scientific reports

OPEN

Machine learning predicts clinically significant health related quality of life improvement after sensorimotor rehabilitation interventions in chronic stroke

Wan-Wen Liao[1], Yu-Wei Hsieh[2,3,4], Tsong-Hai Lee[5,6], Chia-ling Chen[4,7] & Ching-yi Wu[2,3,4✉]

Health related quality of life (HRQOL) reflects individuals perceived of wellness in health domains and is often deteriorated after stroke. Precise prediction of HRQOL changes after rehabilitation interventions is critical for optimizing stroke rehabilitation efficiency and efficacy. Machine learning (ML) has become a promising outcome prediction approach because of its high accuracy and easiness to use. Incorporating ML models into rehabilitation practice may facilitate efficient and accurate clinical decision making. Therefore, this study aimed to determine if ML algorithms could accurately predict clinically significant HRQOL improvements after stroke sensorimotor rehabilitation interventions and identify important predictors. Five ML algorithms including the random forest (RF), k-nearest neighbors (KNN), artificial neural network, support vector machine and logistic regression were used. Datasets from 132 people with chronic stroke were included. The Stroke Impact Scale was used for assessing multi-dimensional and global self-perceived HRQOL. Potential predictors included personal characteristics and baseline cognitive/motor/sensory/functional/HRQOL attributes. Data were divided into training and test sets. Tenfold cross-validation procedure with the training data set was used for developing models. The test set was used for determining model performance. Results revealed that RF was effective at predicting multidimensional HRQOL (accuracy: 85%; area under the receiver operating characteristic curve, AUC-ROC: 0.86) and global perceived recovery (accuracy: 80%; AUC-ROC: 0.75), and KNN was effective at predicting global perceived recovery (accuracy: 82.5%; AUC-ROC: 0.76). Age/gender, baseline HRQOL, wrist/hand muscle function, arm movement efficiency and sensory function were identified as crucial predictors. Our study indicated that RF and KNN outperformed the other three models on predicting HRQOL recovery after sensorimotor rehabilitation in stroke patients and could be considered for future clinical application.

Health related quality of life (HRQOL) refers to the way an individual feels and reacts to his/her health status affected by medical conditions[1]. Compared to quality of life that covers all aspects of well-beings of human life, HRQOL focuses more on well-beings related to health domains such as physical, functional and mental health and it has been regarded as an important outcome of treatments[1,2].

Stroke remains a leading cause of long-term disability[3]. It has a wide-ranging impact not only on physical and daily function but also on HRQOL[4]. Most patients still suffered from deteriorated HRQOL even in the chronic phase of stroke, which makes HRQOL an crucial target for stroke rehabilitation[4,5]. To improve patients' HRQOL, healthcare professionals have to provide rehabilitation interventions that are most effective for each

[1]Department of Gerontological Health Care, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan. [2]Department of Occupational Therapy and Graduate Institute of Behavioral Sciences, College of Medicine, Chang Gung University, No. 259, Wenhua 1st Rd., Taoyuan, Taiwan. [3]Healthy Aging Research Center, Chang Gung University, No. 259, Wenhua 1st Rd., Taoyuan, Taiwan. [4]Department of Physical Medicine and Rehabilitation, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan. [5]Department of Neurology, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan. [6]College of Medicine, Chang Gung University, Taoyuan, Taiwan. [7]Graduate Institute of Early Intervention, College of Medicine, Chang Gung University, Taoyuan, Taiwan. ✉email: cywu@mail.cgu.edu.tw

patient based on his/her responses to that rehabilitation therapy. Building accurate prediction models for forecasting patients' HRQOL improvements after rehabilitation interventions and identifying predictors relevant for HRQOL improvements in stroke patients are thus imperative for providing insights to healthcare professionals on making accurate clinical decision.

Machine learning (ML) has become a popular prediction analytic approach. Machine learning uses automatic computerized algorithms to discover patterns in the data and builds prediction models to forecast future events. Machine learning is particularly suitable for predicting health outcomes because it can process large volumes of data, analyze the complex relationship between various different features/variables and easily incorporate new variables into prediction models without re-adjusting the preprogrammed rules[6]. In addition, the feature selection procedure can be incorporated into machine learning procedures to help identify important predictors[7]. These advantages make machine learning a potentially ideal tool for realizing accurate outcome prediction in patient populations.

In stroke, machine learning has been primarily used for predicting motor and activities of daily living (ADL) recovery and has achieved an overall positive result[8–12]. However, to our knowledge, only one study to date has applied machine learning algorithms in predicting stroke-specific HRQOL recovery[13]. In that study, the authors incorporated six demographic factors into machine learning models and built a preliminary system to forecast HRQOL changes of chronic stroke patients. Small prediction errors (i.e., the root mean square errors) were found between the data derived from the prediction model and the actual data collected from the patient, suggesting that machine learning might be feasible for predicting HRQOL changes in chronic stroke patients[13].

Despite this positive evidence, the previous study only included demographic attributes into the machine learning prediction model[13]; nevertheless, HRQOL has been shown to be affected by factors across multiple domains including demographic as well as health-related domains such as physical and functional domains[4,5,14]. Including only demographic attributes in the machine learning model may not be sufficient for optimizing prediction accuracy. In addition, the previous study only examined prediction errors (e.g., the mean squared error) of the machine learning model[13]. Important clinical performance metrics such as prediction accuracy and the ability of machine learning models to distinguish between responders and non-responders to rehabilitation interventions remain largely unexplored[15]. A comprehensive examination of machine learning prediction performance along with factors across health domains is required for determining the efficacy of machine learning on predicting HRQOL recovery of stroke patients after rehabilitation interventions.

Stroke sensorimotor rehabilitation interventions including the robot-assisted therapy (RT), mirror therapy (MT) and transcranial direct current stimulation (tDCS) have become popular approaches for improving stroke recovery in the recent decade. These three approaches (i.e., RT, MT and tDCS) use modern equipment/modalities (e.g., robotic arms, mirror boxes and electrical stimulators) to modulate peripheral and/or central sensorimotor systems (e.g., visuomotor and sensorimotor systems and cortical areas) to augment stroke recovery[16–18]. Several studies have demonstrated that these three sensorimotor interventions (i.e., RT, MT and tDCS) not only facilitated functional recovery but also improved participation and HRQOL in stroke patients[19–25]. The rationale of why these three sensorimotor interventions (i.e., RT, MT and tDCS) could improve HRQOL is that these interventions could reduce arm/hand impairment, restore arm/hand function, which would allow stroke patients to participate in daily activities and accomplish essential daily tasks[19–25]. Most daily tasks such as bathing, dressing, dining and grocery shopping all involve use of the arm/hand to manipulate objects to accomplish tasks. Good arm/hand function would lead to successful participation in daily tasks and subsequently may increase stroke patients' subjective feeling of well-beings and satisfaction toward daily life[19–25]. Thus, these three interventions (i.e., RT, MT and tDCS) may have potentials to be incorporated into current clinical practice to facilitate not only functional recovery but also HRQOL in stroke patients. Machine learning may be a potentially useful tool for predicting HRQOL changes after these three interventions, which may help identify responders to these three interventions and facilitate clinical application[6,7].

