

RESEARCH

Open Access



The application of machine learning for identifying frailty in older patients during hospital admission

Yin-Yi Chou^{1,2}, Min-Shian Wang^{3,4}, Cheng-Fu Lin^{1,5,6,10}, Yu-Shan Lee^{1,7}, Pei-Hua Lee¹, Shih-Ming Huang⁸, Chieh-Liang Wu^{9,10} and Shih-Yi Lin^{1,10,11*}

Abstract

Background Early identification of frail patients and early interventional treatment can minimize the frailty-related medical burden. This study investigated the use of machine learning (ML) to detect frailty in hospitalized older adults with acute illnesses.

Methods We enrolled inpatients of the geriatric medicine ward at Taichung veterans general hospital between 2012 and 2022. We compared four ML models including logistic regression, random forest (RF), extreme gradient boosting, and support vector machine (SVM) for the prediction of frailty. The feature window as well as the prediction window was set as half a year before admission. Furthermore, Shapley additive explanation plots and partial dependence plots were used to identify Fried's frailty phenotype for interpreting the model across various levels including domain, feature, and individual aspects.

Results We enrolled 3367 patients. Of these, 2843 were frail. We used 21 features to train the prediction model. Of the 4 tested algorithms, SVM yielded the highest AUROC, precision and F1-score (78.05%, 94.53% and 82.10%). Of the 21 features, age, gender, multimorbidity frailty index, triage, hemoglobin, neutrophil ratio, estimated glomerular filtration rate, blood urea nitrogen, and potassium were identified as more impactful due to their absolute values.

Conclusions Our results demonstrated that some easily accessed parameters from the hospital clinical data system can be used to predict frailty in older hospitalized patients using supervised ML methods.

Keywords Frailty, Machine learning, Elderly, Hospital admission

*Correspondence:

Shih-Yi Lin

sylin@vghtc.gov.tw

¹Center for Geriatrics & Gerontology, Taichung Veterans General Hospital, Taichung, Taiwan

²Division of Allergy, Immunology and Rheumatology, Taichung Veterans General Hospital, Taichung, Taiwan

³Smart Healthcare Committee, Taichung Veterans General Hospital, Taichung, Taiwan

⁴Department of Applied Mathematics, National Chung Hsing University, Taichung, Taiwan

⁵Department of Family Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

⁶Division of Occupational Medicine, Department of Emergency, Taichung Veterans General Hospital, Taichung, Taiwan

⁷Division of Neurology, Taichung Veterans General Hospital, Taichung, Taiwan

⁸Department of Pharmacy, Taichung Veterans General Hospital, Taichung, Taiwan

⁹Department of Critical Care Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

¹⁰Department of Post-Baccalaureate Medicine, College of Medicine, National Chung Hsing University, Taichung, Taiwan

¹¹Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Frailty is a common clinical condition in the elderly. It is caused by the gradual loss of physiological reserve capacity of several interconnected organs and the gradual but cumulative capacity loss [1, 2]. It is characterized by the loss of health reserves, reduced physiological functions, and impaired homeostasis, which in turn, increases the individual's vulnerability. The presence of frailty is associated with an increased risk of negative health outcomes, such as functional decline, falls, fractures, disability, morbidity, dependence, recurrent hospitalizations, institutionalization, reduced quality of life, and death [3–6]. In high-income countries, frailty affects an estimated 10% of community-dwelling individuals over the age of 65 years (65+), and 25–50% of those aged 85+ [7, 8]. In acute care hospitals, the prevalence of frailty among hospitalized patients, over 65 years of age, is even higher, at nearly 50% [9]. Research has demonstrated that frailty can be delayed or ameliorated through various interventions, including nutritional supplementation and enhanced physical activity [10]. Primary health care plays a vital role in accurately identifying and managing frailty among patients to facilitate early initiation of interventional strategies [11]. Moreover, advanced frailty screening tests [12] can improve the quality of personal care. These may potentially reduce hospitalizations and healthcare costs [11].

In 2001, two key frailty assessment models were introduced. Fried's frailty phenotype (FFP) assesses physical frailty using five clinical indicators: weak grip, slow walking speed, unintended weight loss, low physical activity, and self-reported exhaustion [13]. The frailty index (FI), based on Rockwood and Mitnitski's age-related health deficits model, measures age-related health problems including diseases, symptoms, signs, disability, and abnormal lab results. Additionally, the FI-comprehensive geriatric assessment (CGA) was created by selecting items from the CGA questionnaire to establish a clinically useful FI based on cumulative health deficits. Various other frailty instruments, such as the FRAIL scale, clinical frailty scale (CFS), and the Groningen frailty indicator, provide alternative approaches to assess frailty [14].

Currently, there are no universally agreed-upon definitions or tools for diagnosing frailty. As a result, the estimated prevalence of frailty varies significantly depending on the diagnostic tool used [7]. Williamson et al. reported on using supervised machine learning (ML) for analyzing vulnerabilities in electronic medical record (EMR) profiles by using data from the Canadian province of Alberta [15]. This study used the CFS [16] to define frailty and obtained a sensitivity of 0.28 and a specificity of 0.94. Other studies have employed various frailty diagnostic tools to diagnose frailty, while ML methods have been used to classify the frailty extent. Additionally,

Hassler et al. used frailty phenotypes to identify frailty and supervised ML methods to infer frailty. Sensitivity estimates for this study ranged from 65.7 to 86.7%, and specificity ranged from 58.1 to 85.6% [17]. Anbagt Sher et al. used the electronic FI [18] to identify frailty, and used EMR data to make frailty inferences using supervised ML methods [19]. The best-performing model has a sensitivity of 97.8% and a specificity of 89.1%. Currently, there have been many studies on prediction models for frailty in primary care settings. In contrast, prediction models for frailty in hospitalized populations are limited.

This study aimed to develop a validated case definition of frailty for the elderly patient admitted to the hospital using ML. Our intention is to improve the quality of treatment and reduce future medical expenses through early diagnosis and timely intervention.

Materials and methods

Data source

The Institutional Review Board (IRB) of the Taichung Veterans General Hospital (TCVGH) (IRB no.: CE22109A) approved this study, ensuring that all procedures adhered strictly to the approved study protocol, following standard regulations and the principles of the Declaration of Helsinki. The study utilized de-identified Electronic Medical Record (EMR) data from TCVGH, allowing the IRB to waive the need for informed consent. The study focused on frailty diagnosis using the FFP criteria [13], which include factors such as fatigue, unintentional weight loss, weakness, sluggishness, and low physical activity.

