



Artificial Intelligence (AI)-Driven Frailty Prediction Using Electronic Health Records in Hospitalized Patients With Cardiovascular Disease

Masashi Yamashita, PT, PhD; Kentaro Kamiya, PT, PhD; Kazuki Hotta, PT, PhD;
Anna Kubota, PhD; Kenji Sato, MD; Emi Maekawa, MD, PhD;
Hiroaki Miyata, PhD; Junya Ako, MD, PhD, FJCS

Background: This study aimed to create a deep learning model for predicting phenotypic physical frailty from electronic medical record information in patients with cardiovascular disease.

Methods and Results: This single-center retrospective study enrolled patients who could be assessed for physical frailty according to cardiovascular health study criteria (25.5% [691/2,705] of the patients were frail). Patients were randomly separated for training (Train set: 80%) and validation (Test set: 20%) of the deep learning model. Multiple models were created using LightGBM, random forest, and logistic regression for deep learning, and their predictive abilities were compared. The LightGBM model had the highest accuracy (in a Test set: F1 score 0.561; accuracy 0.726; area under the curve of the receiver operating characteristics [AUC] 0.804). These results using only commonly used blood biochemistry test indices (in a Test set: F1 score 0.551; accuracy 0.721; AUC 0.793) were similar. The created models were consistently and strongly associated with physical functions at hospital discharge, all-cause death, and heart failure-related readmission.

Conclusions: Deep learning models derived from large sample sizes of phenotypic physical frailty have shown good accuracy and consistent associations with prognosis and physical functions.

Key Words: Cardiovascular disease; Deep learning; Physical frailty; Prognosis

As the population ages, the number of frail older adults, one of the most significant challenges of aging, is increasing. Aging is a multifactorial and multimodal process characterized by an increased incidence of health-related problems and non-communicable diseases, with Japan leading the world in life expectancy and aging rate.¹ Frailty is one of the most common non-communicable diseases,² increasing at a rate of 4.3% per year³ and affecting 8.7% of community-dwelling older adults in Japan.⁴ Similar problems occur in patients with cardiovascular disease (CVD), with a high rate of concomitant frailty (15–63%), requiring adequate clinical management.⁵ Therefore, it is important to detect not only community-dwelling frail older adults but also hospitalized frail patients at an early stage and to provide appropriate care.^{6,7}

The concept of frailty is commonly known as a phenotypic

model;⁸ it is defined as a reversible condition between being independent and dependent physiologically, and is a term that refers to subjects who benefit from geriatric intervention to return to a healthy state. Assessment of frailty is useful in determining the feasibility of invasive treatments, and identifying high-risk individuals for prognostic purposes; early detection of frailty is critical for implementing interventions aimed at reversing frailty. Indeed, we have previously reported the need to assess physical frailty and to intervene appropriately in patients with CVD.^{9–15} Although the importance of frailty assessment is already well recognized in clinical practice, it is not easy to assess muscle strength, walking speed and physical activity required for standard frailty assessment for all older patients in routine clinical practice. As a solution to this problem, it may be useful to determine frailty risk using indicators that can be

Received September 23, 2024; accepted September 23, 2024; J-STAGE Advance Publication released online October 29, 2024 Time for primary review: 1 day

Division of Research, ARCE Inc., Sagamihara (M.Y.); Department of Rehabilitation Sciences, Graduate School of Medical Sciences, Kitasato University, Sagamihara (M.Y., K.K., K.H.); Department of Rehabilitation, Kitasato University School of Allied Health Sciences, Sagamihara (K.K., K.H.); Department of Health Policy and Management, School of Medicine, Keio University, Tokyo (A.K., H.M.); Director, Sado General Hospital, Niigata (K.S.); and Department of Cardiovascular Medicine, Kitasato University School of Medicine, Sagamihara (E.M., J.A.), Japan

K.K., J.A. are members of *Circulation Reports*' Editorial Team.

Mailing address: Kentaro Kamiya, PT, PhD, Department of Rehabilitation, School of Allied Health Sciences, Kitasato University, 1-15-1 Kitasato, Minami-ku, Sagamihara, Kanagawa 252-0373, Japan. email: k-kamiya@kitasato-u.ac.jp

All rights are reserved to the Japanese Circulation Society. For permissions, please email: cr@j-circ.or.jp

ISSN-2434-0790



Table 1. Association Between Both Phenotypic Frail Criteria and Physical Function Separated by Sex

	Women				Men			
	n	Frail n=215	Non-frail n=584	P value	n	Frail n=415	Non-frail n=1,220	P value
CHS criteria								
SPPB (point)	779	9.10 (2.89)	10.88 (1.87)	<0.001	1,588	9.80 (2.68)	11.40 (1.27)	<0.001
Grip strength (kg)	790	14.5 (4.4)	18.5 (5.2)	<0.001	1,613	23 (6)	31 (8)	<0.001
AC (cm)	791	23.9 (4.2)	25.2 (3.7)	<0.001	1,612	24.8 (3.4)	27.0 (3.4)	<0.001
QIS/BM (%)	772	30 (10)	38 (12)	<0.001	1,551	38 (16)	50 (16)	<0.001
CC (cm)	791	30.5 (4.3)	32.2 (3.7)	<0.001	1,611	32.2 (4.0)	35.0 (4.1)	<0.001
MGS (m/s)	736	1.02 (0.37)	1.32 (0.33)	<0.001	1,504	1.23 (0.40)	1.58 (0.38)	<0.001
CGS (m/s)	775	0.81 (0.27)	1.03 (0.25)	<0.001	1,583	0.92 (0.30)	1.17 (0.26)	<0.001
6MWD (m)	777	257 (118)	346 (110)	<0.001	1,571	307 (129)	423 (114)	<0.001
J-CHS criteria		n=309	n=590			n=382	n=1,423	
SPPB (point)	875	8.75 (2.88)	11.16 (1.55)	<0.001	1,755	9.06 (2.77)	11.50 (1.11)	<0.001
Grip strength (kg)	896	13.8 (3.9)	19.1 (5.1)	<0.001	1,788	22 (6)	31 (7)	<0.001
AC (cm)	893	24.0 (4.1)	25.1 (3.7)	<0.001	1,787	24.7 (3.4)	27.0 (3.5)	<0.001
QIS/BM (%)	875	29 (10)	39 (12)	<0.001	1,716	36 (13)	50 (17)	<0.001
CC (cm)	893	30.6 (4.2)	32.3 (3.7)	<0.001	1,787	31.9 (3.9)	34.9 (4.1)	<0.001
MGS (m/s)	827	0.94 (0.32)	1.39 (0.30)	<0.001	1,671	1.08 (0.40)	1.60 (0.36)	<0.001
CGS (m/s)	876	0.76 (0.26)	1.06 (0.23)	<0.001	1,758	0.84 (0.28)	1.18 (0.25)	<0.001
6MWD (m)	877	235 (102)	365 (104)	<0.001	1,738	271 (123)	427 (108)	<0.001

Data are shown to mean (standard deviation) or n (%). 6MWD, 6-min walking distance; AC, arm circumference; CC, calf circumference; CGS, comfortable gait speed; CHS, Cardiovascular Health Study; J-CHS, revised Japanese version of the Cardiovascular Health Study; MGS, maximum gait speed; QIS/BM, quadriceps isometric strength/body mass; SPPB, short physical performance battery.

assessed in daily clinical practice, such as electronic health record (EHR) data.

In recent years, there have been some reports on using artificial intelligence (AI) to create predictive models of frailty.¹⁶ However, studies with adequate sample sizes often use data such as mortality and prognosis as training data, and there are few models that use standard phenotypic frailty criteria, including assessments of muscle strength and walking speed, as training data. In addition, there are yet to be any reports of prediction models using only commonly used interview data and blood biochemistry indices that can be generalized, such as collecting general health checkups. Therefore, in the present study, we created a model for predictive phenotypic physical frailty and verified its usefulness based on information obtained from EHRs for patients with CVD, and checked whether the model would be equally valuable with only commonly used blood biochemistry test indices.

Methods

Study Design and Population

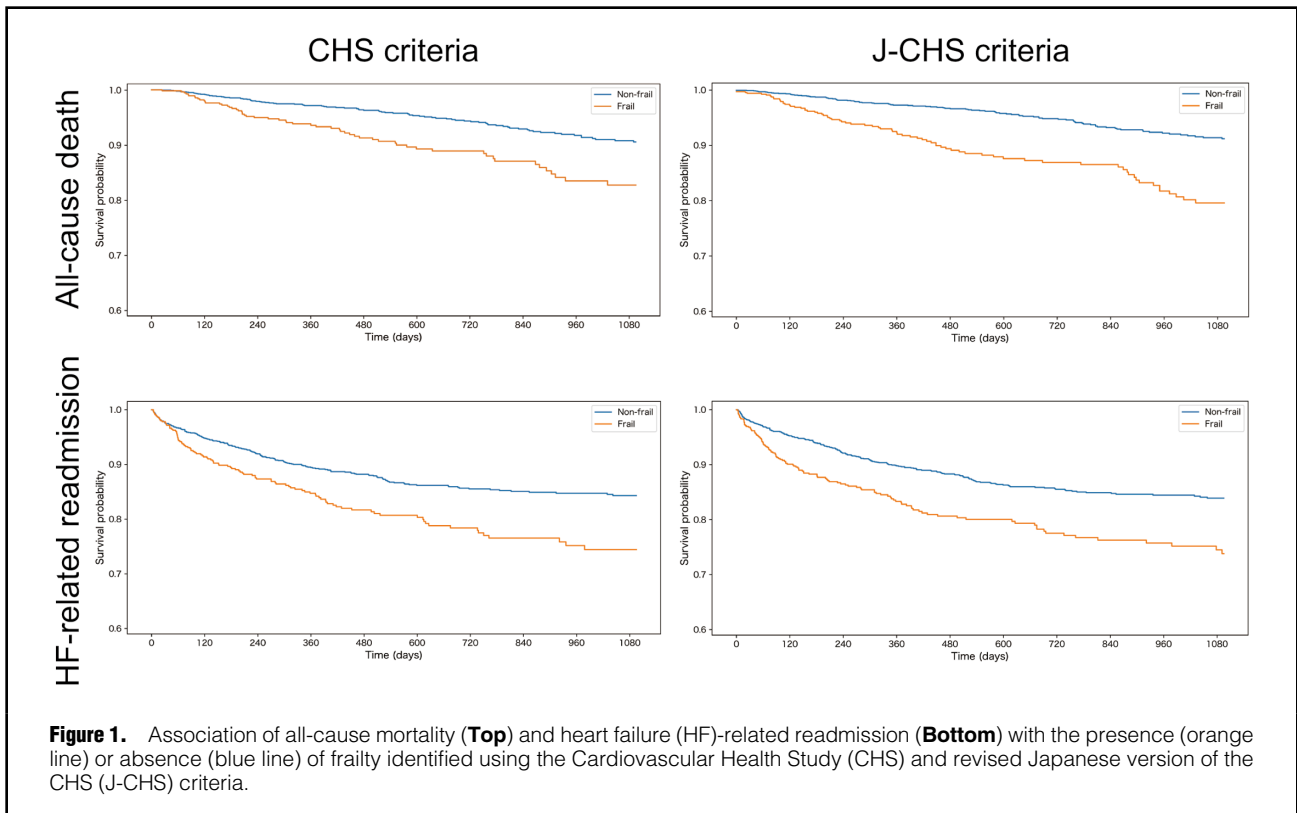
The study was performed retrospectively at a Kitasato University Hospital. Data were obtained from EHRs and physical function assessments, and data analysis, including deep learning, was outsourced to a vendor (Open Health Initiative, Minato-ku, Japan; <https://www.openhealth-i.com/>). Patients who underwent inpatient rehabilitation at the Cardiovascular Center of Kitasato University Hospital between January 2008 and December 2020 were included. Patients who could not be assessed for physical frailty for any reason were excluded from the study. The study was performed according to the Declaration of Helsinki, and the study protocol, including matters of sharing data and materials with vendors, was approved by the Ethics Committee of Kitasato University Hospital (B21-170). The

study's information was posted on plainly worded opt-out materials, and a guide for withdrawing from the study was presented.

Clinical Data Collection and Assessment of Frailty, Physical Function, and Prognosis

We extracted data necessary for defining frailty and to create a predictive model using items that were considered likely to be related to frailty, as follows: patient information (e.g., age, sex, body weight, blood pressure, and heart rate), blood test values, history of disease, and medications. Data used to predict physical frailty were collected from EHRs and selected based on the time of hospital discharge. The value of the corresponding period was used for items missing at discharge or measured only once during the hospitalization period. General-purpose indicators were targeted at commonly used daily blood biochemistry test indices that could be collected throughout various settings (**Supplementary Table 1**). When collecting blood test results and medication data, we exported the data from the EHR system. Disease information and frailty status were manually entered by the researchers while reviewing the EHR, with a careful double-checking process to ensure accuracy.

In this study, 2 criteria based on phenotype from Fried et al.⁸ were used to measure physical frailty: Cardiovascular Health Study (CHS) criteria,¹⁷ and the revised Japanese version of the CHS (J-CHS) criteria.¹⁸ Both sets of criteria consisted of muscle weakness, slow gait speed, physical inactivity, loss of body weight, and fatigue. Details of each criterion are presented in **Supplementary Table 2**. A physical therapist assessed the measured frailty at the end of the in-hospital rehabilitation before the patient was discharged with stable disease. For both criteria, applying 3 or more items was considered physically frail and used as the correct label for subsequent analyses.



A short physical performance battery, grip strength, mid-upper arm circumference, quadriceps isometric strength, calf circumference, maximum gait speed, comfortable gait speed, and 6-min walking distance were assessed as physical function parameters at hospital discharge.^{19–21} As a primary outcome, all-cause death during 5 years was followed up using the date of discharge as the baseline. In cases where death could not be confirmed from EHRs, follow up was censored at the last date of survival confirmation. As a secondary outcome, heart failure (HF)-related rehospitalization events during 5 years were investigated. The definition of censoring was the same as for all-cause deaths up to the last survival date that could be verified in the EHR.

Preparation for Deep Learning

Initially, descriptive statistics were used to confirm the characteristics of the 2 frailty criteria. For confirmation, physical functions were measured at the same time as the identification of physical frailty for men and women. Comparisons between frail and non-frail were examined using the Wilcoxon rank-sum test. To test the association between the 2 frailty criteria sets and prognosis, we classified the patients into 2 groups according to the presence or absence of physical frailty in both sets. We drew Kaplan-Meier survival curves for all-cause death and HF-related rehospitalization. Last, we examined which features were associated with the presence or absence of both frailty criteria sets using the Wilcoxon rank-sum test, the χ -square test, or Fisher's exact probability test.

The data were then randomly split into a training set (Train set: 80%) and a test set (Test set: 20%). After splitting, the data were compared for feature bias between the Train

and Test sets. The Wilcoxon rank-sum test, χ -square test, or Fisher's exact probability test were used to verify the comparison. Kaplan-Meier survival curves were then drawn to check whether the Train and Test sets had a similar association with prognosis.

Deep Learning

The LightGBM,²² logistic regression, and random forest²³ methods were used for deep learning, and their accuracies were compared. To create the learning model using the Train set, we tried several combinations of features: a statistically significant difference item (All EHR model), and commonly used indices (Simple EHR model). To evaluate the accuracy of the predicting model made using the Train set, we used the values of precision (how well the model fits the prediction that it is a positive example), recall (how well the model picks up a positive example), F1 score (which shows the trade-off relationship between precision and recall), and accuracy; the area under the curve of the receiver operating characteristics (AUC) values were also calculated. Using the 2×2 division confusion matrix (true positive [TP], true negative [TN], false positive [FP], and false negative [FN]), the formulas for each value are as follows:

$$\text{Precision} = (\text{TP}) / (\text{TP} + \text{FP})$$

$$\text{Recall} = (\text{TP}) / (\text{TP} + \text{FN})$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FN} + \text{FP} + \text{TN})$$

$$\text{F1 score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

Using the Test set of cases, we then performed the prediction of frailty decisions based on the models with the best accuracy and investigated the degree of agreement

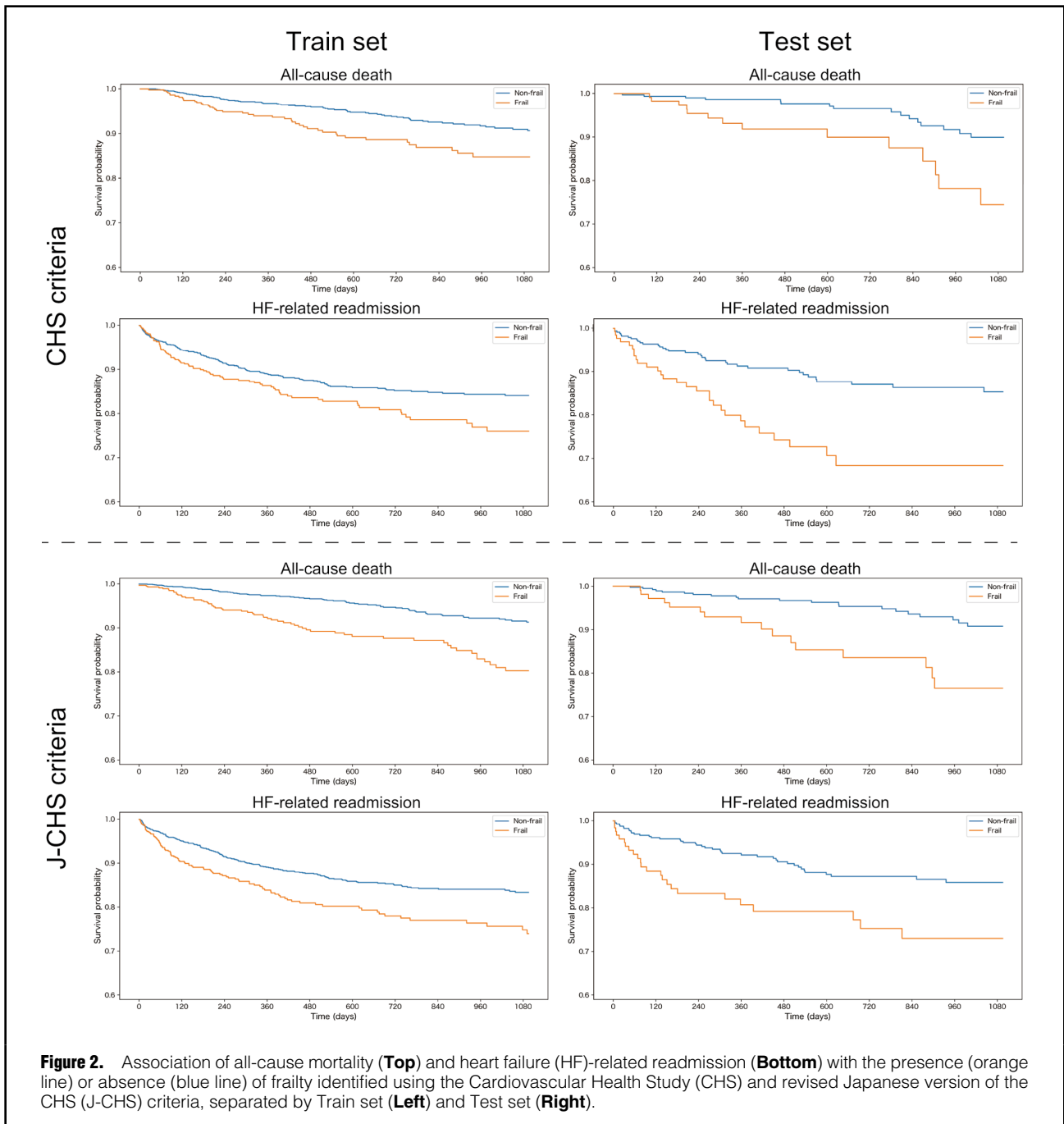


Figure 2. Association of all-cause mortality (**Top**) and heart failure (HF)-related readmission (**Bottom**) with the presence (orange line) or absence (blue line) of frailty identified using the Cardiovascular Health Study (CHS) and revised Japanese version of the CHS (J-CHS) criteria, separated by Train set (**Left**) and Test set (**Right**).

with the actual results of frailty decisions. The reliability evaluation was based on the F1 score, recall, precision, accuracy, AUC value, and confusion matrix. In addition, the patients in the Test set were divided into 2 groups according to frailty and non-frailty predicted by the model, and their relationship to each physical function and prognosis after discharge was investigated. Last, the clinical validity of the generated frailty prediction model was visually confirmed using feature importance, partial dependence plots with the main features, and the shapely additive explanation (SHAP) value.

Python 3.9 software was used for both analyses, and the statistical significance level was $<5\%$ with a 2-tailed test.

Results

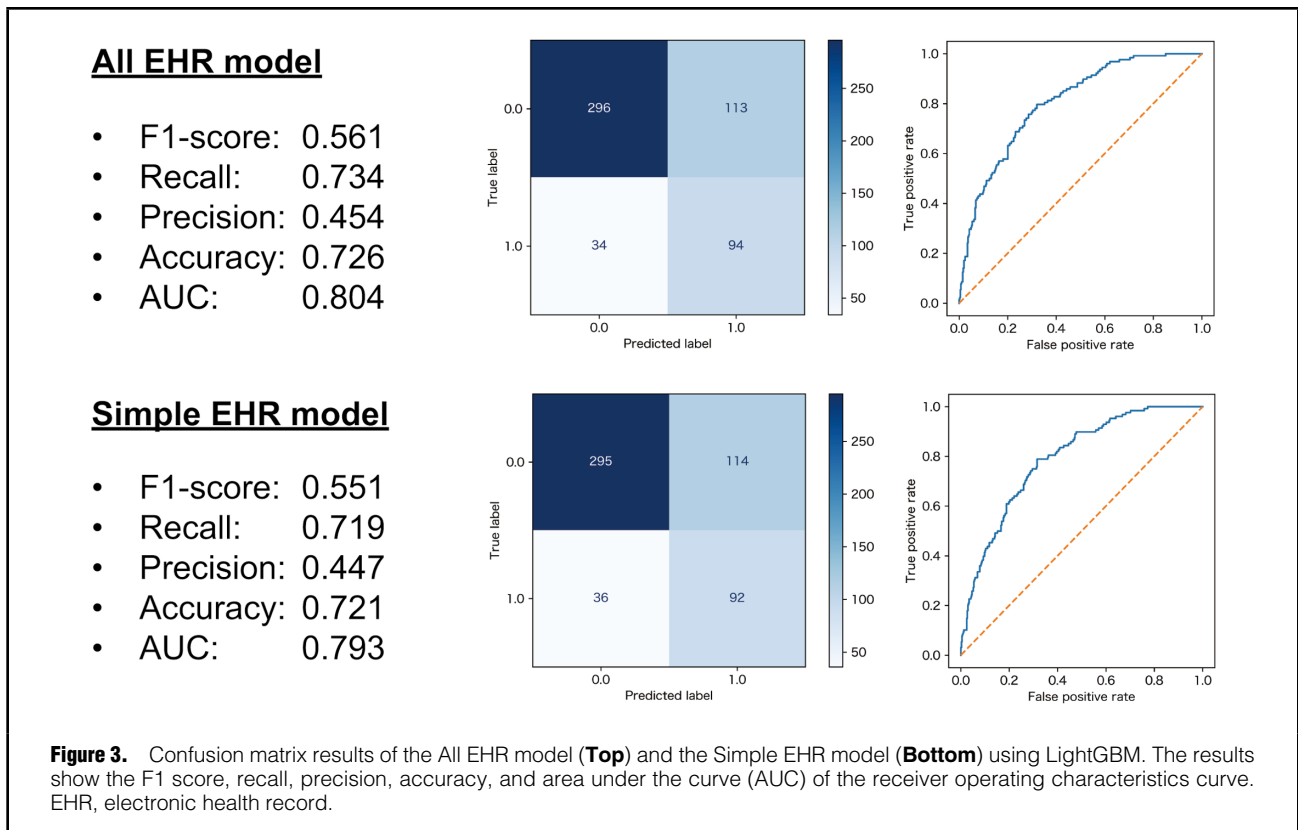
Results of the Descriptive Statistics of Patients in the Present Study

During the inclusion period, 8,507 patients underwent in-hospital rehabilitation. Of these, 2,434 subjects were able to be identified using the CHS criteria, and 2,705 subjects were able to be identified using the J-CHS criteria. A total of 25.9% (630/2,434) was determined to be frail using the CHS criteria and 25.5% (691/2,705) using the J-CHS criteria.

Table 1 shows the physical functions of the 2 groups, determined by the presence or absence of phenotypic physical frailty as assessed by each criterion. For both criteria sets,

	CHS criteria					J-CHS criteria				
	F1 score	Recall	Precision	Accuracy	AUC	F1 score	Recall	Precision	Accuracy	AUC
LightGBM										
All EHR model	0.507	0.639	0.420	0.685	0.714	0.546	0.687	0.454	0.705	0.743
Simple EHR model	0.506	0.646	0.416	0.680	0.690	0.535	0.664	0.449	0.701	0.730
Logistic										
All EHR model	0.471	0.623	0.379	0.647	0.688	0.522	0.683	0.423	0.675	0.735
Simple EHR model	0.490	0.670	0.387	0.647	0.694	0.517	0.706	0.409	0.658	0.722
Random forest										
All EHR model	0.478	0.643	0.381	0.646	0.707	0.519	0.692	0.417	0.668	0.731
Simple EHR model	0.473	0.644	0.374	0.637	0.684	0.504	0.663	0.408	0.663	0.721

AUC, area under the curve; EHR, electronic health record. Other abbreviations as in Table 1.



all physical function indices were significantly lower in the frail group, confirming the robustness of the frail phenotype criterion. Similarly, all survival analyses showed that the frail group had a poorer prognosis (all-cause death and HF-related readmission), confirming the robustness of the frail phenotype criterion (**Figure 1**).

The relationship between the presence or absence of phenotypic frailty and each characteristic measured is shown in **Supplementary Table 3**. Both criteria showed significant differences in age, body mass index, and other factors associated with frailty. In contrast, we also observed that there were factors, such as sex and history of dementia, which differed in the tendency of group differences in the characteristics between the 2 frailty criteria sets. The items that showed significant differences were used in the model

design for each frailty criterion.

Training in Deep Learning

We compared the characteristics (**Supplementary Table 4**) and prognoses (**Figure 2**) of the patients classified into the Train set (80%) and Test set (20%) for each phenotypic frailty criterion. We confirmed that there was no difference between the 2 groups in all characteristics in the CHS criteria, and no difference between the 2 groups in the J-CHS criteria except for the history of HF and smoking history. In addition, there were significant difference in the incidence of all-cause deaths or HF-related readmissions in all categories ($P < 0.001$).

Table 2 shows the results of the model creation using the Train set. First, the All EHR model created using

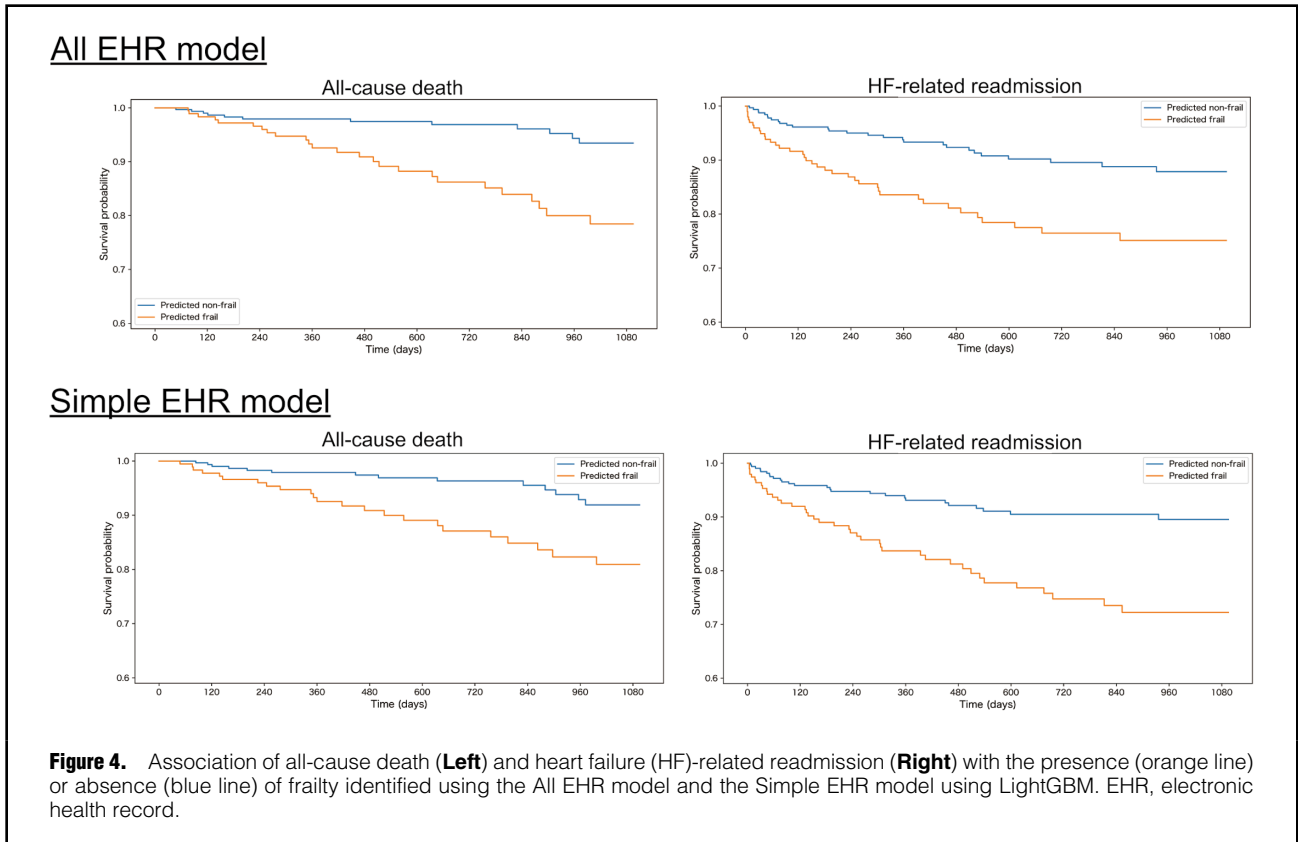


Figure 4. Association of all-cause death (Left) and heart failure (HF)-related readmission (Right) with the presence (orange line) or absence (blue line) of frailty identified using the All EHR model and the Simple EHR model using LightGBM. EHR, electronic health record.

	Women				Men			
	n	Frail n=89	Non-frail n=82	P value	n	Frail n=118	Non-frail n=248	P value
All EHR model								
SPPB (point)	170	9.48 (2.70)	11.32 (1.36)	<0.001	349	9.88 (2.50)	11.51 (1.16)	<0.001
Grip strength (kg)	171	15.3 (4.4)	19.8 (5.0)	<0.001	360	23 (6)	31 (7)	<0.001
AC (cm)	170	23.8 (3.6)	25.8 (3.1)	<0.001	358	24.4 (3.1)	27.4 (3.4)	<0.001
QIS/BM (%)	167	31 (11)	40 (14)	<0.001	347	38 (13)	50 (16)	<0.001
CC (cm)	170	31.0 (3.7)	33.2 (2.8)	<0.001	358	32.0 (3.3)	35.0 (4.1)	<0.001
MGS (m/s)	157	1.10 (0.33)	1.45 (0.30)	<0.001	327	1.19 (0.39)	1.62 (0.37)	<0.001
CGS (m/s)	169	0.87 (0.27)	1.14 (0.22)	<0.001	351	0.92 (0.28)	1.20 (0.25)	<0.001
6MWD (m)	170	281 (112)	388 (95)	<0.001	345	298 (127)	443 (99)	<0.001
Simple EHR model		n=91	n=80			n=115	n=251	
SPPB (point)	170	9.69 (2.66)	11.12 (1.64)	<0.001	349	9.89 (2.58)	11.47 (1.15)	<0.001
Grip strength (kg)	171	15.2 (4.2)	20.1 (5.0)	<0.001	360	23 (6)	31 (7)	<0.001
AC (cm)	170	23.6 (3.6)	26.1 (3.0)	<0.001	358	24.5 (3.2)	27.4 (3.4)	<0.001
QIS/BM (%)	167	31 (12)	40 (14)	<0.001	347	38 (13)	50 (16)	<0.001
CC (cm)	170	30.9 (3.6)	33.3 (2.9)	<0.001	358	32.1 (3.5)	34.9 (4.1)	<0.001
MGS (m/s)	157	1.12 (0.34)	1.43 (0.31)	<0.001	327	1.20 (0.39)	1.61 (0.37)	<0.001
CGS (m/s)	169	0.89 (0.28)	1.13 (0.23)	<0.001	351	0.93 (0.29)	1.19 (0.25)	<0.001
6MWD (m)	170	286 (113)	384 (99)	<0.001	345	302 (128)	440 (104)	<0.001

Data are shown to mean (standard deviation). Abbreviations as in Tables 1,2.

LightGBM showed the high prediction accuracies for both the CHS (F1 score 0.507; AUC 0.714) and J-CHS criteria (F1 score 0.546; AUC 0.743). Second, the Simple EHR model created with commonly used indices were compared. The F1 scores and AUC were 0.506 and 0.690, respectively,

for CHS criteria, which was lower than those for the All EHR model. In contrast, the frailty model defined by the J-CHS criteria showed good accuracy, with an F1 score of 0.535 and an AUC of 0.730, almost as accurate as the All EHR model, even when commonly used indices were used.

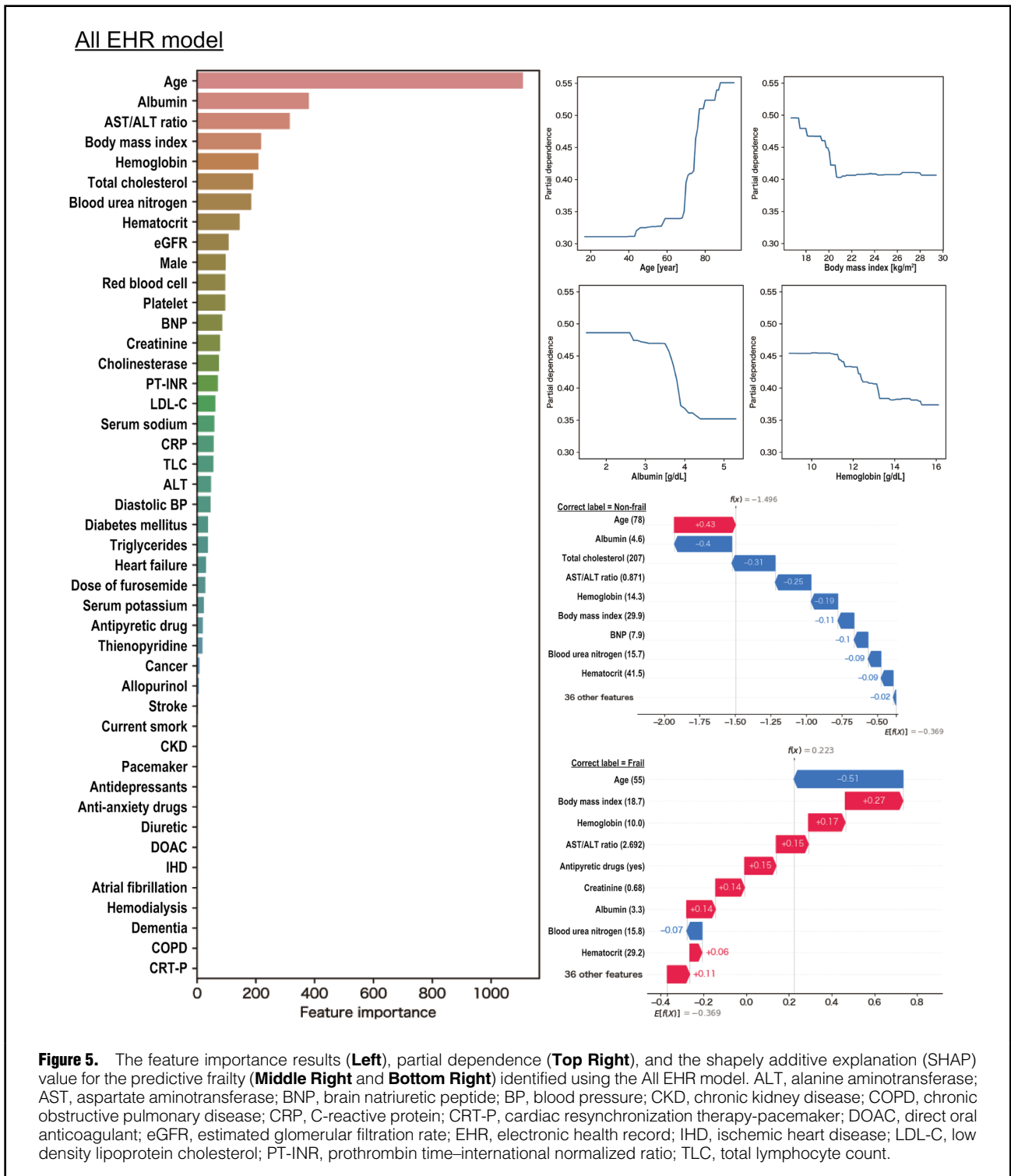


Figure 5. The feature importance results (Left), partial dependence (Top Right), and the shapely additive explanation (SHAP) value for the predictive frailty (Middle Right and Bottom Right) identified using the All EHR model. ALT, alanine aminotransferase; AST, aspartate aminotransferase; BNP, brain natriuretic peptide; BP, blood pressure; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; CRP, C-reactive protein; CRT-P, cardiac resynchronization therapy-pacemaker; DOAC, direct oral anticoagulant; eGFR, estimated glomerular filtration rate; EHR, electronic health record; IHD, ischemic heart disease; LDL-C, low density lipoprotein cholesterol; PT-INR, prothrombin time–international normalized ratio; TLC, total lymphocyte count.

Third, we compared the prediction accuracy of the models created by logistic regression and random forest with that of the LightGBM model. In both the All EHR model and the Simple EHR model, the prediction accuracy of the model obtained with LightGBM was the same or better than the others. Based on the above analysis, it was determined that the model created with LightGBM showed the best accuracy, and it was decided to use LightGBM for

subsequent validation. Also, the J-CHS criterion was used for validation by the Test set because of its higher prediction accuracy for the J-CHS criterion compared with the CHS criterion in all analyses using the Train set.

Confirming Validity Using the Test Set

Figure 3 shows the confusion matrix results using the LightGBM model (All EHR model and Simple EHR

model). In both models, F1 scores ranged from 0.551 to 0.561, and AUC ranged from 0.793 to 0.804, showing the same high accuracy as when models were created using the Train set.

Figure 4 shows the relationship between the with or without frailty groups predicted by the created models and prognosis. As with the Train set, the incidence of all-cause death and HF-related readmission was worse in the frailty-labeled group in all models. Similarly, in all prediction models, physical function was significantly lower in the frailty-labeled group (**Table 3**).

Last, we present the results of a visual verification of the feature trends. The feature importance results indicate that age, albumin, aspartate aminotransferase to alanine aminotransferase ratio, body mass index, and hemoglobin are the most important features, in that order (**Figure 5 Left**). Partial dependence of these major characteristics showed that confidence in judging frailty increased non-linearly after age exceeded approximately 65 years, after body mass index fell below approximately 20 kg/m², after albumin fell below approximately 4.0 mg/dL, and after hemoglobin fell below approximately 13 mg/dL (**Figure 5 Top Right**). In addition, the contribution of each feature was visually verified using SHAP values. The case in **Figure 5 Middle Right** is an older patient (age 78 years), and the parameters are trending towards predicting frailty, but considering the results of the other parameters, the model predicts non-frail. In contrast, the case in **Figure 5 Bottom Right** is comparatively young (age 55 years), but the model predicts frailty because of a low body mass index, hemoglobin concentration values, and other worse factors.

Discussion

The goal of this work was to create and establish the utility of an AI model for predicting phenotypic frailty models in patients with CVD using common indicators available from EHRs. The results showed that (1) the model created by Light Gradient Boosting was more accurate, (2) models assessing commonly used blood biochemistry test indices (Simple EHR model) showed almost the same discrimination ability, and (3) all of the frailty prediction models were consistently associated with prognosis and physical function. The main strengths of this study, which produced these results, are that we built a model to predict actual measured phenotypic physical frailty in patients with CVD, and that the model was generated using commonly used blood biochemistry test indices.

To our knowledge, this is the first report to use machine learning for phenotypic frail prediction with EHR data as features in patients with CVD. CVD is widely known to have a high rate of frailty complications and a poor prognosis compared with the community-dwelling older population.^{5,17,24} A previous study using machine learning with an unsupervised clustering approach to predict frailty in 37,431 veterans' EHR data, composed of frailty index, ejection fraction, laboratory values, blood pressure, and demographic information from EHRs in HF, is associated with mortality, similar to the results in this study.²⁵ Another notable strength is that, to our knowledge, this is the largest sample size of any report on machine learning using EHRs with Fried et al.'s phenotype-defined frailty as the correct label. Two reports on machine learning with phenotypic frailty as a supervised label validated the model using EHR data. One report assessed the EHR data of 474

patients using physique, blood, cardiac, disease, self-reported disease, consumption, and medical test attributes.²⁶ Another study used discharge data from 469 hospitalized patients using age, sex, cumulative length of stay in acute care and the intensive care unit, presence of at least 1 emergency admission, diagnostic code, and an electrical frailty index.²⁷ In contrast, some reports use frailty, defined by other frailty assessment tools, as the correct label. One report validated the Gradient Boosting model using the same method as in the present study;²⁸ a decision model was created for 5,466 primary care patients using frailty defined by the Rockwood clinical frailty scale as the correct label. Features in this previous study included age, sex, diagnosis, chronic conditions, biometrics, province, medications, physique, and blood pressure. Despite these previous studies, there have been no reports on the development of a deep learning model to predict phenotypic frailty using blood test data obtained in daily clinical practice, and the present study is the first attempt to do so.

One of the strengths of the frailty prediction model developed in this study for clinical use is its ability to visualize which features contribute to the degree of frailty using SHAP. The main features extracted (aging, underweight, undernutrition, liver dysfunction, and anemia) were familiar factors associated with frailty.^{17,29–31} Moreover, the thresholds at which the confidence level to determine frailty increased non-linearly were also generally consistent with the clinical cut-off values (age 65 years; body mass index 20 kg/m²; albumin 4.0 g/dL; hemoglobin 13–14 g/dL). These results could provide a helpful guide to the question of which aspects of the multifactorial and diverse clinical profile of frail patients should be focused on and intensively treated. In other words, we successfully modeled the individually reported frailty-related indicators, with weighting according to their characteristics. Thus, it became possible to evaluate and infer the pathophysiology of frailty from multiple perspectives for each patient. This strength point is not only original and unique compared with previous studies, but also is expected to be used as a predictive model of frailty that can be applied to clinical practice.

In this study, we performed the same validation using only commonly used blood biochemistry test indices for general purposes in addition to basic information, such as age, sex, and body mass index, so that the results could be used not only in acute care settings but also in the community and for health checkups. In addition, the model was designed for broad application in various healthcare institutions to automatically assess frailty risk, and therefore we deliberately chose not to incorporate specific score-based indices such as nutritional scores. Additionally, we excluded feature factors like gait speed and grip strength, which are not routinely measured in some facilities. However, many of the variables used in developing the algorithm are well-known indicators related to physical frailty and malnutrition. As a result, a good accuracy prediction model was created with an AUC of approximately 0.8. In recent years, health checkups have been conducted in Japan to screen for high-risk frail individuals in community-dwelling older adults, and the usefulness of these checkups is reported.³² For example, implementing the frailty determination model developed in this study could lead to identifying detailed high-risk cases and early intervention to correct frailty. Thus, in the future, it will be necessary to focus on evaluating and reinforcing frailty measures based on the frailty prediction model developed in this study, and

to determine whether frailty measures lead to improved patient outcomes.³³ In addition, given the inherent trade-off between sensitivity and specificity in frailty prediction, it may be necessary to adjust the cut-off value based on the target population and the clinical context in which the model is applied. Lowering the cut-off value could reduce the occurrence of FNs, minimizing the risk of missing frailty cases. When applying this model in clinical practice, careful consideration should be given to adjusting the cut-off value to ensure optimal balance between the risks of missing frailty and the potential for overdiagnosis, tailored to the specific population and clinical circumstances.

Study Limitations

While this study has the strengths mentioned above, it also has some limitations. First, this study included only single-center, retrospective, cross-sectional data, including only patients with CVD, and it did not examine longitudinal changes in the frailty prediction model. Further validation is required to determine whether the model developed in this study can be applied to community-dwelling older adults and individuals with other diseases. In addition, future studies need to follow up to see how changes in the prediction model due to improved or worsened parameters contribute to clinical and other outcomes. Second, this study only deals with information that can be collected directly from EHRs and does not cover other indicators that may be useful for predicting frailty, such as text data³⁴ and medical images.³⁵ Therefore, although the predictive ability was high, the AUC was only approximately 0.8. Solving these problems would enable us to create a model for predicting frailty with a higher ability. Last, because this study focused on cases in which the phenotypic model could be measured, it is not clear whether it can be adapted to patients with functional impairment to the extent that measurement is not possible.

Conclusions

In the present study, using a phenotypic model of frailty as the correct label, we developed a model to predict frailty in patients with CVD using data only from EHRs. Models derived from very large sample sizes of frailty assessment data, based on actual measurements, such as gait speed and grip strength, have shown consistent associations with frailty and prognosis, creating a robust model. The main features were widely known to be associated with frailty, and only commonly used blood biochemistry test indices, in addition to basic information such as age and sex, also had sufficient predictive accuracy. For social implementation of the obtained model, it is necessary to confirm external validity, report the results through trial operations in the clinical setting or any other field, and track the longitudinal changes in the frailty model due to changes in the features.

Acknowledgments

The authors thank Mr. Takemura of the Open Health Initiative staff for data analysis, and Health Care Relations Co., Ltd staff for their contributions to social implementation. This work was supported by JST-OPERA Program (Grant no. JPMJOP1842), Japan. The manuscript of this study was written in its original form, checked for grammar, rewritten using AI, and then proofread by a professional grammar checker.

Disclosures

M.Y. has no conflicts of interest to disclose in the conduct of the present

study, although he belongs to ARCE Inc. and receives a salary as one of the directors of an employer. K.K., J.A. are members of *Circulation Reports'* Editorial Team. The other authors have no conflicts of interest to declare.

IRB Information

The present study was approved by the Ethics Committee of Kitasato University Hospital (B21-170).

Data Availability

The data is not available. The original data can be provided by contacting the corresponding author, but the data used for deep learning and the socially implemented files are not available.

References

1. Pison G, Couppié E, Caporali A. The population of the world, 2022. *Population & Societies* 2022; **603**: 1–8, doi:10.3917/popsoc.603.0001.
2. Walston J, Hadley EC, Ferrucci L, Guralnik JM, Newman AB, Studenski SA, et al. Research agenda for frailty in older adults: Toward a better understanding of physiology and etiology: Summary from the American Geriatrics Society/National Institute on Aging Research Conference on Frailty in Older Adults. *J Am Geriatr Soc* 2006; **54**: 991–1001, doi:10.1111/j.1532-5415.2006.00745.x.
3. Ofori-Asenso R, Chin KL, Mazidi M, Zomer E, Ilomaki J, Zullo AR, et al. Global incidence of frailty and prefrailty among community-dwelling older adults: A systematic review and meta-analysis. *JAMA Netw Open* 2019; **2**: e198398, doi:10.1001/jamanetworkopen.2019.8398.
4. Murayama H, Kobayashi E, Okamoto S, Fukaya T, Ishizaki T, Liang J, 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.
5. Kim DH, Zhong L, Rich MW. Frailty-guided management of cardiovascular disease-from clinical trials to clinical practice. *JAMA Cardiol* 2023; **8**: 897–898, doi:10.1001/jamacardio.2023.1634.
6. Rezaei-Shahsavarloo Z, Atashzadeh-Shoorideh F, Gobbens RJJ, Ebadi A, Ghaedamini Harouni G. The impact of interventions on management of frailty in hospitalized frail older adults: A systematic review and meta-analysis. *BMC Geriatr* 2020; **20**: 526, doi:10.1186/s12877-020-01935-8.
7. Kitzman DW, Whellan DJ, Duncan P, Pastva AM, Mentz RJ, Reeves GR, et al. Physical rehabilitation for older patients hospitalized for heart failure. *N Engl J Med* 2021; **385**: 203–216, doi:10.1056/NEJMoa2026141.
8. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: Evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 2001; **56**: M146–M156.
9. Kamiya K, Sato Y, Takahashi T, Tsuchihashi-Makaya M, Kotooka N, Ikegame T, et al. Multidisciplinary cardiac rehabilitation and long-term prognosis in patients with heart failure. *Circ Heart Fail* 2020; **13**: e006798, doi:10.1161/circheartfailure.119.006798.
10. Tanaka S, Kamiya K, Hamazaki N, Matsuzawa R, Nozaki K, Nakamura T, et al. Short-term change in gait speed and clinical outcomes in older patients with acute heart failure. *Circ J* 2019; **83**: 1860–1867, doi:10.1253/circj.CJ-19-0136.
11. Hamazaki N, Kamiya K, Yamamoto S, Nozaki K, Ichikawa T, Matsuzawa R, et al. Changes in respiratory muscle strength following cardiac rehabilitation for prognosis in patients with heart failure. *J Clin Med* 2020; **9**: 952, doi:10.3390/jcm9040952.
12. Tanaka S, Kamiya K, Matsue Y, Yonezawa R, Saito H, Hamazaki N, et al. Effects of electrical muscle stimulation on physical function in frail older patients with acute heart failure: A randomized controlled trial. *Eur J Prev Cardiol* 2022; **29**: e286–e288, doi:10.1093/eurjpc/zwac022.
13. Yamashita M, Kamiya K, Hamazaki N, Uchida S, Noda T, Maekawa E, et al. Effects of Acute Phase Intensive Physical Activity (ACTIVE-PA) monitoring and education for cardiac patients: Pilot study of a randomized controlled trial. *J Med Internet Res* 2023; **25**: e42235, doi:10.2196/42235.
14. Yamashita M, Kamiya K, Kitamura T, Hamazaki N, Ichikawa T, Shibuya M, et al. Trajectory of abdominal skeletal muscle changes during cardiac rehabilitation in patients with aortic

- disease. *Am J Phys Med Rehabil* 2024; **103**: 158–165, doi:10.1097/phm.0000000000002322.
15. Shakuta S, Noda T, Kamiya K, Hamazaki N, Nozaki K, Yamashita M, et al. Clinical impact of improvement in sarcopenia through cardiac rehabilitation in patients with heart failure. *J Am Med Dir Assoc* 2024; **25**: 514–520.e2, doi:10.1016/j.jamda.2023.10.035.
 16. Velazquez-Diaz D, Arco JE, Ortiz A, Pérez-Cabezas V, Lucena-Anton D, Moral-Munoz JA, et al. Use of artificial intelligence in the identification and diagnosis of frailty syndrome in older adults: Scoping review. *J Med Internet Res* 2023; **25**: e47346, doi:10.2196/47346.
 17. Matsue Y, Kamiya K, Saito H, Saito K, Ogasahara Y, Maekawa E, et al. Prevalence and prognostic impact of the coexistence of multiple frailty domains in elderly patients with heart failure: The FRAGILE-HF cohort study. *Eur J Heart Fail* 2020; **22**: 2112–2119, doi:10.1002/ejhf.1926.
 18. Satake S, Arai H. The revised Japanese version of the Cardiovascular Health Study criteria (revised J-CHS criteria). *Geriatr Gerontol Int* 2020; **20**: 992–993, doi:10.1111/ggi.14005.
 19. Chen LK, Woo J, Assantachai P, Auyeung TW, Chou MY, Iijima K, et al. Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. *J Am Med Dir Assoc* 2020; **21**: 300–307.e2, doi:10.1016/j.jamda.2019.12.012.
 20. Kamiya K, Masuda T, Matsue Y, Hamazaki N, Matsuzawa R, Tanaka S, et al. Prognostic usefulness of arm and calf circumference in patients ≥ 65 years of age with cardiovascular disease. *Am J Cardiol* 2017; **119**: 186–191, doi:10.1016/j.amjcard.2016.09.040.
 21. Kamiya K, Masuda T, Tanaka S, Hamazaki N, Matsue Y, Mezzani A, et al. Quadriceps strength as a predictor of mortality in coronary artery disease. *Am J Med* 2015; **128**: 1212–1219, doi:10.1016/j.amjmed.2015.06.035.
 22. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: A highly efficient gradient boosting decision tree. In: *Advances in Neural Information Processing Systems* 30. Annual Conference on Neural Information Processing Systems. 2017; 3146–3154.
 23. Breiman L. Random forests. *Machine Learning* 2001; **45**: 5–32.
 24. Yamada S, Adachi T, Izawa H, Murohara T, Kondo T. Prognostic score based on physical frailty in patients with heart failure: A multicenter prospective cohort study (FLAGSHIP). *J Cachexia Sarcopenia Muscle* 2021; **12**: 1995–2006, doi:10.1002/jcsm.12803.
 25. Razjouyan J, Horstman MJ, Orkaby AR, Virani SS, Intrator O, Goyal P, et al. Developing a parsimonious frailty index for older, multimorbid adults with heart failure using machine learning. *Am J Cardiol* 2023; **190**: 75–81, doi:10.1016/j.amjcard.2022.11.044.
 26. Hassler AP, Menasalvas E, García-García FJ, Rodríguez-Mañás L, Holzinger A. Importance of medical data preprocessing in predictive modeling and risk factor discovery for the frailty syndrome. *BMC Med Inform Decis Mak* 2019; **19**: 33, doi:10.1186/s12911-019-0747-6.
 27. Le Pogam MA, Seematter-Bagnoud L, Niemi T, Assouline D, Gross N, Trächsel B, et al. 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, doi:10.1016/j.eclinm.2021.101260.
 28. Aponte-Hao S, Wong ST, Thandi M, Ronksley P, McBrien K, Lee J, et al. Machine learning for identification of frailty in Canadian primary care practices. *Int J Popul Data Sci* 2021; **6**: 1650, doi:10.23889/ijpds.v6i1.1650.
 29. Yamashita M, Kamiya K, Hamazaki N, Nozaki K, Uchida S, Maekawa E, et al. Predictive value of cholinesterase in patients with heart failure: A new blood biochemical marker of undernutrition. *Nutr Metab Cardiovasc Dis* 2023; **33**: 1914–1922, doi:10.1016/j.numecd.2023.06.005.
 30. Noda T, Kamiya K, Hamazaki N, Nozaki K, Ichikawa T, Nakamura T, et al. Prognostic value of cardio-hepatic-skeletal muscle syndrome in patients with heart failure. *Sci Rep* 2021; **11**: 3715, doi:10.1038/s41598-020-80641-2.
 31. Tavenier J, Leng SX. Inflammatory pathways to anemia in the frail elderly. *Clin Geriatr Med* 2019; **35**: 339–348, doi:10.1016/j.cger.2019.03.005.
 32. Hori N, Ishizaki T, Masui Y, Yoshida Y, Inagaki H, Ito K, et al. Criterion validity of the health assessment questionnaire for the national screening program for older adults in Japan: The SONIC study. *Geriatr Gerontol Int* 2023; **23**: 437–443, doi:10.1111/ggi.14585.
 33. Best K, Todd O, Clegg A. Are frailty measurements derived using electronic health records fit for clinical use? *Age Ageing* 2024; **53**: afae001, doi:10.1093/ageing/afae001.
 34. Shao Y, Mohanty AF, Ahmed A, Weir CR, Bray BE, Shah RU, et al. Identification and use of frailty indicators from text to examine associations with clinical outcomes among patients with heart failure. *AMIA Annu Symp Proc* 2016; **2016**: 1110–1118.
 35. Liu J, Su S, Liu W, Xie E, Hu X, Lin W, et al. The impact of machine-learning-derived lean psoas muscle area on prognosis of type B aortic dissection patients undergoing endovascular treatment. *Eur J Cardiothorac Surg* 2022; **62**: ezac482, doi:10.1093/ejcts/ezac482.

Supplementary Files

Please find supplementary file(s);
<https://doi.org/10.1253/circrep.CR-24-0112>