Therefore, the purpose of this study was to determine the performance of machine learning algorithms on predicting clinically significant HRQOL improvements of chronic stroke patients after stroke sensorimotor rehabilitation interventions including the RT, MT and tDCS. We examined the performance of five commonly used machine learning algorithms and identified important predictors for building machine learning prediction models.

Methods

Study design.
This study was an observational cohort study that used secondary analysis of data from our previous randomized controlled or cluster-controlled trials and ongoing projects[24,26–28]. Data screening was done by three investigators (Liao WW, Wu CY and Hsieh YW). The three investigators determined the eligibility and completeness of the data. Patients that completed the interventions and outcome measurements at pre- and post-intervention were included for analysis.

Participants.
One hundred and thirty-two chronic stroke patients (N = 132) were included. Participants were recruited from three hospitals in the northern part of Taiwan. Table 1 outlines the characteristics of participants. The inclusion criteria were (1) a first-ever unilateral ischemic or hemorrhagic stroke, (2) more than 6 months post stroke, (3) Fugl-Meyer assessment scale of upper extremity (FMA) scores between 18 and 60, suggesting mild to moderate arm hemiparesis[29], (4) no excessive spasticity in upper limb joints (Modified Ashworth Scale, MAS ≤ 3)[30], (5) ability to follow study instructions (Mini-Mental State Examination ≥ 22), and (6) no concomitant neurological disorders (e.g., brain tumor and dementia). The exclusion criteria were (1) participation in any drug or rehabilitation projects/experiments in the past 6 months, (2) had Botulinum toxin injections in the past 3 months, (3) severe vision or visual perception impairments (e.g., neglect and poor visual field) as

Baseline variables	Participants (N = 132)
Age (years)	55.28 ± 11.66
Gender (male/female)	98/34
Side of lesion (right/left)	73/59
Time since stroke (months)	26.89 ± 22.81
Education (years)	11.29 ± 4.68
MOCA	24.94 ± 4.27
FMA total scores	34.75 ± 9.58
MAS mean scores	0.35 ± 0.6
MRC mean scores	3.83 ± 1.26
WMFT-TIME mean time (s)	11.12 ± 5.58
WMFT-FAS mean scores	2.66 ± 0.6
MAL-AOU mean scores	1.27 ± 0.8
MAL-QOM mean scores	0.91 ± 0.71
BBT-paretic side	9.45 ± 11.3
RNSA-tactile	74.71 ± 28.45
RNSA-proprioception	16.74 ± 5.25
RNSA-stereognosis	13.36 ± 8.65
FIM total scores	111.6 ± 9.63
SIS mean score (%)	65.71 ± 9.94
SIS recovery (%)	50.33 ± 16.48

Table 1. Clinical characteristics of participants. *MOCA* Montreal Cognitive Assessment assessing global cognitive function (total scores:30), *FMA* Fugl-Meyer Assessment Scale of Upper Extremity assessing motor impairment of the upper extremity (total scores:66), *MAS* Modified Ashworth Scale assessing muscle tone level of the upper extremity (item score range:0–4, the MAS mean scores were the average scores of all parts of upper extremity), *MRC* Medical Research Council Scale for muscle strength assessing muscle strength of the upper extremity (item score range:0–5, the MRC mean scores were the average scores of all parts of upper extremity), *WMFT-TIME* Wolf Motor Function Test TIME-representing movement efficiency of the upper extremity (WMFT-TIME mean scores are the average time of all test items, Unit = seconds), *WMFT-FAS* Wolf Motor Function Test-functional ability scale representing motor function of the upper extremity (item score range 0–5; WMFT-FAS mean scores are the average scores of all test items), *MAL AOU* motor activity log-amount of use representing participants' self-perceived amount of use of the paretic arm (item score range:0–5; MAL-AOU mean scores are the average scores of all test item), *MAL QOM* motor activity log-quality of movement representing participants' self-perceived quality of paretic arm movements (item score range:0–5; MAL-AOU mean scores are the average scores of all test item), *BBT-Paretic* Box and Block Test-Paretic representing the paretic hand function (total scores = 150 (150 cubes), *RNSA* Revised Nottingham Sensation Assessment assessing the tactile, proprioception, and stereognosis sensation of the paretic side of the body (item score range: 0–2; unable to test = 9), *FIM* functional independence measure assessing participants' levels of disability (item score range: 1–7; total score:18–126), *SIS* stroke impact scale. Value is mean ± standard deviation.

assessed by the National Institutes of Health Stroke Subscale, and (4) any contradictions to non-invasive brain stimulation (for participants receiving tDCS)[31]. The institutional review boards of participating hospitals including the Linkou Chang Gung Memorial Hospital and Taipei Tzu Chi Hospital approved the trials. All participants provided written informed consents before enrolled into clinical trials. All study procedures were conducted in accordance with the Declaration of Helsinki.

Stroke sensorimotor rehabilitation interventions. All participants received interventions for 1.5 to 2 h per session with a total of approximately 30 h of training across 3 to 4 weeks. The frequency and duration of training were similar to those of most rehabilitation interventions studies[19–25]. Participants received interventions at the hospitals where they were recruited from. The interventions were administered by certified occupational therapists that were properly trained by the senior therapists and the principal investigators (Wu CY). Among these participants, 70 received RT, 32 received MT and 30 received tDCS/MT.

For the RT, participants practiced unilateral paretic movements by using the InMotion robotic systems (the InMotion ARM and InMotion WRIST)[28]. For the MT, participants imagined that the mirror reflection of the non-paretic arm was the paretic arm and performed bilateral movements as simultaneously as possible[26,27]. For the tDCS/MT, participants received 2 mA anodal tDCS on the ipsilesional primary motor cortex for 20 min followed by another 20 min of MT[24]. For all trainings (i.e., RT, MT and tDCS/MT), participants performed an additional 15–30 min of functional task training in each session. Participants were assessed within 1 week before and after interventions by the evaluators that were blinded to the study purpose and treatment allocation of participants.

