

Original Paper

Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study

Natthanaphop Isaradech^{1,2*}, MD; Wachiranun Sirikul^{1,3,4*}, MD, MSc; Nida Buawangpong⁵, MD; Penprapa Siviroj¹, DrPH; Amornphat Kitro^{1,2}, MD, MCTM

¹Department of Community Medicine, Faculty of Medicine, Chiang Mai University, Meaung, Thailand

²Environmental and Occupational Medicine Excellence Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

³Center of Data Analytics and Knowledge Synthesis for Health Care, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

⁴Department of Biomedical Informatics and Clinical Epidemiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

⁵Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

*these authors contributed equally

Corresponding Author:

Wachiranun Sirikul, MD, MSc
Department of Community Medicine, Faculty of Medicine
Chiang Mai University
110, Intrawarorot Road
Meaung, 50200
Thailand
Phone: 66 53935472
Fax: 66 935476
Email: wachiranun.sir@cmu.ac.th

Related Article:

This is a corrected version. See correction statement in: <https://aging.jmir.org/2025/1/e75690>

Abstract

Background: Frailty is defined as a clinical state of increased vulnerability due to the age-associated decline of an individual's physical function resulting in increased morbidity and mortality when exposed to acute stressors. Early identification and management can reverse individuals with frailty to being robust once more. However, we found no integration of machine learning (ML) tools and frailty screening and surveillance studies in Thailand despite the abundance of evidence of frailty assessment using ML globally and in Asia.

Objective: We propose an approach for early diagnosis of frailty in community-dwelling older individuals in Thailand using an ML model generated from individual characteristics and anthropometric data.

Methods: Datasets including 2692 community-dwelling Thai older adults in Lampang from 2016 and 2017 were used for model development and internal validation. The derived models were externally validated with a dataset of community-dwelling older adults in Chiang Mai from 2021. The ML algorithms implemented in this study include the k-nearest neighbors algorithm, random forest ML algorithms, multilayer perceptron artificial neural network, logistic regression models, gradient boosting classifier, and linear support vector machine classifier.

Results: Logistic regression showed the best overall discrimination performance with a mean area under the receiver operating characteristic curve of 0.81 (95% CI 0.75–0.86) in the internal validation dataset and 0.75 (95% CI 0.71–0.78) in the external validation dataset. The model was also well-calibrated to the expected probability of the external validation dataset.

Conclusions: Our findings showed that our models have the potential to be utilized as a screening tool using simple, accessible demographic and explainable clinical variables in Thai community-dwelling older persons to identify individuals with frailty who require early intervention to become physically robust.

Keywords: aged care; gerontology; geriatric; old; aging; clinical decision support; delivering health information and knowledge to the public; diagnostic systems; digital health; epidemiology; surveillance; diagnosis; frailty; machine learning; prediction; predictive; AI; artificial intelligence; Thailand; community dwelling; health care intervention; patient care

Introduction

The world population is moving toward an aging society. As health care technology improves, people are expected to live longer and healthier [1]. According to the World Health Organization, the population aged ≥ 60 years will increase from 1 billion in 2020 to 2.1 billion in 2050 and the number of people aged ≥ 80 years will reach 426 million in 2050 [2]. Researchers predicted that the proportion of people in Thailand aged ≥ 60 years would be more than 20% of the population in 2025 and more than 30% in 2031 [3,4].

The prevalence of frailty is high among older adults aged ≥ 60 years [5]. Global frailty prevalence ranges from approximately 10% to 12% [6-11]. The percentage varies by age, gender, and frailty classification tool. In Thailand, frailty prevalence was 22.1%, which is twice the global frailty prevalence, according to the Thai National Health Examination Survey cohort in 2018. Specifically, Thailand's northern region frailty prevalence was found to be 15% [12,13]. This creates concerns about the increasing aging population in Thailand.

Frailty is defined as a clinical state of increased vulnerability due to the age-associated decline of an individual's body resulting in increased morbidity and mortality when exposed to everyday or acute stressors [14,15]. This clinical syndrome is associated with decreased quality of life [7], slow gait speed [16], more depressive symptoms, higher BMI, reduced cognitive function [17], decreased strength [12], and increased risk of fall, hospital re-admission, and all-cause mortality in the older adult population [13,18-20]. Frailty has become a crucial research topic because this clinical syndrome can be reversed. Studies have shown that early detection and intervention can revert individuals from a frail to a fit state [21-23].

However, incorporating frailty evaluation into clinical practice in a primary care context is challenging due to increased administrative tasks, time limitations, and a lack of diagnostic effectiveness [15,24,25]. In the age of technology, health informatics has become an important role in health care research [26,27]. Several studies have applied information technology to frailty detection in primary care settings using machine learning models and artificial intelligence [28-30]. A study from Canada showcased an efficient frailty identification tool using the XGBoost machine learning model. Features used for the model were medication, medical billing codes, and other primary care clinical data [31]. Another example showed the development of a predictive machine learning model for frailty conditions based on a database of demographic data and clinical characteristics [32]. In China, a study simplified the Frailty Index assessment for older individuals using machine learning techniques and showed that logistic regression was the best performing and

most interpretable model, with a mean area under the receiver operating characteristic curve (AUC) of 0.974 in the internal validation dataset [33]. Another study in China also developed and validated models using data from 6997 older adult participants to predict frailty risk, with random forest (RF) and logistic regression (LR) achieving AUC values of 0.77 and 0.76, respectively [34].

Despite the abundance of evidence of conducting frailty assessments using machine learning globally and in Asia, we have found no integration of data science tools and frailty screening and surveillance studies in Thailand. Most studies focused on the risk factors and their association with frailty syndrome but failed to show application in real-world settings [12,13,35,36]. This leads to our research question, "Can machine learning models predict frailty in community-dwelling older adults in Northern Thailand?" Therefore, we propose an approach to frailty detection using a machine learning model in the community-dwelling population from Lampang and Chiang Mai, Thailand, to effectively screen frailty status among Thai community-dwelling older adults and to help decrease clinicians' burden of work.

Methods

Source of Data

Development and Internal Validation Datasets

The datasets were derived from a cross-sectional study carried out in Lampang in 2016 and 2017; it is a northern Thai province with one of the highest aging indexes [37]. This study included older adults aged ≥ 60 years. Those with dementia (as determined by the Thai Mental State Examination), blindness, deafness, bedridden status, disabilities, or severe acute diseases were excluded.

To represent urban (8 villages), semiurban (8 villages), and rural (8 villages) communities, 24 villages in 3 districts were chosen. The records of the primary care unit were used to compile lists of community-dwelling older adults. The details of the data collection in this study are published elsewhere [36]. A total of 2228 older adults from this study were derived for model development and internal validation.

External Validation Dataset

The datasets used for external validation were derived from a cross-sectional study in Kuamung (suburban), Sankampang, Chiang Mai, Thailand, in 2021. The participants were included and excluded from the study with the same criteria as described in the internal validation datasets. A total of 464 older adults from this study were used for external validation of the derived models.

Predictors

Characteristics and Demographics

Assessments were conducted through a questionnaire at the participant's residence. The questionnaire included questions about sociodemographic information (age, household living arrangement, gender, and education level), self-reported medical diagnoses (such as hypertension, diabetes mellitus, and heart disease), level of physical activity per week, and exhaustion.

Anthropometric Variables

Anthropometric measures included BMI, waist circumference, and calf circumference (CC). A handheld dynamometer (Takei TKK5001) was used to assess handgrip strength. Height and calf circumference were measured with standard tape (Tajima brand, PIT-20BL model), and weight was measured with a calibrated weighing scale (Shaper Disney). The CC was measured over the unclothed area at the maximum diameter on the left leg. The tape was wrapped snugly around the calf and measured to the nearest 0.1 cm. All measurements were administered by 10 qualified field investigators, and the measurements were standardized by the principal investigator.

Outcome Variable

Frailty was evaluated based on Fried's phenotype [38], which includes the following five criteria:

1. Unintentional weight loss: The participant will be asked "In the last ten years, have you lost more than 10 pounds intentionally (not due to diet or exercise)?" If yes, then frail for weight loss criterion. At follow-up, weight loss was calculated as: $(\text{weight in previous year} - \text{current measured weight}) / (\text{weight in previous year}) = K$. If $K \geq 0.05$ and the subject did not report that he/she was trying to lose weight (ie, unintentional weight loss of at least 5% of previous year's body weight), then frail for weight loss=yes.
2. Exhaustion, assessed with self-report using Fried's method of assessment. The participant was first asked to self-assess whether she/he felt exhausted. If yes, she/he would be asked to rate the severity of the exhaustion. Ratings of 2-4 suggested a positive assessment.
3. Physical activity, based on the short version of the Minnesota Leisure time Activity questionnaire, which asked about walking, chores (moderately strenuous), mowing the lawn, raking, gardening, hiking, jogging, biking, exercise cycling, golf, single tennis, doubles tennis, racquetball, and calisthenics. For men, those with <383 kcals of physical activity per week were frail. For women, those with <270 kcals of physical activity per week were frail.
4. Walk time, stratified by gender and height (gender-specific cutoff for medium height).
 - Height ≤ 173 cm and walk time ≥ 7 seconds for men.