In this study, the diagnosis of frailty was determined using the FFP criteria, which include fatigue, unintentional weight loss, weakness, sluggishness, and low physical activity. Participants who met any two criteria were defined as “pre-frail”. Patients who met ≥ 3 criteria were classified as “frail” [13].

Constructing a dataset for supervised learning

This retrospective study included patients aged 65 years and older who were admitted to the geriatric ward of the medical center between 2012 and 2022. From an initial cohort of 7,421 patients, we excluded those with incomplete information (42 patients), repeated assessments before discharge (3,353 patients), hospital stays of less than three days (65 patients), and cases where the assessment was conducted more than four days after admission (400 patients). The final dataset comprised 3,367 patients, of which 2,843 were classified as frail and 524 as non-frail. This dataset was notably imbalanced, with a significantly higher number of frail cases.

Fig. 1 presents the study flowchart and sample selection.

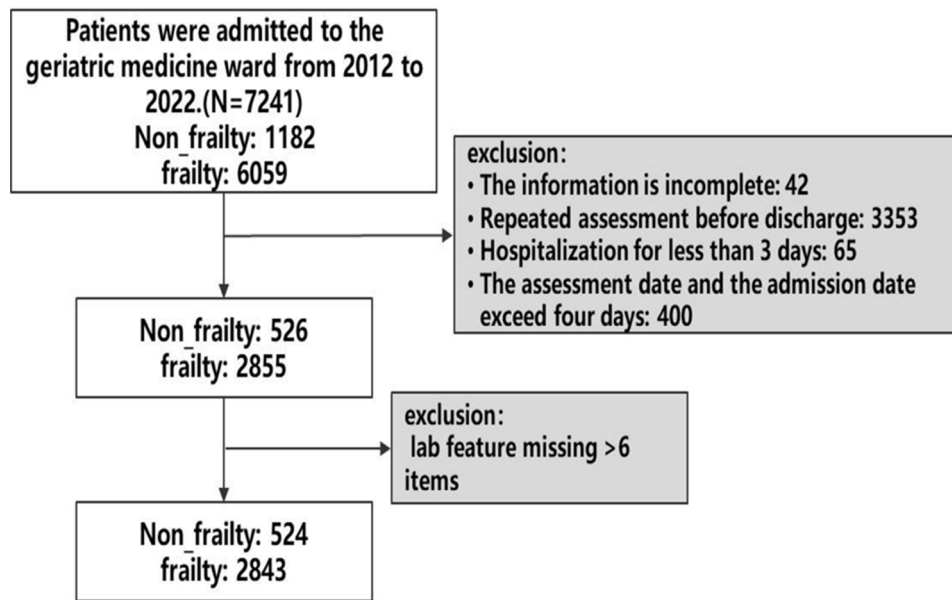


Fig. 1 Study flowchart

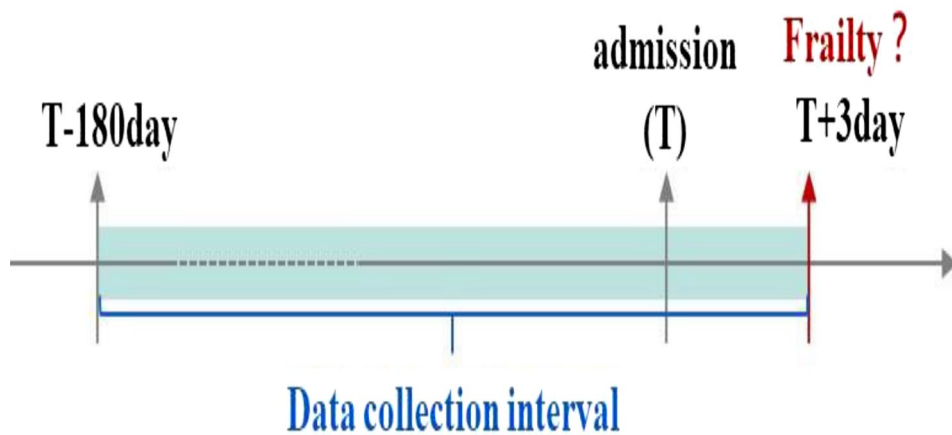


Fig. 2 Frailty prediction task

Data description

Data collected included information from six months before to three days after the hospitalization date. This encompassed blood test reports, diagnosis codes, the number of hospitalizations, emergency visits, multimorbidity frailty index (mFI) [20], injury classification, and treatment orders. The study stipulated that hospitalization must exceed three days to avoid assessment duplications, requiring a minimum of four days between the assessment and previous admission. The primary aim of this study was to develop a binary classification prediction model to identify frailty in patients admitted to the geriatric ward.

Fig. 2 illustrates the frailty prediction task.

Feature selection and data preprocessing

Feature selection for this study was guided by a review of existing frailty prediction models and adjusted based on clinical relevance and data availability [1, 19, 21, 22]. An initial set of 71 features was selected, including diagnosis codes, drug codes, and doctor’s order codes. To address missing data, features with a missing rate greater than 20% were removed, following the guidelines from Rubin and Allison [23, 24]. Various imputation methods were explored, including mean imputation, k-nearest neighbor (KNN) imputation, and multiple imputation. The impact of these imputation methods on the performance of different classifiers, such as extreme gradient boosting (Xgboost), random forest (RF), logistic regression (LR), and support vector machine (SVM), was analyzed. The results indicated that mean imputation effectively

addressed missing data without significantly affecting model performance metrics like sensitivity, specificity, accuracy, AUROC, precision, and F1-score. Consequently, 21 clinical variables were retained for analysis, including demographic data, blood test results, diagnosis codes, and emergency data [25, 26]. For detailed results of each imputation method, see Supplementary Table 2.

Table 1 presents the demographic data including age and gender. Blood test reports included white blood cell count, neutrophil ratio, hemoglobin (Hgb), platelet count, blood urea nitrogen (BUN), creatinine, estimated glomerular filtration rate (eGFR), glucose, sodium, potassium (K), and c-reactive protein (CRP) levels. Diagnosis codes are represented by mFI. Emergency data comprises the number of visits, injury classification and triage, the number of hospitalizations, and treatment orders, such as nasogastric (NG) tube, restriction, Foley catheter, and nothing by mouth.

