Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development and temporal validation of a model to identify high risk of oral frailty in mechanically ventilated intensive care unit patients after extubation

Yongjun Chen
Wannan Medical University
Huixian Tang
Wannan Medical University
Bo Li
Wannan Medical University
Liangfeng Wu
First Affiliated Hospital, Wannan Medical University
Ruixiang Sun
First Affiliated Hospital, Wannan Medical University
Li-Ping Yuan
1144747045@qq.com

First Affiliated Hospital, Wannan Medical University

Research Article

Keywords: Oral frailty, Intensive care unit, Mechanical ventilation, Extubation, Prediction model, Temporal validation, Oral health

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11428701/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Development and temporal validation of a model to identify high risk of oral frailty in mechanically ventilated intensive care unit patients after extubation

Yongjun Chen[1], Huixian Tang[1], Bo Li[1], Liangfeng Wu[2], Ruixiang Sun[2], Li-Ping Yuan[3*]

[1] Graduate School, Wannan Medical University, Wuhu, Anhui 241002, China

[2] Intensive Care Unit, First Affiliated Hospital, Wannan Medical University, Wuhu, Anhui 241001, China

[3] Department of Nursing, First Affiliated Hospital, Wannan Medical University, Wuhu, Anhui 241001, China

*** Correspondence:**

Li-Ping Yuan

Department of Nursing, First Affiliated Hospital, Wannan Medical University, Wuhu, Anhui 241001, China

Email: 1144747045@qq.com

Main text word count: 2,706 words (excluding headings, tables, figure legends and references)

Number of figures: 4

Number of tables: 5

Abstract

Background: To develop and temporally validate a model for identifying current screening-defined high risk of oral frailty within 24 hours after extubation in intensive care unit (ICU) patients.

Methods: This single-centre study in China used cross-sectional data from adults assessed after ≥48 h of invasive mechanical ventilation. High-risk status was defined as an Oral Frailty Index-8 (OFI-8) score ≥4 within 24 hours after extubation. Development and temporal validation cohorts comprised 232 and 100 patients, respectively. Least absolute shrinkage and selection operator regression selected predictors for logistic refitting. Bootstrap validation repeated tuning, selection and refitting, conditional on existing predictor forms. The locked model was compared with an Oral Health Assessment Tool (OHAT)-only model for temporal

discrimination, calibration and Brier scores. Decision curves explored a hypothetical oral health review strategy.

Results: Among 332 participants, 135 (40.7%) met the OFI-8 high-risk criterion. The final model included age, Charlson Comorbidity Index, Acute Physiology and Chronic Health Evaluation II, and OHAT scores. The apparent area under the receiver operating characteristic curve (AUC) was 0.857, with an optimism-corrected AUC of 0.829. Temporal AUC was 0.867 (95% confidence interval [CI], 0.788–0.934), with a Brier score of 0.145 and calibration slope of 1.106 (95% CI, 0.706–1.853). Compared with OHAT alone, the AUC difference was 0.185 (95% CI, 0.069–0.303), and the Brier score difference was −0.072 (95% CI, −0.117 to −0.028).

Conclusions: The four-variable model improved identification of current OFI-8-defined high-risk status over OHAT alone in same-centre temporal validation, although calibration estimates were imprecise. Multicentre validation and evaluation against direct OFI-8 screening are needed before routine nursing use.

Keywords: Oral frailty; Intensive care unit; Mechanical ventilation; Extubation; Prediction model; Temporal validation; Oral health

1 Background

Oral health may deteriorate during invasive mechanical ventilation in intensive care unit (ICU) patients (1). Swallowing recovery varies during the first 48 hours after extubation (2). A randomized trial found that combined swallowing and oral care increased the proportion of patients resuming oral feeding within seven days (3). A qualitative study involving former ICU patients, family members, and clinicians highlighted communication difficulties and a need for professional guidance on oral frailty (4). In alert patients able to communicate after extubation,

assessment can combine oral examination with patient-reported difficulties in oral function. Considering oral health and function together may help nurses determine the need for further evaluation and tailor oral care to individual patients.

Oral frailty is characterized by mild declines in oral function, including difficulties with chewing, swallowing, and communication (5). In community-dwelling older adults, accumulated oral deficits have been associated with subsequent physical frailty, disability, and mortality (6). In older outpatients attending a frailty clinic, accumulated oral problems were also associated with malnutrition and sarcopenia (7). The Oral Frailty Index-8 (OFI-8) is an eight-item questionnaire for identifying individuals at risk of oral frailty, with scores ≥4 indicating high risk (8). Its Chinese version has undergone reliability and validity testing in community-dwelling older adults (9). A recent study in community-dwelling adults in Singapore reported moderate reliability and limited structural validity for OFI-8 (10). The applicability of the ≥4 threshold to ICU patients shortly after extubation requires further evaluation. In this study, an OFI-8 score ≥4 defines a screening outcome and does not establish a clinical diagnosis of oral frailty.

In ICU patients assessed after extubation, age, intubation duration, and oral health status have been associated with OFI-8 scores (11). These associations suggest potential predictors, but do not establish the accuracy of individual probability estimates. A recent systematic review of oral frailty models in older adults found incomplete calibration assessment and limited external validation (12). Model evaluation requires both discrimination and calibration, including agreement between estimated probabilities and observed frequencies (13). Performance also needs evaluation in subsequent patients, as changes in patient characteristics and care can affect predictive accuracy over time (14). Whether clinical and oral health

information can provide reliable estimates of OFI-8-defined high-risk status in the early post-extubation setting remains uncertain.

We aimed to develop a model estimating the probability of current OFI-8-defined high-risk status within 24 hours after extubation in ICU patients following invasive mechanical ventilation. The model incorporated both clinical information and oral health assessment. We evaluated its discrimination and calibration through internal validation and temporal validation in a subsequent cohort from the same centre.

2 Methods

2.1 Study design and participants

This single-centre model development and temporal validation study used cross-sectional data. Data were collected in the general ICU of Yijishan Hospital, Wannan Medical University, Wuhu, China, from January to September 2026. Patients were recruited using convenience sampling. The inclusion criteria were: (1) age ≥18 years; (2) invasive mechanical ventilation via an oral endotracheal tube for at least 48 hours; and (3) ability to understand and respond to the Oral Frailty Index-8 (OFI-8) items within 24 hours after extubation. The exclusion criteria were: (1) tracheostomy; and (2) a previous diagnosis of dementia. The study was approved by the Ethics Committee of the First Affiliated Hospital of Wannan Medical University. All participants provided informed consent; when physical limitations prevented signing, a family member signed on their behalf.

The study was not publicly registered. Reporting was checked against TRIPOD+AI; the checklist accompanies this submission.

