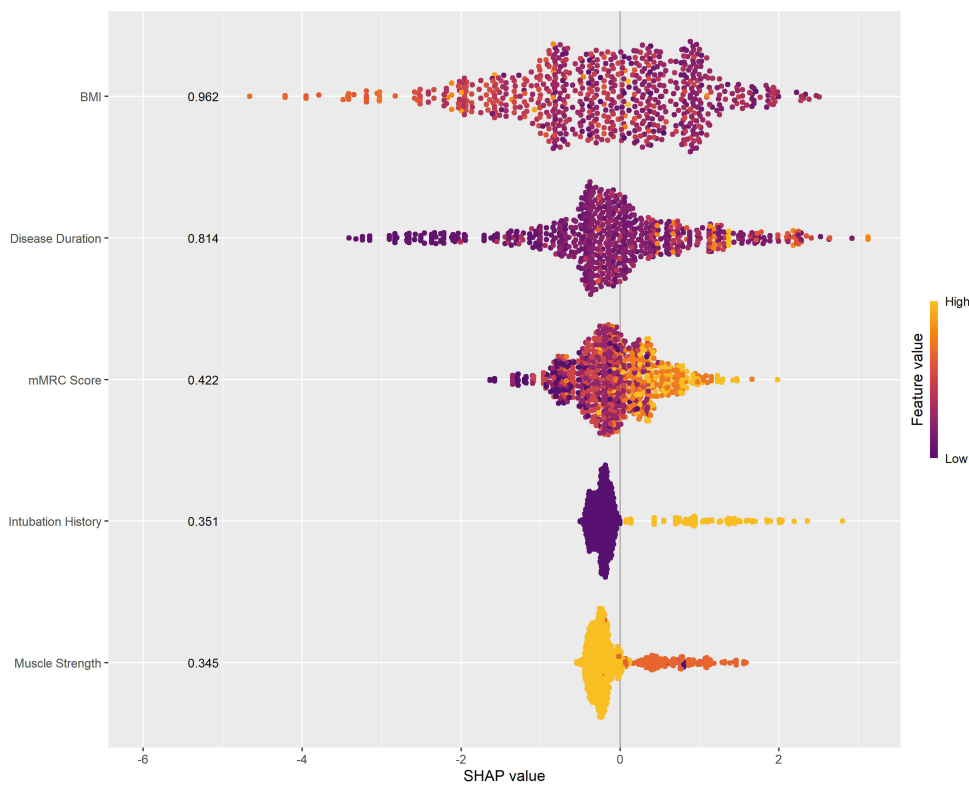


Figure 8 SHAP bee swarm plot illustrating the impact and distribution of individual feature contributions to the XGBoost model predictions. Each point represents one patient sample, and color indicates the feature value magnitude.

were disease duration, mMRC, and tracheal intubation history. As the values of these features increased (or whether they were “present”), the sample points were more concentrated in the positive region of the SHAP axis. This finding meant that these features drive the model’s predictions in a positive direction, thus increasing the likelihood of swallowing dysfunction.

Developing a Dynamic Calculator

On the basis of the five key variables, we developed a deployable web platform using Shiny to visualize and use the prediction model (<https://dysphagiamodel.shinyapps.io/COPD-DP/>). This platform provides an online tool for assessing dysphagia risk for patients with COPD. Users can easily obtain risk assessment results by entering clinical feature data into the specified text fields on the webpage. (Figure 9)

Discussion

We found that the prevalence of swallowing dysfunction for patients with COPD was 29.3%, which was comparable to the approximately 33% prevalence reported in a meta-analysis of observational studies conducted by Li et al,⁶ suggesting that the present study population was reasonably representative. We developed and validated a web-based interactive tool for risk assessment. We used eight different machine learning algorithms to build the models, with XGBoost performing the best. Using the SHAP method for model interpretation, we identified the top five key variables most strongly associated with swallowing dysfunction: BMI, disease duration, mMRC score, tracheal intubation history, and muscle strength. On the basis of this model, we created an online risk calculator that provides real-time, personalized assessments.