- Height > 173 cm and walk time ≥ 6 seconds for men.
 - Height ≤ 159 cm and walk time ≥ 7 seconds for women.
 - Height > 159 cm and walk time ≥ 6 seconds for women.
5. Grip strength, stratified by gender and BMI quartiles.
 - a. For men:
 - BMI ≤ 24 and grip strength ≤ 29 for men
 - BMI 24.1-26 and grip strength ≤ 30
 - BMI 26.1-28 and grip strength ≤ 30
 - BMI > 28 and grip strength ≤ 32
 - b. For women:
 - BMI ≤ 23 and grip strength ≤ 17
 - BMI 23.1-26 and grip strength ≤ 17.3
 - BMI 26.1-29 and grip strength ≤ 18
 - BMI < 29 and grip strength ≤ 21

Older adults with 3 or more phenotypes were considered to have physical frailty, and those who had 1 or 2 phenotypes were classified as prefrailty. According to the study design of the derived datasets, the outcome was concurrently measured with the predictors.

Sample Size Calculation

The sample size for model development was calculated using the *pmsampsize* package via STATA (version 16; StataCorp). A minimum sample size required 16.41 events per predictor, which was 1519 patients with 263 frailty cases based on a maximum candidate predictor of 16, frailty prevalence of 17.29% in the dataset for model development, small overfitting defined as expected shrinkage of predictor effects by 10% or less, a small difference in the developed model's apparent and optimism-adjusted values of 0.15 ($R^2_{\text{Nagelkerke}}$), as suggested by Riley et al [39]. For the external validation, the sample size was calculated using the *pmvalsampsize* package via STATA. From the available data of 464 patients with 192 frailty cases, it was able to precisely estimate confidence interval widths of observed/expected statistic of 0.22, calibration slope of 0.60, and concordance statistic of 0.10 [40].

Statistical Analysis Methods and Synthesis of the Results

Data exploration was performed using descriptive statistics to determine the data quantity, quality, and distribution. A frequency and a percentage were used to describe categorical variables. For continuous variables, a mean with SD was used for parametrically distributed data, and a median with an IQR was used for nonparametrically distributed data. The comparison of characteristics between the frail and nonfrail populations was performed using the chi-square test for categorical variables, the independent *t* test for parametric variables, and the rank-sum test for nonparametric variables. The methods for feature selection, model development, and validation are described in the section "Model Development and Validation". Estimates of model discrimination and optimism are reported as the mean AUC with 95% CI across all repetitions of cross-validation for the

internal validation and the AUC with 95% CI from 1000 bootstrapping samples for the external validation. To further explain model performance, we also created model calibration plots and calculated secondary metrics of prediction models, including the confusion matrix and specificity, sensitivity, and predictive values.

Missing Data and Imputation

There was only missing data in participants' age in the internal validation dataset (4/2228, 0.18%); therefore, a complete case analysis was performed on the dataset. The external validation dataset contained missing data, including age (2/464, 0.43%), BMI (11/464, 2.37%), waist and calf circumference (3/464, 0.64%), handgrip strength (2/464, 0.43%), and frailty status (4/464, 0.86%). We performed multiple imputations of missing data, except for frailty status (the outcome of interest), using the predictive mean matching imputation with 5 nearest neighbors via the *KNNImputer* from the *Scikit-Learn library 1.1.2*. Four participant records that did not have frailty status were removed (list-wise deletion).

Model Development and Validation

Feature Selection

The classification models were developed using the variables from the derived datasets. The variables for model development were selected using both a data-driven method and domain expertise. For a data-driven selection, a multivariable logistic regression with stepwise backward elimination was performed to determine statistically significant variables (P value less than .20). For the P value threshold of .20, we used a higher threshold to give priority to clinical reasoning in selecting variables by domain expertise and associated factors from the previous studies, along with statistical significance, which would allow more important variables to be entered into the model.

Rebalancing Data Strategy

We resolved unbalanced classification for the dataset's minority class by using the Synthetic Minority Oversampling Technique (SMOTE) to enhance model decision boundaries via the *imblearn.over_sampling.SMOTE* package [41,42].

Model Development and Internal Validation

The model development and validation were performed using Python (version 3.9; Python Software Foundation). The machine learning algorithms implemented in this study include the k-nearest neighbors (KNN) algorithm, RF machine learning algorithms, multilayer perceptron artificial neural network (MLP), gradient boosting classifier (GBC), linear support vector machine classifier (SVM), and LR models via the *Scikit-Learn library 1.1.2*. The hyperparameters were determined by using a grid search via the *Grid-Search CV* package with 10-fold cross-validation on the derived dataset to determine the parameters of each model that led to the best discriminative performance. For 10-fold cross-validation, the derived dataset was divided into 10 folds of data and repeated 10 times to perform model training

and testing. For each iteration, 9 folds of data were used to train the model and then it was tested with the remaining fold to ensure that almost all the derived data were used to train and test the models. The discriminative performance of the derived models was assessed by computing a confusion matrix and sensitivity, specificity, and predictive values, as well as AUC with a 95% CI.

The model calibration was evaluated using the calibration plot, which indicated the congruence between the observed proportion of the actual probability of outcome and the mean predicted probability (MPP) from the derived models.

External Validation

The derived models were validated again with the external validation dataset to determine the model optimism and calibration. The discriminative performance and the model optimism were re-evaluated and presented by discriminative performance matrices and a 95% CI of AUC from the 1000 bootstrapping samples, respectively. The model calibration using the external validation dataset was re-evaluated using the calibration plot.

Ethical Considerations

This study complies with the research with exemption category and has been certified by the Research Ethics Committee of the Faculty of Medicine, Chiang Mai University (study code: COM-2565-09159, number EXEMPTION 9159/2022). We requested a waiver of informed consent because this study is a retrospective analysis that exclusively utilizes anonymized secondary data from our research database, without collecting any additional information from medical records or other sources. All participants have previously provided informed consent for the primary data collection as described elsewhere [35,36,43]. All personal patient data were anonymized by removing citizen ID numbers, hospital numbers, addresses, and contact information from the dataset. The investigator cannot trace or identify individuals. The study results were reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis plus Artificial Intelligence (TRIPOD+AI) statement [44].

Results

Baseline Characteristics of Participants in Each Dataset

Baseline characteristics of participants in this study are shown in Table 1. The participants in the internal validation dataset had a mean age of 71.0 years. Most of the participants were male and had finished primary school. Among 2228 old adults, 2160 lived with either their spouse, relative, or children while the others were living alone. The average BMI of the participants was 32.6 (SD 7.4) kg/m². The prevalence of hypertension, dyslipidemia, type 2 diabetes mellitus, and heart disease were 45.26%, 19.67%, 16.21%, and 4.31%, respectively. The averages of waist circumference, calf circumference, handgrip strength, and walk time

were 83.40 (SD 11.23) cm, 32.40 (SD 4.35) cm, 32.40 (SD 6.68) kg, and 6.42 (SD 2.01) minutes, respectively. Overall, 9.79% of the participants were exhausted and 16.3% had

physical activity higher than 150 minutes per week. The prevalence of frailty was 17.3% (n=385).

Table 1. The characteristics of participants in the development and internal validation datasets and the external validation dataset.

Characteristics	Development and internal validation datasets (Lampang, 2016-2017; N=2228)	External validation datasets (Chiang Mai, 2021; N=464)	<i>P</i> value
Age (years), mean (SD)	70.96 (7.49)	70.68 (5.58)	.45
Gender, n (%)			
Male	1569 (70.45)	193 (41.59)	<.001
Female	658 (29.55)	271 (58.41)	
Household living arrangement, n (%)			
Living alone	160 (7.18)	42 (9.05)	<.001
Living with spouse	1177 (52.85)	9 (1.94)	
Living with children	823 (36.96)	283 (60.99)	
Living with relatives or others	67 (3.01)	130 (28.02)	
BMI (kg/m ²), mean (SD)	32.64 (7.40)	22.73 (3.89)	.001
Education, n (%)			
No education	192 (8.62)	12 (2.61)	<.001
Primary school	1745 (78.36)	398 (86.52)	
Secondary school or higher	290 (13.02)	50 (10.87)	
Underlying diseases, n (%)			
Hypertension	1008 (45.26)	237 (51.08)	.02
Dyslipidemia	438 (19.67)	79 (17.03)	.19
Type 2 diabetes mellitus	361 (16.21)	78 (16.81)	.75
Heart diseases	96 (4.31)	18 (3.88)	.68
Anthropometric variables, mean (SD)			
Waist circumference (cm)	83.40 (11.23)	81.16 (10.80)	<.001
Calf circumference (cm)	32.40 (4.35)	32.85 (4.55)	.04
Walk time (min), mean (SD)	6.42 (2.01)	8.50 (5.38)	<.001
Exhaustion, n (%)	218 (9.79)	189 (40.74)	<.001
Adequate level of physical activity as defined by the World Health Organization, n (%)	363 (16.30)	104 (22.42)	<.001
Grip strength (kg), mean (SD)	32.40 (6.68)	19.87 (7.29)	<.001
Frailty, n (%)	385 (17.29)	192 (41.74)	<.001

In the external validation dataset, the participants had a mean age of 70.68 years. Most of the participants were male and had finished primary school. Among 464 persons, 422 lived with either their spouse, relative, or children, while the others were living alone. The average BMI of the participants was 22.73 (SD 3.89) kg/m². The prevalence of hypertension, dyslipidemia, type 2 diabetes mellitus, and heart disease were 51.08%, 17.03%, 16.81%, and 3.88%, respectively. The averages of waist circumference, calf circumference, grip strength, and walk time were 81.16 (SD 10.80) cm, 32.85 (SD 4.55) cm, 19.87 (SD 7.29) kg, and 8.50 (SD 5.38) minutes, respectively. The prevalence of participants with exhausted state was 40.74%, and 22.42% of the participants participated in physical activity more than 150 minutes per week.

