frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 08 March 2023
DOI 10.3389/fneur.2023.1130236

Check for updates

OPEN ACCESS

EDITED BY
Mahesh P. Kate,
University of Alberta Hospital, Canada

REVIEWED BY
Zhixin Huang,
Guangdong Second Provincial General Hospital, China
James Chow,
University of Toronto, Canada

*CORRESPONDENCE
Yun-Hee Kim
✉ yunkim@skku.edu;
✉ yun1225.kim@samsung.com

SPECIALTY SECTION
This article was submitted to Stroke, a section of the journal Frontiers in Neurology

RECEIVED 23 December 2022
ACCEPTED 15 February 2023
PUBLISHED 08 March 2023

CITATION
Shin S, Chang WH, Kim DY, Lee J, Sohn MK, Song M-K, Shin Y-I, Lee Y-S, Joo MC, Lee SY, Han J, Ahn J, Oh G-J, Kim Y-T, Kim K and Kim Y-H (2023) Clustering and prediction of long-term functional recovery patterns in first-time stroke patients.
Front. Neurol. 14:1130236.
doi: 10.3389/fneur.2023.1130236

COPYRIGHT
© 2023 Shin, Chang, Kim, Lee, Sohn, Song, Shin, Lee, Joo, Lee, Han, Ahn, Oh, Kim, Kim and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Clustering and prediction of long-term functional recovery patterns in first-time stroke patients

Seyoung Shin[1], Won Hyuk Chang[1], Deog Young Kim[2], Jongmin Lee[3], Min Kyun Sohn[4], Min-Keun Song[5], Yong-Il Shin[6], Yang-Soo Lee[7], Min Cheol Joo[8], So Young Lee[9], Junhee Han[10], Jeonghoon Ahn[11], Gyung-Jae Oh[12], Young-Taek Kim[13], Kwangsu Kim[14] and Yun-Hee Kim[1,15]*

[1]Department of Physical and Rehabilitation Medicine, Center for Prevention and Rehabilitation, Heart Vascular Stroke Institute, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea, [2]Department and Research Institute of Rehabilitation Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea, [3]Department of Rehabilitation Medicine, Konkuk University School of Medicine, Seoul, Republic of Korea, [4]Department of Rehabilitation Medicine, College of Medicine, Chungnam National University, Daejeon, Republic of Korea, [5]Department of Physical and Rehabilitation Medicine, Chonnam National University Medical School, Gwangju, Republic of Korea, [6]Department of Rehabilitation Medicine, Pusan National University School of Medicine, Pusan National University Yangsan Hospital, Yangsan-si, Republic of Korea, [7]Department of Rehabilitation Medicine, School of Medicine, Kyungpook National University, Kyungpook National University Hospital, Daegu, Republic of Korea, [8]Department of Rehabilitation Medicine, Wonkwang University School of Medicine, Iksan, Republic of Korea, [9]Department of Rehabilitation Medicine, Jeju National University Hospital, Jeju National University School of Medicine, Jeju-si, Republic of Korea, [10]Department of Statistics, Hallym University, Chuncheon-si, Republic of Korea, [11]Department of Health Convergence, Ewha Womans University, Seoul, Republic of Korea, [12]Department of Preventive Medicine, School of Medicine, Wonkwang University, Iksan, Republic of Korea, [13]Department of Preventive Medicine, Chungnam National University Hospital, Daejeon, Republic of Korea, [14]College of Computing, Sungkyunkwan University, Suwon-si, Republic of Korea, [15]Department of Health Sciences and Technology, Department of Medical Device Management and Research, Department of Digital Health, Samsung Advanced Institute for Health Sciences & Technology (SAIHST), Sungkyunkwan University, Seoul, Republic of Korea

Objectives: The purpose of this study was to cluster long-term multifaceted functional recovery patterns and to establish prediction models for functional outcome in first-time stroke patients using unsupervised machine learning.

Methods: This study is an interim analysis of the dataset from the Korean Stroke Cohort for Functioning and Rehabilitation (KOSCO), a long-term, prospective, multicenter cohort study of first-time stroke patients. The KOSCO screened 10,636 first-time stroke patients admitted to nine representative hospitals in Korea during a three-year recruitment period, and 7,858 patients agreed to enroll. Early clinical and demographic features of stroke patients and six multifaceted functional assessment scores measured from 7 days to 24 months after stroke onset were used as input variables. K-means clustering analysis was performed, and prediction models were generated and validated using machine learning.

Results: A total of 5,534 stroke patients (4,388 ischemic and 1,146 hemorrhagic; mean age 63·31 ± 12·86; 3,253 [58.78%] male) completed functional assessments 24 months after stroke onset. Through K-means clustering, ischemic stroke (IS) patients were clustered into five groups and hemorrhagic stroke (HS) patients into four groups. Each cluster had distinct clinical characteristics and functional recovery patterns. The final prediction models for IS and HS patients achieved relatively high prediction accuracies of 0.926 and 0.887, respectively.

Conclusions: The longitudinal, multi-dimensional, functional assessment data of first-time stroke patients were successfully clustered, and the prediction

models showed relatively good accuracies. Early identification and prediction of long-term functional outcomes will help clinicians develop customized treatment strategies.

KEYWORDS

stroke, functional recovery, artificial intelligence, machine learning, clustering, prediction

Introduction

Although early stroke management and rehabilitation protocols have improved during the past decade, stroke remains the most common cause of adult physical disability worldwide (1). The number of stroke survivors and the related overall global burden of stroke are both increasing (2). The ability to predict long-term recovery and prognosis of functional deficits after stroke is of interest. If clinicians could foresee the long-term functional recovery prospects for a certain patient, they could devise better treatment strategies.

Previous studies have attempted to develop algorithms to predict the prognosis for recovery after stroke (3). However, such algorithms considered only a single time point, making it difficult to establish overall recovery patterns. Douiri et al. (4) suggested creating decision curves to produce dynamic, time-dependent, multivariate, patient-specific predictive models that could overcome those limitations. Focusing on this attempt and the results of previous studies, it seems to be necessary to include the multifaceted functional outcomes for future prediction models. Because stroke patients often suffer from distinct motor, language, cognitive, and swallowing dysfunction, measuring only their ability to perform activities of daily living (ADL) is insufficient to classify patient-specific recovery patterns.

According to a previous study, lifestyle also is an important factor in stroke outcome (5). Therefore, when designing a prognosis prediction model for stroke patients, lifestyle factors need to be considered. The Korean Stroke Cohort for Functioning and Rehabilitation (KOSCO) (2) includes clinical characteristics; serial data of various functional domains; and lifestyle factors such as alcohol, smoking, and education level.

Therefore, in this study, we used a clustering analysis based on an unsupervised machine learning method that is suitable for classifying large, real-world KOSCO datasets containing longitudinal, multi-dimensional, functional assessments. Our primary aim in this study was to identify multifaceted, functional recovery patterns among first-time stroke patients using an unsupervised machine learning algorithm. Our secondary aim was to generate a prediction model for those recovery pattern clusters and examine the accuracy of the models.

Methods

Study populations

This study used data from the KOSCO study (6), a long-term, prospective, multicenter cohort study of residual disability and functional independence among Korean stroke patients following their first stroke episode.