Currently, research on predicting swallowing dysfunction for patients with COPD is relatively limited; most studies rely on traditional statistical methods to develop assessment models. For example, in a study from Guangdong Provincial Hospital of Chinese Medicine, Xiong et al included 100 AECOPD patients and used LASSO regression to identify four predictors for swallowing dysfunction: age, mMRC score, hospitalization days, and the use of BIPAP-assisted ventilation. The model achieved a high discriminative ability with an AUC of 0.909.⁷ In another study involving 405 patients with COPD, Fan et al used logistic regression to develop a swallowing dysfunction prediction model that included factors such as age, cerebrovascular disease, chronic pulmonary heart disease, home non-invasive ventilation, acute exacerbations, dyspnea, and dry mouth. This model had an AUC of 0.879, and the results were presented in a visual Nomogram format.¹⁸ Additionally, a recent study of elderly patients with COPD by Chen et al showed that a decision tree-based model outperformed the traditional logistic regression model in prediction performance (AUC 0.747 vs. 0.682).²⁰ Although these studies provide valuable insights into the risk assessment of swallowing dysfunction in patients with COPD, most of them were based primarily on single statistical models and may therefore have limited ability to capture complex nonlinear relationships and interactions among clinical variables.⁴¹ In contrast, we systematically compared various machine learning algorithms in a larger patient cohort, and the results demonstrated that the XGBoost model achieved the best predictive performance, suggesting that machine learning approaches may offer greater potential for swallowing dysfunction risk identification in patients with COPD.

In the feature contribution analysis, BMI ranked first in the XGBoost model, highlighting the central effect of nutritional status in developing swallowing dysfunction. As COPD progresses and becomes more severe, patient BMI tends to decrease progressively.^{42,43} This change not only indicates malnutrition but also the change serves as a key marker of disease progression and poor prognosis. Several previous studies have similarly reported that low BMI or malnutrition is an important influencing factor for swallowing dysfunction in patients with COPD. These findings are generally consistent with our results and further support the importance of BMI in swallowing risk assessment among patients with COPD.^{20,44} Moreover, low BMI is often linked with a systemic inflammatory state driven by chronic inflammation and abnormal repair mechanisms in COPD. As described by Barnes et al, persistent inflammatory mediators can circulate through the body and affect peripheral tissues, leading to skeletal muscle dysfunction and metabolic disturbances.⁴⁵ Additionally, COPD-related metabolic changes involve abnormalities in glucose and lipid metabolism; these abnormalities sustain chronic inflammation and exacerbate muscle wasting.^{46,47}

Our SHAP analysis also identified muscle strength as a significant factor, ranking 5th among all variables. Beyond overall nutritional status, muscle function has a crucial effect in swallowing safety. Jones et al report that low BMI is frequently associated with decreased muscle mass and strength, with about 15% of stable COPD patients showing varying degrees of muscle loss.⁴⁸ Swallowing is a complex process that depends on the fine coordination of multiple muscle groups. As described in detail by Garand et al, when skeletal muscle atrophy and weakness affect swallowing-

COPD Dysphagia Predictor
☰

Prediction Calculator

User Guide

About

Enter patient information on the right panel to calculate dysphagia risk.

Patient Information

BMI (kg/m²)

Duration of COPD (years)

Dyspnea Severity (mMRC)

History of Tracheal Intubation

Muscle Strength Grade

📄 Calculate Risk Probability

Prediction Result

71.6%
High Risk

Interpretation ⓘ

Based on the patient information provided, the predicted probability of dysphagia is 71.6%. Key risk factors include: history of tracheal intubation.

Recommendations 📄

- Recommend immediate professional swallowing assessment
- Consider rehabilitation consultation
- Modify diet texture and eating methods
- Increase supervision during meals
- Prevent aspiration-related complications

Risk Level Description

Low Risk

Predicted probability < 30%

Recommendation: Routine care

Medium Risk

Predicted probability 30-60%

Recommendation: Enhanced monitoring

High Risk

Predicted probability > 60%

Recommendation: Immediate assessment

Figure 9 A screenshot of the Online prediction tool page. Users can input patient characteristics into the designated fields, and the tool automatically calculates and displays the predicted risk probability based on the selected machine learning model.

related muscles, the weakness can limit muscle power and endurance, leading to reduced swallowing efficiency, impaired respiratory-swallowing coordination, and an increased risk of swallowing dysfunction.⁴⁴ Therefore, the relations between BMI, nutritional status, and muscle strength form a risk pathway with biological relevance. This relationship explains the high importance of BMI in the predictive model and underscores the need to consider muscle strength as a critical functional indicator in swallowing dysfunction risk assessment. The importance of BMI also suggests that, in managing swallowing risk for patients with COPD, both nutritional assessment and muscle strength monitoring should be integrated into a comprehensive, dynamic framework for nursing interventions.