Modeling

Once data preprocessing was completed, the dataset was divided into a training set (80%) and a test set (20%) for model development. The model was trained using Python version 3.6. Four machine learning (ML) algorithms were employed: Xgboost, RF, LR, and SVM. Hyperparameter optimization was conducted using a grid search approach with 5-fold cross-validation to identify optimal parameter configurations, specifically targeting improvements in the AUROC. Given the significant class imbalance between frail and non-frail groups, class weights were applied to improve model performance. The models were implemented using widely recognized Python libraries: Xgboost for Xgboost, imbalanced-learn for RF, and scikit-learn for both LR and SVM. The Xgboost and RF models represented integrated learning and tree-based methods, respectively [27, 28]. The LR model represented a linear model [29], while SVM represented a non-probabilistic binary linear classifier [30].

Table 1 The demographic data of the study population

	All N=3367	Non-frailty N=524	Frailty N=2843	p-value
Demographic data				
Age (years)	81.81 ± 8.75	76.20 ± 7.92	82.84 ± 8.50	< 0.001
Sex (male)	1876 (55.72%)	316 (60.31%)	1560 (54.87%)	0.0242
Order				
NG	484 (14.37%)	7 (1.34%)	477 (16.78%)	< 0.001
Foley	1239 (36.8%)	102 (19.47%)	1137 (39.99%)	< 0.001
Rest	175 (5.2%)	4 (0.76%)	171 (6.01%)	< 0.001
NPO	111 (3.3%)	20 (3.82%)	91 (3.2%)	0.5535
Laboratory				
K (mEq/L)	3.95 ± 0.56	3.95 ± 0.50	3.95 ± 0.57	0.7949
NA (mEq/L)	137.14 ± 5.37	137.40 ± 4.10	137.09 ± 5.57	0.1351
Creat (mg/dL)	1.69 ± 1.84	1.40 ± 1.49	1.75 ± 1.89	< 0.001
WBC (/μL)	9830.23 ± 4437.93	9797.00 ± 4443.46	9836.35 ± 4437.66	0.8522
NEUT (%)	77.13 ± 10.96	76.64 ± 11.16	77.22 ± 10.93	0.2813
CRP (mg/dL)	6.21 ± 6.85	6.41 ± 7.02	6.18 ± 6.83	0.5429
BUN (mg/dL)	28.98 ± 22.97	22.52 ± 15.69	30.10 ± 23.84	< 0.001
Hgb (g/dL)	11.22 ± 2.07	12.13 ± 2.08	11.05 ± 2.03	< 0.001
eGFR (mL/min/1.73m ²)	61.99 ± 33.93	67.41 ± 28.45	60.99 ± 34.76	< 0.001
PLT (*10 ³ /μL)	207.28 ± 85.96	204.15 ± 76.44	207.85 ± 87.60	0.3205
Glu (mg/dL)	154.04 ± 77.66	148.38 ± 70.03	155.02 ± 78.87	0.0649
Others				
H_times	0.35 ± 0.74	0.18 ± 0.49	0.38 ± 0.77	< 0.001
ER_times	1.33 ± 1.50	1.02 ± 1.01	1.38 ± 1.57	< 0.001
mFI	3.24 ± 3.33	2.03 ± 2.60	3.46 ± 3.40	< 0.001
Triage_0	784 (23.28%)	147 (28.05%)	637 (22.41%)	< 0.001
Triage_1	1 (0.03%)	0 (0.0%)	1 (0.04%)	
Triage_2	34 (1.01%)	7 (1.34%)	27 (0.95%)	
Triage_3	1922 (57.08%)	318 (60.69%)	1604 (56.42%)	
Triage_4	545 (16.19%)	51 (9.73%)	494 (17.38%)	
Triage_5	81 (2.41%)	1 (0.19%)	80 (2.81%)	

Table 2 The predictive performance of frailty using different algorithms

	Classifier	Sensitivity	specificity	Accuracy	AUROC	Precision	F1-score
Cross-validation	Xgboost	71.77% (3.56%)	67.07% (3.90%)	71.03% (2.75%)	77.50% (2.04%)	92.15% (0.78%)	80.64% (2.25%)
	RF	68.91% (3.35%)	71.57% (4.16%)	69.33% (2.63%)	77.44% (2.07%)	92.89% (0.89%)	79.07% (2.23%)
	LR	68.25% (3.88%)	72.06% (4.65%)	68.84% (3.26%)	78.02% (2.61%)	92.93% (1.12%)	78.64% (2.71%)
	SVM	69.18% (3.66%)	71.34% (4.27%)	69.51% (3.14%)	78.05% (2.11%)	92.84% (1.06%)	79.23% (2.56%)
Testing	Xgboost	71.68%	75.49%	72.26%	80.94%	94.25%	81.43%
	RF	70.28%	76.47%	71.22%	81.72%	94.37%	80.56%
	LR	71.50%	76.47%	72.26%	80.93%	94.46%	81.39%
	SVM	72.55%	76.47%	73.15%	81.06%	94.53%	82.10%