Between August 2012 and May 2015, the KOSCO study recruited 10,636 Korean patients. The inclusion criteria were (1) first-time acute IS or HS with a corresponding lesion on computed tomography or magnetic resonance imaging/angiography, (2) at least 19 years of age at stroke onset, and (3) onset of symptoms within 7 days prior to study enrollment. Patients with any of the following criteria were excluded: (1) transient ischemic attack, (2) history of previous stroke, and (3) traumatic intracerebral hemorrhage. Of the 10,636 first-time stroke patients (8,210 IS patients and 2,426 HS patients) admitted to nine representative hospitals in Korea during the recruitment period, 7,858 (6,253 IS patients and 1,605 HS patients) agreed to enroll after exclusion of patients who died or declined to participate. Among them, 5,534 patients (4,388 IS patients and 1,146 HS patients) who completed their follow-up assessments through 24 months after stroke onset were used in this analysis (Figure 1).

Written informed consent was obtained from all patients prior to inclusion, and the study protocol was approved by the ethics committees of the involved hospitals (Supplementary material).

Measurements

Demographic and clinical characteristics

We considered the following demographic and clinical characteristics: sex, age, obesity (body mass index $\geq$ 26), education level (high: more than 9 years, low: <9 years), and stroke location (right, left, or both). Stroke severity was measured by the National Institutes of Health Stroke Scale (NIHSS) for 7 days after stroke onset for both IS and HS because the time from stroke onset to emergency department admission was different for each patient. A previous study showed that NIHSS is a reliable tool for clinical monitoring not only IS, but also HS patients (7). Combined condition- and age-related score (CCAS) according to Charlson Comorbidity Index, smoking, and history of alcohol consumption also was assessed. History of patient risk factors such as hypertension (systolic blood pressure > 160 mm Hg, diastolic blood pressure > 90 mm Hg, or history of hypertension or medical treatment), diabetes mellitus (DM; blood glucose level >126 mg/d or history of DM or medical treatment), hyperlipidemia (elevated low-density lipoprotein cholesterol level >160 mg/dL, elevated total cholesterol level > 240 mg/dL, or history of hy-perlipidemia or medical treatment), and atrial fibrillation (documented by standard electrocardiogram [ECG], long-term ECG, or history of atrial fibrillation or medical

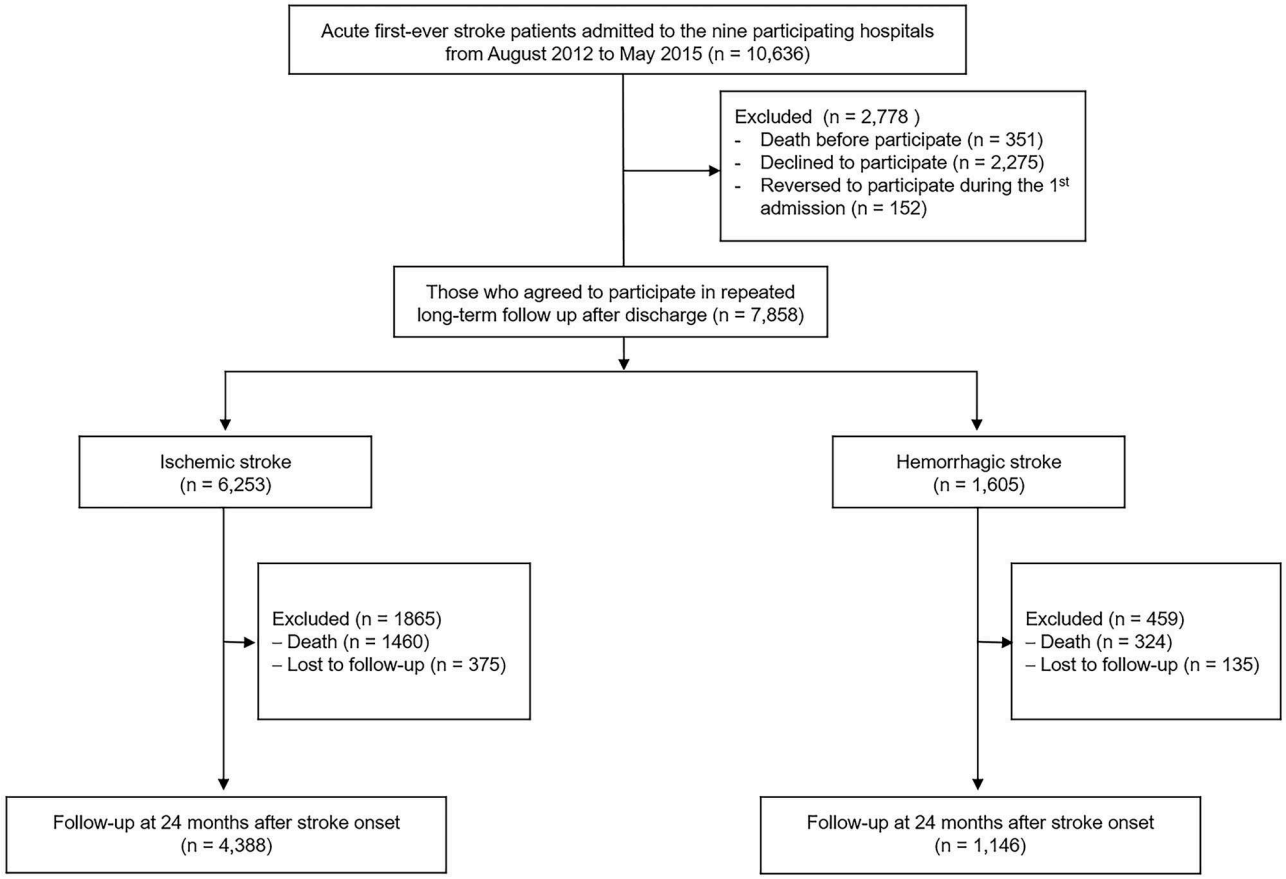

FIGURE 1
Flow chart of participant inclusion.

treatment) was assessed. Medical complications such as pneumonia and urinary tract infection during admission period were also included.

Functional assessments

The multifaceted functional assessments included six international tools: the Fugl-Meyer Assessment (FMA; range 0–100, higher score means higher motor function) (8) for motor function, the Functional Ambulation Classification (FAC; range 0–5, higher score means higher ambulatory function) (9) for mobility and gait, the Korean Mini-mental State Examination (K-MMSE; range 0–30, higher score means higher cognitive function) (10) for cognition, the short version of the Korean version of the Frenchay Aphasia Screening Test (short K-FAST; range 0–20, higher score means higher language function) (11) for language function, the American Speech-Language-Hearing Association's National Outcomes Measurement System (ASHA-NOMS; range 1–7, higher score means higher swallowing function) (12) for swallowing function, and the Korea Modified Barthel Index (K-MBI; range 0–100, higher score means higher activities of daily living performance independence) (13) for ADL function. Serial data from face-to-face functional assessments were gathered 7 days and 3, 6, 12, 18, and 24 months after stroke onset for all measures except K-MBI. K-MBI was not assessed at 7 days after stroke because most patients remained in the stroke unit for intensive care during the first week of admission.