Comparing the variables between the 2 datasets, we found that gender, household living arrangement, BMI, education,

waist circumference, walk time, exhaustion, grip strength, and level of physical activity were significantly different.

Model Development

We used a dataset from Lampang (2016-2017; N=2228) for model development. The association between candidate predictors and frailty by univariate analysis is reported in [Table 2](#). Feature selection for model development was selected by a backward elimination approach via a multivariable logistic regression and expert judgment. We chose the following features as model predictors: age, gender, status, underlying diseases (hypertension and dyslipidemia), BMI, waist and calf circumference, and level of exhaustion. Finally, the derived data included 385 participants with frailty and 1842 participants without frailty, and all candidate predictors, as shown in [Table 2](#), were used in model development.

Table 2. The comparison of participants' characteristics and their associations with frailty in the development and internal validation datasets.

Characteristics	Adjusted odds ratio – full model (95% CI)	P value	Adjusted odds ratio – reduced model (95% CI)	P value
Age (years)	1.04 (1.02-1.07)	<.001	1.09 (1.06-1.11)	<.001
Gender				
Male	Reference		Reference	
Female	1.93 (1.26-2.94)	.002	0.90 (0.63-1.27)	.55
Status				
Living alone	Reference			
Living with others	0.92 (0.46-1.86)	.82	0.87 (0.45-1.68)	.68
BMI (kg/m ²)	0.92 (0.90-0.95)	<.001	0.91 (0.88-0.93)	<.001
Education				
No education	Reference			
Primary school	1.00 (0.59-1.70)	.99		
Secondary school or higher	1.34 (0.65-2.77)	.42		
Underlying diseases				
Hypertension	1.47 (1.03-2.10)	.03	1.27 (0.92-1.75)	.15
Dyslipidemia	1.22 (0.78-1.91)	.38	1.45 (0.97-2.17)	.07
Type 2 diabetes mellitus	0.84 (0.54-1.31)	.44		
Heart disease	1.03 (0.45-2.34)	.95		
Anthropometric variables				
Waist circumference (cm)	1.02 (1.00-1.04)	.046	1.02 (1.00-1.04)	.02
Calf circumference (cm)	1.05 (1.01-1.09)	.01	1.04 (1.01-1.08)	.02
Walk time (min)	1.89 (1.71-2.10)	<.001	1.95 (1.78-2.13)	<.001
Exhaustion	20.39 (12.61-32.94)	<.001	28.23 (18.22-43.73)	<.001
Adequate level of physical activity as defined by the World Health Organization	4.49 (3.06-6.59)	<.001		
Grip strength	0.83 (0.79-0.85)	<.001		

However, the model appeared to be poor in frailty prediction performance and imbalanced as the classifiers intended to classify only the majority class (accuracy paradox). Therefore, a rebalancing strategy by SMOTE was applied to counter this problem. The oversampling data were generated and rebalanced the minority group in a 1:1 ratio. The details of the final model's hyperparameters using *GridSearch CV* are presented in Table S1 in [Multimedia Appendix 1](#).

Adjusted odds ratios (aORs) of the full model were obtained from a multivariable logistic regression with all features. aORs of the reduced model were obtained from a multivariable logistic regression with stepwise backward elimination ($P < .10$) and feature selection based on the domain expertise.

Discrimination Performance of Internal Validated Models

We evaluated the model's performance by 10-fold cross-validation. The discrimination performances of the models are

presented in [Table 3](#) and [Figure 1](#). The overall discrimination performance was presented by mean AUC 10-fold cross-validation. The KNN model achieved the highest overall performance with a mean AUC of 0.85 (95% CI 0.82-0.88), followed by MLP (AUC 0.81, 95% CI 0.72-0.89), LR (AUC 0.81, 95% CI 0.75-0.86), and SVM (AUC 0.75, 95% CI 0.75-0.86), respectively. In addition, the KNN model had the highest sensitivity and specificity (89% and 76%, respectively). RF had the lowest discrimination performance with almost all metrics. Other metrics that were not affected by data rebalancing, like positive predictive value (PPV) and negative predictive value (NPV), were also used to express the performance of the models. The best performances for both PPV and NPV were found in the KNN model (0.79 and 0.87, respectively).

Table 3. Discrimination and optimism of internal validated models with 95% CIs.

Models	AUC ^a , mean (95% CI)	Predictive values		Sensitivity	Specificity
		Positive	Negative		
LR ^b	0.81 (0.75-0.86)	0.74 (0.68-0.80)	0.74 (0.69-0.78)	0.74 (0.69-0.79)	0.73 (0.65-0.81)
KNN ^c	0.85 (0.82-0.88)	0.79 (0.76-0.82)	0.88 (0.86-0.90)	0.89 (0.87-0.91)	0.76 (0.71-0.81)
RF ^d	0.70 (0.60-0.79)	0.65 (0.58-0.73)	0.70 (0.63-0.76)	0.79 (0.74-0.83)	0.55 (0.39-0.71)
MLP ^e	0.81 (0.72-0.89)	0.73 (0.64-0.81)	0.75 (0.68-0.81)	0.78 (0.71-0.84)	0.69 (0.57-0.81)
GBC ^f	0.74 (0.72-0.89)	0.66 (0.59-0.74)	0.70 (0.63-0.78)	0.78 (0.73-0.83)	0.58 (0.43-0.73)
SVM ^g	0.75 (0.75-0.86)	0.74 (0.68-0.80)	0.73 (0.69-0.78)	0.74 (0.69-0.79)	0.73 (0.65-0.81)

^aAUC: area under the receiver operating characteristic curve.

^bLR: logistic regression.

^cKNN: k-nearest neighbors.

^dRF: random forest.

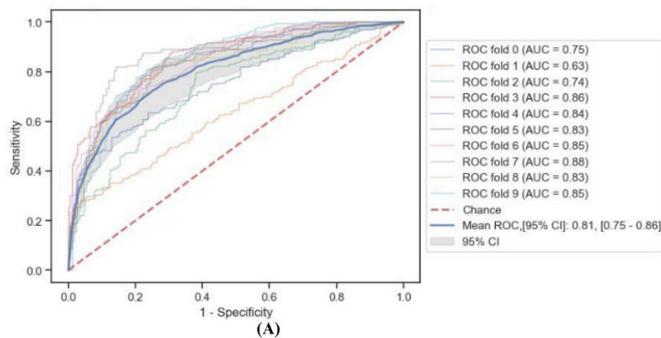
^eMLP: multilayer perceptron artificial neural network.

^fGBC: gradient boosting classifier.

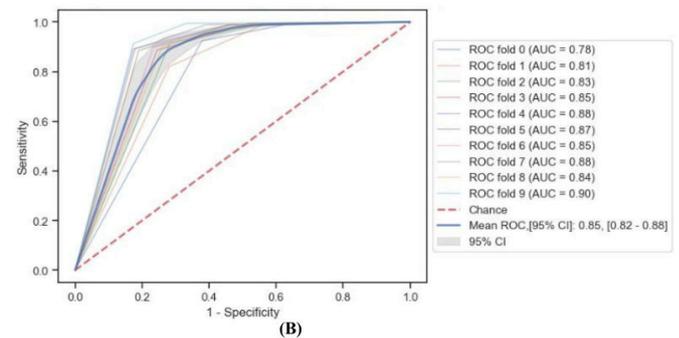
^gSVM: linear support vector machine classifier.

Figure 1. Receiver operating characteristic curves from 10-fold cross-validation of the rebalanced learning classifiers by SMOTE: (A) logistic regression model: mean AUC 0.81 (95% CI 0.75-0.86); (B) k-nearest neighbors model: mean AUC 0.85 (95% CI 0.82-0.88); (C) random forest model: mean AUC 0.70 (95% CI 0.60-0.79); (D) multilayer perceptron model: mean AUC 0.81 (95% CI 0.72-0.89); (E) gradient boosting classifier model: mean AUC 0.74 (95% CI 0.66-0.82); (F) linear support vector machine classifier model: mean AUC 0.75 (95% CI 0.75-0.86). ROC: receiver operating characteristic curve; SMOTE: synthetic minority oversampling technique.