The duration of illness, identified as the second most important factor in this study, suggests that the long-term progression of COPD may have a significant effect in swallowing dysfunction. As COPD advances, the gradual worsening of lung hyperinflation and dyspnea can alter intrathoracic pressure, increasing the risk of respiratory-swallowing coordination problems.^{9,49} Additionally, prolonged airflow limitation keeps respiratory muscles under constant strain, which can lead to muscle fatigue and reduced neuromuscular coordination, further worsening swallowing function.^{50,51} Patients with COPD generally have impaired swallowing function compared with healthy individuals,^{52,53} and this dysfunction worsens as the disease progresses,⁵⁴ a fact for which our study agrees.

The mMRC score, commonly used to assess the severity of dyspnea, also proved to be a valuable assessment tool in our study. According to the logistic regression analysis, patients with an mMRC score ≥ 3 who show clear respiratory limitations in daily activities may have persistent dyspnea that forces them to prioritize ventilation during eating and swallowing, disrupting the balance between breathing and swallowing. An increased respiratory rate and stronger inspiratory drive can reduce the swallowing pause time and increase the likelihood of overlap between swallowing and inspiration,^{8,10} which significantly elevates the risk of swallowing dysfunction. Additionally, patients with severe dyspnea often experience fatigue while eating, resulting in faster eating, insufficient preparation for swallowing, or forced interruptions in their meals; these compensatory behaviors may further weaken the coordination and safety of swallowing.⁵⁵ An epidemiological study by Gonzalez et al showed a consistent increase in the reporting of swallowing difficulties among patients with COPD as their dyspnea worsened.⁵⁶ On the basis of these mechanisms and the results of this study, we suggest that the mMRC score is not only a marker of respiratory limitation, but also the score is an important clinical indicator for identifying patients at high risk for swallowing dysfunction.

We recognize that whether patients with COPD have undergone tracheal intubation may exert a significant influence on their swallowing function. Tracheal intubation can cause mechanical irritation and mucosal injury that affects the vocal cords and laryngeal structures,^{57,58} potentially leading to reduced sensation in the larynx and a delayed swallowing reflex that weakens airway protection during swallowing. Additionally, Wallace and McGrath reported that any damage to the recurrent laryngeal nerve or sensory pathways during intubation may persist after extubation, making patients more susceptible to swallowing dysfunction.⁵⁹ Studies of critically ill patients and COPD populations have also shown a strong link between a history of tracheal intubation and the development of swallowing dysfunction after extubation, indicating that the impact on swallowing safety may be long-lasting.^{7,60} For patients with COPD, a history of tracheal intubation often suggests a previous severe illness or acute exacerbation. These patients are already at risk of impaired coordination between breathing and swallowing, and the added effects of intubation-related damage may further increase their swallowing dysfunction risk. Moreover, some patients may develop compensatory swallowing behaviors or even fear of swallowing following intubation and extubation, which can disrupt normal swallowing patterns.¹¹ Our findings emphasize that a history of tracheal intubation should be viewed as a critical marker for identifying high-risk patients during nursing assessments, and more attention should be given to swallowing management and monitoring for patients with an intubation history.

In the evaluation of model performance, we systematically compared eight machine learning algorithms. Although the XGBoost model showed strong performance in the training set, its performance significantly declined during the Bootstrap internal validation. Several factors may have contributed to this decline. First, overfitting was likely a major factor. XGBoost, being a robust machine learning algorithm, can fit the training data very accurately, including noise and outliers; as a result, the model overfits the details of the training data, which reduces its ability to generalize to new data.⁶¹ Additionally, the small patients sample size may have limited the model's ability to learn the full diversity of the data, impacting its performance on unseen data.⁶² Data imbalance could also result in the model's inability to effectively

predict the minority class, negatively affecting performance on the validation set.⁶³ Other factors that might have contributed to the performance drop are feature redundancy or noise,⁶⁴ inadequate model parameter tuning,⁶⁵ and differences in data distributions between the training and validation sets during the Bootstrap process.⁶⁶ Therefore, to improve the model's generalization capabilities, further work should focus on increasing the sample size, addressing data imbalance, reducing feature redundancy, refining parameter tuning, and using more stable validation techniques.

Based on the developed prediction model, we further established a web-based interactive risk assessment tool to enhance its applicability in clinical nursing settings. The interactive web interface enables healthcare professionals to input relevant patient data and quickly generate personalized risk assessments for swallowing dysfunction. This approach may assist nursing staff in rapidly identifying high-risk patients at the bedside and facilitate the implementation of early interventions. This ability not only helps healthcare providers identify high-risk patients more efficiently, but also the tool supports early interventions in clinical practice.⁶⁷ For patients identified by the model as high-risk, targeted interventions such as swallowing exercises and respiratory training can be implemented to prevent aspiration and choking. In addition, individualized dietary management strategies may also be introduced to further reduce the risk of aspiration, aspiration pneumonia, and choking-related adverse events. Furthermore, the online tool enables patients to receive real-time swallowing dysfunction risk assessments, raising their awareness of their condition. This patient education feature improves understanding and adherence to swallowing function management plans and encourages active patient participation in personalized treatments and interventions, ultimately enhancing overall treatment experience and prognosis. Overall, transforming machine learning models into visualized web-based tools may facilitate the translation of swallowing dysfunction risk assessment from research settings into routine clinical practice, thereby providing new technical support for precision nursing in patients with COPD.

The strengths of this study lie in the development of a comprehensive and clinically applicable risk prediction framework for swallowing dysfunction in patients with COPD. We integrated traditional regression analysis with eight machine learning algorithms and systematically compared their performance, ensuring a robust methodological foundation beyond single-model approaches. The optimal model was further interpreted using SHAP analysis, which enhanced transparency by identifying key contributors such as BMI, disease duration, mMRC score, history of tracheal intubation, and muscle strength, thereby improving clinical interpretability. In addition, the best-performing model was transformed into a user-friendly web-based interactive tool, facilitating rapid individualized risk assessment and supporting bedside clinical decision-making. Importantly, this study addresses a clinically significant yet often under-recognized complication in COPD, and the proposed framework holds potential value for early identification, risk stratification, and timely intervention in high-risk patients, ultimately contributing to improved swallowing safety and clinical outcomes.

This study has several limitations that should be considered when interpreting the findings. First, it is a single-center retrospective study, and the accuracy and completeness of the data depend on the original medical records, which may be subject to information bias. In addition, the study sample was drawn from a single hospital, and the case composition may be subject to selection bias, limiting the generalizability of the model, which requires external validation in multicenter studies. Second, owing to clinical feasibility constraints, swallowing function was assessed using bedside screening tools rather than gold-standard instrumental examinations, which may have introduced a certain degree of misclassification bias. Third, the assessment of certain predictors was limited; COPD severity was evaluated using the mMRC dyspnea scale without incorporating pulmonary function parameters, and medication use data were restricted to the history of dry powder inhaler use and home non-invasive ventilation, lacking detailed information on specific drug types, dosages, and systemic corticosteroid exposure. These missing data may constitute residual confounding factors, potentially affecting the study results. Finally, although an online risk prediction tool was developed in this study, its clinical utility has not yet been externally validated or prospectively assessed. Future multicenter prospective studies using standardized pulmonary function testing and comprehensive medication data are warranted to further verify the robustness and clinical applicability of the model.

Conclusion

We successfully developed an interpretable machine learning model to evaluate swallowing function for patients with COPD and created an online calculator for easy clinical use. The model is capable of identifying high-risk patients, significantly improving the early screening efficiency for dysphagia. Compared with traditional risk assessment methods, it integrates multiple clinically relevant predictors and uses interpretable algorithms to reveal key influencing factors,

highlighting its novelty. Clinical application of this model can assist healthcare professionals in implementing early interventions, reducing the risk of aspiration and malnutrition, and providing important support for individualized care and clinical decision-making in patients with COPD. The findings of this study enrich the toolkit for dysphagia risk assessment in COPD patients and offer a valuable reference for future research and clinical practice.

Generative AI Statement

During the preparation of this manuscript, the authors consulted ChatGPT to obtain suggestions for refining the title. The authors reviewed and edited the output and take full responsibility for the final content of the manuscript.

Data Sharing Statement

The dataset generated and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

The Ethics Committee of Jinhua University of Vocational Technology approved the study protocol (202502) and performed in accordance with the Declaration of Helsinki.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Detailed contributions are as follows: Conceptualization, Chiteng Zhou and Huiying Pan; Methodology, Chiteng Zhou and Shuangwei Hong; Software, Chiteng Zhou and Shuangwei Hong; Formal analysis, Chiteng Zhou and Huiying Pan; Investigation, Min Zang, Jun Fang, and Ningfei Tang; Data curation, Ximei Yan and Lingwei Bao; Writing – original draft, Chiteng Zhou and Shuangwei Hong; Writing – review & editing, Shuangwei Hong and Huiying Pan; Funding acquisition, Chiteng Zhou.

Funding

This study was funded by the Jinhua City Public Welfare Technology Application Research Program (Grant No. 2024-4-033). The authors declare that no other funds, grants, or other support were received during the preparation of this manuscript.

Disclosure

The authors declare no competing interests.

References

1. Zhou H, Dong Z, Ye X. Global and national burden of chronic obstructive pulmonary disease and tracheal, bronchus, and lung cancer from 1990 to 2021: comorbidity burden analysis based on the global burden of disease study 2021. *Cancer Control*. 2026;33:10732748251407363. doi:10.1177/10732748251407363
2. Li Y, Sun P, Yin Y, et al. Global, regional, and national burden attributed to particulate matter pollution, 1990–2021: a systematic analysis for the global burden of disease study 2021. *Annals Global Health*. 2026;92(1):22. doi:10.5334/aogh.4965
3. Madani NA, Carpenter DO. Patterns of emergency room visits for respiratory diseases in New York State in relation to air pollution, poverty and smoking. *Int J Environ Res Public Health*. 2023;20(4):3267. doi:10.3390/ijerph20043267
4. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet*. 2020;396(10258):1204–1222. doi:10.1016/S0140-6736(20)30925-9
5. Yi FL, Yi ST. Epidemiological investigation and prevention measures of chronic obstructive pulmonary disease. *J Prevent Med Chin People's Lib Army*. 2018;36(02):171–3+80.
6. Li W, Gao M, Liu J, et al. The prevalence of oropharyngeal dysphagia in patients with chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Expert Rev Respir Med*. 2022;16(5):567–574. doi:10.1080/17476348.2022.2086123
7. Xiong S, Zhou Y, He W, et al. Study on predictive models for swallowing risk in patients with AECOPD. *BMC Pulm Med*. 2024;24(1):95. doi:10.1186/s12890-024-02908-y

8. Nagami S, Oku Y, Yagi N, et al. Breathing-swallowing discoordination is associated with frequent exacerbations of COPD. *BMJ Open Resp Res.* 2017;4(1):e000202. doi:10.1136/bmjresp-2017-000202
9. Yoshimatsu Y, Tobino K, Nagami S, et al. Breathing-swallowing discoordination and inefficiency of an airway protective mechanism puts patients at risk of COPD exacerbation. *Int J Chronic Obstr.* 2020;15:1689–1696. doi:10.2147/COPD.S257622
10. Cvejic L, Bardin PG. Swallow and aspiration in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 2018;198(9):1122–1129. doi:10.1164/rccm.201804-0704PP
11. Lin TF, Shune S. Chronic obstructive pulmonary disease and dysphagia: a synergistic review. *Geriatrics.* 2020;5(3):45. doi:10.3390/geriatrics5030045
12. Nativ-Zeltzer N, Nachalon Y, Kaufman MW, et al. Predictors of aspiration pneumonia and mortality in patients with dysphagia. *Laryngoscope.* 2022;132(6):1172–1176. doi:10.1002/lary.29770
13. Poulsen SH, Rosenvinge PM, Modlinski RM, et al. Signs of dysphagia and associated outcomes regarding mortality, length of hospital stay and readmissions in acute geriatric patients: observational prospective study. *Clin Nutr ESPEN.* 2021;45:412–419. doi:10.1016/j.clnesp.2021.07.009
14. de Deus Chaves R, Chiarion Sassi F, Davison mangilli L, et al. Swallowing transit times and valleculae residue in stable chronic obstructive pulmonary disease. *BMC Pulm Med.* 2014;14:62. doi:10.1186/1471-2466-14-62
15. Langen RC, Gosker HR, Remels AH, et al. Triggers and mechanisms of skeletal muscle wasting in chronic obstructive pulmonary disease. *Int J Biochem Cell Biol.* 2013;45(10):2245–2256. doi:10.1016/j.biocel.2013.06.015
16. Epiu I, Jenkins CR, Bulamu NB, et al. Cost effectiveness of a novel swallowing and respiratory sensation assessment and a modelled intervention to reduce acute exacerbations of COPD. *BMC Pulm Med.* 2025;25(1):165. doi:10.1186/s12890-025-03615-y
17. Yoshimatsu Y, Tobino K, Sueyasu T, et al. Repetitive saliva swallowing test and water swallowing test may identify a COPD phenotype at high risk of exacerbation. *Clin Resp J.* 2019;13(5):321–327. doi:10.1111/crj.13014
18. Fan Y, Shi Y, Wu Y, et al. A nomogram-based prediction model for dysphagia in patients with chronic obstructive pulmonary disease: a cross-sectional study. *J Clin Nurs.* 2025;34(4):1325–1337. doi:10.1111/jocn.17208
19. Su S, Su Q, Zhong H, et al. Construction and validation of an ultrasound-based nomogram model for predicting dysphagia in patients with chronic obstructive pulmonary disease. *Front Med.* 2025;12:1533165. doi:10.3389/fmed.2025.1533165
20. Chen YP, Ding MZ, Shao SY, et al. Construction of risk prediction model of dysphagia in elderly patients with COPD. *Chin Nurs Res.* 2025;39(2):204–210.
21. Lienhart AM, Kramer D, Jauk S, et al. Multivariable risk prediction of dysphagia in hospitalized patients using machine learning. *Stud Health Technol Inform.* 2020;271:31–38. doi:10.3233/SHTI200071
22. Karabacak M, Margetis K. A machine learning-based online prediction tool for predicting short-term postoperative outcomes following spinal tumor resections. *Cancers.* 2023;15(3):812. doi:10.3390/cancers15030812
23. Ye F, Cheng LL, Li WM, et al. A machine-learning model based on clinical features for the prediction of severe dysphagia after ischemic stroke. *Int J Gene Med.* 2024;17:5623–5631. doi:10.2147/IJGM.S484237
24. Sakai K, Gilmour S, Hoshino E, et al. A machine learning-based screening test for sarcopenic dysphagia using image recognition. *Nutrients.* 2021;13(11):4009. doi:10.3390/nu13114009
25. Rajpoot NK, Singh PD, Pant B. A novel framework for COPD management in cyber-physical systems using machine learning. *Sci Rep.* 2025;15(1):36517. doi:10.1038/s41598-025-08932-0
26. Agrawal R, Gupta T, Gupta S, et al. Fostering trust and interpretability: integrating explainable AI (XAI) with machine learning for enhanced disease prediction and decision transparency. *Diagn Pathol.* 2025;20(1):105. doi:10.1186/s13000-025-01686-3
27. Feng G, Xu H, Wan S, et al. Twelve practical recommendations for developing and applying clinical predictive models. *Innovation Med.* 2024;2(4):100105. doi:10.59717/j.xinn-med.2024.100105
28. Liang W, Liang H, Ou L, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Intern Med.* 2020;180(8):1081–1089. doi:10.1001/jamainternmed.2020.2033
29. Sankar A, Beattie WS, Wijeyesundera DN. How can we identify the high-risk patient? *Current Opinion Crit Care.* 2015;21(4):328–335. doi:10.1097/MCC.0000000000000216
30. Soleimanpour N, Bann M. Clinical risk calculators informing the decision to admit: a methodologic evaluation and assessment of applicability. *PLoS One.* 2022;17(12):e0279294. doi:10.1371/journal.pone.0279294
31. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the prevention, diagnosis, and management of COPD: 2026 report; 2025. Available from: https://goldcopd.org/wp-content/uploads/2026/01/GOLD-REPORT-2026-v1.3-8Dec2025_WMV2.pdf. Accessed February 19, 2026.
32. Royston P, Moons KG, Altman DG, et al. Prognosis and prognostic research: developing a prognostic model. *BMJ.* 2009;338:b604. doi:10.1136/bmj.b604
33. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ.* 2020;368:m441. doi:10.1136/bmj.m441
34. Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med.* 2019;38(7):1276–1296. doi:10.1002/sim.7992
35. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol.* 1996;49(12):1373–1379. doi:10.1016/S0895-4356(96)00236-3
36. O'Rourke F, Vickers K, Upton C, et al. Swallowing and oropharyngeal dysphagia. *Clin Med.* 2014;14(2):196–199. doi:10.7861/clinmedicine.14-2-196
37. Le KHN, Low EE, Yadlapati R. Evaluation of esophageal dysphagia in elderly patients. *Current Gastroenterol Rep.* 2023;25(7):146–159. doi:10.1007/s11894-023-00876-7
38. Ji Q, Han M, Xu K, et al. Research on the application of feedback method combined with diversified health education in elderly patients with COPD complicated with dysphagia. *Sci Rep.* 2025;16(1):3658. doi:10.1038/s41598-025-33757-2
39. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell.* 2020;2(1):56–67. doi:10.1038/s42256-019-0138-9
40. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ.* 2024;385:e078378. doi:10.1136/bmj-2023-078378