The investigators of the KOSCO study were expert occupational therapists and underwent a standardized training program every 3 months to maintain inter-rater reliability.

Clustering of functional recovery patterns

Clustering of functional recovery patterns was performed in first-time IS and HS stroke survivors. In order to select the most suitable clustering algorithm for the KOSCO dataset, we performed prior clustering using three well-known algorithms of K-means clustering (14), the Gaussian Mixture Model (15), and the Agglomerative clustering algorithm (15) and compared the functional scores. With these three algorithms, prior clustering was performed for cluster numbers 2–15, and the Silhouette Index (SI) (16) was estimated (Supplementary Figure 1). Among the three clustering algorithms, the K-means method was chosen for its higher SI for both IS and HS than the others.

The K-means clustering algorithm is one of the most popular unsupervised machine learning algorithms that partitions a dataset into a given number of clusters. The algorithm gathers each data

point to the nearest centroid according to the number of clusters (17). To choose the optimal number of clusters that is not only able to explain the clinical features of patients, but also suitable for practical use, we used three functional scores: SI, Davies-Boulding Index (DBI) (18), and Calinski-Harabasz Index (CHI) (19). Cluster sizes from $k = 2$ to $k = 15$ were tested (Supplementary Figure 2). To avoid dividing into the highest, middle, and lowest groups, we considered only $k > 3$. In IS, $k = 5$ showed tolerable functional scores (SI 0.47, DBI 1.43, and CHI 2487.07). In HS, $k = 4$ showed similar SI scores to that of IS (SI 0.42, DBI 1.36, CHI 875.67).

Finally, *K*-means clustering was performed with 100% of the dataset (4,388 IS and 1,146 HS patients). In this step, early clinical and demographic features of stroke patients and the repeated multifaceted functional assessment scores until 24 months were used as input variables. Missing or incomplete data were imputed using the *k*-nearest neighbor-5 (kNN-5) method (20). To confirm proper clustering, we visualized the clustered groups in low-dimensional images derived by t-Distributed Stochastic Neighbor Embedding (t-SNE) (21), which is widely used to convert high-dimensional data into a two- or three-dimensional map.

Prediction model for long-term functional recovery

After clustering, we generated models that predict the cluster of new first-time stroke patients based on basic demographic data and functional scores from 7 days to 3 months after stroke onset. We used 70% of the dataset (3,071 IS and 802 HS patients) to generate the models and the remaining 30% (1,317 IS and 344 HS patients) for validation. Prediction models were simultaneously generated by eight machine learning algorithms: Light Gradient Boosting Machine (Light GBM) (22), extended version of Light GBM (Light GBM-XT) (22), Random Forest (RF) (23), CatBoost (CB) (24), extreme gradient boosting (XGBoost) (25), Weighted Ensemble (26), Neural Network (27), and Extra Trees (28). The performance metrics true positive (*TP*), true negative (*TN*), false positive (*FP*), and false negative (*FN*) were calculated. The mathematical expressions for F1 score, precision, and recall were as follows:

$$\text{Precision (PR) was given by}: \ \mathrm{PR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

$$\text{Recall (RC) was given by}: \ \mathrm{RC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

$$\text{F1 score was given by}: \ \mathrm{F}_\beta = \frac{\left(1 + \beta^2\right) \cdot \mathrm{PR} \cdot \mathrm{RC}}{\left(\beta^2 \left(\mathrm{PR} + \mathrm{RC}\right)\right)}$$

where β represents the weighted value between precision and recall. In this case, $\beta = 1$.

The accuracy of the overall prediction model for each IS and HS was calculated as follows:

$$\text{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

Computational details

Descriptive statistical analyses and within-group comparison were implemented in R (version 4.0.3). The independent *t*-test was used for comparison of continuous variables, and the chi-square test was used for comparison of categorical variables between IS and HS. The level of significance was set as two-sided $p < 0.05$. All machine learning steps of data analysis, preprocessing of model training, and visualization were performed using open source libraries in Python (version 3.9.0). Pandas (version 1.5.2) was used for data analysis and preprocessing, and scikit-learn (version 1.2.0) was used to impute missing values and to establish the clustering model. Visualization was conducted through plotly (version 5.9.0) and matplitlib (version 3.6.2). Model predictions were performed using Autogluon (version 0.6.2).

Results

Patient characteristics

The demographic and clinical characteristics of the participants are provided in Table 1. Of the 5,534 patients who underwent functional assessment 24 months after stroke onset, 4,388 were IS (2,700 males) and 1,146 were HS (553 males) patients. The mean age (standard deviation, SD) of IS patients was 64.8 (SD, 12.4) years, and their mean NIHSS score at 7 days was 3.5 (SD, 5.1). There were significant differences between IS and HS patients in demographic and clinical characteristics except obesity and alcohol history ($p < 0.001$).

Clustering of long-term functional outcomes in survivors of first-time stroke

Regarding the *K*-means clustering algorithm, the optimal number of clusters was five for IS patients and four for HS patients. In both the IS and HS groups, the clusters differed in mean age, initial stroke severity as measured by NIHSS scores, complications, and comorbidities. Figure 2 visualizes IS (Figure 2A) and HS (Figure 2B) *K*-means clustering using t-SNE.

The functional recovery characteristics of the final five clusters of IS patients are presented in Supplementary Table 1 and Figure 3. Cluster 1, which contained 3,346 patients (60.46%), was characterized by a mean age of 63.53 years (SD, 12.25) and a low 7-day stroke severity with a mean of 1.44 (SD, 1.81). These patients showed minimal deficits in every functional domain. Cluster 2, comprising 405 patients (7.32%), was characterized by a mean age similar to cluster 1 (mean [SD]; 64.47 years [12.12]) and moderate initial severity (7.07 [5.24]). This cluster showed low motor and ambulatory functions at onset but rapidly improved during the subacute phase. Cluster 3, comprising 232 patients (4.19%), was characterized by a mean age of 64.46 years (SD, 10.87), with a low but moderately severe NIHSS score of 11.88 (SD, 5.82). This group showed significant motor and ambulatory dysfunction compared to cognitive and language functioning; however, they showed continuous improvement during the 24-month study period. In contrast, ADL and cognitive and language

TABLE 1 Demographic and clinical characteristics of the participants.

	All stroke patients (n = 5,534)	Ischemic stroke (n = 4,388)	Hemorrhagic stroke (n = 1,146)	P- value[a]
Age, years	63.31 ± 12.86 (19–98)	64.79 ± 12.38 (19–98)	57.64 ± 13.10 (20–92)	<0.001***
Sex, male	3,253 (58.78%)	2,700 (61.53%)	553 (48.26%)	<0.001***
Education				
<9 years	3,776 (68.23%)	2,915 (66.43%)	861 (75.13%)	<0.001***
≥9 years	1,758 (31.77%)	1,473 (33.57%)	285 (24.87%)	<0.001***
Stroke location				
Right	2,297 (41.51%)	1,975 (45.01%)	322 (28.10%)	<0.001***
Left	2,431 (43.93%)	2,125 (48.43%)	306 (26.70%)	<0.001***
Both	806 (14.56%)	288 (6.56%)	518 (45.20%)	<0.001***
Risk factors				
Hypertension, yes	2,937 (53.07%)	2,446 (55.74%)	491 (42.84%)	<0.001***
Diabetes mellitus, yes	1,160 (20.96%)	1,048 (23.88%)	112 (9.77%)	<0.001***
Atrial fibrillation, yes	388 (7.01%)	371 (8.45%)	17 (1.48%)	<0.001***
Hyperlipidemia, yes	510 (9.22%)	460 (10.48%)	50 (4.36%)	<0.001***
Obesity (BMI ≥ 26), yes	615 (11.11%)	502 (11.44%)	113 (9.86%)	0.14
CCAS ≥ 6, yes	2,319 (41.90%)	1,989 (45.33%)	330 (28.80%)	<0.001***
Smoking, yes	1,505 (27.20%)	1,244 (28.35%)	261 (22.77%)	<0.001***
Alcohol, yes	2,228 (40.26%)	1,751 (39.90%)	477 (41.62%)	0.31
Complications				
Pneumonia, yes	137 (2.48%)	78 (1.78%)	59 (5.15%)	<0.001***
Urinary tract infection, yes	135 (2.44%)	80 (1.82%)	55 (4.80%)	<0.001***
NIHSS at 7 days	4.05 ± 6.23 (0–42)	3.53 ± 5.11 (0–42)	6.71 ± 8.85 (0–42)	<0.001***
Functional assessments at 7 days				
FMA	78.77 ± 32.41 (0–100)	81.89 ± 29.73 (0–100)	66.84 ± 38.83 (0–100)	<0.001***
FAC	2.89 ± 1.91 (0–5)	3.17 ± 1.80 (0–5)	1.82 ± 1.95 (0–5)	<0.001***
K-MMSE	21.61 ± 8.79 (0–30)	22.64 ± 7.98 (0–30)	17.65 ± 10.45 (0–30)	<0.001***
Short K-FAST	13.01 ± 6.35 (0–20)	13.63 ± 5.92 (0–20)	10.65 ± 7.30 (0–20)	<0.001***
ASHA-NOMS	5.92 ± 1.95 (1–7)	6.13 ± 1.76 (1–7)	5.11 ± 2.38 (1–7)	<0.001***
Functional assessments at 3 months				
K-MBI	86.37 ± 25.35 (0–100)	87.89 ± 23.55 (0–100)	80.58 ± 30.63 (0–100)	<0.001***

Values are mean ± standard deviation (range) or number (%).
[a]The independent t-test was used for continuous variables and the chi-square test was used for categorical variables between ischemic and hemorrhagic stroke types.
***$p < 0.001$.
BMI, Body mass index; CCAS, Combined condition- and age-related score; NIHSS, National Institutes of Health Stroke Scale; FMA, Fugl-Meyer Assessment; FAC, Functional Ambulatory Category; K-MMSE, Korean Mini-Mental State Examination; Short K-FAST, Short Korean-language Frenchay Aphasia Screening Test; ASHA-NOMS, American Speech-Language-Hearing Association National Outcome Measurement System Swallowing Scale; K-MBI, Korean Modified Barthel Index.

functions showed little decline after 12 months. Cluster 4, which contained 204 patients (3.69%), was characterized by a relatively older mean age of 76.24 years (SD, 8.44) and a 7-day NIHSS score of 5.35 (SD, 4.74). All functional domains showed dysfunction, especially the motor and ambulatory domains, which showed little improvement during the first 6 months and then decreased later. Cluster 5, which contained 201 patients (3.63%), was characterized by relatively older age, with a mean of 75.10 years (SD, 8.92), and a higher mean initial NIHSS score of 15.90 (SD, 8.07). This cluster showed the worst functional recovery over all six functional domains.

The final four clusters of HS patients and their functional recovery characteristics are presented in Supplementary Table 1 and Figure 4. The 710 patients (61.95%) in cluster 1 were

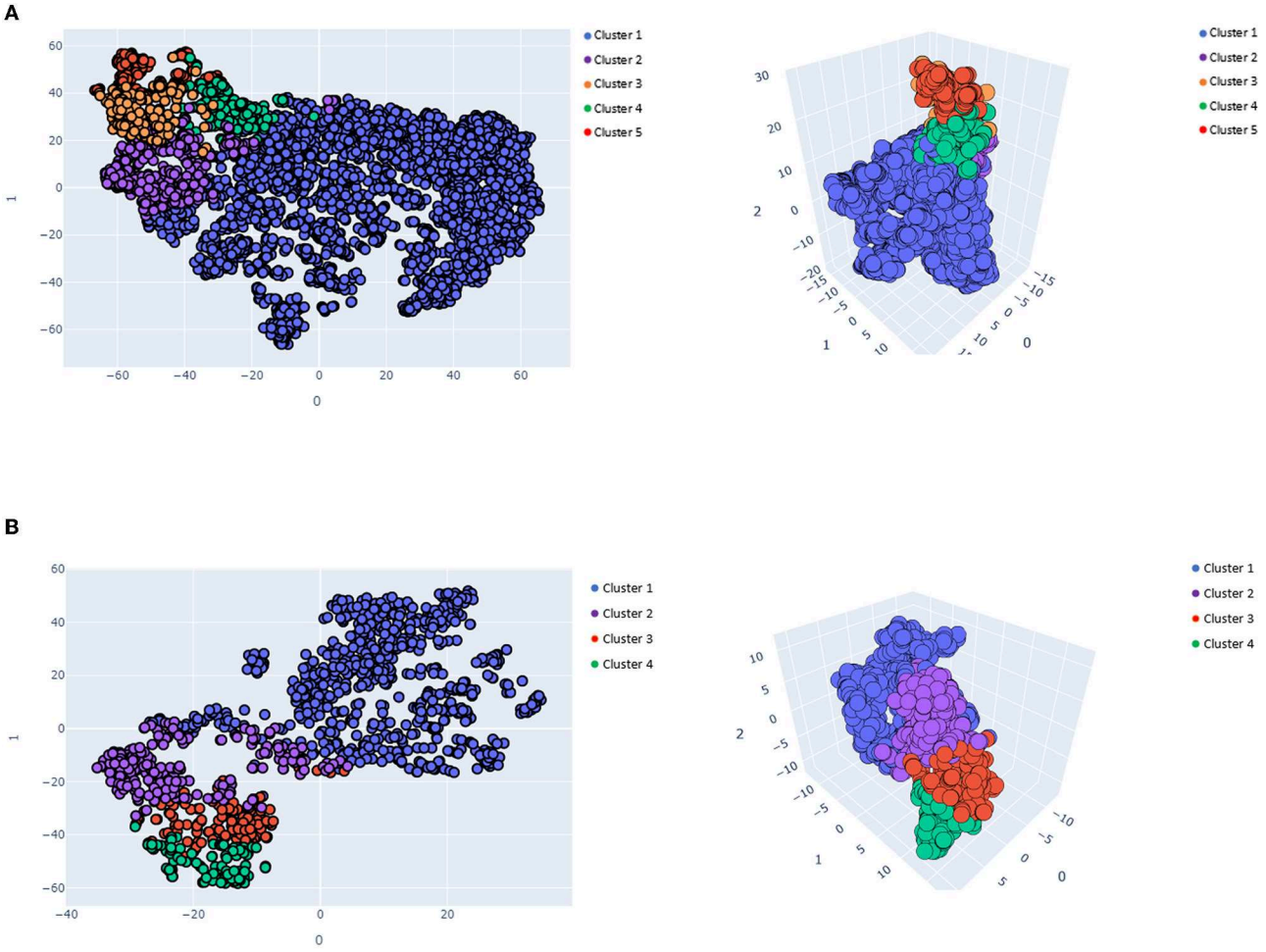

FIGURE 2
Visualized *K*-means clusters results in 2- and 3-dimensional spaces. **(A)** Visualized ischemic *K*-means clusters results ($k = 5$), **(B)** Visualized hemorrhagic *K*-means cluster results ($k = 4$).

characterized by young age (55.99 years [12.64]) and mild initial severity (1.61 [2.61]), similar to IS cluster 1. Cluster 2, which contained 208 patients (18.15%), was characterized by a mean age of 59.02 years (SD, 13.40) and a 7-day NIHSS of 12.51 (SD, 8.68). These patients had low scores in motor, cognitive, language, ambulatory, and swallowing functions at 7 days after stroke onset, but they recovered rapidly, especially in the motor and ambulatory domains. Cluster 3, comprising 128 patients (11.17%), was characterized by a mean age of 57.56 years (SD, 12.66) and a 7-day NIHSS of 15.44 (SD, 7.71). The average age was younger than in cluster 2, but the initial stroke severity was slightly higher. All functional domains showed low scores at 7 days after stroke and improved significantly during the first 3 months. Those patterns were similar to those of cluster 2. However, motor and ambulatory functions were much lower in cluster 3. Cluster 4, which contained 100 patients (8.73%), was characterized by older age (66.52 years [12.44]) and the highest initial severity (19.59 [9.92]). All domains showed the lowest initial scores with a slight improvement in the first 3 months; however, these patients showed little improvement or even worse performance over time.

Predicting recovery patterns after first-time stroke

The predictive model was developed to predict the recovery pattern for up to 24 months with only patient information up to the subacute phase of stroke. The input variables were demographic features and functional scores assessed 7 days and 3 months after stroke onset. The accuracy, F1 score, precision, and recall scores of the final IS cluster ($k = 5$) are demonstrated in Table 2. Among the eight machine learning models, CatBoost and Light GBM-XT showed the best performance (accuracy, 0.926 and 0.925, respectively). All other models showed accuracies higher than 0.90 for IS (XGBoost, 0.920; RF, 0.919; Light GBM, 0.917; Weighted Ensemble, 0.917; Neural Net, 0.912; Extra trees, 0.912).

The performance scores of the final HS cluster ($k = 4$) are described in Table 2. Both CatBoost and Light GBM-XT showed the highest accuracy of 0.887, and the Extra trees model showed the lowest accuracy of 0.861. All other models showed accuracies higher than 0.85 for HS (Light GBM, 0.887; XGBoost, 0.883; Weighted Ensemble, 0.883; RF, 0.883; Neural Net, 0.870; Extra trees, 0.861).

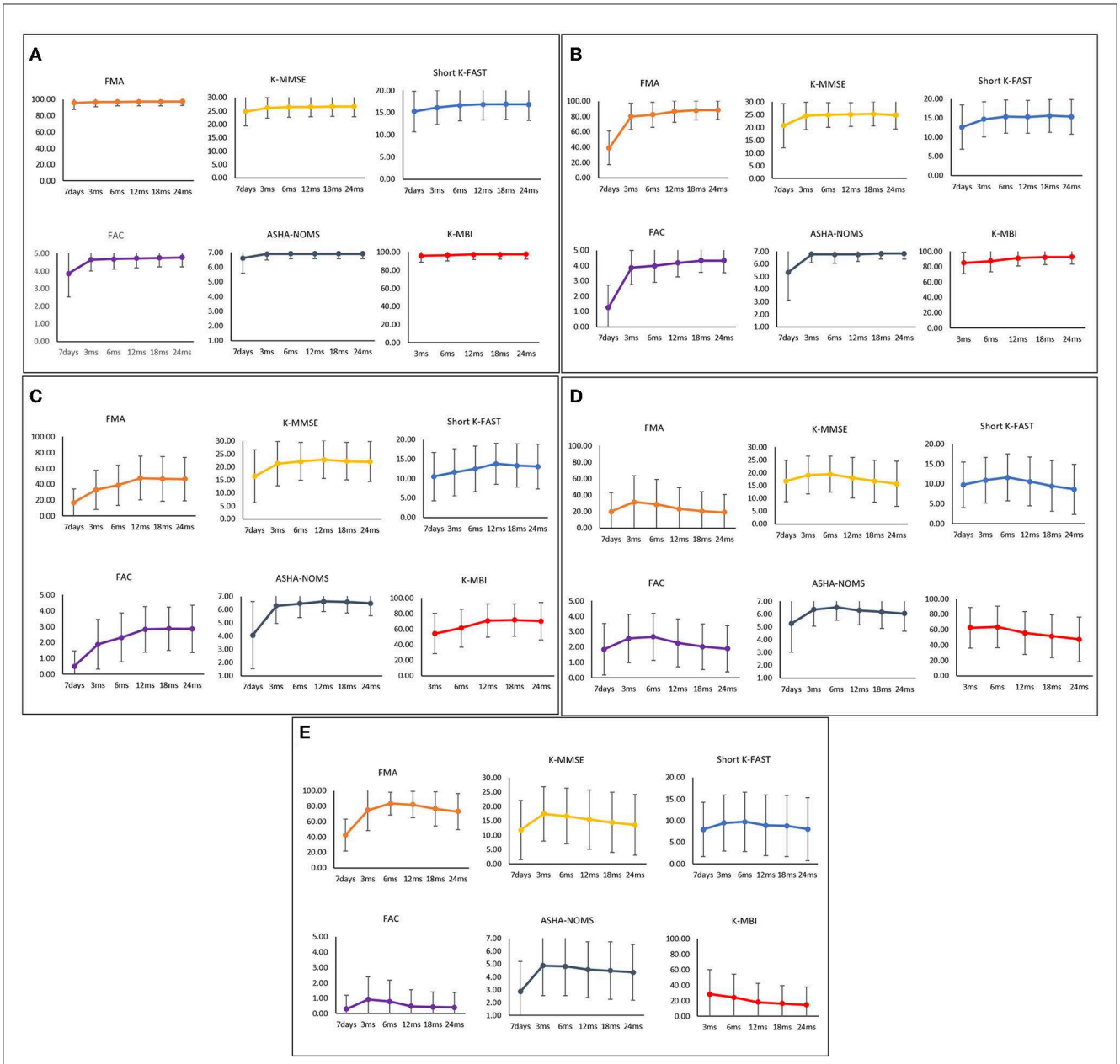

FIGURE 3
Functional recovery patterns of clusters of ischemic stroke patients until 24 months after onset. **(A)** Cluster 1 (n = 3,346); minimal functional deficit in all domains, **(B)** Cluster 2 (n = 405); rapid improvement of motor and ambulatory functions during the subacute phase, **(C)** Cluster 3 (n = 232); significant motor and ambulatory dysfunctions with continuous improvement, **(D)** Cluster 4 (n = 204); moderate dysfunctions with late phase decrement, **(E)** Cluster 5 (n = 201); severe dysfunctions in all domains. FMA, Fugl-Meyer Assessment; FAC, Functional Ambulatory Category; K-MMSE, Korean Mini-Mental State Examination; Short K-FAST, Short Korean version of the Frenchay Aphasia Screening Test; AHSA-NOMS, American Speech-Language-Hearing Association National Outcome Measurement System Swallowing Scale; K-MBI, Korean modified Barthel Index.

The detailed parameters of CatBoost and Light GBM-XT models, which showed the highest accuracies in this study, are described in the Supplementary material.

Discussion

In this study, we used an unsupervised machine learning algorithm to extract clusters of long-term, multifaceted functional recovery patterns in first-time stroke patients. After identifying the most suitable algorithm and number of clusters for our data, we identified five distinct IS clusters and four HS clusters based on clinical and demographic features. All prediction models for the IS and HS clusters achieved accuracies of 0.90 and 0.85, respectively, when using demographic, 7-day, and 3-month functional data after stroke. Among the models evaluated, CatBoost and Light GBM-XT showed the best performance in both IS and HS.

Recently, machine learning has played an increasing role in medical research. The number of publications about machine learning in the medical field is increasing annually. PubMed

10.3389/fneur.2023.1130236

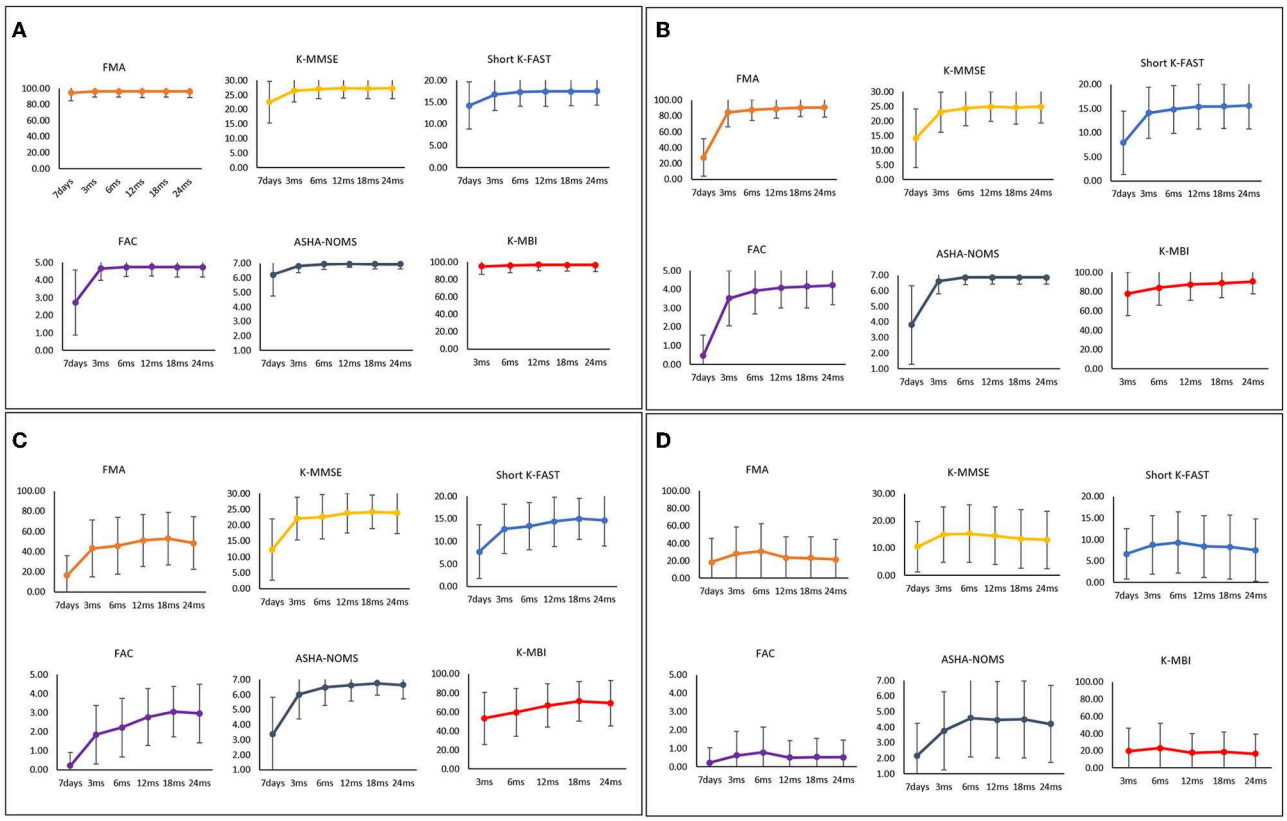

FIGURE 4
Functional recovery patterns of clusters of hemorrhagic stroke patients until 24 months after onset. **(A)** Cluster 1 ($n = 710$); minimal functional deficit in all domains, **(B)** Cluster 2 ($n = 208$); early rapid recovery of all functions, **(C)** Cluster 3 ($n = 128$); significant improvement in all domains over 24 months after stroke with lower motor and ambulatory function, **(D)** Cluster 4 ($n = 100$); severe dysfunctions in all domains. FMA, Fugl-Meyer Assessment; FAC, Functional Ambulatory Category; K-MMSE, Korean Mini-Mental State Examination; Short K-FAST, Short Korean version of the Frenchay Aphasia Screening Test; AHSA-NOMS, American Speech-Language-Hearing Association National Outcome Measurement System Swallowing Scale; K-MBI, Korean modified Barthel Index.

showed only 370 articles using machine learning in 2007; however, the number increased to 3,978 articles in 2017 (29). Machine learning has advantages in personalized medicine, handling large data sets, and design of prediction models. Indeed, multiple studies have attempted to predict prognoses after stroke using methods that showed relatively acceptable levels of accuracy (3, 30, 31). Among them, the Predicting Recovery Potential (PREP) (32) and Time to Walking Independently After Stroke (TWIST) (33) algorithms used decision trees, which is a machine learning method, to predict upper limb or walking abilities. The PREP algorithm for prognosis of upper limb motor recovery had a positive predictive power of 88%, specificity of 88%, and sensitivity of 73%. The TWIST algorithm for prognosis of independent gait at 3 months after stroke showed prediction accuracy for 95% of patients. Scrutinio et al. (34) compared three tree-based machine learning algorithms to predict whether a patient who suffered a severe stroke would be dead or alive 3 years later. The machine learning model that showed the highest performance score had an area under the curve (AUC) of 0.928 and an accuracy of 86.1%. Another recent study used five types of machine learning algorithms to predict favorable outcomes (modified Rankin Scale 0 or 1) for acute IS patients at 3 months (35) and revealed that all five algorithms had an AUC >0.8. All those studies suggest the possibilities and usefulness of machine learning in clinical medicine.

The KOSCO data are characterized by multi-time point longitudinal and multivariate assessments. A previous study using KOSCO data suggested that long-term functional recovery patterns varied by patient baseline characteristics (36). This indicates that it would be difficult to predict prognosis using only fragmentary information. Moreover, our primary aim was not just to predict a binary classification at a specific time point, but to predict functional recovery patterns of six domains over 24 months.

The KOSCO data indicated that 60.46% of IS patients (cluster 1) and 61.95% of HS patients (cluster 1) showed minimal dysfunction after stroke. Another 11.51% of IS patients (clusters 2 and 3) and 29.32% of HS patients (clusters 2 and 3) showed significant improvement over 24 months, reaching near full or significant recovery from their dysfunction, though some exceptions showed unsatisfactory recovery. In addition, 201 (4.58%) IS patients (cluster 5) and 100 (8.73%) HS patients (cluster 4) showed severe dysfunction from the onset of stroke. Another 3.69% of IS patients (cluster 4) showed a decline in functional outcomes during the late phase of follow-up. Overall, clusters with older patients presented

TABLE 2 Performance scores of the prediction models.

Ischemic stroke					Hemorrhagic stroke				
Model	Accuracy	F1[a]	Precision[a]	Recall[a]	Model	Accuracy	F1[a]	Precision[a]	Recall[a]
CatBoost	0.926	0.751	0.780	0.737	CatBoost	0.887	0.778	0.796	0.771
Light GBM-XT	0.925	0.747	0.789	0.731	Light GBM-XT	0.887	0.793	0.802	0.789
XGBoost	0.920	0.728	0.809	0.708	Light GBM	0.887	0.786	0.795	0.785
Random Forest	0.919	0.722	0.764	0.711	XGBoost	0.883	0.779	0.837	0.765
Light GBM	0.917	0.725	0.764	0.718	Weighted Ensemble	0.883	0.775	0.799	0.765
Weighted Ensemble	0.917	0.725	0.764	0.718	Random Forest	0.883	0.77	0.786	0.765
Neural Network	0.912	0.714	0.809	0.699	Neural Network	0.870	0.749	0.750	0.749
Extra Trees	0.912	0.702	0.757	0.683	Extra Trees	0.861	0.735	0.746	0.732

[a]Values are presented as mean values for every cluster.
Light GBM, Light Gradient Boosting Machine; Light GBM-XT, extended version of Light GBM; XGBoost, extreme gradient boosting.

worse outcomes than those of younger patients. The reason for the larger number of clusters of poor prognoses in IS (cluster 4 and 5) than HS (cluster 4) likely is not only the larger number of patients, but also their older mean age.

For both stroke types, CatBoost and Light GBM-XT were the most suitable of the eight machine learning algorithm prediction models for the KOSCO data. Whereas previous studies targeted only one or two time points and used few types of functional outcomes in their predictions, our study provides stronger evidence. We used six functional assessments to characterize motor, mobility, cognitive, language, and swallowing functions and ADL independence at five to six time points. Therefore, our study can describe patients in a more detailed and accurate manner than previous studies and shows time-course changes. Although we analyzed high-dimensional clinical data from a large, long-term cohort of patients, the accuracy of our best prediction model was 0.926 for IS and 0.887 for HS.

Limitations and conclusions

This study has some limitations. First, because the KOSCO dataset contains long-term repetitive assessment data, some data were missing and were handled by statistical methods. First, because the KOSCO dataset contains long-term repetitive assessment data, some data were missing. All subjects included in this study were followed for up to 24 months after stroke, but some cases had missing data at some time point. In such cases, we imputed data using the k-NN5 method as described in the Method section. Second, we analyzed data only for those who survived at 24 months after stroke. Therefore, the subjects of this study may not represent all stroke patients in Korea. Finally, imaging biomarkers such as dynamic nomogram (37) or diffusion tensor image and functional MRI (38) also are useful predictors. However, the KOSCO study did not include such imaging biomarkers because there were limitations in the study design as a multicenter national study with a large number of subjects and many time points assessment.

Despite the above limitations, this study successfully clustered long-term functional recovery patterns in IS and HS patients using machine learning. Machine learning algorithms are increasing in efficacy and overcoming their limitations. Early identification and accurate prediction of long-term functional outcomes using machine learning will help clinicians to develop customized management strategies for stroke patients.

Data availability statement

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Ethics statement

The studies involving human participants were reviewed and approved by Samsung Medical Center, 2012-06-016 Severance Hospital, 4-2012-0341 Konkuk University Medical Center, 1180-01-700 Chungnam National University Hospital, 2012-06-011 Chonnam National University Hospital, CNUH-2012-127 Pusan National University Yangsan Hospital, 05-2012-057 Kyungpook National University Hospital, 2013-03-029 Wonkwang University Hospital, and 1515 Jeju National University Hospital, 2013-02-001. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Y-HK had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Conceptualization, project administration, and funding acquisition: Y-HK. Data curation and Investigation: Y-HK, SS, WHC, MKS, JL, DYK, Y-IS, G-JO, Y-SL, MCJ, SL, M-KS, JH, JA, and Y-TK. Methodology, writing—review, and editing: Y-HK, SS, WHC, MKS, JL, DYK, Y-IS, G-JO, Y-SL, MCJ, SL, M-KS, JH, JA, Y-TK, and KK. Formal analysis: SS, JH, and KK. Writing—original draft: SS. Resources: Y-HK, WHC, MKS, JL, DYK, Y-IS, G-JO, Y-SL, MCJ, SL, M-KS, JH, JA, and Y-TK. Supervision: Y-HK, MKS, JJ, DYK, Y-IS, G-JO, Y-SL, MCJ, SL, M-KS, JH, JA, Y-TK, and KK. All

authors have read and approved the final manuscript and revised and approved the manuscript.

Funding

This study was supported by the Research Program funded by the Korea Disease Control and Prevention Agency (2022-11-006) and by a Korea Medical Device Development Fund grant of the Korean Government (Ministry of Science and ICT, Ministry of Trade, Industry, and Energy, Ministry of Health and Welfare, and Ministry of Food and Drug Safety) (KMDF-RS-2022-00140478).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2023.1130236/full#supplementary-material

References

1. Mukherjee D, Patil CG. Epidemiology and the global burden of stroke. *World Neurosurg.* (2011) 76:S85–90. doi: 10.1016/j.wneu.2011.07.023

2. Feigin VL, Forouzanfar MH, Krishnamurthi R, Mensah GA, Connor M, Bennett DA, et al. Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. *Lancet.* (2014) 383:245–54. doi: 10.1016/S0140-6736(13)61953-4

3. Stinear CM, Smith MC, Byblow WD. Prediction tools for stroke rehabilitation. *Stroke.* (2019) 50:3314–22. doi: 10.1161/STROKEAHA.119.025696

4. Douiri A, Grace J, Sarker SJ, Tilling K, McKevitt C, Wolfe CD, et al. Patient-specific prediction of functional recovery after stroke. *Int J Stroke.* (2017) 12:539–48. doi: 10.1177/1747493017706241

5. Huang ZX, Lin XL, Lu HK, Liang XY, Fan LJ, Liu XT. Lifestyles correlate with stroke recurrence in Chinese inpatients with first-ever acute ischemic stroke. *J Neurol.* (2019) 266:1194–202. doi: 10.1007/s00415-019-09249-5

6. Chang WH, Sohn MK, Lee J, Kim DY, Lee SG, Shin YI, et al. Korean Stroke Cohort for functioning and rehabilitation (KOSCO): study rationale and protocol of a multi-centre prospective cohort study. *BMC Neurol.* (2015) 15:42. doi: 10.1186/s12883-015-0293-5

7. Finocchi C, Balestrino M, Malfatto L, Mancardi G, Serrati C, Gandolfo C. National Institutes of Health Stroke Scale in patients with primary intracerebral hemorrhage. *Neurol Sci.* (2018) 39:1751–5. doi: 10.1007/s10072-018-3495-y

8. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1 a method for evaluation of physical performance. *Scand J Rehabil Med.* (1975) 7:13–31.

9. Holden MK, Gill KM, Magliozzi MR, Nathan J, Piehl-Baker L. Clinical gait assessment in the neurologically impaired. Reliability and meaningfulness. *Phys Ther.* (1984) 64:35–40. doi: 10.1093/ptj/64.1.35

10. Kang Y, Na DL, Hahn S. A validity study on the Korean Mini-Mental State Examination (K-MMSE) in dementia patients. *J Korean Neurol Assoc.* (1997) 15:300–8.

11. Pyun SB, Hwang YM, Ha JW Yi H, Park KW, Nam K. Standardization of Korean version of Frenchay Aphasia screening test in normal adults. *J Korean Acad Rehabil Med.* (2009) 33:436–40.

12. Wesling M, Brady S, Jensen M, Nickell M, Statkus D, Escobar N. Dysphagia outcomes in patients with brain tumors undergoing inpatient rehabilitation. *Dysphagia.* (2003) 18:203–10. doi: 10.1007/s00455-002-0098-8

13. Jung HY, Park BK, Shin HS, Kang YK, Pyun SB, Paik NJ, et al. Development of the Korean Version of Modified Barthel Index (K-MBI): Multi-center Study for Subjects with Stroke. *J Korean Acad Rehabil Med.* (2007) 31:283–97.

14. Arthur D, Vassilvitskii S. k-means++: the advantages of careful seeding. In: *SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms January*. Society for Industrial and Applied Mathematics, United States (2007) p. 1027–35.

15. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. *J Mach Learn Res.* (2011) 12:2825–30.

16. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math.* (1987) 20:53–65. doi: 10.1016/0377-0427(87)90125-7

17. Ogbuabor G, Ugwoke FN. Clustering algorithm for a healthcare dataset using silhouette score value. *Int J Comput Sci Inf Technol.* (2018) 10:27–37. doi: 10.5121/ijcsit.2018.10203

18. Davies DL, Bouldin DW. A cluster separation measure. *IEEE Trans Pattern Anal Mach Intell.* (1979) 1:224–7. doi: 10.1109/TPAMI.1979.4766909

19. Caliński T, Harabasz J. A dendrite method for cluster analysis. *Commun Stat Theory Methods.* (1974) 3:1–27. doi: 10.1080/03610927408827101

20. Kenyhercz MW, Passalacqua NV. Missing data imputation methods and their performance with biodistance analyses. In: Pilloud MA, Hefner JT, editors. *Biological Distance Analysis: Forensic and Bioarchaeological Perspectives*. Amsterdam: Elsevier (2016) p. 181–94. doi: 10.1016/B978-0-12-801966-5.00009-3

21. Van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res.* (2008) 9:2579–605.

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 (NIPS 2017)*. Available online at: https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf. (accessed October 1, 2022).

23. Breiman L. Random forests. *Mach Learn.* (2001) 45:5–32. doi: 10.1023/A:1010933404324

24. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. In: *NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Red Hook, NY: Curran Associates Inc., (2017) p. 6639–49.

25. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining.* (2016):785-94. doi: 10.1145/2939672.2939785

26. Caruana R, Niculescu-Mizil A, Crew G, Ksikes A. Ensemble selection from libraries of models. In: *Proceedings of the Twenty-First International Conference on Machine Learning*. (2004) p. 18. doi: 10.1145/1015330.1015432

27. Erickson N, Mueller J, Shirkov A, Zhang H, Larroy P, Li M, et al. *Autogluon-Tabular: Robust and Accurate Automl for structured data.* (2020). Available online at: https://arxiv.org/abs/2003.06505 (accessed October 1, 2022).

28. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn.* (2006) 63:3–42. doi: 10.1007/s10994-006-6226-1

29. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med.* (2018) 284:603–19. doi: 10.1111/joim.12822

30. Matsumoto K, Nohara Y, Soejima H, Yonehara T, Nakashima N, Kamouchi M. Stroke prognostic scores and data-driven prediction of clinical outcomes after acute ischemic stroke. *Stroke.* (2020) 51:1477–83. doi: 10.1161/STROKEAHA.119.027300

31. Fahey M, Crayton E, Wolfe C, Douiri A. Clinical prediction models for mortality and functional outcome following ischemic stroke: a systematic review and meta-analysis. *PLoS One.* (2018) 13:e0185402. doi: 10.1371/journal.pone.0185402

32. Stinear CM, Barber PA, Petoe M, Anwar S, Byblow WD. The PREP algorithm predicts potential for upper limb recovery after stroke. *Brain.* (2012) 135:2527–35. doi: 10.1093/brain/aws146

33. Smith MC, Barber PA, Stinear CM. The TWIST algorithm predicts time to walking independently after stroke. *Neurorehabil Neural Repair.* (2017) 31:955–64. doi: 10.1177/1545968317736820

34. Scrutinio D, Ricciardi C, Donisi L, Losavio E, Battista P, Guida P, et al. Machine learning to predict mortality after rehabilitation among patients with severe stroke. *Sci Rep.* (2020) 10:20127. doi: 10.1038/s41598-020-77243-3

35. Park D, Jeong E, Kim H, Pyun HW, Kim H, Choi YJ, et al. Machine learning-based three-month outcome prediction in acute ischemic stroke: a single cerebrovascular-specialty hospital study in South Korea. *Diagnostics (Basel).* (2021) 11:1909. doi: 10.3390/diagnostics11101909

36. Shin S, Lee Y, Chang WH, Sohn MK, Lee J, Kim DY, et al. Multifaceted assessment of functional outcomes in survivors of first-time stroke. *JAMA Netw Open.* (2022) 5:e2233094. doi: 10.1001/jamanetworkopen.2022.33094

37. Huang ZX Li YK, Li SZ, Huang XJ, Chen Y, Hong QL, et al. A Dynamic nomogram for 3-month prognosis for acute ischemic stroke patients after endovascular therapy: a pooled analysis in Southern China. *Front Aging Neurosci.* (2021) 13:796434. doi: 10.3389/fnagi.2021.796434

38. Lee J, Kim H, Kim J, Chang WH, Kim Y-H. Multimodal imaging biomarker-based model using stratification strategies for predicting upper extremity motor recovery in severe stroke patients. *Neurorehabil Neural Repair.* (2022) 36:217–26. doi: 10.1177/15459683211070278