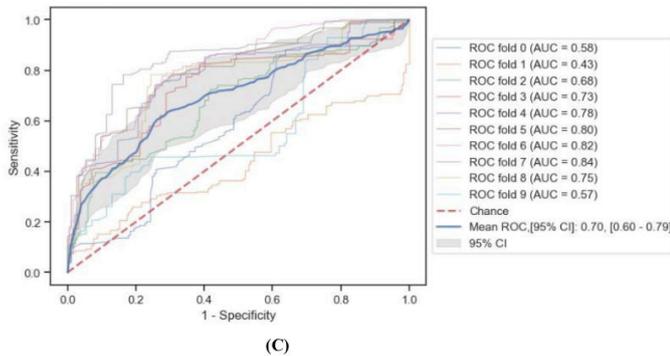
Receiver operating characteristic logistic regression using rebalanced data by SMOTE



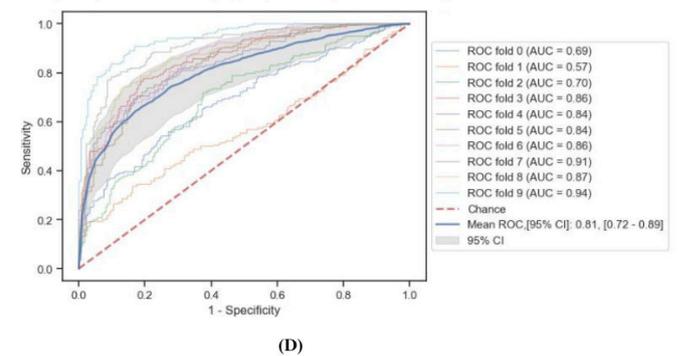
Receiver operating characteristic k-nearest neighbors using rebalanced data by SMOTE



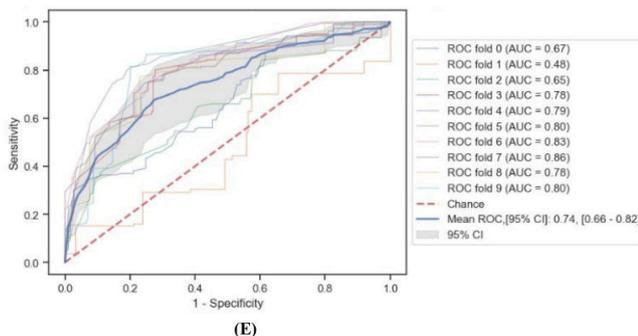
Receiver operating characteristic random forest using rebalanced data by SMOTE



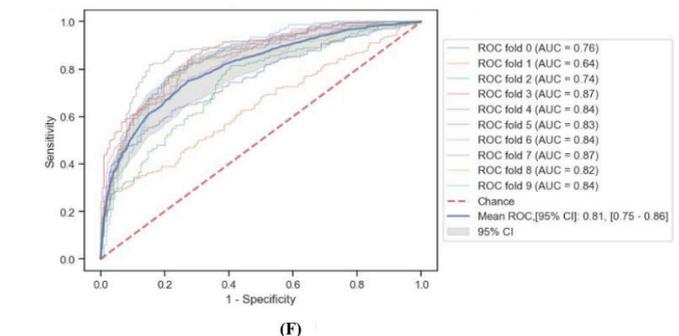
Receiver operating characteristic multilayer perceptrons using rebalanced data by SMOTE



Receiver operating characteristic gradient boosting classifier using rebalanced data by SMOTE



Receiver operating characteristic linear support vector classifier using rebalanced data by SMOTE



Discrimination Performance of External Validated Models

We validated our trained model with an external validation dataset to evaluate the bias and variance of our trained

models. SMOTE was not applied to the dataset since it was already well-balanced between participants with and without frailty. The performance of machine learning models validated by the external validation dataset is shown in [Table 4](#).

Table 4. Discrimination and optimism of external validated models.

Models and model prediction	True label (frailty/ nonfrailty)	AUC ^a , mean (95% CI)	Predictive values		Sensitivity	Specificity
			Positive	Negative		
Logistic regression						
Frailty	140/65	0.75 (0.71-0.78)	0.68	0.80	0.73	0.76
Nonfrailty	52/207					
K-nearest neighbors						
Frailty	41/35	0.54 (0.51-0.57)	0.54	0.61	0.21	0.87
Nonfrailty	151/237					
Random forest						
Frailty	137/63	0.75 (0.71-0.78)	0.69	0.79	0.71	0.77
Nonfrailty	55/209					
Multilayer perceptron artificial neural network						
Frailty	98/38	0.68 (0.65-0.72)	0.72	0.71	0.51	0.86
Nonfrailty	94/234					
Gradient boosting classifier						
Frailty	89/31	0.69 (0.65-0.72)	0.74	0.70	0.46	0.89
Nonfrailty	103/241					
Linear support vector machine classifier						
Frailty	131/60	0.73 (0.70-0.77)	0.69	0.78	0.68	0.78
Nonfrailty	61/212					

^aAUC: area under the receiver operating characteristic curve.

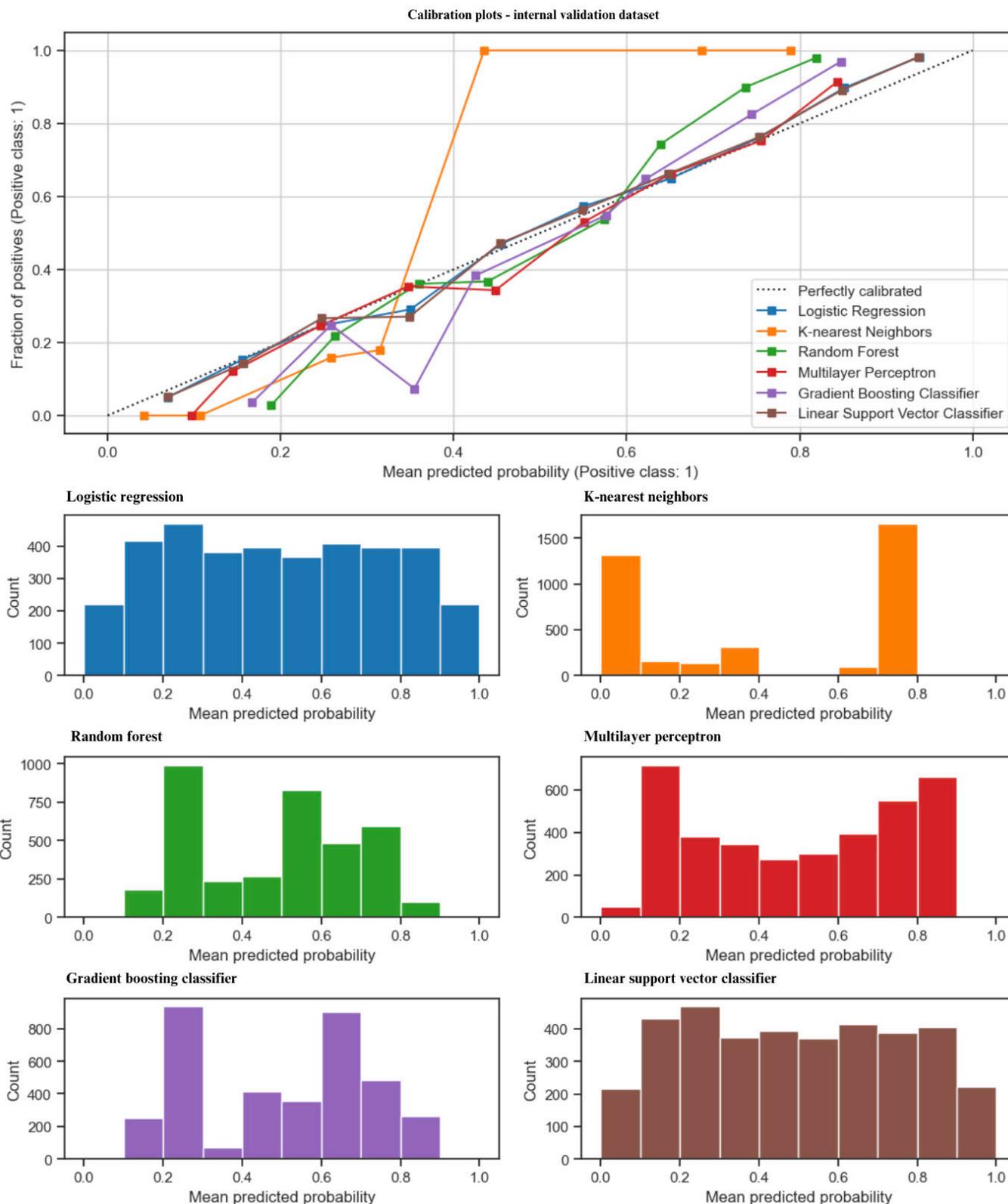
The overall discrimination performance was presented by mean AUC and 95% CI via 1000-fold bootstrapping. The LR and RF model achieved the highest overall performance with mean AUC 0.75 (95% CI 0.71-0.78), followed by SVM (mean AUC 0.73, 95% CI 0.70-0.77), and GBC (mean AUC 0.69, 95% CI 0.65-0.72), respectively. The LR model had the highest sensitivity (73%) and the MLP model had the highest specificity (86%). KNN had the lowest discrimination performance with all metrics. Other metrics that were not affected by data rebalancing, like PPV and NPV, were also used to express the performance of the models. The best performance for PPV was found in the GBC model (0.74); for NPV, it was found in the LR model (0.80).