2.2 Variables and instruments

2.2.1 Outcome assessment

Oral frailty risk was assessed within 24 hours after extubation using the Chinese version of the Oral Frailty Index-8 (OFI-8) (8,9). The instrument comprises eight items covering denture use, swallowing, chewing, oral health behaviours, and social participation. Responses concerning toothbrushing, frequency of going out, and dental visits referred to patients' usual practices before intubation. Total scores range from 0 to 11, with higher scores indicating greater oral frailty risk (9). Item-level coding, including reverse scoring of annual dental attendance, is detailed in Additional file 1. For model development and validation, the outcome was coded as 1 for OFI-8 scores ≥4 and 0 for scores <4, using the published threshold (8). This outcome represented screening-defined high-risk status at the post-extubation assessment and did not establish a clinical diagnosis of oral frailty.

2.2.2 Candidate predictors

Twenty-three candidate predictors were considered, covering demographic and lifestyle characteristics, treatment-related variables, and clinical and oral health assessments (Supplementary Table S1). Treatment durations, events, and exposures were recorded up to the OFI-8 assessment. Mechanical ventilation duration included invasive and post-extubation non-invasive ventilation.

The Charlson Comorbidity Index (CCI) used the original disease weights and included age points (15). Acute Physiology and Chronic Health Evaluation II (APACHE II) scores reflected the first 24 hours after ICU admission (16). The first recorded Sequential Organ Failure Assessment (SOFA) (17) and Richmond Agitation–Sedation Scale (RASS) (18) scores after ICU admission were used. Nutritional Risk Screening 2002 (NRS-2002) (19) and Chinese FRAIL scale (20) scores also reflected the first assessment after admission. These admission

scores were retrieved at the post-extubation assessment. Oral health was assessed within 24 hours after extubation using the Chinese Oral Health Assessment Tool (OHAT) (21,22). Its total score (0–16) was included as a candidate predictor, with higher scores indicating poorer oral health.

2.3 Data collection

Data were collected by two trained authors using standardized instructions. For each patient, the same assessor conducted the OFI-8 interview followed by OHAT assessment within 24 hours after extubation. OFI-8 total scores and risk classifications had not been calculated at the time of OHAT assessment. Demographic and lifestyle information was obtained through patient interviews, and clinical data were extracted from electronic medical, nursing and ventilator records. Data were checked for completeness and internal consistency, with medical records reviewed when clarification was needed.

2.4 Sample size and temporal allocation

The sample size was determined by the available patients during the study period, without an a priori calculation for model development. The final analytic sample comprised 332 eligible patients who completed OFI-8 and OHAT assessment. Allocation followed enrolment date: 232 patients enrolled from January to July 2026 formed the development cohort, and 100 enrolled from August to September formed the temporal validation cohort. The model specification, predictor coding and coefficients were fixed before temporal validation performance was evaluated.

2.5 Statistical analysis

Statistical analyses used SPSS 27.0 and R 4.4.1. Continuous variables were summarised using means and standard deviations (SDs), or medians and interquartile ranges. Categorical

variables were reported as frequencies and percentages. Absolute standardized mean differences (SMDs) described cohort differences. All 23 candidate predictors and the OFI-8 outcome were complete for all 332 analysis participants; no imputation was performed.

Development analyses included univariable logistic regression, likelihood ratio tests for nonlinearity and variance inflation factors (VIFs) for multicollinearity. NRS-2002 was entered as 0–2, 3 and ≥4 following the nonlinearity assessment. Least absolute shrinkage and selection operator (LASSO) logistic regression selected predictors using stratified 10-fold cross-validation and the one-standard-error rule (23). Selected predictors were refitted using unpenalized logistic regression and presented in a nomogram. Elastic net regression provided a sensitivity analysis. Further details are provided in Additional file 1.

Performance was assessed using the area under the receiver operating characteristic (ROC) curve (AUC), calibration plots, calibration intercept and slope, and the Brier score (24). Internal validation used 1,000 bootstrap samples, repeating standardization, LASSO tuning, selection and logistic refitting. Candidate coding and functional forms remained fixed, leaving corrected performance conditional on these choices. The apparent AUC confidence interval (CI) used bootstrap resampling with fixed predictions. The final model and an OHAT-only comparator were fitted in the development cohort and evaluated without updating in the temporal cohort. Paired bootstrap resampling compared temporal performance between models; computational details are provided in Additional file 1.

Exploratory decision curve analysis compared a hypothetical strategy of model-guided oral health review with reviewing all or no patients (25). Threshold probabilities of 0.05–0.70 were explored; no clinical action threshold was established.

ChatGPT (OpenAI, including GPT-6.1-sol) assisted with language editing, analysis-script development and manuscript checking. Claude (Anthropic) assisted with pre-submission review. The authors remain responsible for the accuracy and integrity of the work.

3 Results

3.1 Participant characteristics

Between January and September 2026, 348 patients were assessed for eligibility. Sixteen were excluded before OFI-8 and OHAT assessment: eight had received invasive mechanical ventilation via an oral endotracheal tube for <48 h, one had undergone tracheostomy, and seven could not respond to the questionnaire within 24 h after extubation because cognitive or communication function had not sufficiently recovered. The remaining 332 patients formed the analytic sample, comprising 232 in the development cohort and 100 in the temporal validation cohort.

The mean age was 61.30 years (SD, 15.84), and 191 patients (57.5%) were male. Overall, 135 patients (40.7%) had OFI-8-defined high-risk status, including 97 (41.8%) in the development cohort and 38 (38.0%) in the temporal validation cohort.

Cohort characteristics are presented in Table 1. Age and the CCI, APACHE II and OHAT scores differed little between cohorts (absolute SMDs, 0.003–0.041). The largest differences were observed in monthly income (SMD, 0.269), parenteral nutrition (0.255), marital status (0.245) and RASS scores (0.232).

[Table 1 near here]

3.2 Univariable analysis and predictor selection

In univariable analyses, age, CCI, APACHE II, SOFA, RASS, NRS-2002 and OHAT were associated with OFI-8-defined high-risk status (all $P < 0.001$; Table 2). Associations were also

observed for marital status, educational level, monthly income, employment status and parenteral nutrition (all $P < 0.05$).

[Table 2 near here]

After Benjamini–Hochberg correction for 13 nonlinearity tests, evidence of nonlinearity remained for NRS-2002 ($q = 0.0024$). It was entered as three categories (0–2, 3 and ≥4). Candidate-predictor VIFs ranged from 1.131 to 7.051. LASSO selected age, CCI, APACHE II and OHAT at $\lambda = 0.0564$. All four were retained in 100/100 repeated fold assignments and in the elastic net and NRS-2002 coding sensitivity analyses. Their individual selection frequencies in patient bootstrap resampling ranged from 88.8% to 99.9%, while the exact four-variable set recurred in 371/1,000 samples (37.1%) (Additional file 1, Tables S2–S4).

3.3 Multivariable analysis and model development

After unpenalized logistic refitting, all four predictors were positively associated with OFI-8-defined high-risk status (Table 3).