Classification of HRQOL improvement. The Stroke Impact Scale (SIS) 3.0 was selected as the major outcome for classifying HRQOL improvements[32]. It is a self-reported questionnaire used for evaluating HRQOL in stroke patients. The reliability and validity of SIS have been well established[33,34]. The SIS consists of two parts. The first part is the main scale of SIS that assesses multidimensional HRQOL including the strength, hand function, ADL/instrumental ADL, mobility, communication, emotion, memory/thinking, and social participation[32]. It has 59 items and each item was scored subjectively by stroke patients based on the difficulty they perceived in that item during the past two weeks. The scores of each domain were transformed to a score out of 100 and the mean scores of all domains were used to represent multidimensional HRQOL of stroke patients[35,36]. The second part is a global rating scale that evaluates stroke patients' self-perceived HRQOL recovery. The scores range from 0 (no recovery) to 100 (full recovery). Both parts were included in this study to comprehensively represent multi-dimensional HRQOL and global self-perceived HRQOL changes of chronic stroke patients.

To facilitate clinical use of our ML prediction models, the minimal clinical important differences (MCID) were selected as the criterion to classify participants into high and low responders. The MCID is the smallest change in scores that were considered clinically important and meaningful in health status perceived by the patient[37]. Previous studies have set the MCID as 10–15% of total scores in patient populations and received clinically beneficial results[38–40]. As a result, based on the literatures, we defined the MCID as 10% of changes in the scores of main SIS scale and global rating scale. Participants that had SIS change scores greater than or equal to 10 were classified as high responders and participants with SIS change scores less than 10 were classified as low responders to stroke sensorimotor rehabilitation interventions.

Candidate predictors. We selected thirty-two potential predictors based on stroke HRQOL literatures and the International Classification of Functioning, Disability and Health (ICF) framework to include "Body function and structures", "Activity" and "Participation" attributes[41–43]. These predictors included (1) personal characteristic attributes: age, gender, education, time since stroke and side of hemiplegia, (2) baseline cognitive and motor function attributes: The Montreal Cognitive Assessment (MOCA) scores[44], FMA proximal/distal/total scores[45], Wolf Motor Function Test (time and functional ability scale) scores[46], Medical Research Council (MRC) muscle strength scale scores (paretic shoulder, elbow, wrist, and finger muscle strength scores and MRC mean scores)[47], MAS scores (paretic shoulder, elbow, forearm, wrist, finger scores and MAS mean scores)[47], Motor Activity Log (MAL) amount of use (AOU) and quality of movement (QOM) scores[48] and Box and Block test scores[49], (3) baseline sensory function attributes: the revised Nottingham Sensory Assessment (RNSA) scores (paretic side tactile sensation, proprioception and stereognosis)[50], (4) baseline ADL and instrumental ADL attributes: Functional Independence Measure (FIM) scores[51] and the Nottingham Extended Activities of Daily Living (NEADL) scale scores[52], and (5) baseline HRQOL attributes: the SIS mean scores and global rating scores[32]. These attributes are commonly used in research and clinical settings to represent the motor/sensory impairment, functional ability and participation of stroke patients, and therefore were selected as potential predictors[41–43].

Machine learning algorithms. Five ML algorithms, which were the random forest (RF), k-nearest neighbors (KNN), artificial neural network (ANN), support vector machine (SVM), and logistic regression (LG) were used for developing prediction models. The RF uses the ensemble learning method for outcome prediction. It combines the results of multiple decision trees and generates a final overall result to augment prediction accuracy[53]. It is a flexible method that can be used in categorical and continuous data. In addition, the RF has a low probability of overfitting and is therefore suitable for use in the clinical setting[53]. The KNN is a distance-based method. It predicts that similar objects would exist in close proximity. As a result, it labels the class of the target based on the majority of classes of its surrounding k neighbors[54]. The KNN classification pattern is similar to the clinical decision-making process made by the clinicians/therapists, where similar treatments would be prescribed to patients with similar responses and characteristics[55]. The ANN is inspired by the neurological network of the brain[56]. It consists of several neurons/nodes in layers including the input, hidden and output layers. The input layer receives the data, and transfers it to the hidden layer, where the computations (e.g. the activation function) are primarily taken places. After the computations are done, the hidden layer generates the final output to the output layer and finishes the prediction model. The ANN can process complex health informatics data and is therefore a potentially useful tool for outcome prediction in stroke patients[57]. The feedforward back propagation method was used in this study. The SVM uses the binary classification technique, where it projects data onto a high-dimensional plane by using the kernel function first, and then finds the maximum-margin hyperplane that best separates data into two classes[58]. The SVM is efficient in high dimensional planes and suitable for modeling complicated medical data. The LG is a binary classification technique that uses the logistic sigmoid function to predict the probability of observed data that would belong to one of the two possible classes[59]. It is a commonly used algorithm for building prediction models in stroke patients.

These 5 ML algorithms were selected because they are widely used modeling techniques for outcome prediction in patients and have been shown to have good prediction performance[6].

Feature selection procedure. The feature selection procedure was performed to remove redundant attributes and identify the essential ones for prediction accuracy[7]. A popular feature selection method called "information gain ratio" was implemented[11]. This feature selection method evaluates the influence (i.e., the information gain) of each attribute to the output classes (i.e., the SIS classes) using the ranker search method[60–62]. A higher gain ratio of the attribute indicates a greater contribution of this attribute to prediction accuracy[62,63]. In this study, attributes with gain ratio greater than zero were used for developing ML prediction models.
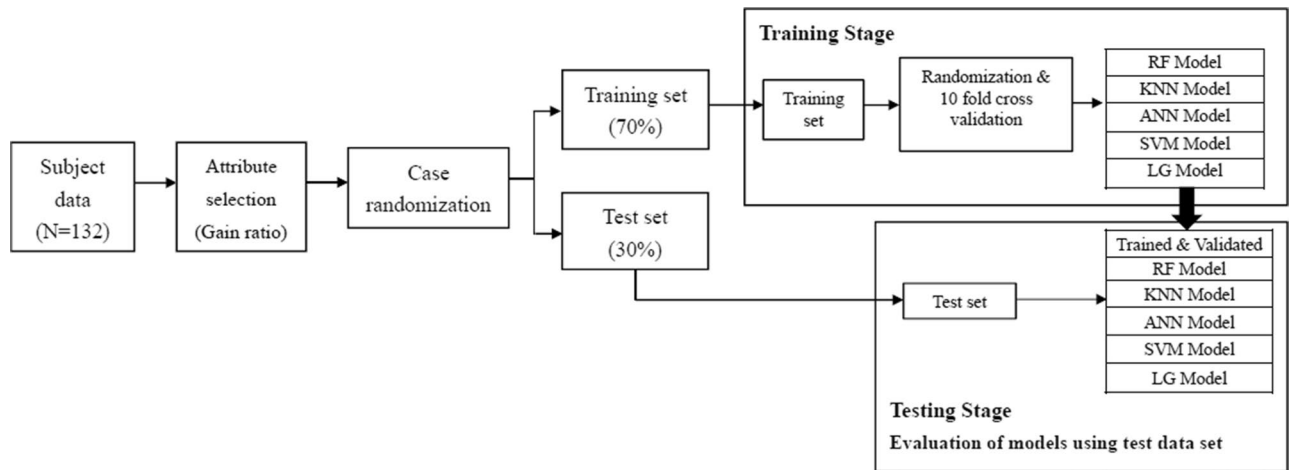
Figure 1. The flow chart of model development and validation process. Subject data were randomized into a training set and a test set. The training set was 70% of the data and the test set was 30% of the data. For the training data set, the tenfold cross validation procedure was used to train and build 5 machine learning models (i.e., the RF, KNN, ANN, SVM and LG) in which the data was randomly split into 10 groups (9 groups for training and 1 group for validation). The tenfold cross validation process repeated until all 10 groups of data were trained and validated. The tenfold cross validation process was performed for all 5 machine learning models. After the 5 models were built, the test data set was entered into the 5 models to determine the model performance.