Model Calibration

The model calibration was visualized with the calibration plot, which compared the expected probability of frailty, and

the mean 10-fold cross-validation predicted the probability of each model. In the internal validation dataset ([Figure 2](#)), the LR, SVM, and MLP were well-calibrated with the expected probability of the data. The distribution of predicted probabilities of the models was well-balanced between 0.00 and 1.00. The KNN classifier aligned well with the MPP between 0.00 and 0.40 but it overestimated the high predicted probability. In contrast, the GBC underestimated MPP between 0.20 and 0.40 but aligned well with the rest of the expected probability. These results suggest that LR, SVM, and MLP are the most reliable for balanced predictions, while KNN, GBC, and RF require careful consideration depending on the probability range of interest.

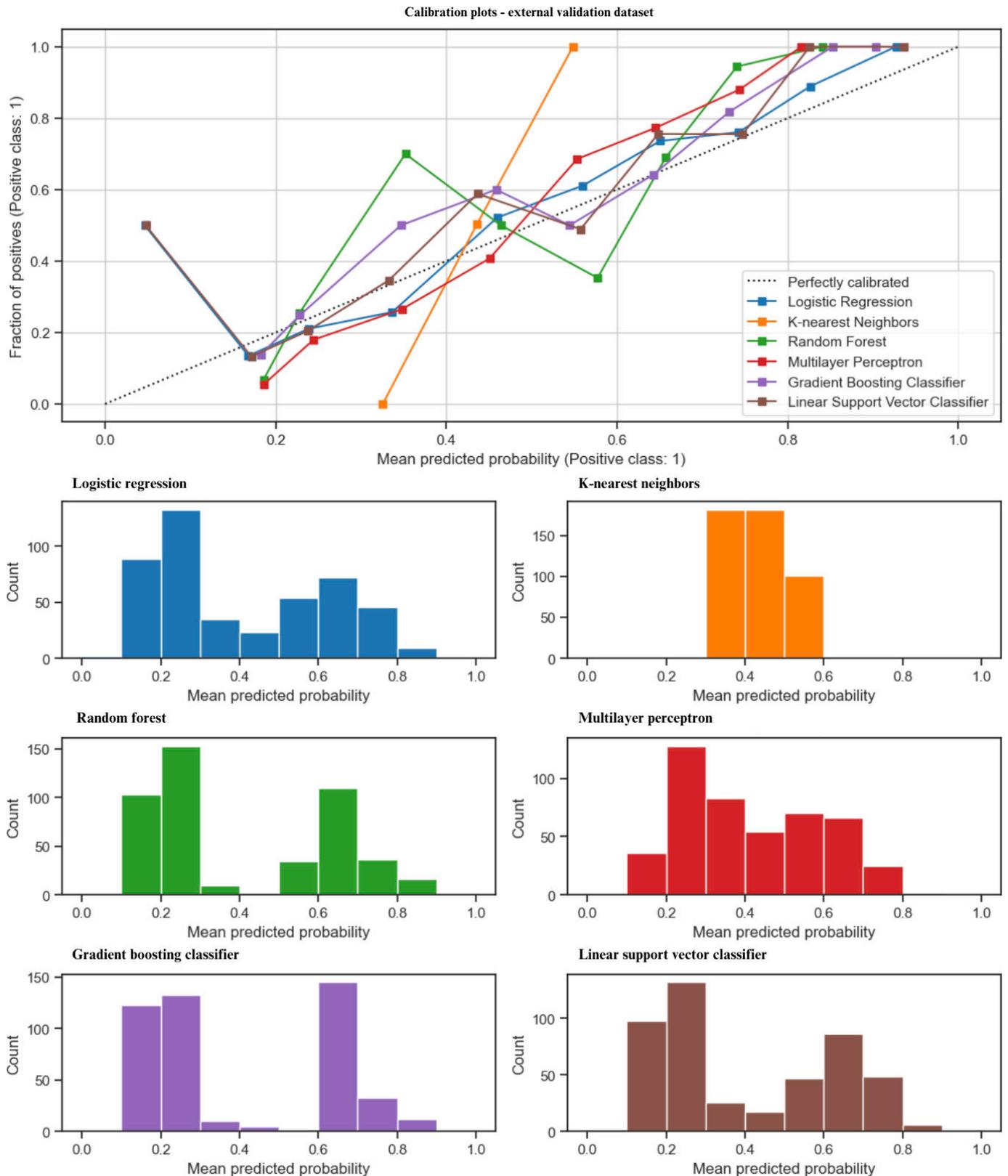
Figure 2. Model calibration of the developed models with internal validation.



In the external validation dataset (Figure 3), the LR, SVM, and MLP were almost perfectly calibrated to the expected probability of the data with a slight overestimation in the 0.40-0.80 MPP for MLP and an overestimation in the 0.00-0.40 MPP for LR and SVM. These 3 models have a similarly balanced distribution of MPP. RF aligned poorly

with an overestimation in the low and high MPP, and it underestimated in the range between 0.40 and 0.70. The GBC aligned well with the calibration plot after 0.60 but overestimated before 0.40. The KNN classifier was poorly calibrated to the plot, giving only MPP in the range of 0.30-0.60.

Figure 3. Model calibration of the developed models with external validation.



Discussion

Model Performance

The internal validated models all showed remarkable discrimination performance, with the KNN model having

the highest overall discrimination performance. However, the KNN model performed poorly in external validation datasets, indicating overfitting and the inability to generalize its results to other populations. In terms of model optimism, the LR, SVM, and MLP models had better discrimination capabilities, despite a slight decline in the models' performance

during external validation. The mean AUCs of these models remained in the range of fair to good performance. When considering all aspects of model performance, the LR model was the most preferable because it provided good discrimination (AUC), an optimal range of predicted probabilities, and good calibration in both the internal and external validation datasets. On the other hand, other models had predicted probabilities that were not extreme enough, resulting in poor to fair calibrations, and they lacked consistency between the internal and external validation datasets. In comparing the performance and validation aspects of previous frailty prediction models, most studies showed better performance than our best model (LR with a mean AUC of 0.81 in the internal validation set) [45]. The ELSA cohort study (2023) [46] reported the best performance from an RF model with an internal AUC of 0.92. Similarly, Bu et al (2023) [47] presented a multivariate LR model for predicting frailty in diabetic patients that had an AUC of 0.88 in the internal validation set. Additionally, a study on predictive modeling for frailty in older people using various machine learning methods found that artificial neural networks and SVM had the best performance in predicting mortality, with accuracies around 0.78-0.79 [32]. Although these studies reported high internal performance metrics, the absence of external validation raises concerns about their generalizability, especially for our focus on the community-dwelling older adult population in general. Our LR showed a mean AUC of 0.75 for the external validation dataset, which dropped slightly from the internal validated results, demonstrating a better edge in the community setting for our focused population. Furthermore, the predictors included in the previously reported model played an important role in prediction performance. Particularly, the utilization of handgrip strength in the study conducted by Bu et al [47] and the assessment of balance and chair stand in the ELSA cohort study [46] could potentially explain the high reported performance in both studies. These predictors either formed part of the frailty phenotype or served as surrogates for physical performance. However, incorporating these predictors into the models necessitated a trade-off between their value added to model performance and the difficulties of model utilization due to the time and skill required for the assessments. In our study, we proposed more parsimonious models using simpler predictions, which still achieved satisfactory performance for frailty screening.

Feature Selection and Findings Explanation

This study's findings suggest that machine learning models can be effective in classifying frailty status among community-dwelling older adults in Northern Thailand using simple predictors including, age, gender, household living arrangement, underlying diseases (hypertension and dyslipidemia), BMI, waist and calf circumference, and level of exhaustion. Age, BMI, waist circumference, and calf circumference are all potent risk factors for frailty in older adults. When the human body deviates from the normal physiologic process of aging, our levels of estrogen and testosterone gradually decline. These hormones play a vital role in maintaining

muscle and bone mass, enhancing strength, and promoting optimal nervous system function [48,49]. As a result, the aging process can accelerate the decline of muscle, bone, and the nervous system, transforming an individual from fit to frail. Additionally, the female population is at higher risk of frailty because the normal bone turnover cycle is disrupted by estrogen deficiency during menopause, increasing bone resorption over deposition, and resulting in net bone loss in women [50]. Household living arrangement was also added to model features as we found that there were studies that showed an association of social adversity and support with frailty status [51,52]. Lastly, we selected level of exhaustion as a predictor in our models, as it had the highest crude aOR among all other features, and we found that it is highly feasible and time-efficient to acquire this data in real clinical settings, compared to other Fried's phenotypes, which involve multiple anthropometric and physical performance tests. These predictors also showed a significant contribution to the frailty prediction in the previous studies [46,47]. Most of our features used in the model are easily collectible, which makes it highly feasible and time-efficient to acquire all data in real clinical settings without depending on a high-level professional, making our model highly applicable for early frailty screening.

Limitations

The most important limitation of this study's models is the generalizability of the models. Our internal and external validation datasets were collected from a community-dwelling population in Lampang and Chiang Mai, respectively, which we confidently believe means that our models have high generalizability to the general population in Thailand. From our perspective, we assume that the models could also be implemented in other Asian countries, as studies showed similar frailty prevalence and population characteristics such as anthropometric measurements, age, household living arrangement, and underlying diseases of hypertension and diabetes [53-58].