Table 3. Final multivariable logistic regression model

Predictor	β	Standard error	Odds ratio (95% CI)	P value
Age, per year	0.030	0.013	1.030 (1.004–1.057)	0.024
CCI, per point	0.202	0.068	1.224 (1.071–1.399)	0.003
APACHE II, per point	0.143	0.032	1.154 (1.084–1.228)	<0.001
OHAT, per point	0.389	0.103	1.476 (1.207–1.806)	<0.001
Intercept	−7.156	1.106	—	<0.001

β, regression coefficient.

The final model was expressed as:

$$\text{logit}(P) = -7.156219 + 0.029515 \times \text{age} + 0.202337 \times \text{CCI} + 0.143138 \times \text{APACHE II} + 0.389442 \times \text{OHAT}.$$

P denotes the probability of OFI-8-defined high-risk status, with $P = 1/[1 + \exp(-\text{logit}(P))]$. Age is expressed in years, and CCI, APACHE II and OHAT are entered as scores. The model was presented as a nomogram (Fig. 1).

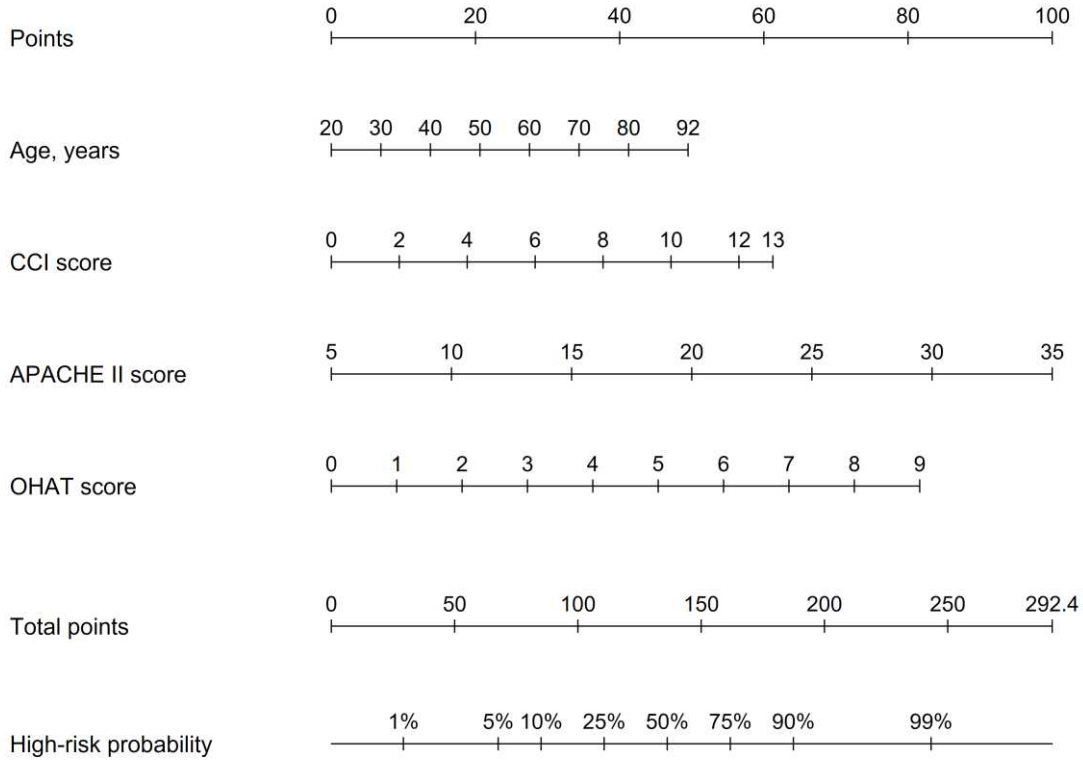

Fig. 1 Nomogram for estimating OFI-8-defined high-risk status.
Read the points for each predictor from the top scale, sum them, and locate the corresponding probability on the bottom scale.

3.4 Model performance and validation

In the development cohort, the apparent AUC was 0.857 (95% CI, 0.810–0.902), with a Brier score of 0.151. Bootstrap-corrected AUC and Brier score were 0.829 and 0.171, respectively (Table 4). Fixed-predictor bootstrap results and selection frequencies are provided in Additional file 1 (Tables S4 and S5).

In the temporal validation cohort, the locked model had an AUC of 0.867 (95% CI, 0.788–0.934) and a Brier score of 0.145. The calibration intercept was −0.147 and the calibration slope was 1.106 (Table 4). Mean predicted probability was 40.3%, compared with an observed high-risk proportion of 38.0%. ROC and calibration curves are presented in Figs. 2 and 3.

Table 4. Model performance in the development and temporal validation cohorts

Performance measure	Development apparent	Development optimism-corrected	Temporal validation estimate	Temporal validation 95% CI
AUC	0.857	0.829	0.867	0.788–0.934
Brier score	0.151	0.171	0.145	0.108–0.182
Calibration intercept	0.000	−0.006	−0.147	−0.649–0.332
Calibration slope	1.000	0.844	1.106	0.706–1.853

Ideal calibration intercept and slope are 0 and 1, respectively. Corrected estimates are conditional on existing candidate coding and functional forms.

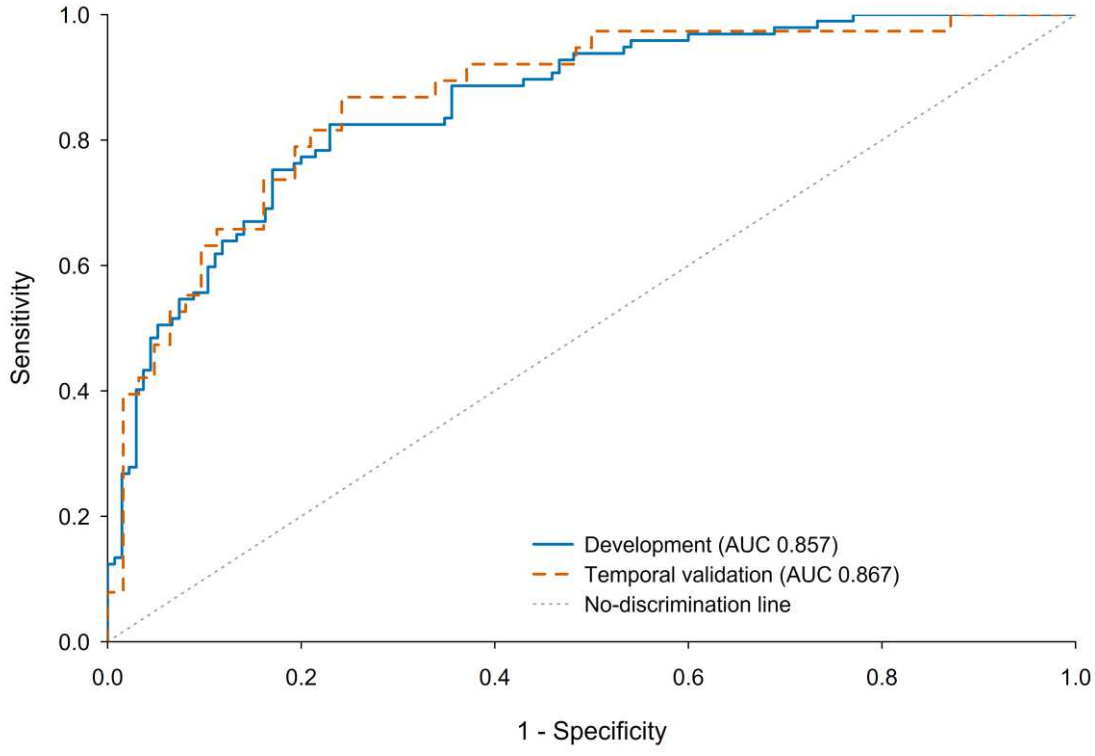

Fig. 2 ROC curves in the development and temporal validation cohorts.
Development performance is apparent; temporal performance uses fixed model predictions.