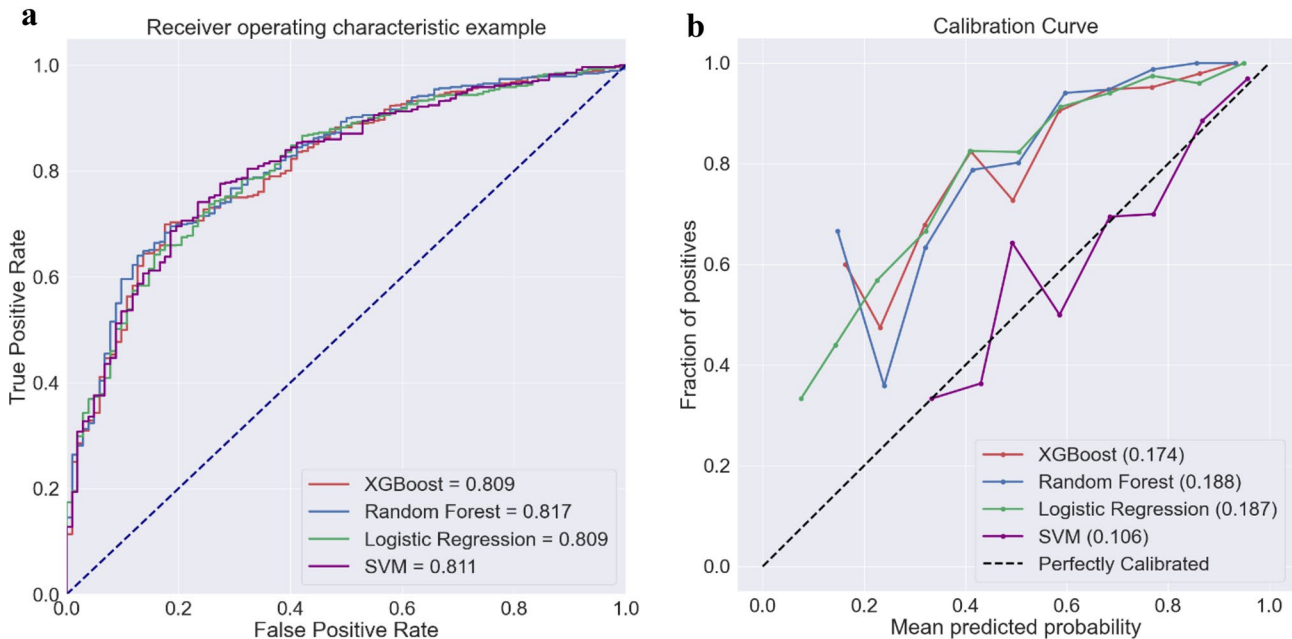


Fig. 3 Comparative performance evaluation of machine learning models. (a) AUROC curves for model discrimination analysis. (b) Calibration curves for model probability calibration

Statistical analysis

Traditional statistical methods were employed to analyze the data, with independent t-tests used to compare continuous variables between positive and negative outcomes. The study evaluated the performance of the ML models using sensitivity, specificity, accuracy, AUROC, precision, and F1-score as metrics. AUROC values were categorized as follows: 0.7 to 0.8 as acceptable, 0.8 to 0.9 as good, and greater than 0.9 as excellent. To further interpret model predictions, Shapley Additive Explanations (SHAP) plots were utilized to determine the importance of each feature, while partial dependence plots (PDPs) were used to explore trends related to individual features. The study’s goal was to develop a robust frailty prediction model that could effectively identify frail patients in various clinical settings.

Results

Evaluation of the different models

The present study compared four different algorithms and evaluated the performance of each using our dataset. Table 2; Fig. 3 present the predictive performance of frailty. We used the predictive results from the ML algorithm to represent the risk probability of frailty. The default value was set at 0.5, meaning if the model’s threshold exceeds 0.5, the patient was defined as ‘frail’. When comparing sensitivity and specificity, it was evident that the LR classifier had the highest specificity for both the validation and testing datasets (72.06% and 76.47%, respectively). The SVM showed the highest AUROC on the validation dataset, with a value of 78.05%. Additionally, SVM showed highest precision and F1-score on the testing dataset (94.53% and 82.10%).

The predictive performance was further assessed based on the AUROC and calibration from the testing data. The AUROC curve provides a clear visualization of a classifier’s diagnostic capability as the discrimination threshold

is adjusted. In our study, all models demonstrated robust predictive performance, with AUROC values around 0.8, effectively distinguishing between frail and non-frail individuals. Notably, the RF and SVM models slightly outperformed the others, with the SVM achieving the highest AUROC at 0.811. This indicates that the SVM model is particularly effective in detecting frailty within this dataset. These findings highlight the SVM's potential in targeted health assessments. Among the evaluated models, the SVM demonstrates the best calibration, making it exceptionally reliable for applications that require precise predictions and risk assessments. Its close adherence to the ideal calibration line ensures that its probability estimates accurately reflect true event likelihoods. In contrast, the Xgboost, RF, and LR models tend to predict higher probabilities than actual outcomes across a wide range of probabilities, consistently positioning above the 45-degree line. This tendency indicates an overestimation of event likelihoods, which could result in a higher rate of false positives in practical settings. Despite adjustments made using class weight to address data imbalance, these models still exhibit a tendency to overestimate, suggesting that other factors might be contributing to this issue.

Clinical features: importance and visualization

Figure 4 presents the feature importance as the strongest predictor to effect frailty. The most important feature was age, followed by Hgb, Foley catheter, NG tube, and so on. By analyzing the SHAP plot [31], we can gain insights into how each feature influences the model's predictions.

A positive SHAP value indicates that as the feature value increases, the model's predicted outcome also increases. Conversely, a negative SHAP value means that higher feature values lead to a lower predicted outcome. The SHAP summary plot (Fig. 5) illustrates the strength and direction of each feature's impact in this study. Older age, the presence of NG or Foley catheters, male, higher mFI, higher triage, elevated BUN levels, and lower K levels, were more strongly associated with frailty. On the contrary, being male, higher Hgb levels, higher eGFR, higher K levels, and a higher neutrophil ratio were less strongly associated with frailty.

After identifying the positive and negative influences of each feature on the model using the SHAP plot, we further analyzed the relationship between a feature's SHAP value and its actual value using a dependence plot. By setting the SHAP value of 0 as a threshold, we can identify the turning point of the feature value. This turning point serves as a clinical reference, indicating that if the feature value exceeds or falls below this point, it may positively or negatively affect the model's predictions. Upon reviewing the dependence plots for all features, we found that the turning points of seven features have a significant impact on the model. We will now explain how each of these seven key features influences the probability of failure in the ML model. The seven key features are: including age, gender, BUN, eGFR, Hgb, mFI, and neutrophil ratio. As shown in Fig. 6, ages approximately 78 years and older, BUN level higher than nearly 47, Hgb lower than 12, eGFR lower than 50, neutrophil ratio lower than 70,