Model development and testing. Figure 1 shows the model development and testing process. Data were randomized and divided into a training data set (70%) and a test data set (30%)[64]. The training data set was used for training and developing the model. The test data set was used for final evaluation of model performance. The tenfold cross validation procedure was performed to train the models[65]. During the tenfold cross validation process, the training data set was split into 10 groups, where 9 of them were used for training the model while the remaining one was used for validating the model. This process was repeated until all groups of data had been trained and validated. After the model was built, the testing data set was entered into the model to assess the model performance.

The hyper-parameters of the prediction models were determined according to the procedures used in the ML literatures. For the RF model, the numbers of trees to build were 100 and the numbers of features to consider at a node were the first integer less than $\log_2 M + 1$ (M is the number of inputs)[66]. For the KNN and ANN models, the tenfold cross validation was employed to tune the value of hyper-parameters (i.e., the k value of the KNN; the numbers of hidden neurons in the hidden layer of the ANN)[10]. We found that k = 5 and the numbers of hidden neurons = 3 in one hidden layer (the main SIS scale) and k = 9 and the numbers of hidden neurons = 2 in one hidden layer (the SIS global rating scale) had the best prediction accuracy. As a result, these hyper-parameters were used in the KNN and ANN models. For the SVM model, the polynomial kernel function was employed because it provided the best prediction accuracy[58,59].

Model performance metrics. The performance of ML models was evaluated using the standard ML performance metrics including (1) accuracy, (2) recall, (3) precision, (4) F1 scores, and (5) area under the receiver operating characteristic curve (AUC-ROC)[15]. Accuracy is an overall index of prediction performance. Accuracy was computed as the sum of true positive (TP) and true negative (TN) divided by the sum of TP, TN, false positive (FP) and false negative (FN). Recall is the ratio of participants that were correctly identified as positive by the model to those whom were actually positive. Recall was computed as the TP divided by the sum of TP and FN. Precision is the ratio of participants that were correctly identified as positive by the model to those were labelled as positive by the model. Precision was calculated as TP divided by the sum of TP and FP. F1 scores is a combined index of precision and recall. It was calculated as the harmonic mean of precision and recall. The AUC-ROC is the ratio of area under the ROC curve to the total area. It represents the ability of the model to distinguish between classes.

Statistical analysis. The continuous variables were standardized and the categorical variables were coded before developing the ML models. The Waikato Environment for Knowledge Analysis (Weka) 3.8.3 developed by the University of Waikato, New Zeeland was employed for model development, training and testing[67]. The Weka has been extensively used for constructing ML prediction models in various fields and in different patient populations in the medical field[11,68,69].

Model	Accuracy (%)	Precision	Recall	F1 scores	AUC-ROC
SIS main scale					
RF	85	0.88	0.85	0.85	0.86
KNN	75	0.76	0.75	0.75	0.8
ANN	75	0.77	0.75	0.74	0.87
SVM	72	0.73	0.73	0.73	0.71
LG	72.82	0.73	0.73	0.72	0.77
SIS global rating scale					
RF	80	0.78	0.8	0.78	0.75
KNN	82.5	0.82	0.83	0.81	0.76
ANN	77.5	0.77	0.78	0.77	0.75
SVM	77.5	0.77	0.78	0.77	0.68
LG	77.5	0.78	0.78	0.78	0.75

Table 2. Performance metrics of SIS prediction models. *SIS* Stroke Impact Scale, *RF* random forest, *KNN* k-nearest neighbors, *ANN* artificial neural network, *SVM* support vector machine, *LG* logistic regression, *AUC-ROC* area under the receiver operating characteristic curve.

Results

Five most important attributes were identified by the feature selection procedure for the SIS HRQOL main scale, which were the baseline SIS mean scores (gain ratio = 0.1), baseline MRC finger metacarpophalangeal (MP) extensors (gain ratio = 0.15) and flexors (gain ratio = 0.06) scores, baseline MAS wrist flexors (gain ratio = 0.1) scores and baseline WMFT time (gain ratio = 0.14). The gain ratio of the other 31 attributes was 0. As a result, these 5 attributes were used for developing the SIS multidimensional HRQOL prediction model.

Four most important attributes were identified by the feature selection procedure for the SIS global rating scale, including age (gain ratio = 0.16), gender (gain ratio = 0.06), baseline RNSA stereognosis (gain ratio = 0.1) and proprioception (gain ratio = 0.07) scores. The gain ratio of the other 32 attributes was 0. Therefore, these 4 attributes were used for developing the stroke global self-perceived HRQOL recovery prediction model.

Table 2 summarizes the performance metrics of the five ML models. For the SIS multidimensional HRQOL scale, the prediction performance was the best in the RF model. The accuracy of the RF model was 85%, precision was 0.88, recall was 0.85 the F1 scores were 0.85 and the AUC-ROC was 0.86. The prediction performance was similar between the other 4 models (KNN, ANN, SVM and LG). The accuracy ranged from 72 to 75%, the precision was from 0.73 to 0.77, the recall was from 0.73 to 0.75, the F1 scores were from 0.72 to 0.75 and the AUC-ROC was from 0.71 to 0.87.

For the SIS global rating scale, the prediction performance was the best and similar between the RF and KNN model. The accuracy of the RF model was 80%, the precision was 0.78, the recall was 0.8, the F1 scores were 0.78 and the AUC-ROC was 0.75. The accuracy of KNN model was 82.5%, the precision was 0.82, the recall was 0.83, the F1 scores were 0.81 and the ACU-ROC was 0.76. The prediction performance was similar between the other three models (ANN, SVM and LG). The accuracy was all 77.5%, the precision was from 0.77 to 0.78, the recall was all 0.78, the F1 scores were from 0.77 to 0.78 and the AUC-ROC was from 0.68 to 0.75.