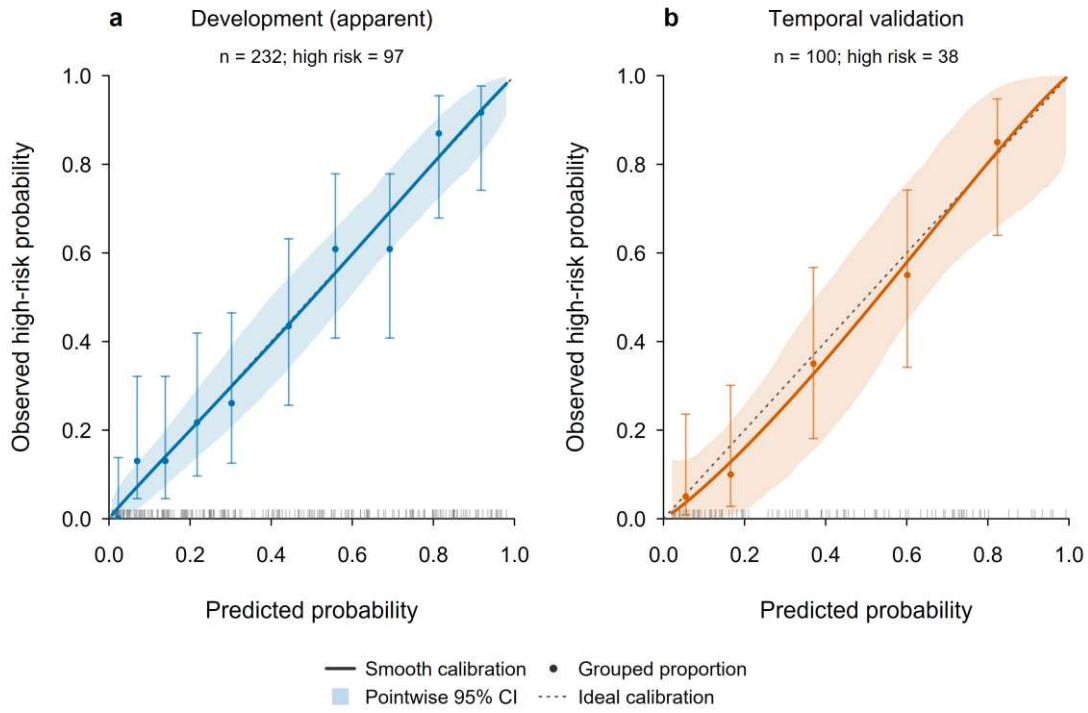

Fig. 3 Calibration plots in the development (a) and temporal validation (b) cohorts. Solid lines show smooth calibration curves with pointwise 95% confidence bands. Points and error bars show grouped proportions and 95% Wilson intervals. Bottom ticks show individual predicted probabilities; dotted lines indicate ideal calibration. Development curves show apparent calibration.

Compared with OHAT alone, the four-variable model had a higher temporal AUC (0.867 versus 0.682; paired difference, 0.185; 95% CI, 0.069–0.303) and a lower Brier score (0.145 versus 0.217; difference, −0.072; 95% CI, −0.117 to −0.028). Calibration estimates are reported in Table 5. The comparator equation and calibration plots are provided in Additional file 1 (Supplementary methods and Fig. S1).

Table 5. Comparison with an OHAT-only model in the temporal validation cohort

Performance measure	Four-variable model Estimate (95% CI)	OHAT-only model Estimate (95% CI)	Paired difference (95% CI)
AUC	0.867 (0.788 to 0.934)	0.682 (0.575 to 0.787)	0.185 (0.069 to 0.303)
Brier score	0.145 (0.108 to 0.182)	0.217 (0.181 to 0.251)	−0.072 (−0.117 to −0.028)
Calibration intercept	−0.147 (−0.649 to 0.332)	−0.154 (−0.615 to 0.280)	—
Calibration slope	1.106 (0.706 to 1.853)	0.785 (0.279 to 1.559)	—

Differences are four-variable minus OHAT-only, with paired bootstrap 95% CIs. Positive AUC and negative Brier differences favour the four-variable model.

3.5 Exploratory decision curve analysis

The four-variable model showed higher estimated net benefit than the review-all and review-none strategies across threshold probabilities of 0.05–0.70 in development and 0.08–0.70 in temporal validation (Fig. 4). Curves for the OHAT-only model were included for comparison.

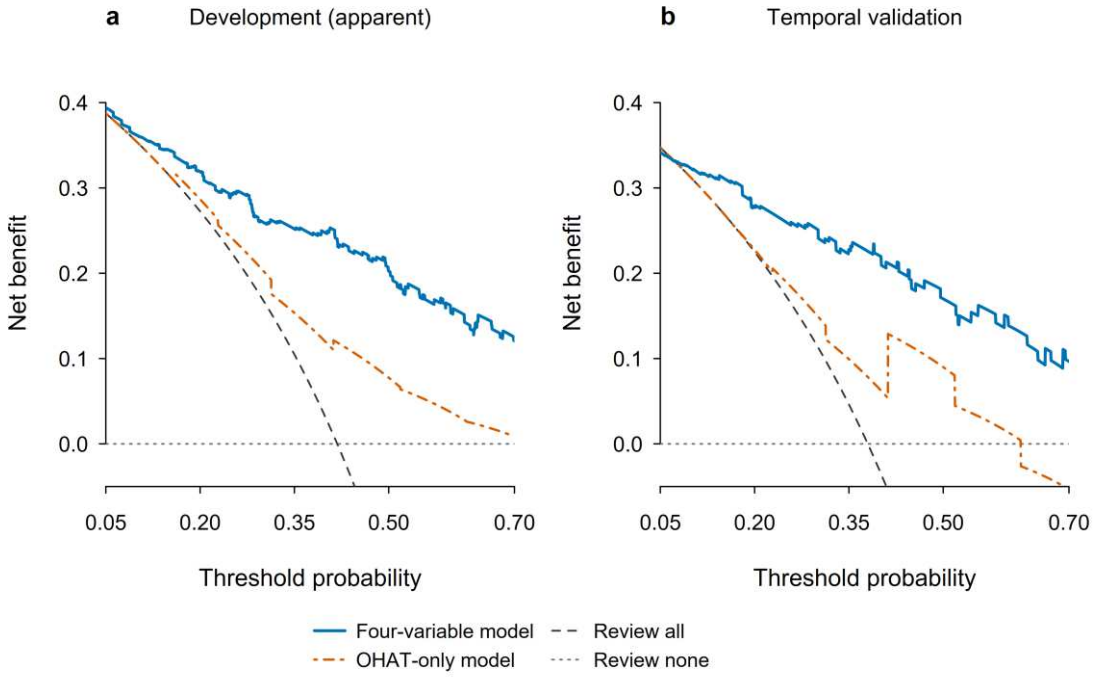

Fig. 4 Exploratory decision curves in the development (a) and temporal validation (b) cohorts.
Curves compare both models with review-all and review-none strategies. Development estimates are apparent; temporal estimates use fixed predictions. Curves are point estimates without CIs.