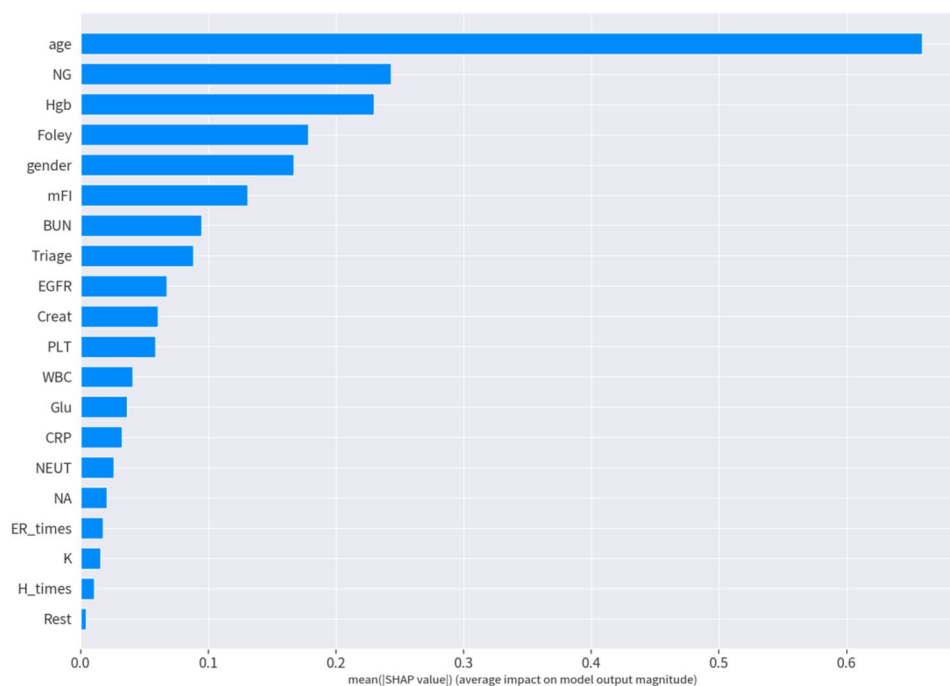


Fig. 4 Model interpretation and visualization according to feature importance

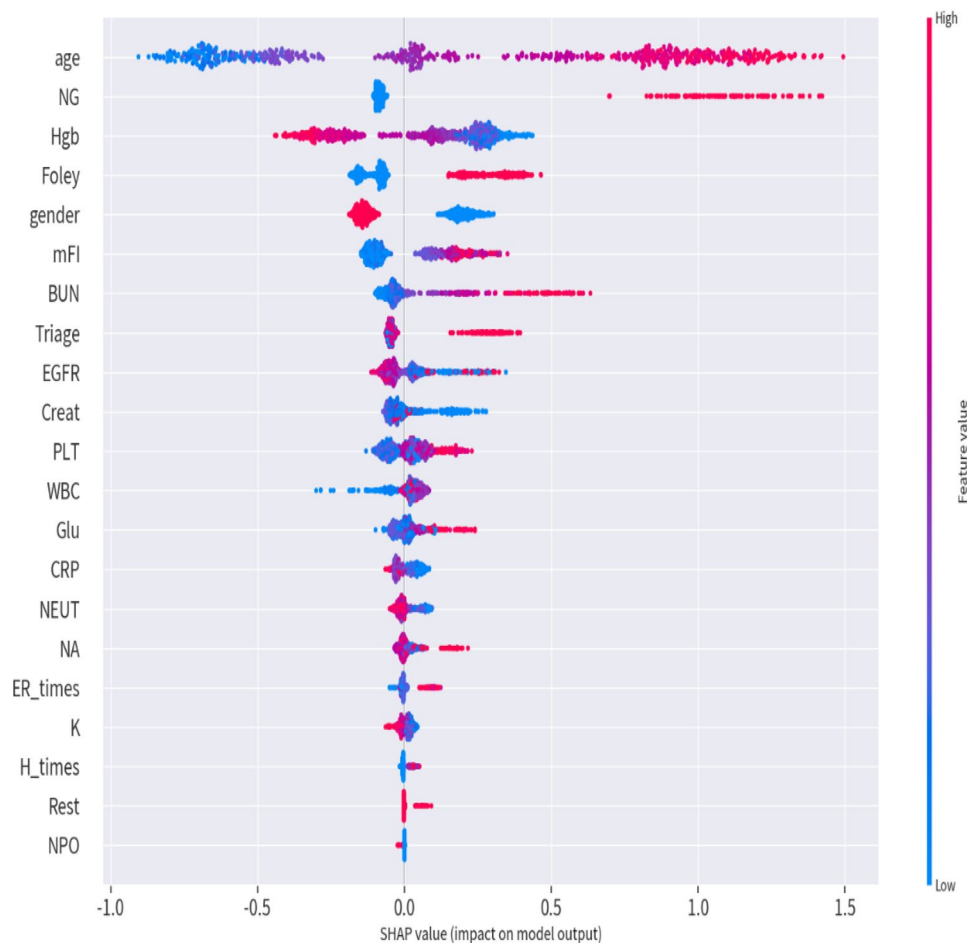


Fig. 5 The relationship between different features and frailty

and mFI higher than 5, were consistently associated with an increased risk of frailty.

Discussion

This study investigated 4 algorithms and 21 features, and utilized a six-month window from the time of hospital admission for both features and predictions to build a frailty detection model. Among the five algorithms, SVM showed the best predictive results: accuracy (73.15%), Precision, F1-score (82.10%), and AUROC (0.811%). To gain a deeper understanding, the SHAP and PDP were utilized to interpret the model across various dimensions. Of the 21 features, age was the most predictive of frailty, followed by Hgb, Foley catheter, NG tube, and so on.

The incidence of frailty varies with different diagnostic tools, patient ethnicity, and population. It is estimated that 10% of community residents over 65 years are frail, and the frailty rate is as high as 25–50% among community residents over 85 years of age [13, 32]. In one of our previous retrospective studies, the proportion of frailty among elderly hospitalized patients was as high as 44.7% [33].

Due to the high prevalence and negative health effects of frailty in the elderly, it is very important to clinically diagnose frailty as early as possible and provide timely frailty interventions. In our geriatric medicine ward, we screen patients to identify those who are frail and provide them with CGA. These evaluations are conducted by geriatricians who assess the causes of frailty and determine appropriate interventions, including dental care, nutrition, and rehabilitation. We also guide both patients and caregivers on post-discharge care and health education to ensure tailored support for improved health and well-being.

Due to the detailed nature of these assessments, which required 40 to 60 min to be complete specialized staff, such resources were predominantly available to geriatric wards. However, our newly developed frailty module can detect frailty risks early in the admission process, allowing for timely interventions that can significantly reduce frailty levels in elderly hospitalized patients.