Discussion

Our results demonstrated that machine learning could accurately predict HRQOL improvements after stroke sensorimotor rehabilitation interventions in chronic stroke patients. In particular, the RF and the KNN models had better performance than the other three algorithms (i.e., ANN, SVM and LG). The RF model had 85% accuracy on predicting multidimensional HRQOL changes and 80% accuracy on forecasting global self-perceived recovery. It could also accurately distinguish between high and low responders on the multidimensional HRQOL outcome with 86% chances and on the global self-perceived recovery with 75% chances. The KNN model had good prediction performance on global self-perceived recovery only, where it had 82.5% prediction accuracy and could distinguish between high and low responders with 76% chances. Furthermore, we identified important attributes for predicting multidimensional HRQOL improvements, which were the baseline HRQOL, baseline paretic finger muscle strength, wrist muscle tone and arm movement efficiency, and also key attributes for forecasting global self-perceived recovery including age, gender, baseline hand stereognosis and limb proprioception.

To our knowledge, this study is the first to comprehensively evaluate the performance of ML algorithms on predicting HRQOL improvements after stroke sensorimotor rehabilitation interventions in stroke patients. In addition, we identified the two more effective algorithms, which were the RF and KNN among 5 commonly used ML algorithms. The RF algorithm had good prediction performance on both multidimensional HRQOL and global self-perceived recovery. This good prediction performance may be contributed by the unique ensemble method that it employed in the modeling process. During the ensemble modeling process, the RF creates as many base models (i.e., the decision trees) as many as possible and combines these base models into a final one[66]. Therefore, instead of creating one model and hoping this model would be the best, the RF takes a myriad of multiple models into account to optimize its prediction accuracy. Furthermore, these base models (i.e., the decision trees) are designed to be uncorrelated with each other, thus reducing the probability of overfitting[66]. Our finding of the superior performance of RF was in line with several previous studies demonstrating that ML

algorithms with ensemble methods such as the RF and the Adaboost algorithms outperformed other types of ML algorithms, for example the decision trees or LG[63,70–73]. Future studies could examine the performance of ML algorithms with different types of ensemble methods such as the boosting, bagging and stacking on HRQOL outcome prediction to determine the best one for use in stroke patients.

To our surprise, the KNN algorithm has similar and slightly better prediction performance than the RF algorithm on predicting global self-perceived HRQOL recovery in chronic stroke patients. Compared to the RF, the KNN is a simpler and more straightforward model because it is a distance-based method and does not involve ensemble learning processes[54,55]. Our study revealed that despite its simplicity, the KNN may be as powerful as the RF model when predicting a single-item outcome such as the SIS global self-perceived HRQOL recovery scale in stroke patients. In contrast, the KNN may not be the best algorithm for processing multidimensional health outcome data due to its weaker prediction performance on multidimensional than single item SIS HRQOL outcomes. As a result, we recommend using the KNN model for predicting simple HRQOL outcome changes in chronic stroke patients. Future studies could compare and contrast the performance of KNN on single domain and multidimensional health outcome data to validate findings of this study.

Our study showed that 5 attributes related to initial HRQOL, muscle function (the muscle strength and muscle tone) and movement efficiency were important predictors for forecasting multidimensional HRQOL improvements after sensorimotor rehabilitation interventions. Indeed, studies have found that baseline HRQOL was associated with HRQOL restoration in stroke patients[74,75]. It is thus not surprising to find baseline SIS scores important for HRQOL improvements after interventions. In addition, studies have shown that paretic arm/hand muscle strength and muscle tone significantly affected functional recovery after stroke[43,76,77]. The movement efficiency of the paretic arm was also associated with HRQOL recovery post stroke[5]. Similarly, in the present study, we found that muscle function (i.e., baseline MRC finger MP extensors/flexors, MAS wrist flexors) and the movement efficiency (i.e., baseline WMFT time scores) of the arm associated with prediction of HRQOL improvements. However, compared to previous studies, we further identified the most important components, which were the muscle function of "the wrist/hand" and the movement efficiency of "the whole arm". These components are highly involved in daily routines. For example, stroke patients have to be able to appropriately open/close their hands, grasp/release objects and move their arms efficiently in time to perform most essential daily tasks such as dressing, bathing and cooking. The inability to accomplish essential daily tasks may result in a source of distress and consequently affect participation and life satisfaction after stroke[78]. Therefore, even though muscle function and movement efficiency are commonly regarded as the basic level (i.e., body functions and structures) of the ICF model, they could substantially affect the multi-dimensional HRQOL recovery after stroke sensorimotor interventions and therefore should be considered when assigning these interventions to stroke patients.

In addition, our study also found another four attributes imperative for forecasting global self-perceived HRQOL recovery, which were age, gender and the paretic hand stereognosis and limb proprioception. Age and gender have been demonstrated to be associated with HRQOL recovery post stroke in previous studies[5,79]. Our results support theses previous findings and further suggest that additional somatosensory impairments, especially the sensory components including the hand stereognosis and limb proprioception are also important for predicting global self-perceived HRQOL recovery. Indeed, reduced sensation has been found to be related to slower recovery, decreased motor function (e.g., motor control and activity) and lesser rehabilitation outcomes[43,80,81]. Sensory deficits, such as impaired hand stereognosis and limb proprioception may cause difficulties in sensing arm/hand position and recognizing objects in the hand and hard to control arm/hand movements. This may result in decreased confidence in using the arm/hand in daily activities, fear of safety, thus affecting stroke patients' participation and satisfaction in daily life[80–82]. Our results were also in line with one previous study that found paretic limb proprioception associated with reduced HRQOL and increased feeling of social isolation in stroke patients[14]. Nonetheless, most studies did not assess the contributions of sensory impairments on HRQOL outcome prediction after rehabilitation interventions in stroke patients[83]. Our study suggested that sensory impairment may contribute to global self-perceived HRQOL recovery post stroke intervention and could be considered in the data collection and HRQOL model building process.

Taken together, we found that attributes in the three major categories: the personal characteristics (i.e., age and gender), initial life satisfaction (baseline SIS mean scores) and arm/hand sensorimotor components (i.e., muscle function, movement efficiency, hand stereognosis and limb proprioception) were important predictors for HRQOL restoration after stroke sensorimotor rehabilitation. Based on our findings, healthcare professionals could at least assess these attributes in the three categories before assigning the three sensorimotor interventions to stroke patients. In addition, these attributes may have potential to be used as indicators to determine which stroke patient may be suitable for receiving stroke sensorimotor rehabilitation interventions to improve HRQOL. This may help to improve clinical rehabilitation efficacy and potentially save workload in the hospitals/clinics.