4 Discussion

This study developed a model to identify OFI-8-defined high-risk status assessed in ICU patients within 24 h after extubation. The model combined age with CCI, APACHE II and OHAT scores, integrating clinical characteristics with oral health assessment. Its predictors and coefficients were fixed before temporal validation in a subsequent cohort from the same centre. In this cohort, the model showed good discrimination, with an AUC of 0.867. Its clinical role requires further evaluation.

In our cohort, 40.7% of patients met the OFI-8 threshold for high-risk status. A previous ICU study reported 36.7% and linked OFI-8 scores to age, intubation duration and oral health (11). Models in older stroke patients and geriatric inpatients also incorporated clinical and oral health characteristics (26,27). Differences in patient selection and assessment timing, however, limit direct comparisons of high-risk proportions and model performance. Our study extends this evidence to the identification of current screening high-risk status within 24 h after extubation.

Older patients and those with higher CCI scores were more likely to meet the OFI-8 high-risk threshold. The age association is consistent with findings in ICU and older stroke populations (11,26). Greater comorbidity burden has also been associated with OFI-8 screening outcomes in older adults with multimorbidity (28). Medication-related oral symptoms may partly explain the comorbidity association. In acutely ill older adults, the use of more xerogenic medications was associated with xerostomia (29). Whether medication exposure contributed to the association observed in our cohort remains uncertain. Because CCI included age points, age and CCI provided partly overlapping information. Their inclusion in the model supports their use in combination for identification, but does not establish distinct biological effects.

The APACHE II finding links the screening outcome to illness severity recorded at ICU admission. APACHE II reflects acute physiological abnormalities together with age and chronic health status (16). A large ICU cohort linked admission illness severity to post-extubation dysphagia (30). A recent meta-analysis found that higher APACHE II scores were associated with post-extubation dysphagia (31). Associations with swallowing recovery patterns have also been reported, although they differed across trajectory groups (32). ICU-acquired weakness may also be relevant: a prospective cohort of critically ill respiratory patients found an association between weakness and post-extubation dysphagia (33). These studies provide indirect support for our findings because swallowing dysfunction and OFI-8-defined high-risk status are different outcomes.

At the screening assessment, patients with OFI-8-defined high-risk status had poorer oral health, consistent with previous ICU findings (11). OHAT assesses oral tissues, saliva, dentition, cleanliness and pain (21), with good scoring agreement reported between trained community nurses (34). OFI-8 combines reported oral difficulties with denture use, health behaviours and social participation (8). The tools therefore capture related aspects of oral health. Their administration by the same assessor within the same assessment window may also have contributed to the observed association. In a longitudinal stroke cohort, however, baseline OHAT scores did not differ significantly across subsequent oral frailty trajectories (35). An OHAT-guided oral-care protocol was associated with better oral health in a quasi-experimental ICU study (36), illustrating how assessment findings can inform nursing care.

The model retained discrimination in subsequent patients from the same centre, although calibration estimates remained imprecise (14,24). Compared with OHAT alone, it provided additional information for identifying high-risk patients. Clinical usefulness also depends on

calibration, outcome frequency and the consequences of missed high-risk patients and unnecessary reviews (24,37). Decision curves suggested potential value in prioritising oral health review (25). Its practical advantage over direct OFI-8 screening remains unproven. Prospective comparisons should assess care decisions and nursing workload before routine implementation (38).

Convenience sampling at one centre and requiring patients to answer OFI-8 after extubation limit generalisability (24). The OFI-8 ≥4 threshold requires validation in this acute setting (8,10). Having the same assessor administer OFI-8 before OHAT may have influenced OHAT ratings despite the total score remaining uncalculated (39). Cross-sectional assessment could not distinguish pre-existing oral difficulties from changes during critical illness. The development sample and number of candidate predictors leave scope for overfitting (40). Bootstrap validation retained data-informed coding and functional forms, leaving potential optimism from these decisions uncorrected (41). The two-month temporal cohort of 100 patients yielded imprecise estimates, particularly for calibration, and could not assess longer-term performance (42). Multicentre studies should validate OFI-8 and the model using independent predictor and outcome assessors. Future development should prespecify coding and functional-form rules and repeat their application within internal validation (41).

5 Conclusions

The model combined age, CCI, APACHE II and OHAT to estimate the probability of current OFI-8-defined high-risk status within 24 hours after extubation in ICU patients. In same-centre temporal validation, it showed higher discrimination and a lower Brier score than an OHAT-only model, although calibration estimates remained imprecise. Multicentre validation and comparison with direct OFI-8 screening are needed before its use to prioritise oral health review.

List of abbreviations

APACHE II, Acute Physiology and Chronic Health Evaluation II; AUC, area under the receiver operating characteristic curve; BMI, body mass index; CCI, Charlson Comorbidity Index; CI, confidence interval; CNY, Chinese yuan; FRAIL, Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight; ICU, intensive care unit; LASSO, least absolute shrinkage and selection operator; NRS-2002, Nutritional Risk Screening 2002; OFI-8, Oral Frailty Index-8; OHAT, Oral Health Assessment Tool; OR, odds ratio; RASS, Richmond Agitation–Sedation Scale; ROC, receiver operating characteristic; SD, standard deviation; SMD, standardized mean difference; SOFA, Sequential Organ Failure Assessment; VIF, variance inflation factor.

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of the First Affiliated Hospital of Wannan Medical University. Written informed consent was obtained from all participants. When a participant had consented but was physically unable to sign, a family member signed on the participant's behalf. The approval number was 2026-Lunshen-Yan-153.

Consent for publication

Not applicable.

Availability of data and materials

De-identified data, the data dictionary and analysis scripts are available from the corresponding author upon reasonable request. The study protocol submitted for ethics review is also available on request.

Competing interests

The authors declare that they have no competing interests.

Funding

No specific funding was received for this study.

Authors' contributions

YC designed the study, collected and analysed the data, and drafted the manuscript. HT, LW and RS contributed to data collection and study coordination. BL and LY were responsible for quality control and manuscript review. All authors read and approved the final manuscript.

Acknowledgements

We thank all participants for their contribution to this study and acknowledge the support provided by the staff at Yijishan Hospital, Wannan Medical University.