In a recent review, Liotta et al. emphasized the importance of identifying elements that contribute to the effectiveness of health and social care interventions and the

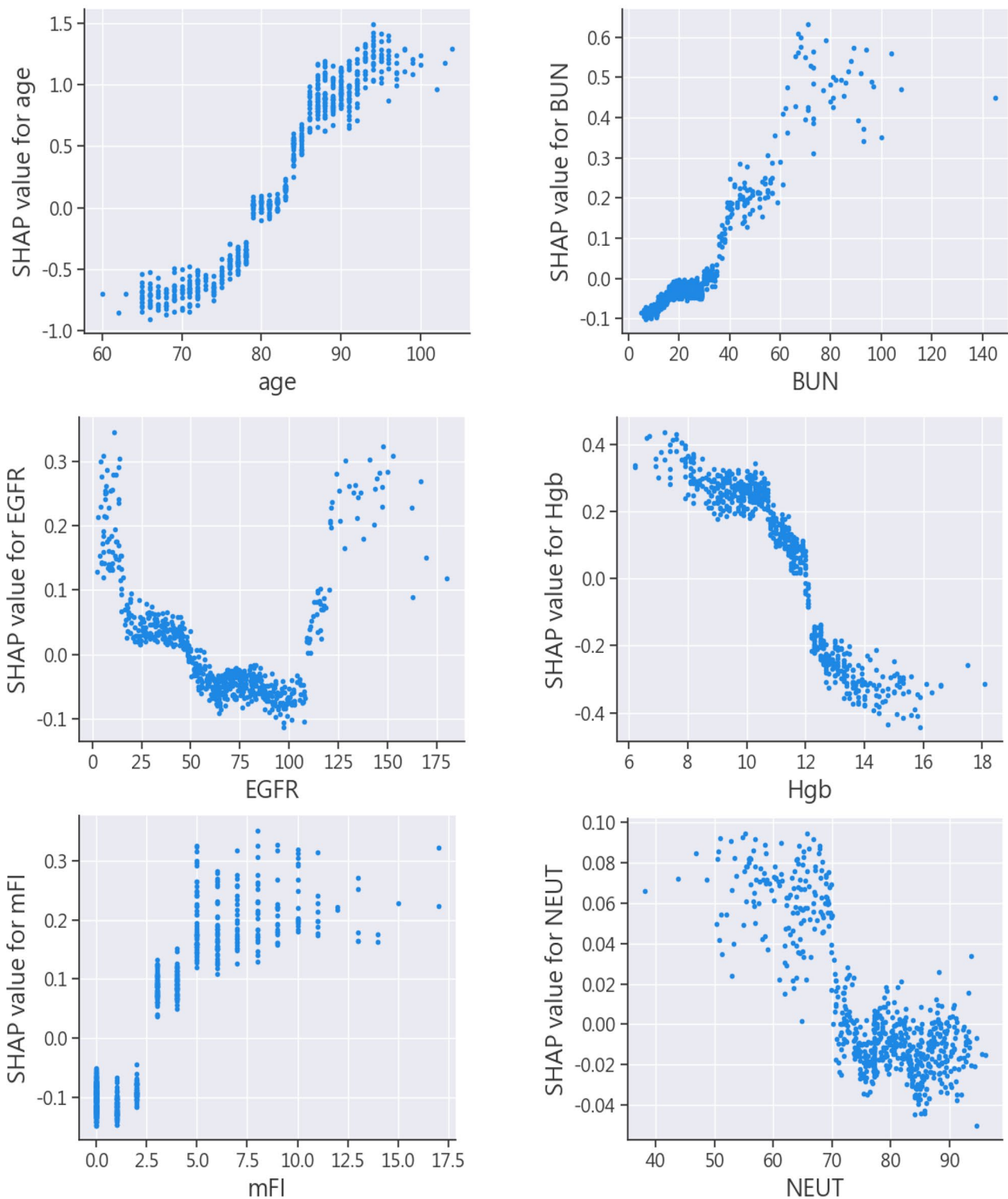


Fig. 6 The dependence plots

long-term viability of healthcare systems, with a focus on public health considerations [34]. A prospective study by Wang et al. showed that providing rehabilitation and nutritional counseling to patients who are hospitalized

and frail can improve activities of daily living, and CFS during hospitalization, shorten hospitalization time, reduce medical costs, and lower the 30-day and 90-day readmission rates [35]. Once the module detects frailty,

the attending physician can consult with a geriatric specialist. The geriatric medicine team will then perform a comprehensive assessment of the patient. Using the assessment results, along with the patient's medical and medication history, the geriatric doctor will identify the primary cause of frailty. Based on the patient's rehabilitation potential and nutritional status, they will implement a plan that includes exercise interventions and nutritional adjustments. After discharge, follow-up visits will be scheduled to monitor improvements in frailty risk and evaluate the effectiveness of the intervention. The plan will be continuously adjusted as needed based on the patient's progress.

Previous research on ML frailty detection models has mostly targeted community residents and institutional residents. Ambagtsheer et al. designed a frailty prediction system for 10 residential elderly care facilities in Queensland, Australia, based on frailty identification across the performance of 18 specific scenarios using three ML algorithms (SVM, decision tree, and KNN). The results showed that the overall accuracy of ML based on SVM was greater than 75% [19]. Eskandari et al. used a long-short term memory model to detect frailty based on the institutional resident data of the Arizona frailty and fall cohort. Its sensitivity and specificity were both higher than 80% [22]. Sajeev et al. used three algorithms (discriminant analysis, SVM, and RF) to estimate the accuracy of frailty in community-dwelling older adults. When phenotypes are used to define frailty, RF demonstrated the best AUROC (0.722 in selected features, 0.701 in all 63 features) [14]. Aponte-Hao et al. used the Canadian primary care sentinel surveillance network data to predict frailty. The prevalence of frailty in their sample was 18.4%. The Xgboost module demonstrated the best frailty prediction ability (sensitivity 78.14% and specificity 74.41%) [21]. A unique aspect of our study is its focus on detecting frailty in hospitalized patients. This is especially important because frail older adults admitted to hospitals face a higher risk of poor outcomes. Additionally, we make use of "early" and "easily accessible" clinical data during hospitalization to predict frailty, which enhances its practical application in real-world clinical settings. Our research significantly contributes to the field by demonstrating how machine learning can be effectively utilized in hospital environments. This opens up new possibilities for integrating machine learning into hospital screening and intervention programs, addressing a gap in current research on frailty detection in hospitals.