However, we encourage conducting validation studies for other regions and populations before clinical application as frailty risk factors do vary across countries as well as in regions with different socioeconomic and health care contexts [59-61]. Another limitation of our study is spectrum bias, as our models were only able to distinguish between frail and robust older adults, despite Fried's criteria having 3 stages of frailty: robust, prefrail, and frail [62]. Nonetheless, we do not believe this bias will significantly affect our study's primary objective, which is the early detection of frailty in older adults to effectively prescribe nutrition and exercise interventions.

Implications

Identifying frailty early helps patient gain access to interventions like nutritional support and exercise programs faster, improving outcomes for older adults by preventing frailty progression, reducing falls, hospitalizations, and mortality, and enhancing quality of life [63,64]. We see 2 possibilities for the practical application of our developed and validated models. One option is to develop a web application that

serves as a frailty screening tool that could be self-assessed by individuals or be used in outpatient clinic settings to screen patients. This would help health care providers efficiently identify patients with frailty who require closer monitoring or interventions. We have implemented the validated models to run on our web application, which can be accessed at the link in the Data Availability section.

Another option is to incorporate our machine learning models into electronic medical record or health surveillance systems. The machine learning models could be integrated with the electronic medical record system to provide automated frailty probability scores for each patient. This would enable health care professionals to identify patients who require closer monitoring or interventions and could help optimize treatment plans for the community-dwelling population at risk of frailty. Furthermore, this could help future researchers retrieve and analyze frailty data from the hospital easier, leading to a better understanding of the factors that contribute to frailty and the development of more effective interventions, which also promotes more efficient use of resources within the health care system.

Conclusion

Machine learning models were fairly good at classifying frailty status among Thai community-dwelling older adults using age, gender, household living arrangement, underlying diseases (hypertension and dyslipidemia), BMI, waist and calf circumference, and level of exhaustion as predictors. The LR and RF models demonstrated the best discrimination performance and model calibration in both the internal and external validation datasets.

There are 2 potential practical applications for utilizing the study findings. These include creating a web application for self-screening or individual screening and incorporating machine learning models into electronic medical record or surveillance systems to provide automated frailty probability scores for individual patients. We advocate for further research on model external validation and temporal recalibration to enhance the model's practicality and applicability to the specific context in which it is used.

Acknowledgments

This work was supported by Chiang Mai University.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The validated model's prediction and application interface can be accessed through the web application [65].

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary materials

[DOCX File (Microsoft Word File), 15 KB-Multimedia Appendix 1]

References

1. World population ageing, 2019: highlights. United Nations Digital Library. URL: <http://digitallibrary.un.org/record/3846855> [Accessed 2025-03-23]
2. Ageing and health. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health> [Accessed 2022-08-02]
3. Population ageing in Thailand. ERIA. URL: <http://www.eria.org/publications/population-ageing-in-thailand/> [Accessed 2022-08-02]
4. Tantirat P, Suphanchaimat R, Rattanathumsakul T, Noree T. Projection of the number of elderly in different health states in Thailand in the next ten years, 2020-2030. *Int J Environ Res Public Health*. Nov 23, 2020;17(22):8703. [doi: [10.3390/ijerph17228703](https://doi.org/10.3390/ijerph17228703)] [Medline: [33238588](https://pubmed.ncbi.nlm.nih.gov/33238588/)]
5. O'Caomh R, Sezgin D, O'Donovan MR, et al. Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies. *Age Ageing*. Jan 8, 2021;50(1):96-104. [doi: [10.1093/ageing/afaa219](https://doi.org/10.1093/ageing/afaa219)] [Medline: [33068107](https://pubmed.ncbi.nlm.nih.gov/33068107/)]
6. Siriwardhana DD, Hardoon S, Rait G, Weerasinghe MC, Walters KR. Prevalence of frailty and prefrailty among community-dwelling older adults in low-income and middle-income countries: a systematic review and meta-analysis. *BMJ Open*. Mar 1, 2018;8(3):e018195. [doi: [10.1136/bmjopen-2017-018195](https://doi.org/10.1136/bmjopen-2017-018195)] [Medline: [29496895](https://pubmed.ncbi.nlm.nih.gov/29496895/)]
7. Nguyen AT, Nguyen LH, Nguyen TX, et al. Frailty prevalence and association with health-related quality of life impairment among rural community-dwelling older adults in Vietnam. *Int J Environ Res Public Health*. Oct 12, 2019;16(20):3869. [doi: [10.3390/ijerph16203869](https://doi.org/10.3390/ijerph16203869)] [Medline: [31614836](https://pubmed.ncbi.nlm.nih.gov/31614836/)]
8. Siriwardhana DD, Weerasinghe MC, Rait G, Falcaro M, Scholes S, Walters KR. Prevalence of frailty in rural community-dwelling older adults in Kegalle district of Sri Lanka: a population-based cross-sectional study. *BMJ Open*. Jan 25, 2019;9(1):e026314. [doi: [10.1136/bmjopen-2018-026314](https://doi.org/10.1136/bmjopen-2018-026314)] [Medline: [30782757](https://pubmed.ncbi.nlm.nih.gov/30782757/)]

9. Wu C, Smit E, Xue QL, Odden MC. Prevalence and correlates of frailty among community-dwelling Chinese older adults: the China Health and Retirement Longitudinal Study. *J Gerontol A Biol Sci Med Sci*. Dec 12, 2017;73(1):102-108. [doi: [10.1093/gerona/glx098](https://doi.org/10.1093/gerona/glx098)] [Medline: [28525586](https://pubmed.ncbi.nlm.nih.gov/28525586/)]
10. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. *J Am Geriatr Soc*. Aug 2012;60(8):1487-1492. [doi: [10.1111/j.1532-5415.2012.04054.x](https://doi.org/10.1111/j.1532-5415.2012.04054.x)] [Medline: [22881367](https://pubmed.ncbi.nlm.nih.gov/22881367/)]
11. Vaingankar JA, Chong SA, Abdin E, et al. Prevalence of frailty and its association with sociodemographic and clinical characteristics, and resource utilization in a population of Singaporean older adults. *Geriatr Gerontol Int*. Oct 2017;17(10):1444-1454. [doi: [10.1111/ggi.12891](https://doi.org/10.1111/ggi.12891)] [Medline: [27576598](https://pubmed.ncbi.nlm.nih.gov/27576598/)]
12. Semmarath W, Seesen M, Yodkeeree S, et al. The association between frailty indicators and blood-based biomarkers in early-old community dwellers of Thailand. *Int J Environ Res Public Health*. Sep 17, 2019;16(18):3457. [doi: [10.3390/ijerph16183457](https://doi.org/10.3390/ijerph16183457)] [Medline: [31533354](https://pubmed.ncbi.nlm.nih.gov/31533354/)]
13. Srinonprasert V, Chalerm Sri C, Aekplakorn W. Frailty index to predict all-cause mortality in Thai community-dwelling older population: a result from a National Health Examination Survey cohort. *Arch Gerontol Geriatr*. 2018;77:124-128. [doi: [10.1016/j.archger.2018.05.002](https://doi.org/10.1016/j.archger.2018.05.002)] [Medline: [29751290](https://pubmed.ncbi.nlm.nih.gov/29751290/)]
14. Xue QL. The frailty syndrome: definition and natural history. *Clin Geriatr Med*. Feb 2011;27(1):1-15. [doi: [10.1016/j.cger.2010.08.009](https://doi.org/10.1016/j.cger.2010.08.009)] [Medline: [21093718](https://pubmed.ncbi.nlm.nih.gov/21093718/)]
15. Dent E, Morley JE, Cruz-Jentoft AJ, et al. Physical frailty: ICFSR International Clinical Practice Guidelines for Identification and Management. *J Nutr Health Aging*. 2019;23(9):771-787. [doi: [10.1007/s12603-019-1273-z](https://doi.org/10.1007/s12603-019-1273-z)] [Medline: [31641726](https://pubmed.ncbi.nlm.nih.gov/31641726/)]
16. Setiati S, Laksmi PW, Aryana I, et al. Frailty state among Indonesian elderly: prevalence, associated factors, and frailty state transition. *BMC Geriatr*. Jul 3, 2019;19(1):182. [doi: [10.1186/s12877-019-1198-8](https://doi.org/10.1186/s12877-019-1198-8)] [Medline: [31269921](https://pubmed.ncbi.nlm.nih.gov/31269921/)]
17. Robertson DA, Savva GM, Kenny RA. Frailty and cognitive impairment--a review of the evidence and causal mechanisms. *Ageing Res Rev*. Sep 2013;12(4):840-851. [doi: [10.1016/j.arr.2013.06.004](https://doi.org/10.1016/j.arr.2013.06.004)] [Medline: [23831959](https://pubmed.ncbi.nlm.nih.gov/23831959/)]
18. End of life care in frailty: falls. British Geriatrics Society. URL: <https://www.bgs.org.uk/resources/end-of-life-care-in-frailty-falls> [Accessed 2022-08-02]
19. Cheng MH, Chang SF. Frailty as a risk factor for falls among community dwelling people: evidence from a meta-analysis. *J Nurs Scholarsh*. Sep 2017;49(5):529-536. [doi: [10.1111/jnu.12322](https://doi.org/10.1111/jnu.12322)] [Medline: [28755453](https://pubmed.ncbi.nlm.nih.gov/28755453/)]
20. Davis-Ajami ML, Chang PS, Wu J. Hospital readmission and mortality associations to frailty in hospitalized patients with coronary heart disease. *Aging and Health Research*. Dec 2021;1(4):100042. [doi: [10.1016/j.ahr.2021.100042](https://doi.org/10.1016/j.ahr.2021.100042)]
21. Travers J, Romero-Ortuno R, Bailey J, Cooney MT. Delaying and reversing frailty: a systematic review of primary care interventions. *Br J Gen Pract*. Jan 2019;69(678):e61-e69. [doi: [10.3399/bjgp18X700241](https://doi.org/10.3399/bjgp18X700241)] [Medline: [30510094](https://pubmed.ncbi.nlm.nih.gov/30510094/)]
22. Bray NW, Smart RR, Jakobi JM, Jones GR. Exercise prescription to reverse frailty. *Appl Physiol Nutr Metab*. Oct 2016;41(10):1112-1116. [doi: [10.1139/apnm-2016-0226](https://doi.org/10.1139/apnm-2016-0226)] [Medline: [27649859](https://pubmed.ncbi.nlm.nih.gov/27649859/)]
23. Marcucci M, Damanti S, Germini F, et al. Interventions to prevent, delay or reverse frailty in older people: a journey towards clinical guidelines. *BMC Med*. Oct 29, 2019;17(1):193. [doi: [10.1186/s12916-019-1434-2](https://doi.org/10.1186/s12916-019-1434-2)] [Medline: [31660959](https://pubmed.ncbi.nlm.nih.gov/31660959/)]
24. Turner G, Clegg A, British Geriatrics Society, Age UK, Royal College of General Practitioners. Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. *Age Ageing*. Nov 2014;43(6):744-747. [doi: [10.1093/ageing/afu138](https://doi.org/10.1093/ageing/afu138)] [Medline: [25336440](https://pubmed.ncbi.nlm.nih.gov/25336440/)]
25. Bergman H, Ferrucci L, Guralnik J, et al. Frailty: an emerging research and clinical paradigm--issues and controversies. *J Gerontol A Biol Sci Med Sci*. Jul 2007;62(7):731-737. [doi: [10.1093/gerona/62.7.731](https://doi.org/10.1093/gerona/62.7.731)] [Medline: [17634320](https://pubmed.ncbi.nlm.nih.gov/17634320/)]
26. Savel TG, Foldy S, Centers for Disease Control and Prevention. The role of public health informatics in enhancing public health surveillance. *MMWR Suppl*. Jul 27, 2012;61(3):20-24. [Medline: [22832993](https://pubmed.ncbi.nlm.nih.gov/22832993/)]
27. Gamache R, Kharrazi H, Weiner JP. Public and population health informatics: the bridging of big data to benefit communities. *Yearb Med Inform*. Aug 2018;27(1):199-206. [doi: [10.1055/s-0038-1667081](https://doi.org/10.1055/s-0038-1667081)] [Medline: [30157524](https://pubmed.ncbi.nlm.nih.gov/30157524/)]
28. Yang H, Chang J, He W, Wee CF, Yit JST, Feng M. Frailty modeling using machine learning methodologies: a systematic review with discussions on outstanding questions. *IEEE J Biomed Health Inform*. 2024;29(1):631-642. [doi: [10.1109/JBHI.2024.3430226](https://doi.org/10.1109/JBHI.2024.3430226)]
29. Bian C, Ye B, Chu CH, McGilton KS, Mihailidis A. Technology for home-based frailty assessment and prediction: a systematic review. *Gerontechnology*. Nov 2, 2020;19(3):1-13. [doi: [10.4017/gt.2020.19.003.06](https://doi.org/10.4017/gt.2020.19.003.06)]
30. Mohanty SD, Lekan D, McCoy TP, Jenkins M, Manda P. Machine learning for predicting readmission risk among the frail: explainable AI for healthcare. *Patterns (N Y)*. Jan 14, 2022;3(1):100395. [doi: [10.1016/j.patter.2021.100395](https://doi.org/10.1016/j.patter.2021.100395)] [Medline: [35079714](https://pubmed.ncbi.nlm.nih.gov/35079714/)]
31. Aponte-Hao S, Wong ST, Thandi M, et al. Machine learning for identification of frailty in Canadian primary care practices. *Int J Popul Data Sci*. 2021;6(1):1650. [doi: [10.23889/ijpds.v6i1.1650](https://doi.org/10.23889/ijpds.v6i1.1650)] [Medline: [34541337](https://pubmed.ncbi.nlm.nih.gov/34541337/)]

32. Tarekegn A, Ricceri F, Costa G, Ferracin E, Giacobini M. Predictive modeling for frailty conditions in elderly people: machine learning approaches. *JMIR Med Inform.* Jun 4, 2020;8(6):e16678. [doi: [10.2196/16678](https://doi.org/10.2196/16678)] [Medline: [32442149](https://pubmed.ncbi.nlm.nih.gov/32442149/)]
33. Huang L, Chen H, Liang Z. Enhancing the convenience of frailty index assessment for elderly Chinese people with machine learning methods. *Sci Rep.* Oct 5, 2024;14(1):23227. [doi: [10.1038/s41598-024-74194-x](https://doi.org/10.1038/s41598-024-74194-x)] [Medline: [39369089](https://pubmed.ncbi.nlm.nih.gov/39369089/)]
34. Zhang W, Wang J, Xie F, et al. Development and validation of machine learning models to predict frailty risk for elderly. *J Adv Nurs.* Dec 2024;80(12):5064-5075. [doi: [10.1111/jan.16192](https://doi.org/10.1111/jan.16192)] [Medline: [38605460](https://pubmed.ncbi.nlm.nih.gov/38605460/)]
35. Seesen M, Sirikul W, Ruangsuriya J, Griffiths J, Siviroj P. Cognitive frailty in Thai community-dwelling elderly: prevalence and its association with malnutrition. *Nutrients.* Nov 25, 2021;13(12):4239. [doi: [10.3390/nu13124239](https://doi.org/10.3390/nu13124239)] [Medline: [34959791](https://pubmed.ncbi.nlm.nih.gov/34959791/)]
36. Thinuan P, Siviroj P, Lerttrakarnnon P, Lorga T. Prevalence and potential predictors of frailty among community-dwelling older persons in Northern Thailand: a cross-sectional study. *Int J Environ Res Public Health.* Jun 8, 2020;17(11):4077. [doi: [10.3390/ijerph17114077](https://doi.org/10.3390/ijerph17114077)] [Medline: [32521642](https://pubmed.ncbi.nlm.nih.gov/32521642/)]
37. Haque MN. Active ageing level of older persons: regional comparison in Thailand. *J Aging Res.* 2016;2016:9093018. [doi: [10.1155/2016/9093018](https://doi.org/10.1155/2016/9093018)] [Medline: [27375903](https://pubmed.ncbi.nlm.nih.gov/27375903/)]
38. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci.* Mar 2001;56(3):M146-56. [doi: [10.1093/gerona/56.3.m146](https://doi.org/10.1093/gerona/56.3.m146)] [Medline: [11253156](https://pubmed.ncbi.nlm.nih.gov/11253156/)]
39. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ.* Mar 18, 2020;368:m441. [doi: [10.1136/bmj.m441](https://doi.org/10.1136/bmj.m441)] [Medline: [32188600](https://pubmed.ncbi.nlm.nih.gov/32188600/)]
40. Riley RD, Snell KIE, Archer L, et al. Evaluation of clinical prediction models (part 3): calculating the sample size required for an external validation study. *BMJ.* Jan 22, 2024;384:e074821. [doi: [10.1136/bmj-2023-074821](https://doi.org/10.1136/bmj-2023-074821)] [Medline: [38253388](https://pubmed.ncbi.nlm.nih.gov/38253388/)]
41. Lemaître G, Nogueira F, Aridas CK. Imbalanced-learn: a Python toolbox to tackle the curse of imbalanced datasets in machine learning. *J Mach Learn Res.* 2017;18(17):1-5.
42. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *Jair.* Jun 1, 2002;16:321-357. [doi: [10.1613/jair.953](https://doi.org/10.1613/jair.953)]
43. Chittrakul J, Siviroj P, Sungkarat S, Sapbamrer R. Physical frailty and fall risk in community-dwelling older adults: a cross-sectional study. *J Aging Res.* 2020;2020:3964973. [doi: [10.1155/2020/3964973](https://doi.org/10.1155/2020/3964973)] [Medline: [32714612](https://pubmed.ncbi.nlm.nih.gov/32714612/)]
44. 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.* Apr 16, 2024;385:e078378. [doi: [10.1136/bmj-2023-078378](https://doi.org/10.1136/bmj-2023-078378)] [Medline: [38626948](https://pubmed.ncbi.nlm.nih.gov/38626948/)]
45. Velazquez-Diaz D, Arco JE, Ortiz A, et al. Use of artificial intelligence in the identification and diagnosis of frailty syndrome in older adults: scoping review. *J Med Internet Res.* Oct 20, 2023;25(1):e47346. [doi: [10.2196/47346](https://doi.org/10.2196/47346)] [Medline: [37862082](https://pubmed.ncbi.nlm.nih.gov/37862082/)]
46. Leme D da C, de Oliveira C. Machine learning models to predict future frailty in community-dwelling middle-aged and older adults: the ELSA cohort study. *J Gerontol A Biol Sci Med Sci.* Oct 28, 2023;78(11):2176-2184. [doi: [10.1093/gerona/glad127](https://doi.org/10.1093/gerona/glad127)] [Medline: [37209408](https://pubmed.ncbi.nlm.nih.gov/37209408/)]
47. Bu F, Deng XH, Zhan NN, et al. Development and validation of a risk prediction model for frailty in patients with diabetes. *BMC Geriatr.* Mar 27, 2023;23(1):172. [doi: [10.1186/s12877-023-03823-3](https://doi.org/10.1186/s12877-023-03823-3)] [Medline: [36973658](https://pubmed.ncbi.nlm.nih.gov/36973658/)]
48. Dobbs MB, Buckwalter J, Saltzman C. Osteoporosis: the increasing role of the orthopaedist. *Iowa Orthop J.* 1999;19:43-52. [Medline: [10847516](https://pubmed.ncbi.nlm.nih.gov/10847516/)]
49. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing.* Jul 1, 2019;48(4):601-601. [doi: [10.1093/ageing/afz046](https://doi.org/10.1093/ageing/afz046)] [Medline: [31081853](https://pubmed.ncbi.nlm.nih.gov/31081853/)]
50. Truumees E. Chapter 22 - Osteoporosis. In: Errico TJ, Lonner BS, Moulton AW, editors. *Surg Manag Spinal Deform.* W.B. Saunders; 2009:351-372. [doi: [10.1016/B978-141603372-1.50025-1](https://doi.org/10.1016/B978-141603372-1.50025-1)] ISBN: 978-1-4160-3372-1
51. Gomes CDS, Guerra RO, Wu YY, et al. Social and economic predictors of worse frailty status occurrence across selected countries in North and South America and Europe. *Innov Aging.* Sep 2018;2(3):igy037. [doi: [10.1093/geroni/igy037](https://doi.org/10.1093/geroni/igy037)] [Medline: [30569024](https://pubmed.ncbi.nlm.nih.gov/30569024/)]
52. Lee Y, Chon D, Kim J, Ki S, Yun J. The predictive value of social frailty on adverse outcomes in older adults living in the community. *J Am Med Dir Assoc.* Oct 2020;21(10):1464-1469. [doi: [10.1016/j.jamda.2020.03.010](https://doi.org/10.1016/j.jamda.2020.03.010)] [Medline: [32362535](https://pubmed.ncbi.nlm.nih.gov/32362535/)]
53. Yu R, Wu WC, Leung J, Hu SC, Woo J. Frailty and its contributory factors in older adults: a comparison of two Asian regions (Hong Kong and Taiwan). *Int J Environ Res Public Health.* Sep 21, 2017;14(10):1096. [doi: [10.3390/ijerph14101096](https://doi.org/10.3390/ijerph14101096)] [Medline: [28934150](https://pubmed.ncbi.nlm.nih.gov/28934150/)]
54. Liu ZY, Wei YZ, Wei LQ, et al. Frailty transitions and types of death in Chinese older adults: a population-based cohort study. *Clin Interv Aging.* 2018;13:947-956. [doi: [10.2147/CIA.S157089](https://doi.org/10.2147/CIA.S157089)] [Medline: [29805253](https://pubmed.ncbi.nlm.nih.gov/29805253/)]

55. To TL, Doan TN, Ho WC, Liao WC. Prevalence of frailty among community-dwelling older adults in Asian countries: a systematic review and meta-analysis. *Healthcare (Basel)*. May 12, 2022;10(5):895. [doi: [10.3390/healthcare10050895](https://doi.org/10.3390/healthcare10050895)] [Medline: [35628034](https://pubmed.ncbi.nlm.nih.gov/35628034/)]
56. Huynh TQH, Pham TLA, Vo VT, Than HNT, Nguyen TV. Frailty and associated factors among the elderly in Vietnam: a cross-sectional study. *Geriatrics (Basel)*. Aug 20, 2022;7(4):85. [doi: [10.3390/geriatrics7040085](https://doi.org/10.3390/geriatrics7040085)] [Medline: [36005261](https://pubmed.ncbi.nlm.nih.gov/36005261/)]
57. Norazman CW, Adznam SN, Jamaluddin R. Physical frailty among urban-living community-dwelling older adults in Malaysia. *Int J Environ Res Public Health*. Sep 9, 2020;17(18):6549. [doi: [10.3390/ijerph17186549](https://doi.org/10.3390/ijerph17186549)] [Medline: [32916809](https://pubmed.ncbi.nlm.nih.gov/32916809/)]
58. Murayama H, Kobayashi E, Okamoto S, et al. National prevalence of frailty in the older Japanese population: findings from a nationally representative survey. *Arch Gerontol Geriatr*. 2020;91:104220. [doi: [10.1016/j.archger.2020.104220](https://doi.org/10.1016/j.archger.2020.104220)] [Medline: [32810734](https://pubmed.ncbi.nlm.nih.gov/32810734/)]
59. Morkphrom E, Srinonprasert V, Sura-amonrattana U, et al. Severity of frailty using modified Thai frailty index, social factors, and prediction of mortality among community-dwelling older adults in a middle-income country. *Front Med (Lausanne)*. 2022;9:1060990. [doi: [10.3389/fmed.2022.1060990](https://doi.org/10.3389/fmed.2022.1060990)] [Medline: [36569139](https://pubmed.ncbi.nlm.nih.gov/36569139/)]
60. Qin Y, Hao X, Lv M, Zhao X, Wu S, Li K. A global perspective on risk factors for frailty in community-dwelling older adults: a systematic review and meta-analysis. *Arch Gerontol Geriatr*. Feb 2023;105:104844. [doi: [10.1016/j.archger.2022.104844](https://doi.org/10.1016/j.archger.2022.104844)] [Medline: [36335672](https://pubmed.ncbi.nlm.nih.gov/36335672/)]
61. Zeng XZ, Meng LB, Li YY, et al. Prevalence and factors associated with frailty and pre-frailty in the older adults in China: a national cross-sectional study. *Front Public Health*. 2023;11:1110648. [doi: [10.3389/fpubh.2023.1110648](https://doi.org/10.3389/fpubh.2023.1110648)] [Medline: [37554734](https://pubmed.ncbi.nlm.nih.gov/37554734/)]
62. Willis BH. Spectrum bias--why clinicians need to be cautious when applying diagnostic test studies. *Fam Pract*. Oct 2008;25(5):390-396. [doi: [10.1093/fampra/cmn051](https://doi.org/10.1093/fampra/cmn051)] [Medline: [18765409](https://pubmed.ncbi.nlm.nih.gov/18765409/)]
63. Rockwood K, Song X, Mitnitski A. Changes in relative fitness and frailty across the adult lifespan: evidence from the Canadian National Population Health Survey. *CMAJ*. May 17, 2011;183(8):E487-94. [doi: [10.1503/cmaj.101271](https://doi.org/10.1503/cmaj.101271)] [Medline: [21540166](https://pubmed.ncbi.nlm.nih.gov/21540166/)]
64. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. *Lancet*. Mar 2, 2013;381(9868):752-762. [doi: [10.1016/S0140-6736\(12\)62167-9](https://doi.org/10.1016/S0140-6736(12)62167-9)] [Medline: [23395245](https://pubmed.ncbi.nlm.nih.gov/23395245/)]
65. Frailty Classification Using Machine Learning Model. URL: <https://frailty-commmed-cmu.streamlit.app> [Accessed 2025-03-23]

Abbreviations

- aOR:** adjusted odds ratio
- AUC:** area under the receiver operating characteristic curve
- CC:** calf circumference
- GBC:** gradient boosting classifier
- KNN:** k-nearest neighbors
- LR:** logistic regression
- MLP:** multilayer perceptron artificial neural network
- MPP:** mean predicted probability
- NPV:** negative predictive value
- PPV:** positive predictive value
- RF:** random forest
- SMOTE:** synthetic minority oversampling technique
- SVM:** linear support vector machine classifier
- TRIPOD+AI:** Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis plus Artificial Intelligence

Edited by Jinjiao Wang; peer-reviewed by Yeen Huang, Yiias Huusi; submitted 05.06.2024; final revised version received 26.11.2024; accepted 28.02.2025; published 02.04.2025

Please cite as:

Isaradech N, Sirikul W, Buawangpong N, Siviroj P, Kitro A

Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study

JMIR Aging 2025;8:e62942

URL: <https://aging.jmir.org/2025/1/e62942>

doi: [10.2196/62942](https://doi.org/10.2196/62942)

© Natthanaphop Isaradech, Wachiranun Sirikul, Nida Buawangpong, Penprapa Siviroj, Amornphat Kitro. Originally published in JMIR Aging (<https://aging.jmir.org>), 02.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on <https://aging.jmir.org>, as well as this copyright and license information must be included.