Patient and public involvement

Patients and members of the public were not involved in study design, conduct or dissemination.

Additional file

Additional file 1 (.docx). Supplementary methods, Tables S1–S5, and Fig. S1.

References

1. Causey C, El Karim I, Blackwood B, McAuley DF, Lundy FT. Quantitative oral health assessments in mechanically ventilated patients: a scoping review. Nurs Crit Care. 2023 Sep;28(5):756–72. doi:10.1111/nicc.12789 PubMed PMID: 35771584.

2. Wang W, Wang S, Dong T, Zhou Y, Ma R. Trajectory and influencing factors of swallowing function in adults after orotracheal extubation in an intensive care unit in China: a

prospective cohort study. BMJ Open. 2026 Feb 16;16(2):e105312. doi:10.1136/bmjopen-2025-105312

3. Siao SF, Ku SC, Tseng WH, Wei YC, Chang YC, Hsiao TY, et al. Effects of a swallowing and oral-care program on resuming oral feeding and reducing pneumonia in patients following endotracheal extubation: a randomized, open-label, controlled trial. Crit Care. 2023 Jul 12;27(1):283. doi:10.1186/s13054-023-04568-6 PubMed PMID: 37438759; PubMed Central PMCID: PMC10339550.

4. Zhang H, Xu M, Xu Y, Zhang Y, Zhang Y, Zheng Y, et al. Towards a conceptual definition of ICU-acquired oral frailty: a qualitative study of patients, family members and clinicians. Nurs Crit Care. 2026 Jul;31(4):e70564. doi:10.1111/nicc.70564 PubMed PMID: 42405381.

5. Tanaka T, Hirano H, Ikebe K, Ueda T, Iwasaki M, Minakuchi S, et al. Consensus statement on “oral frailty” from the Japan geriatrics society, the japanese society of gerodontology, and the japanese association on sarcopenia and frailty. Geriatr Gerontol Int. 2024 Nov;24(11):1111–9. doi:10.1111/ggi.14980 PubMed PMID: 39375858; PubMed Central PMCID: PMC11843523.

6. Tanaka T, Takahashi K, Hirano H, Kikutani T, Watanabe Y, Ohara Y, et al. Oral frailty as a risk factor for physical frailty and mortality in community-dwelling elderly. J Gerontol A Biol Sci Med Sci. 2018 Nov 10;73(12):1661–7. doi:10.1093/gerona/glx225 PubMed PMID: 29161342.

7. Kawamura K, Maeda K, Ishino S, Iwase T, Akatsu H, Kagaya H, et al. Association of oral problems with malnutrition and sarcopenia in older adults in a frailty clinic. Nutrition. 2026;148:113213. doi:10.1016/j.nut.2026.113213

8. Tanaka T, Hirano H, Ohara Y, Nishimoto M, Iijima K. Oral frailty index-8 in the risk assessment of new-onset oral frailty and functional disability among community-dwelling older adults. Arch Gerontol Geriatr. 2021;94:104340. doi:10.1016/j.archger.2021.104340 PubMed PMID: 33529863.

9. Chen Z, Tan Y, Liang Y, Zhang H, Jiang Y, Shi G. Chinesization of the Oral Frailty Index-8 and its reliability and validity test [in Chinese]. Chin Nurs Res. 2023 Nov 10;37(21):3808–12. doi:10.12102/j.issn.1009-6493.2023.21.003

10. Natarajan K, Sim YF, Chen H, Lai H, Venkateswaran S, Chen MZ, et al. Reliability and validity of oral frailty index-8: findings among community-dwelling adults in Singapore. J Oral Rehabil. 2026;53(6):1159–68. doi:10.1111/joor.70169

11. Shi M, Li L, Zhai H. Oral frailty and its influencing factors among ICU patients with oral endotracheal intubation: a cross-sectional study. Front Med. 2025 Oct 2;12:1650313. doi:10.3389/fmed.2025.1650313

12. Fan Y, Zhang S, Wang Y, Cheng Z, Yu X, Chen Y, et al. Diagnostic prediction models for oral frailty in older adults: a systematic review and critical appraisal. Patient Prefer Adherence. 2026;20:621510. doi:10.2147/PPA.S621510 PubMed PMID: 42730340; PubMed Central PMCID: PMC13564451.

13. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Topic Group 'Evaluating Diagnostic Tests And Prediction Models' Of The Stratos Initiative. Calibration: the achilles heel of predictive analytics. BMC Med. 2019 Dec 16;17(1):230. doi:10.1186/s12916-019-1466-7 PubMed PMID: 31842878; PubMed Central PMCID: PMC6912996.

14. Collins GS, Dhiman P, Ma J, Schlussel MM, Archer L, Van Calster B, et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ. 2024

Jan 8;384:e074819. doi:10.1136/bmj-2023-074819 PubMed PMID: 38191193; PubMed Central PMCID: PMC10772854.

15. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson comorbidity index: a critical review of clinimetric properties. Psychother Psychosom. 2022;91(1):8–35. doi:10.1159/000521288 PubMed PMID: 34991091.

16. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med. 1985 Oct;13(10):818–29. doi:10.1097/00003246-198510000-00009 PubMed PMID: 3928249.

17. Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score-development, utility and challenges of accurate assessment in clinical trials. Crit Care. 2019 Nov 27;23(1):374. doi:10.1186/s13054-019-2663-7 PubMed PMID: 31775846; PubMed Central PMCID: PMC6880479.

18. Fuchita M, Pattee J, Le D, To T, Mucharraz C, Knippa S, et al. Inter-rater agreement of richmond agitation sedation scale assessments in adult patients receiving mechanical ventilation in the ICU: a cross-sectional study. Crit Care Explor. 2025 Sep 1;7(9):e1302. doi:10.1097/CCE.0000000000001302 PubMed PMID: 40875505; PubMed Central PMCID: PMC12398364.

19. Chinese Society of Critical Care Medicine. Clinical practice guidelines for nutritional assessment and monitoring of adult ICU patients in China [in Chinese]. Zhonghua Wei Zhong Bing Ji Jiu Yi Xue. 2023 Nov;35(11):1121–46. doi:10.3760/cma.j.cn121430-20230829-00710 PubMed PMID: 37987122.

20. Wei Y, Cao Y, Yang X, Xu Y. Reliability and validity of the Chinese version of Fatigue, Resistance, Ambulation, Illness, and Loss for elder inpatients [in Chinese]. Chin J Pract Nurs. 2018;34(20):1526–30. doi:10.3760/cma.j.issn.1672-7088.2018.20.002

21. Chalmers JM, King PL, Spencer AJ, Wright FA, Carter KD. The oral health assessment tool--validity and reliability. Aust Dent J. 2005 Sep;50(3):191–9. doi:10.1111/j.1834-7819.2005.tb00360.x PubMed PMID: 16238218.