41. Jaki T, Chang C, Kuhlemeier A, et al. Predicting individual treatment effects: challenges and opportunities for machine learning and artificial intelligence. *Kunstliche Intell.* 2025;39(1):27–32. doi:10.1007/s13218-023-00827-4
42. Ansari MS. BMI as a marker of severity in patients with COPD. *J Evid Based Med Healthc.* 2017;4:884–886.
43. Javed I, Javed A, Zafar Z, et al. Relationship of BMI with severity of chronic obstructive pulmonary disease (COPD): body mass index with chronic obstructive pulmonary disease. *Pakistan J Health Sci.* 2025;6(3):206–210.
44. Garand KL, Strange C, Paoletti L, et al. Oropharyngeal swallow physiology and swallowing-related quality of life in underweight patients with concomitant advanced chronic obstructive pulmonary disease. *Int J Chronic Obstr.* 2018;13:2663–2671. doi:10.2147/COPD.S165657
45. Barnes PJ, Celli BR. Systemic manifestations and comorbidities of COPD. *Europ Resp J.* 2009;33(5):1165–1185. doi:10.1183/09031936.00128008
46. Zeng S, Zhang Y, Li S, et al. From metabolic alterations to chronic inflammation: mechanisms and immunoregulation of metabolic reprogramming in COPD. *Front Immunol.* 2025;16:1698832. doi:10.3389/fimmu.2025.1698832
47. Liang Q, Wang Y, Li Z. Lipid metabolism reprogramming in chronic obstructive pulmonary disease. *Mol Med.* 2025;31(1):129. doi:10.1186/s10020-025-01191-9
48. Jones SE, Maddocks M, Kon SS, et al. Sarcopenia in COPD: prevalence, clinical correlates and response to pulmonary rehabilitation. *Thorax.* 2015;70(3):213–218. doi:10.1136/thoraxjnl-2014-206440
49. Scelza L, Greco CS, Lopes AJ, et al. Dysphagia in chronic obstructive pulmonary disease. In: *Seminars in Dysphagia*. InTechOpen Ltd London; 2015.
50. Bordoni B, Escher A, Compalati E, et al. The importance of the diaphragm in neuromotor function in the patient with chronic obstructive pulmonary disease. *Int J Chronic Obstr.* 2023;18:837–848. doi:10.2147/COPD.S404190
51. Alexandre F. *Implication du système nerveux central dans la faiblesse musculaire périphérique du patient atteint de broncho-pneumopathie chronique obstructive*. Université Montpellier; 2015.
52. Cassiani RA, Santos CM, Baddini-Martinez J, et al. Oral and pharyngeal bolus transit in patients with chronic obstructive pulmonary disease. *Int J Chronic Obstr Pulmonary Dis.* 2015;10:489–496. doi:10.2147/COPD.S74945
53. Ghannouchi I, Speyer R, Doma K, et al. Swallowing function and chronic respiratory diseases: systematic review. *Respir Med.* 2016;117:54–64. doi:10.1016/j.rmed.2016.05.024
54. FRd S, RGD S, DdL O, et al. Risk of swallowing disorder in chronic obstructive pulmonary disease. *Revista Brasileira de Ciências da Saude.* 2024;36(2):e65562.
55. Lin TF, Shune S. The mind-body-breath link during oral intake in chronic obstructive pulmonary disease: a grounded theory analysis. *Dysphagia.* 2023;38(1):367–378. doi:10.1007/s00455-022-10473-x
56. Gonzalez Lindh M, Malinovsky A, Brandén E, et al. Subjective swallowing symptoms and related risk factors in COPD. *ERJ Open Res.* 2019;5(3):00081–2019. doi:10.1183/23120541.00081-2019
57. Meena V, Gill N. Post-intubation laryngeal injuries: incidence, types, and outcomes in a tertiary setup. *Europ Archiv Oto-Rhino-Laryngol.* 2025;282(9):4721–4725. doi:10.1007/s00405-025-09632-1
58. Menon R, Vasani SS, Widdicombe NJ, et al. Laryngeal injury following endotracheal intubation: have you considered reflux? *Anaesthesia Intensive Care.* 2023;51(1):14–19. doi:10.1177/0310057X221102472
59. Wallace S, McGrath BA. Laryngeal complications after tracheal intubation and tracheostomy. *BJA Educ.* 2021;21(7):250–257. doi:10.1016/j.bjae.2021.02.005
60. Brodsky MB, Nollet JL, Spronk PE, et al. Prevalence, pathophysiology, diagnostic modalities, and treatment options for dysphagia in critically ill patients. *Am J Phys Med Rehab.* 2020;99(12):1164–1170. doi:10.1097/PHM.0000000000001440
61. Lan B, Chen Y, Wu K. A granular XGBoost classification algorithm. *Appl Intell.* 2025;55(13):895.
62. Zou M, Jiang WG, Qin QH, et al. Optimized XGBoost model with small dataset for predicting relative density of Ti-6Al-4V parts manufactured by selective laser melting. *Materials.* 2022;15(15):5298. doi:10.3390/ma15155298
63. Kivrak M, Avci U, Uzun H, et al. The impact of the SMOTE method on machine learning and ensemble learning performance results in addressing class imbalance in data used for predicting total testosterone deficiency in type 2 diabetes patients. *Diagnostics.* 2024;14(23):2634. doi:10.3390/diagnostics14232634
64. Imani M, Beikmohammadi A, Arabnia HRJT. Comprehensive analysis of random forest and XGBoost performance with SMOTE, ADASYN, and GNUS under varying imbalance levels. *Technologies.* 2025;13(3):88.
65. Hidayaturohman QA, Hanada EJB. Impact of data pre-processing techniques on XGBoost model performance for predicting all-cause readmission and mortality among patients with heart failure. *BioMedInformatics.* 2024;4(4):2201–2212.
66. Huang AA, Huang SY. Increasing transparency in machine learning through bootstrap simulation and shapely additive explanations. *PLoS One.* 2023;18(2):e0281922. doi:10.1371/journal.pone.0281922
67. Liu J, Jiang W, Yu Y, et al. Applying machine learning to predict bowel preparation adequacy in elderly patients for colonoscopy: development and validation of a web-based prediction tool. *Annals Med.* 2025;57(1):2474172. doi:10.1080/07853890.2025.2474172

International Journal of Chronic Obstructive Pulmonary Disease

Publish your work in this journal

The International Journal of COPD is an international, peer-reviewed journal of therapeutics and pharmacology focusing on concise rapid reporting of clinical studies and reviews in COPD. Special focus is given to the pathophysiological processes underlying the disease, intervention programs, patient focused education, and self management protocols. This journal is indexed on PubMed Central, MedLine and CAS. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-chronic-obstructive-pulmonary-disease-journal>

Dovepress
Taylor & Francis Group