Study limitations. Six limitations should be considered. First, our HRQOL outcome prediction was focused on stroke sensorimotor rehabilitation interventions that share similar rehabilitation principles. Future studies could examine whether the identified predictors, such as the sensorimotor components could generalize to HRQOL outcome prediction in other types of rehabilitation interventions. Second, our predictions were based on changes immediately after interventions. Future studies could investigate the prediction performance of ML algorithms in the follow-up period. This will help identify patients that could retain improvements after sensorimotor interventions. Third, we examined 5 commonly used ML algorithms. Future studies could evaluate the performance of other types of ML or deep learning algorithms such as those with ensemble methods (e.g., the Adaboost)[84] or the deep learning algorithms (e.g., the deep neural network)[85] and compare the results to our findings to determine the optimal method for predicting HRQOL restoration after rehabilitation interventions in

stroke patients. Fourth, this study included thirty-two potential predictors based on the results of previous prediction model studies. Nevertheless, some stroke-related factors were not included, for example, the size/severity of lesion, due to the incapability to retrieve those data from some of our patients. Future studies could examine if these stroke related factors would also be important predictors for forecasting HRQOL recovery in chronic stroke patients, in addition the predictors identified in this study. Fifth, we used SIS to assess multi-dimensional HRQOL and self-perceived global HRQOL changes in stroke patients. Although SIS is a widely used HRQOL assessment in stroke rehabilitation field, it is an ordinal questionnaire that may still have measurement errors. In addition, the SIS may not cover all aspects of HRQOL related factors such as environmental factors (e.g., the context and time), personal factors (e.g., personality) and social indicators (e.g., economic status). We encourage future researchers to include the above factors into machine learning prediction models and examine whether inclusion of these additional factors would optimize prediction accuracy on HRQOL. Sixth, the machine learning models built in this study were preliminary models that focused on predicting HRQOL recovery. Therefore, these models should not be used for excluding patients from receiving the three stroke sensorimotor interventions at this moment. Future studies could include more participants and develop machine learning models that comprehensively predict motor, functional and HRQOL recovery in stroke patients.

Conclusion

Machine learning may predict clinically significant HRQOL improvements after stroke sensorimotor rehabilitation interventions in chronic stroke patients. In particular, the RF and the KNN algorithms may be more effective than the other 3 ML algorithms (DT, LG and SVM) and could be considered for use in clinical settings. We suggest including at least three categories of predictors, which are age/gender, initial HRQOL and sensorimotor components including sensory function, muscle function and movement time into the ML HRQOL prediction model for optimizing prediction accuracy. Future studies with a different sample of stroke patients are warranted to validate our findings and improve model generalizability.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Lin, X.-J., Lin, I. M. & Fan, S.-Y. Methodological issues in measuring health-related quality of life. *Tzu Chi Med. J.* **25**, 8–12 (2013).
2. Guyatt, G. H. *et al.* Exploration of the value of health-related quality-of-life information from clinical research and into clinical practice. *Mayo Clin. Proc.* **82**, 1229–1239 (2007).
3. Virani, S. S. *et al.* Heart disease and stroke statistics 2021 update. *Circulation* **143**, e254–e743 (2021).
4. Carod-Artal, F. J. & Egido, J. A. Quality of life after stroke: The importance of a good recovery. *Cerebrovasc. Dis.* **27**, 204–214 (2009).
5. Nichols-Larsen, D. S., Clark, P. C., Zeringue, A., Greenspan, A. & Blanton, S. Factors influencing stroke survivors' quality of life during subacute recovery. *Stroke* **36**, 1480–1484 (2005).
6. Bzdok, D. & Ioannidis, J. P. A. Exploration, inference, and prediction in neuroscience and biomedicine. *Trends Neurosci.* **42**, 251–262 (2019).
7. Guyon, I. & Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **3**, 1157–1182 (2003).
8. Heo, J. *et al.* Machine learning-based model for prediction of outcomes in acute stroke. *Stroke* **50**, 1263–1265 (2019).
9. Lin, W. Y. *et al.* Predicting post-stroke activities of daily living through a machine learning-based approach on initiating rehabilitation. *Int. J. Med. Inform.* **111**, 159–164 (2018).
10. Sale, P. *et al.* Predicting motor and cognitive improvement through machine learning algorithm in human subject that underwent a rehabilitation treatment in the early stage of stroke. *J Stroke Cerebrovasc. Dis.* **27**, 2962–2972 (2018).
11. Wang, H. L. *et al.* Automatic machine-learning-based outcome prediction in patients with primary intracerebral hemorrhage. *Front. Neurol.* **10**, 910 (2019).
12. Thakkar, H. K., Liao, W. W., Wu, C. Y., Hsieh, Y. W. & Lee, T. H. Predicting clinically significant motor function improvement after contemporary task-oriented interventions using machine learning approaches. *J. Neuroeng. Rehabil.* **17**, 131 (2020).
13. Tokmakçı, M., Ünalan, D., Soyuer, F. & Öztürk, A. The reevaluate statistical results of quality of life in patients with cerebrovascular disease using adaptive network-based fuzzy inference system. *Expert Syst. Appl.* **34**, 958–963 (2008).
14. Morris, J. H., van Wijck, F., Joice, S. & Donaghy, M. Predicting health related quality of life 6 months after stroke: The role of anxiety and upper limb dysfunction. *Disabil. Rehabil.* **35**, 291–299 (2013).
15. Sokolova, M. & Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf Process Manag.* **45**, 427–437 (2009).
16. Turner, D. L., Ramos-Murguialday, A., Birbaumer, N., Hoffmann, U. & Luft, A. Neurophysiology of robot-mediated training and therapy: A perspective for future use in clinical populations. *Front. Neurol.* **4**, 184 (2013).
17. Deconinck, F. J. *et al.* Reflections on mirror therapy: A systematic review of the effect of mirror visual feedback on the brain. *Neurorehabil. Neural Repair* **29**, 349–361 (2015).
18. Schlaug, G., Renga, V. & Nair, D. Transcranial direct current stimulation in stroke recovery. *Arch. Neurol.* **65**, 1571–1576 (2008).
19. Kutner, N. G., Zhang, R., Butler, A. J., Wolf, S. L. & Alberts, J. L. Quality-of-life change associated with robotic-assisted therapy to improve hand motor function in patients with subacute stroke: A randomized clinical trial. *Phys Ther.* **90**, 493–504 (2010).
20. Mehrholz, J. Is electromechanical and robot-assisted arm training effective for improving arm function in people who have had a stroke?: A cochrane review summary with commentary. *Am. J. Phys. Med. Rehabil.* **98**, 339–340 (2019).
21. Thieme, H. *et al.* Mirror therapy for improving motor function after stroke. *Cochrane Database Syst. Rev.* **7**, Cd008449 (2018).
22. Bornheim, S. *et al.* Evaluating the effects of tDCS in stroke patients using functional outcomes: A systematic review. *Disabil. Rehabil.* **44**, 13–23 (2022).
23. Kang, N., Summers, J. J. & Cauraugh, J. H. Transcranial direct current stimulation facilitates motor learning post-stroke: A systematic review and meta-analysis. *J. Neurol. Neurosurg. Psychiatry* **87**, 345 (2016).
24. Liao, W. W. *et al.* Timing-dependent effects of transcranial direct current stimulation with mirror therapy on daily function and motor control in chronic stroke: A randomized controlled pilot study. *J. Neuroeng. Rehabil.* **17**, 101 (2020).

25. An, T. G., Kim, S. H. & Kim, K. U. Effect of transcranial direct current stimulation of stroke patients on depression and quality of life. *J. Phys. Ther. Sci.* **29**, 505–507 (2017).
26. Wu, C. Y., Huang, P. C., Chen, Y. T., Lin, K. C. & Yang, H. W. Effects of mirror therapy on motor and sensory recovery in chronic stroke: A randomized controlled trial. *Arch. Phys. Med. Rehabil.* **94**, 1023–1030 (2013).
27. Hsieh, Y. W. *et al.* Effects of home-based versus clinic-based rehabilitation combining mirror therapy and task-specific training for patients with stroke: A randomized crossover trial. *Arch. Phys. Med. Rehabil.* **99**, 2399–2407 (2018).
28. Hsieh, Y. W. *et al.* Comparison of proximal versus distal upper-limb robotic rehabilitation on motor performance after stroke: A cluster controlled trial. *Sci. Rep.* **8**, 2091 (2018).
29. Woodbury, M. L., Velozo, C. A., Richards, L. G. & Duncan, P. W. Rasch analysis staging methodology to classify upper extremity movement impairment after stroke. *Arch. Phys. Med. Rehabil.* **94**, 1527–1533 (2013).
30. Gregson, J. M. *et al.* Reliability of the Tone Assessment Scale and the modified Ashworth scale as clinical tools for assessing post-stroke spasticity. *Arch. Phys. Med. Rehabil.* **80**, 1013–1016 (1999).
31. Rossi, S. *et al.* Safety and recommendations for TMS use in healthy subjects and patient populations, with updates on training, ethical and regulatory issues: Expert guidelines. *Clin. Neurophysiol.* **132**, 269–306 (2021).
32. Duncan, P. W., Bode, R. K., Min Lai, S. & Perera, S. Rasch analysis of a new stroke-specific outcome scale: The Stroke Impact Scale. *Arch. Phys. Med. Rehabil.* **84**, 950–963 (2003).
33. Carod-Artal, F. J., Coral, L. F., Trizotto, D. S. & Moreira, C. M. The Stroke Impact Scale 3.0. *Stroke* **39**, 2477–2484 (2008).
34. Lin, K. C. *et al.* Psychometric comparisons of the Stroke Impact Scale 3.0 and stroke-specific quality of life scale. *Qual. Life Res.* **19**, 435–443 (2010).
35. Richardson, M., Campbell, N., Allen, L., Meyer, M. & Teasell, R. The stroke impact scale: Performance as a quality of life measure in a community-based stroke rehabilitation setting. *Disabil. Rehabil.* **38**, 1425–1430 (2016).
36. Duncan, P. W. *et al.* The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. *Stroke* **30**, 2131–2140 (1999).
37. Lang, C. E., Edwards, D. F., Birkenmeier, R. L. & Dromerick, A. W. Estimating minimal clinically important differences of upper-extremity measures early after stroke. *Arch. Phys. Med. Rehabil.* **89**, 1693–1700 (2008).
38. van der Lee, J. H. *et al.* Forced use of the upper extremity in chronic stroke patients: Results from a single-blind randomized clinical trial. *Stroke* **30**, 2369–2375 (1999).
39. Hägg, O., Fritzell, P. & Nordwall, A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur. Spine J.* **12**, 12–20 (2003).
40. Wu, C. Y., Chuang, L. L., Lin, K. C., Lee, S. D. & Hong, W. H. Responsiveness, minimal detectable change, and minimal clinically important difference of the Nottingham Extended Activities of Daily Living Scale in patients with improved performance after stroke rehabilitation. *Arch. Phys. Med. Rehabil.* **92**, 1281–1287 (2011).
41. Lemmens, R. J., Timmermans, A. A., Janssen-Potten, Y. J., Smeets, R. J. & Seelen, H. A. Valid and reliable instruments for arm-hand assessment at ICF activity level in persons with hemiplegia: A systematic review. *BMC Neurol.* **12**, 21 (2012).
42. Chen, C. M. *et al.* Potential predictors for health-related quality of life in stroke patients undergoing inpatient rehabilitation. *Health Qual. Life Outcomes* **13**, 118 (2015).
43. Coupar, F., Pollock, A., Rowe, P., Weir, C. & Langhorne, P. Predictors of upper limb recovery after stroke: A systematic review and meta-analysis. *Clin. Rehabil.* **26**, 291–313 (2012).
44. Chiti, G. & Pantoni, L. Use of Montreal Cognitive Assessment in patients with stroke. *Stroke* **45**, 3135–3140 (2014).
45. Fugl-Meyer, A. R., 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.* **7**, 13–31 (1975).
46. Wolf, S. L. *et al.* Assessing Wolf motor function test as outcome measure for research in patients after stroke. *Stroke* **32**, 1635–1639 (2001).
47. Gregson, J. M. *et al.* Reliability of measurements of muscle tone and muscle power in stroke patients. *Age Ageing* **29**, 223–228 (2000).
48. van der Lee, J. H., Beckerman, H., Knol, D. L., de Vet, H. C. & Bouter, L. M. Clinimetric properties of the motor activity log for the assessment of arm use in hemiparetic patients. *Stroke* **35**, 1410–1414 (2004).
49. Desrosiers, J., Bravo, G., Hébert, R., Dutil, É. & Mercier, L. Validation of the Box and Block Test as a measure of dexterity of elderly people: Reliability, validity, and norms studies. *Arch. Phys. Med. Rehabil.* **75**, 751–755 (1994).
50. Wu, C. Y., Chuang, I. C., Ma, H. I., Lin, K. C. & Chen, C. L. Validity and responsiveness of the Revised Nottingham Sensation Assessment for outcome evaluation in stroke rehabilitation. *Am. J. Occup. Ther.* **70**, 1–8 (2016).
51. Linacre, J. M., Heinemann, A. W., Wright, B. D., Granger, C. V. & Hamilton, B. B. The structure and stability of the functional independence measure. *Arch. Phys. Med. Rehabil.* **75**, 127–132 (1994).
52. Sarker, S. J., Rudd, A. G., Douiri, A. & Wolfe, C. D. Comparison of 2 extended activities of daily living scales with the Barthel Index and predictors of their outcomes: Cohort study within the South London Stroke Register (SLSR). *Stroke* **43**, 1362–1369 (2012).
53. Tin, K. H. The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, 832–844 (1998).
54. Cover, T. & Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **13**, 21–27 (1967).
55. Zhu, M., Chen, W., Hirdes, J. P. & Stolee, P. The K-nearest neighbor algorithm predicted rehabilitation potential better than current clinical assessment protocol. *J. Clin. Epidemiol.* **60**, 1015–1021 (2007).
56. Manning, T., Sleator, R. D. & Walsh, P. Biologically inspired intelligent decision making. *Bioengineered* **5**, 80–95 (2014).
57. Abedi, V. *et al.* Novel screening tool for stroke using artificial neural network. *Stroke* **48**, 1678–1681 (2017).
58. Smola, A. J. & Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **14**, 199–222 (2004).
59. Kim, J. K., Choo, Y. J. & Chang, M. C. Prediction of motor function in stroke patients using machine learning algorithm: Development of practical models. *J. Stroke Cerebrovasc. Dis.* **30**, 105856 (2021).
60. Jiawei, H. M. K. & Jian, P. *Data Mining: Concepts and Techniques* 2nd edn. (Morgan Kaufmann, 2006).
61. Shouman, M., Turner, T. & Stocker, R. Using decision tree for diagnosing heart disease patients. In *Proceedings of the Ninth Australasian Data Mining Conference*, vol. 121, 23–30 (Australian Computer Society, Inc., 2011).
62. Kent, J. T. Information gain and a general measure of correlation. *Biometrika* **70**, 163–173 (1983).
63. Wang, W. *et al.* A systematic review of machine learning models for predicting outcomes of stroke with structured data. *PLoS ONE* **15**, e0234722 (2020).
64. Luo, W. *et al.* Guidelines for developing and reporting machine learning predictive models in biomedical research: A multidisciplinary view. *J. Med. Internet Res.* **18**, e323 (2016).
65. Rodriguez, J. D., Perez, A. & Lozano, J. A. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 569–575 (2010).
66. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
67. Hall, M. *et al.* The WEKA data mining software: An update. *SIGKDD Explor. Newsl.* **11**, 10–18 (2009).
68. Pandey, A. K., Rajpoot, D. S. & Rajpoot, D. S. A comparative study of classification techniques by utilizing WEKA. In *2016 International Conference on Signal Processing and Communication (ICSC)* (2016).
69. Frank, E., Hall, M., Trigg, L., Holmes, G. & Witten, I. H. Data mining in bioinformatics using Weka. *Bioinformatics* **20**, 2479–2481 (2004).

70. Sim, J. A. *et al.* The major effects of health-related quality of life on 5-year survival prediction among lung cancer survivors: Applications of machine learning. *Sci. Rep.* **10**, 10693 (2020).
71. Scrutinio, D. *et al.* Machine learning to predict mortality after rehabilitation among patients with severe stroke. *Sci. Rep.* **10**, 20127 (2020).
72. Apao, N. J., Feliscuzo, L. S., Romana, C. L. S. & Tagaro, J. Multiclass classification using random forest algorithm to prognosticate the level of activity of patients with stroke. *IJSTR* **9**, 1233–1240 (2020).
73. Badriyah, T., Sakinah, N., Syarif, I. & Syarif, D. R. Machine learning algorithm for stroke disease classification. In *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)* (2020).
74. White, J. *et al.* Predictors of health-related quality of life in community-dwelling stroke survivors: A cohort study. *Fam. Pract.* **33**, 382–387 (2016).
75. Katona, M., Schmidt, R., Schupp, W. & Graessel, E. Predictors of health-related quality of life in stroke patients after neurological inpatient rehabilitation: A prospective study. *Health Qual. Life Outcomes* **13**, 58 (2015).
76. Huang, P. C. *et al.* Predictors of motor, daily function, and quality-of-life improvements after upper-extremity robot-assisted rehabilitation in stroke. *Am. J. Occup. Ther.* **68**, 325–333 (2014).
77. Nijenhuis, S. M. *et al.* Strong relations of elbow excursion and grip strength with post-stroke arm function and activities: Should we aim for this in technology-supported training?. *J. Rehabil. Assist. Technol. Eng.* **5**, 2055668318779301 (2018).
78. Clarke, P. & Black, S. E. Quality of life following stroke: Negotiating disability, identity, and resources. *J. Appl. Gerontol.* **24**, 319–336 (2005).
79. Pedersen, S. G. *et al.* Stroke-specific quality of life one-year post-stroke in two Scandinavian country-regions with different organisation of rehabilitation services: A prospective study. *Disabil. Rehabil.* **43**, 3810–3820 (2021).
80. Doyle, S., Bennett, S., Fasoli, S. E. & McKenna, K. T. Interventions for sensory impairment in the upper limb after stroke. *Cochrane Database Syst. Rev*. **6**, CD006331 (2010).
81. Wu, C. W., Seo, H. J. & Cohen, L. G. Influence of electric somatosensory stimulation on paretic-hand function in chronic stroke. *Arch. Phys. Med. Rehabil.* **87**, 351–357 (2006).
82. Turville, M., Carey, L. M., Matyas, T. A. & Blennerhassett, J. Change in functional arm use is associated with somatosensory skills after sensory retraining poststroke. *Am. J. Occup. Ther.* **71**, 1–9 (2017).
83. Meyer, S., Karttunen, A. H., Thijs, V., Feys, H. & Verheyden, G. How do somatosensory deficits in the arm and hand relate to upper limb impairment, activity, and participation problems after stroke? A systematic review. *Phys. Ther.* **94**, 1220–1231 (2014).
84. Rokach, L. Ensemble methods for classifiers. In *Data Mining and Knowledge Discovery Handbook* (eds Maimon, O. & Rokach, L.) 957–980 (Springer, 2005).
85. Hung, C. Y., Chen, W. C., Lai, P. T., Lin, C. H. & Lee, C. C. Comparing deep neural network and other machine learning algorithms for stroke prediction in a large-scale population-based electronic medical claims database. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2017**, 3110–3113 (2017).

Author contributions

W.W.L. performed data analyses and wrote the manuscript. Y.W.H. and C.Y.W. contributed to grant application and project management. C.L.C. and T.H.L. contributed to patient recruitment and project management. All authors involved in revision of this study. All authors read and approved the final manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.W.

Reprints and permissions information is available at www.nature.com/reprints.

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