22. Wang J, Zhu S, Zhan Y, Li J, Liu D. Reliability and validity test of Chinese version of the Oral Health Assessment Tool [in Chinese]. Chin J Mod Nurs. 2019;25(28):3607–10. doi:10.3760/cma.j.issn.1674-2907.2019.28.010

23. Tay JK, Narasimhan B, Hastie T. Elastic Net Regularization Paths for All Generalized Linear Models. J Stat Softw. 2023;106(1):1–31. doi:10.18637/jss.v106.i01 PubMed PMID: 37138589; PubMed Central PMCID: PMC10153598.

24. Riley RD, Archer L, Snell KIE, Ensor J, Dhiman P, Martin GP, et al. Evaluation of clinical prediction models (part 2): how to undertake an external validation study. BMJ. 2024 Jan 15;384:e074820. doi:10.1136/bmj-2023-074820 PubMed PMID: 38224968; PubMed Central PMCID: PMC10788734.

25. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. Diagn Progn Res. 2019;3:18. doi:10.1186/s41512-019-0064-7 PubMed PMID: 31592444; PubMed Central PMCID: PMC6777022.

26. Ma R, Fan X, Tao X, Zhang W, Li Z. Development and validation of risk-predicting model for oral frailty in older adults patients with stroke. BMC Oral Health. 2025 Feb 19;25(1):263. doi:10.1186/s12903-025-05428-1 PubMed PMID: 39972478; PubMed Central PMCID: PMC11841317.

27. Shi X, Zhou X, Yang Y, Yu Q, Pei R, Jing Y, et al. Construction and validation of oral frailty risk assessment model in elderly hospitalized patients. BMC Oral Health. 2026 Apr 28;26(1):1136. doi:10.1186/s12903-026-08359-7

28. Wang X, Xu L, Wang Y, Zhou M, Zhou X, Guo Y, et al. Oral frailty and its influencing factors in elderly patients with multimorbidity: a cross-sectional study. BMC Oral Health. 2026 Jun 26;26(1):1706. doi:10.1186/s12903-026-09060-5 PubMed PMID: 42363210.

29. Ringstad IB, Romskaug R, Santervas LR, Fjeld KG, Hove LH, Skovlund E, et al. Oral health, xerogenic medications, and frailty in acutely ill older adults. J Am Geriatr Soc. 2025 Aug;73(8):2466–74. doi:10.1111/jgs.19564 PubMed PMID: 40458833.

30. Zuercher P, Schenk NV, Moret C, Berger D, Abegglen R, Schefold JC. Risk factors for dysphagia in ICU patients after invasive mechanical ventilation. Chest. 2020 Nov;158(5):1983–91. doi:10.1016/j.chest.2020.05.576 PubMed PMID: 32525018.

31. He Z, Song X, Lei A, Ma Y, Hu Y, Li X, et al. Prevalence and risk factors for postextubation dysphagia in ICU patients with orotracheal intubation: a systematic review and meta-analysis. Front Med. 2026;13:1810274. doi:10.3389/fmed.2026.1810274 PubMed PMID: 42145725; PubMed Central PMCID: PMC13171512.

32. Shan L, Wang H, Yu W, Cai J, Wang Y, Zhang Y, et al. Recovery trajectories and predictors of post-extubation dysphagia in intensive care unit patients following orotracheal intubation: a longitudinal study. Dysphagia. 2026 May 2. doi:10.1007/s00455-026-10941-8 PubMed PMID: 42068342.

33. García-Grimaldo A, Rodríguez-Moguel NC, Godínez-Victoria M, Rodríguez-Llamazares S, Ríos-Ayala MA, Cadeza-Aguilar JD, et al. Associations between intensive care unit acquired weakness with post-extubation dysphagia and other clinical outcomes-a cohort study in critically ill respiratory patients. Clin Nutr ESPEN. 2025 Apr;66:194–201. doi:10.1016/j.clnesp.2025.01.044 PubMed PMID: 39864521.

34. van Noort HHJ, Harderwijk A, Everaars B, Akkermans R, van der Putten GJ, Waal GH de. The dutch version of the oral health assessment tool: translation and inter-rater reliability

among community nurses. Int J Dent Hyg. 2024 Feb;22(1):140–7. doi:10.1111/idh.12678 PubMed PMID: 37093920.

35. Zhao J, Liu Y, Liang S, Zhou M, Zhang Q, Sun H, et al. Oral frailty trajectories and influencing factors in older adults with stroke: a longitudinal study. BMC Nurs. 2025 Nov 18;24(1):1410. doi:10.1186/s12912-025-04042-4 PubMed PMID: 41254618; PubMed Central PMCID: PMC12625639.

36. Sato A, Sasaki Y, Imazu Y. Enhanced moisture-focused, nurse-led oral care improves oral health and suppresses bacterial overgrowth in mechanically ventilated ICU patients: a quasi-experimental study. Nurs Crit Care. 2026 May;31(3):e70428. doi:10.1111/nicc.70428 PubMed PMID: 41958440; PubMed Central PMCID: PMC13067094.

37. Foroutan F, Weiss J, Mayer M, Bache-Mathiesen LK, Agoritsas T, Vandvik PO, et al. Discriminating clinical prediction models: how outcome incidence and decision thresholds shape clinical utility and apt AUROC values. J Clin Epidemiol. 2026 Jul;195:112293. doi:10.1016/j.jclinepi.2026.112293 PubMed PMID: 42044720.

38. Saelmans A, Seinen T, Pera V, Markus AF, Fridgeirsson E, John LH, et al. Implementation and updating of clinical prediction models: a systematic review. Mayo Clin Proc Digit Health. 2025 Sep;3(3):100228. doi:10.1016/j.mcpdig.2025.100228 PubMed PMID: 40599890; PubMed Central PMCID: PMC12212251.

39. Moons KGM, Damen JAA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ. 2025 Mar 24;388:e082505. doi:10.1136/bmj-2024-082505 PubMed PMID: 40127903; PubMed Central PMCID: PMC11931409.

40. Martin GP, Riley RD, Ensor J, Grant SW. Statistical primer: sample size considerations for developing and validating clinical prediction models. Eur J Cardio-Thorac Surg: Off J Eur Assoc Cardio-Thorac Surg. 2025 May 6;67(5):ezaf142. doi:10.1093/ejcts/ezaf142 PubMed PMID: 40279277; PubMed Central PMCID: PMC12106283.

41. Efthimiou O, Seo M, Chalkou K, Debray T, Egger M, Salanti G. Developing clinical prediction models: a step-by-step guide. BMJ. 2024 Sep 3;386:e078276. doi:10.1136/bmj-2023-078276 PubMed PMID: 39227063; PubMed Central PMCID: PMC11369751.

42. Riley RD, Snell KIE, Archer L, Ensor J, Debray TPA, van Calster B, et al. Evaluation of clinical prediction models (part 3): calculating the sample size required for an external validation study. BMJ. 2024 Jan 22;384:e074821. doi:10.1136/bmj-2023-074821 PubMed PMID: 38253388; PubMed Central PMCID: PMC11778934.