Research conducted by Pablo Ferri and colleagues has demonstrated that handling datasets with significant amounts of missing data can be notably improved by using translation and coding interpolation methods that account for missing information. This approach, in conjunction with tree ensemble classifiers such as RF

and gradient boosting, effectively enhances performance [36]. In our study, we adhered to both established practices and specific statistical considerations appropriate for our research context. To preserve the integrity of our statistical analysis and model performance in the face of significant missing data, three geriatric medicine doctors evaluated the importance of each feature before removing any with more than 20% missing data. While we considered various imputation methods, we recognized that the potential bias and variance from significant imputation could outweigh the benefits. Our approach aligns with established statistical guidelines, such as those by Little and Rubin [23] and Allison [24], which suggest removing features with substantial missingness under certain conditions. This cautious strategy ensures more reliable and interpretable results. Our study on frailty detection in hospitalized older adults effectively employs a variety of ML models, including LR, RF, and SVM. This approach enables a detailed analysis of a broad range of clinical data, aiming to significantly enhance early intervention and treatment strategies for this vulnerable population. Our study provides a robust framework, offering a strong comparison point to other research such as that conducted by Sajeev et al. and Blanes-Selva et al., who also utilize ML but with different focuses and contexts. Sajeev et al.'s research targets community-dwelling adults, using techniques like LR and linear discriminant analysis to identify pre-frailty. Their study emphasizes the importance of lifestyle and environmental factors in the progression of frailty, offering crucial insights for preventive measures in non-hospital settings [14]. Conversely, Blanes-Selva et al. employ advanced models such as gradient boosting machines and deep neural networks [37]. Their research extends beyond frailty detection to include mortality predictions, integrating the assessment of palliative care needs into their models, which greatly benefits overall patient care planning. While all three studies aim to use ML to improve healthcare outcomes for older adults, our research is particularly distinguished by its focus on the acute care setting. This specificity facilitates critical, tailored interventions at the point of hospital admission, thus enhancing the management and treatment of frailty in hospitalized older adults. Each study contributes uniquely to the field, collectively advancing the understanding and management of frailty through innovative ML applications across various healthcare environments.

We utilized the `class_weight` parameter in our models to address class imbalance. By adjusting this parameter, we were able to assign a higher weight to the minority class, effectively increasing the cost of misclassifying the underrepresented class. This approach helps to balance the influence of each class on the model's learning process. And, recognizing that accuracy alone can be

misleading in the presence of class imbalance, we incorporated additional performance metrics such as precision, recall, F1-score, and the AUROC. These metrics are more informative and provide a better assessment of model performance in imbalanced settings.

To the best of our knowledge, the models developed by our institute are currently the only models available for predicting frailty in hospitalized elderly patients. Compared with the feature selections in previous studies, we chose a relatively streamlined number of features, and most of the features we used were data that could be easily obtained clinically, including objective emergency injury classification and medical treatment orders. As the included features are conveniently acquired, our prediction module has high reproducibility. Importantly, the prediction results presented by our prediction models performed very well.

Limitations

This study has certain research flaws and limitations. First, our sample population was drawn exclusively from a single hospital and a single ward. In the future, we aim to use other attributes of inpatients to refine the features and models. Second, the study sample provided unbalanced data. The application of post hoc calibration techniques could be considered in the future to further refine the probability estimates. Additionally, collecting more data from non-frail patients may help determine whether if the model or its components require modification. Finally, the detection module developed in this study is specifically designed for patients who have undergone blood tests at a single hospital within the past six months. Patients who have not received treatment at this hospital, or who lack recent blood test reports, will be excluded from the study and are ineligible for assessment by this module. While patients entering through the emergency department typically have complete blood reports available during their admission wait, those admitted through outpatient clinics may not have recent blood tests, making them likely to be excluded from the detection process.

Conclusion

Early diagnosis of frailty can help reduce negative medical impact and reduce overall medical costs. This study shows that using clinically accessible data and ML can predict frailty in patients admitted to the hospital.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12911-024-02684-z>.

Supplementary Material 1

Acknowledgements

Thanks to the smart healthcare committee, Taichung veterans general hospital, Taichung city for helping to train the module, and also to the staff of the Taichung veterans general hospital for helping to export electronic medical data.

Author contributions

C (Yin Yi-Chou), W (Chieh-Liang Wu), L (Shih-Yi Lin) and L (Yu-Shan Lee): Conceive the research architecture. L (Pei-Hua Lee), H (Shih-Ming Hung): Data archiving. W (Min-Shian Wang): Complete module training and chart production. C (Yin Yi-Chou), L (Cheng-Fu Lin): Manuscript Writing. All authors reviewed the Manuscript.

Funding

No finding accepted.

Data availability

This study does not directly provide raw data. However, if there is a reasonable need for us to provide information, please apply to the corresponding author. If the corresponding author agrees, we will provide you with the information for identification in accordance with the regulations of the Human Experimentation Committee.

Declarations

Ethics approval and consent to participate

The institutional review board (IRB) of the Taichung veterans general hospital (TCVGH) (IRB no.: CE22109A) approved this study. All the procedures were conducted in strict accordance with the approved study protocol, adhering to standard regulations and the principles outlined in the Declaration of Helsinki. The study used de-identified electronic medical record data from the TCVGH. Consequently, informed consent was waived by the IRB of TCVGH.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 14 December 2023 / Accepted: 16 September 2024

Published online: 27 September 2024

References

1. Le Pogam MA, Seematter-Bagnoud L, Niemi T, Assouline D, Gross N, Trachsel B, Rousson V, Peytremann-Bridevaux I, Burnand B, Santos-Eggimann B. Development and validation of a knowledge-based score to predict Fried's frailty phenotype across multiple settings using one-year hospital discharge data: the electronic frailty score. *EClinicalMedicine*. 2022;44:101260.
2. Singh M, Stewart R, White H. Importance of frailty in patients with cardiovascular disease. *Eur Heart J*. 2014;35(26):1726–31.
3. Kouroubali A, Kondylakis H, Logothetidis F, Katehakis DG. Developing an AI-Enabled Integrated Care platform for Frailty. *Healthc (Basel)* 2022, 10(3).
4. Soong JTY, Ng SH, Tan KXQ, Kaubryte J, Hopper A. Variation in coded frailty syndromes in secondary care administrative data: an international retrospective exploratory study. *BMJ Open*. 2022;12(1):e052735.
5. Woo J, Leung J, Morley JE. Comparison of frailty indicators based on clinical phenotype and the multiple deficit approach in predicting mortality and physical limitation. *J Am Geriatr Soc*. 2012;60(8):1478–86.
6. Vermeiren S, Vella-Azzopardi R, Beckwee D, Habbig AK, Scafoglieri A, Jansen B, Bautmans I. Gerontopole Brussels Study g: Frailty and the prediction of negative Health outcomes: a Meta-analysis. *J Am Med Dir Assoc*. 2016;17(12):1163.e1161-1163.e1117.
7. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. *J Am Geriatr Soc*. 2012;60(8):1487–92.
8. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. *Lancet*. 2013;381(9868):752–62.

9. Theou O, Squires E, Mallery K, Lee JS, Fay S, Goldstein J, Armstrong JJ, Rockwood K. What do we know about frailty in the acute care setting? A scoping review. *BMC Geriatr*. 2018;18(1):139.
10. Travers J, Romero-Ortuno R, Bailey J, Cooney MT. Delaying and reversing frailty: a systematic review of primary care interventions. *Br J Gen Pract*. 2019;69(678):e61–9.
11. Lacas A, Rockwood K. Frailty in primary care: a review of its conceptualization and implications for practice. *BMC Med*. 2012;10:4.
12. Braithwaite RS, Fiellin D, Justice AC. The payoff time: a flexible framework to help clinicians decide when patients with comorbid disease are not likely to benefit from practice guidelines. *Med Care*. 2009;47(6):610–7.
13. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol Biol Sci Med Sci*. 2001;56(3):M146–156.
14. Sajeev S, Champion S, Maeder A, Gordon S. Machine learning models for identifying pre-frailty in community dwelling older adults. *BMC Geriatr*. 2022;22(1):794.
15. Tyler Williamson P, Aponte-Hao S, Mele B, Lethebe BC, Leduc C, Thandi M, Katz A, Wong ST. Developing and validating a primary Care EMR-based Frailty definition using machine learning. *Int J Popul Data Sci*. 2020;5(1):1344.
16. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, Mitnitski A. A global clinical measure of fitness and frailty in elderly people. *CMAJ*. 2005;173(5):489–95.
17. Hassler AP, Menasalvas E, García-García FJ, Rodríguez-Manas L, Holzinger A. Importance of medical data preprocessing in predictive modeling and risk factor discovery for the frailty syndrome. *BMC Med Inf Decis Mak*. 2019;19(1):33.
18. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, Mohammed MA, Parry J, Marshall T. Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age Ageing*. 2016;45(3):353–60.
19. Ambagtsheer RC, Shafiabady N, Dent E, Seiboth C, Beilby J. The application of artificial intelligence (AI) techniques to identify frailty within a residential aged care administrative data set. *Int J Med Inf*. 2020;136:104094.
20. Wen YC, Chen LK, Hsiao FY. Predicting mortality and hospitalization of older adults by the multimorbidity frailty index. *PLoS ONE*. 2017;12(11):e0187825.
21. Aponte-Hao S, Wong ST, Thandi M, Ronksley P, McBrien K, Lee J, Grandy M, Mangin D, Katz A, Singer A, et al. Machine learning for identification of frailty in Canadian primary care practices. *Int J Popul Data Sci*. 2021;6(1):1650.
22. Eskandari M, Parvaneh S, Ehsani H, Fain M, Toosizadeh N. Frailty Identification using Heart Rate dynamics: a Deep Learning Approach. *IEEE J Biomed Health Inf*. 2022;26(7):3409–17.
23. Little RJ, Rubin DB. Statistical analysis with missing data. Volume 793. Wiley; 2019.
24. Allison PD. Handling missing data by maximum likelihood. In: *SAS global forum*: 2012. San Diego, CA, USA: 1–21.
25. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12:2825–30.
26. Buitinck L, Louppe G, Blondel M, Pedregosa F, Mueller A, Grisel O, Niculae V, Prettenhofer P, Gramfort A, Grobler J. API design for machine learning software: experiences from the scikit-learn project. *arXiv Preprint arXiv:13090238* 2013.
27. Salminen M, Viljanen A, Eloranta S, Viikari P, Wuorela M, Vahlberg T, Isoaho R, Kivelä S-L, Korhonen P, Irjala K, et al. Frailty and mortality: an 18-year follow-up study among Finnish community-dwelling older people. *Aging Clin Exp Res*. 2020;32(10):2013–9.
28. Sutton JL, Gould RL, Daley S, Coulson MC, Ward EV, Butler AM, Nunn SP, Howard RJ. Psychometric properties of multicomponent tools designed to assess frailty in older adults: a systematic review. *BMC Geriatr*. 2016;16(1):55.
29. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inf*. 2002;35(5–6):352–9.
30. Teixeira-Santos L, Bobrowicz-Campos E, Parola V, Coelho A, Gil I, Almeida ML, Apóstolo JL. What is the relationship between Lifestyle and Frailty Status? Data from the Portuguese Multicentre descriptive study. *Nurs Rep*. 2022;12(1):39–49.
31. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst* 2017, 30.
32. Solfrizzi V, Scafato E, Lozupone M, Seripa D, Schilardi A, Custodero C, Sardone R, Galluzzo L, Gandin C, Baldereschi M, et al. Biopsychosocial frailty and the risk of incident dementia: the Italian longitudinal study on aging. *Alzheimers Dement*. 2019;15(8):1019–28.
33. Chou YY, Lee YS, Lin CS, Chen JP, Kuo FH, Lin CF, Chen YM, Lin SY. Gender-specific linkages: frailty, polypharmacy, anti-cholinergic burden, and 5-year mortality risk—a real-world analysis. *PeerJ*. 2023;11:e16262.
34. Liotta G, Ussai S, Illario M, O’Caoimh R, Cano A, Holland C, Roller-Winsberger R, Capanna A, Grecuccio C, Ferraro M, et al. Frailty as the Future Core Business of Public Health: report of the activities of the A3 Action Group of the European Innovation Partnership on active and healthy ageing (EIP on AHA). *Int J Environ Res Public Health*. 2018;15(12):2843.
35. Wang YC, Liang CK, Chou MH, Chiu CF, Lin HC, Hsu YH, Liao MC, Yin CH, Chou MY, Lin YT. The effectiveness of Frailty intervention for older patients with Frailty during hospitalization. *J Nutr Health Aging*. 2023;27(6):413–20.
36. Ferri P, Romero-García N, Badenes R, Lora-Pablos D, Morales TG, de la Cámara AG, García-Gómez JM, Sáez C. Extremely missing numerical data in Electronic Health records for machine learning can be managed through simple imputation methods considering informative missingness: a comparative of solutions in a COVID-19 mortality case study. *Comput Methods Programs Biomed*. 2023;242:107803.
37. Blanes-Selva V, Doñate-Martínez A, Linklater G, García-Gómez JM. Complementary frailty and mortality prediction models on older patients as a tool for assessing palliative care needs. *Health Inf J*. 2022;28(2):14604582221092592.

Publisher’s note

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