532 **Table 1. Characteristics of the overall, development and temporal validation cohorts**

Characteristic	Overall cohort(n = 332)	Development cohort(n = 232)	Temporal validation cohort(n = 100)	SMD
Age, years	61.30 (15.84)	61.15 (16.04)	61.67 (15.41)	0.033
Sex				0.130
Male	191 (57.5)	129 (55.6)	62 (62.0)	
Female	141 (42.5)	103 (44.4)	38 (38.0)	
Marital status				0.245
Unmarried	18 (5.4)	12 (5.2)	6 (6.0)	
Married	287 (86.4)	200 (86.2)	87 (87.0)	
Divorced	15 (4.5)	9 (3.9)	6 (6.0)	
Widowed	12 (3.6)	11 (4.7)	1 (1.0)	
BMI, kg/m²	26.11 (4.15)	26.36 (3.96)	25.53 (4.54)	0.195
Smoking history, yes	154 (46.4)	104 (44.8)	50 (50.0)	0.104
Alcohol use, yes	70 (21.1)	53 (22.8)	17 (17.0)	0.147
Educational level				0.134
Primary school or below	88 (26.5)	62 (26.7)	26 (26.0)	
Junior secondary school	128 (38.6)	93 (40.1)	35 (35.0)	
Senior/vocational secondary school	75 (22.6)	49 (21.1)	26 (26.0)	
College/university or above	41 (12.3)	28 (12.1)	13 (13.0)	
Monthly income, CNY				0.269
<1000	65 (19.6)	52 (22.4)	13 (13.0)	
1000–5000	142 (42.8)	99 (42.7)	43 (43.0)	
>5000	125 (37.7)	81 (34.9)	44 (44.0)	
Employment status				0.176
Employed	113 (34.0)	82 (35.3)	31 (31.0)	
Retired	182 (54.8)	128 (55.2)	54 (54.0)	
Unemployed, freelance or self-employed	37 (11.1)	22 (9.5)	15 (15.0)	
CCI score	4.80 (3.04)	4.83 (3.11)	4.71 (2.87)	0.041
APACHE II score	18.56 (6.50)	18.55 (6.48)	18.57 (6.59)	0.003
SOFA score	7.36 (3.90)	7.47 (3.97)	7.10 (3.74)	0.095
RASS score				0.232
-5	13 (3.9)	8 (3.4)	5 (5.0)	
-4	33 (9.9)	23 (9.9)	10 (10.0)	
-3	52 (15.7)	41 (17.7)	11 (11.0)	
-2	78 (23.5)	53 (22.8)	25 (25.0)	
-1	66 (19.9)	44 (19.0)	22 (22.0)	
0	52 (15.7)	36 (15.5)	16 (16.0)	
1	21 (6.3)	16 (6.9)	5 (5.0)	

Characteristic	Overall cohort(n = 332)	Development cohort(n = 232)	Temporal validation cohort(n = 100)	SMD
2	12 (3.6)	8 (3.4)	4 (4.0)	
3	5 (1.5)	3 (1.3)	2 (2.0)	
Mechanical ventilation duration, h	115.00 (75.00, 176.25)	113.00 (74.00, 173.00)	122.50 (80.00, 187.00)	0.117
Reintubation, yes	98 (29.5)	63 (27.2)	35 (35.0)	0.170
Sedative use, yes	298 (89.8)	206 (88.8)	92 (92.0)	0.109
Neuromuscular blocking agent use, yes	81 (24.4)	55 (23.7)	26 (26.0)	0.053
Enteral nutrition, yes	238 (71.7)	167 (72.0)	71 (71.0)	0.022
Parenteral nutrition, yes	19 (5.7)	17 (7.3)	2 (2.0)	0.255
NRS-2002 score	3.21 (1.25)	3.16 (1.27)	3.34 (1.21)	0.149
ICU stay to assessment, days	8.50 (5.75, 13.00)	8.00 (6.00, 12.25)	9.00 (5.00, 13.00)	0.061
FRAIL score	1.47 (0.82)	1.47 (0.84)	1.47 (0.78)	0.000
OHAT score	2.96 (1.84)	2.98 (1.83)	2.93 (1.88)	0.026
OFI-8-defined high-risk status	135 (40.7)	97 (41.8)	38 (38.0)	—

Values are mean (SD), median (25th, 75th percentiles), or n (%). SMDs are absolute. CCI includes age points. Ventilation duration includes invasive and post-extubation non-invasive ventilation up to assessment. High-risk status: OFI-8 ≥4.

Table 2. Univariable logistic regression analysis in the development cohort

Predictor	Unit / comparison	Odds ratio (95% CI)	P value
Age	Per year	1.05 (1.03–1.07)	<0.001
Sex	Female vs male	0.93 (0.55–1.57)	0.775
Marital status	Reference: Unmarried		0.025
Married		3.70 (0.79–17.31)	0.097
Divorced		1.43 (0.16–12.70)	0.749
Widowed		13.33 (1.78–100.15)	0.012
BMI	Per kg/m²	0.98 (0.92–1.05)	0.575
Smoking history	Yes vs no	1.04 (0.61–1.75)	0.890
Alcohol use	Yes vs no	0.58 (0.31–1.12)	0.104
Educational level	Reference: Primary school or below		<0.001
Junior secondary school		0.62 (0.32–1.19)	0.150
Senior/vocational secondary school		0.29 (0.13–0.64)	0.002
College/university or above		0.12 (0.04–0.39)	<0.001
Monthly income, CNY	Reference: <1000		0.004
1000–5000		0.98 (0.50–1.92)	0.953
>5000		0.37 (0.18–0.77)	0.008
Employment status	Reference: Employed		<0.001
Retired		4.03 (2.15–7.55)	<0.001
Unemployed, freelance or self-employed		3.56 (1.33–9.53)	0.012
CCI	Per point	1.43 (1.28–1.59)	<0.001
APACHE II	Per point	1.19 (1.13–1.25)	<0.001
SOFA	Per point	1.15 (1.07–1.23)	<0.001
RASS	Per point	0.71 (0.60–0.84)	<0.001
Mechanical ventilation duration	Per hour	1.0013 (0.9988–1.0038)	0.304
Reintubation	Yes vs no	0.56 (0.30–1.02)	0.059
Sedative use	Yes vs no	1.41 (0.60–3.31)	0.432
Neuromuscular blocking agent use	Yes vs no	1.00 (0.54–1.85)	0.999
Enteral nutrition	Yes vs no	1.45 (0.80–2.63)	0.217
Parenteral nutrition	Yes vs no	3.67 (1.25–10.79)	0.018
NRS-2002			<0.001
0–2		Reference	—
3		11.09 (1.45–85.01)	0.021
≥4		41.16 (5.23–323.88)	<0.001
ICU stay to assessment	Per day	1.01 (0.97–1.06)	0.529
FRAIL	Per point	1.09 (0.80–1.49)	0.585
OHAT	Per point	1.54 (1.30–1.82)	<0.001

Overall P values for multicategory predictors use likelihood ratio tests; coefficient P values use Wald tests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx