Journal of NeuroEngineering and Rehabilitation

RESEARCH

Open Access

Do predictors of motor recovery differ between robotic and conventional post-stroke rehabilitation?

Silvia Campagnini[1], Piergiuseppe Liuzzi[1], Marco Germanotta[1*], Arianna Pavan[1], Alessio Fasano[1], Francesca Cecchi[1,2], Irene Giovanna Aprile[1†] and Andrea Mannini[1†]

Abstract

Background While research on robotic rehabilitation has largely focused on evaluating device effectiveness, there remains a clear need to investigate patient characteristics that predict response to robot-assisted therapy. Recent advancements in machine learning (ML) techniques support such investigations by enabling the development of data-driven tools to assist in selecting appropriate rehabilitation treatments. This study aimed to develop two ML models, predicting two motor outcomes each, following either robot-assisted or conventional upper limb rehabilitation in post-stroke patients. It further sought to compare how baseline predictors contributed differently across the two treatment modalities.

Methods We conducted a retrospective analysis of data from a previous randomized controlled trial evaluating robotic upper limb rehabilitation in post-stroke patients. Four ML algorithms were trained and validated using a nested cross-validation framework, with models developed separately for robotic and conventional treatment subgroups. Baseline predictors were used to estimate two post-treatment motor outcomes: Fugl-Meyer Assessment (FMA) and Action Research Arm Test (ARAT). SHAP analyses were performed to assess the contribution of each predictor to the models.

Results After data cleaning, 99 patients in the conventional group and 91 in the robotic group were included. Prediction errors for ARAT score after training (expressed in median and absolute error) were 5.0 [6.0] in the robotic group and 3.0 [6.0] in the conventional group. For FMA prediction, results were 5.0 [5.0] and 5.0 [7.0], respectively. Baseline FMA was a strong predictor of ARAT outcomes in both groups. For FMA prediction, the presence of neglect emerged as more influential in the robotic group. Age was a key predictor of both outcomes, but only in the conventional group.

†Andrea Mannini and Irene Giovanna Aprile have equally contributed to this work as last authors.

*Correspondence:
Marco Germanotta
mgermanotta@dongnocchi.it

Full list of author information is available at the end of the article

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Conclusions The differing contributions of baseline predictors across treatment types provide clinically meaningful insights and support the development of clinical decision support systems aimed at optimizing rehabilitation strategies based on individual patient characteristics.

Keywords Clinical decision support systems, Rehabilitation, Robotics, Stroke, Treatment selection

Background

Technological developments over the last few decades have significantly influenced the healthcare field. In particular, in the rehabilitation of post-stroke patients, robotics has been identified as a suitable tool for both treatment and assistance [1], aiding recovery of both upper and lower limbs [2, 3].

Regarding the upper limb, the most recent update of the systematic review by Mehrholz et al. [4] confirmed that robotic treatment yields results comparable to conventional therapy in terms of daily living activities and hand-arm function. However, the authors emphasized the importance of exploring the effects of robot-assisted treatment across different clinical conditions and population characteristics. Evaluating the rehabilitation effects of a device is a fundamental step in clinical validation; nonetheless, it is equally important to understand how its impact differs across diverse clinical profiles. This information is essential for guiding clinical decisions regarding personalized robot-assisted treatments. Although many studies have assessed the effectiveness of robotic rehabilitation devices, analyses of patients' baseline characteristics remain scarce and sometimes inconsistent.

The majority of studies indicate that the severity of motor impairment at baseline is a key predictor of motor recovery [5–9]. In addition, other possible predictors were identified, including baseline cognitive function [8], time since the stroke onset [5, 9], age [5], gender [5], and pain [5]. These variables are often included as potential predictors across different studies; however, while some analyses identify them as predictors, others do not, leading to contrasting findings. Furthermore, the heterogeneity in the outcome scales adopted (the FMA [5, 7, 9], the modified Barthel Index [6], and the manual function test and functional independence measure [8]), as well as the different devices used, may also contribute to these discrepancies.

To facilitate such analyses, statistical modelling of prognostic outcomes may play a key role in identifying predictors associated with more favorable responses across different treatment modalities. More broadly, Machine Learning (ML) techniques are being extensively explored in the healthcare field [10].

Among the available solutions, those designed to support clinical decision-making—Clinical Decision Support Systems (CDSS) [11] —are employed for both diagnostic [12] and prognostic purposes [13]. ML algorithms can learn existing patterns among variables from data and, if properly validated, generalize this knowledge to make predictions on new data [14].

Recent work has explored the use of ML in stroke robotic rehabilitation. According to Iwamoto et al. [15], factors such as initial motor function and early initiation of therapy were significant predictors of recovery. Quattrocelli et al. [16] showed that ML (Gaussian process regression and decision trees, using a leave-one-out cross-validation) can predict outcomes of upper limb robot-assisted therapy with good accuracy. However, a commonly cited drawback of these models is their poor interpretability, particularly when using complex algorithms [17]. The ability to interpret opaque model predictions is crucial in healthcare applications, where clinicians need to understand not only the output but also the reasoning behind a model's predictions in order to trust and effectively integrate these tools into their clinical workflow.

Some classical ML approaches already offer feature ranking capabilities, such as regression coefficients or model-based feature importance. Nonetheless, two significant limitations remain: (i) regression coefficients apply only to linear models (e.g., Linear Regression, Elastic-Net), and (ii) model-based rankings (e.g., Gini index, recursive feature selection orders) are derived from the full dataset and are not patient-specific.

To address this gap, the Shapley Additive exPlanations (SHAP) technique [18], based on Shapley values [19], offers a viable solution. Rooted in game theory, SHAP explains how individual features contribute to a specific outcome estimate. Specifically, SHAP evaluates all possible combinations of input features to assess how the inclusion of each one alters the prediction. For each individual, the contribution of a specific feature is calculated as the average marginal change in the prediction when that feature is added to various subsets of the remaining features. This framework enables precise interpretation of how each factor influences a specific prediction, offering transparency and individualized insights into model behavior.

In this context, our objective was to develop two independent ML models using a dataset of post-stroke patients who underwent either robotic or conventional upper limb rehabilitation to predict two motor outcomes each. Applying the SHAP technique allowed us to identify the most influential predictors—among the patients' admission characteristics—for motor recovery under each treatment type. Our central hypothesis is that

comparing the importance of predictors across the two models can help define personalized therapy strategies for future patients.

The dataset analyzed in this retrospective study was originally collected prospectively during a Randomized Controlled Trial (RCT) comparing conventional physiotherapy and robot-assisted upper limb rehabilitation for post-stroke patients [20]. The random allocation of patients to intervention groups strengthens the design for evaluating differential treatment effects by minimizing selection bias. In a previous study conducted by our research team [21], classical statistical methods were used to compare the impact of baseline characteristics on recovery between the two groups, revealing age as a differential factor. Building on those findings, the present study aims to identify which baseline factors can predict motor recovery, as measured by the Fugl-Meyer Assessment and the Action Research Arm Test, in patients undergoing either conventional or robotic rehabilitation, by implementing nested cross-validated machine learning models.

Methods

Sample

This retrospective prognostic study is based on prospectively collected data from patients enrolled in a multisite RCT [20]. A total of 224 participants were enrolled and randomly assigned to either the experimental group (111 participants) or the control group (113 participants). Of these, 190 patients (99 in the conventional group and 91 in the robotic group) completed the scheduled treatment sessions and were reassessed at the end of the intervention. Recruitment took place between August 2016 and March 2018 across eight rehabilitation centers affiliated with Fondazione Don Carlo Gnocchi, Italy.

Inclusion criteria were as follows:

- Presence of a single ischemic or hemorrhagic stroke (confirmed by MRI or CT);
- Age between 40 and 85 years;
- Stroke onset between two weeks and six months prior to enrollment (i.e., post-acute phase);
- Sufficient cognitive and language abilities to understand the study and follow instructions;
- Upper extremity Fugl-Meyer Assessment (FMA) score (0–66 version) less than or equal to 58;
- Provision of written informed consent.

Exclusion criteria included behavioral or cognitive disorders and/or poor compliance, fixed contracture of the affected limb (ankylosis, modified Ashworth Scale score of 4), and severe visual acuity deficits.

The study adhered to the International Conference on Harmonization Good Clinical Practice guidelines and the Declaration of Helsinki. The study protocol was approved by the Ethics and Experimental Research Committee of the Fondazione Don Carlo Gnocchi on April 6, 2016 (FDG 6.4.2016), and the original trial was registered on ClinicalTrials.gov (NCT02879279).

Intervention

Participants in the experimental group underwent upper-limb robot-assisted rehabilitation for a total of 30 sessions, conducted five days per week, each lasting 45 min. This robotic intervention was provided in addition to conventional rehabilitation sessions. Upper-limb training was delivered using four different electromechanical devices, selected according to individual participant needs:

- Motore (Humanware, Italy): a robotic device enabling passive, active, and active-assistive planar movements of the shoulder and elbow joints;
- Amadeo (Tyromotion, Austria): a robotic device supporting passive, active, and active-assistive finger flexion and extension movements;
- Pablo (Tyromotion, Austria): a sensor-based system that facilitates unsupported three-dimensional movements of the shoulder, elbow, and wrist joints, in both unimanual and bimanual modes;
- Diego (Tyromotion, Austria): an electromechanical system that allows three-dimensional, unimanual and bimanual shoulder joint movements, with integrated arm weight support.

In the control group, participants received conventional therapy focused on functional improvement through task-oriented exercises, sensorimotor reorganization, and spasticity reduction. To stimulate neuroplasticity, exercises included reaching, grasping, and daily living activities, performed in three-dimensional space across active, passive, and assisted modalities.

In both groups, upper limb rehabilitation (either conventional or robotic) was provided in addition to standard lower limb rehabilitation sessions. These standard sessions were conducted six times per week, each lasting 45 min, and included exercises such as sit-to-stand, balance, standing, and gait training. Depending on individual needs, participants in both groups could also receive occupational or speech therapy.

Data collection

Participants' clinical assessments were conducted at the beginning (T0) and at the end (T1) of the treatment period, within a five-day tolerance window for both time points. The primary outcomes of the study were the Fugl-Meyer Assessment (FMA) motor score [22] and the Action Research Arm Test (ARAT) score [23], both

measured at T1. These outcomes were selected to represent two distinct hierarchical levels within the International Classification of Functioning, Disability and Health (ICF) model [24] Specifically, the FMA corresponds to the domain of body functions—particularly neuromusculoskeletal and movement-related functions (b7)—while the ARAT pertains to the domain of activities, specifically general tasks and demands (d2).

Regarding the potential predictors, the following baseline variables (collected at T0) were considered:

- Age.
- Sex.
- Time Post-Onset (TPO), in days.
- Aetiology.
- Side of the lesion.
- Presence/absence of language impairment.
- Presence/absence of neglect.
- FMA upper extremity motor score [22].
- ARAT total score [23].
- Numerical Rating Scale (NRS) score [25].
- Modified Ashworth Scale (MAS) total scores for the shoulder abduction, shoulder intra-rotation, elbow, and wrist [26].
- Medical Research Council (MRC) score for the shoulder, elbow, and wrist [27].
- Physical Composite Score (PCS) and Mental Composite Score (MCS) obtained from the Short Form Health Survey (SF-36) [28].
- Modified Barthel Index (MBI) total score [29].
- Motricity Index (MI) total score [30].
- Frenchay Arm Test (FAT) total score [31].

Data analysis

Participants with missing outcome data were excluded from the analyses, whereas missing values among the predictor variables were imputed using a single imputation method based on k-Nearest Neighbours (kNN). We selected k-Nearest Neighbors (kNN) imputation due to its robustness and suitability for heterogeneous clinical data. kNN is a non-parametric, instance-based imputation method that leverages local similarity structures in the data and has been shown to perform favorably compared to simpler single-value imputation approaches, particularly in settings where relationships between variables are complex and potentially non-linear [32]. Importantly, kNN imputation can be seamlessly embedded within a cross-validation pipeline, as the imputation is performed using only the training data at each fold. This design prevents information leakage from the test set into the training phase, thereby avoiding train–test contamination and ensuring an unbiased evaluation of model performance.

The independent variables were first used in univariate statistical analyses for preliminary exploration. Descriptive and univariate analyses were conducted using IBM SPSS Statistics for Windows, Version 28.0 (IBM Corp., Armonk, NY, released 2020).

Descriptive statistics were presented as relative frequencies and means with standard deviations for categorical variables, and as medians with interquartile ranges (IQR) for numerical variables. Univariate analyses included Pearson's correlation for numerical variables and either the t-test or Mann-Whitney U test, depending on the distribution's normality. The Shapiro-Wilk test was used to assess the normality of distributions. A p-value less than 0.05 was considered statistically significant in all tests. Preliminary statistical analyses were conducted prior to ML modelling to characterize the data distribution, assess relationships among variables, and identify potential sources of bias or instability. This exploratory and inferential step provides critical insight into the structure and quality of the data, supporting informed feature selection and pre-processing choices. Moreover, grounding the machine learning pipeline in statistical analysis facilitates the interpretability and reliability assessment of ML results, enabling consistency checks between data-driven patterns and established statistical evidence.

All variables were included in two ML predictive models for each cohort, targeting the FMA and ARAT scales. This approach produced a 2 × 2 factorial design: FMA_R, FMA_C, ARAT_R, and ARAT_C. Models were developed in Python using the Scikit-learn library [33]. Four algorithms were compared: Lasso regularized regression (LASSO), Support Vector Regression (SVR), Random Forest (RF), and kNN. The selected algorithms were intentionally chosen to maximize methodological diversity and represent distinct modelling paradigms, rather than to exhaustively compare all possible approaches. Specifically, LASSO represents a linear, regularized model with built-in feature selection; Support Vector Regression captures non-linear relationships through kernel-based optimization; Random Forest models complex, non-linear interactions via ensemble learning; and kNN provides a non-parametric, instance-based approach driven by local similarity. Together, these models span a broad range of underlying learning mechanisms, enabling a robust evaluation of performance consistency across fundamentally different algorithmic assumptions.

A nested cross-validation framework was employed [34], with an inner 10-fold cross-validation loop for hyperparameter optimization and imputation, and an outer 5-fold loop for model testing, thereby reducing the

risk of overfitting and spurious associations due to the large number of predictors. During the inner loop, the number of features for forward feature selection (via the Mlxtend library [35]) was also optimized, ranging from 1 to the total number of features (22). Model performance was assessed using the Median Absolute Error (MAE) and interquartile range (IQR). Additionally, the Minimal Clinically Important Difference (MCID) served as the clinical benchmark to evaluate the practical relevance of model predictions. The established MCID values for post-stroke populations were 5.7 points for the ARAT [36] and 5.25 points for the FMA motor score (upper limb) [37].

For each of the four models, the SHAP technique was applied to the best-performing algorithm, defined as the one with the lowest average MAE across validation folds. Shapley values for each feature were concatenated across folds, with null vectors imputed when a feature was not selected in a given fold.

Table 1 Descriptive analyses of the overall sample and of the control and experimental groups, separately

	Enrolled cohort (*N* = 190)	Conventional therapy (*N* = 99)	Robotic therapy (*N* = 91)	*p*-value
A				
Age	70.0 [18.0]	70.0 [19.0]	71.0 [19.0]	0.509
Sex (F)	83 (43.7%)	44 (44.4%)	39 (42.9%)	0.826
TPO	29.5 [54.0]	28.0 [45.0]	31.0 [59.0]	0.796
Ischemic	140 (73.7%)	74 (74.7%)	66 (72.5%)	0.728
Lesion site (Right)	93 (48.9.0%)	53 (53.5%)	40 (44.0%)	0.187
Language impairment (Present)	40 (21.1%)	29 (29.3%)	11 (12.1%)	**0.004**
Neglect (Present)	37 (19.5%)	20 (20.2%)	17 (18.9%)	0.791
FMA	17.0 [28.0]	15.0 [27.0]	19.0 [29.0]	0.125
ARAT	0.0 [19.0]	0.0 [19.0]	0.0 [19.0]	0.552
NRS	2.0 [5.0]	2.0 [5.0]	2.0 [5.0]	0.872
MAS_SA	0.0 [0.0]	0.0 [0.0]	0.0 [0.0]	0.606
MAS_SI	0.0 [1.0]	0.0 [0.1]	0.0 [1.0]	0.476
MAS_E	0.0 [1.0]	0.0 [0.1]	0.0 [1.0]	0.740
MAS_W	0.0 [1.0]	0.0 [0.1]	0.0 [1.0]	0.388
MRC_S	2.0 [3.0]	1.0 [3.0]	2.0 [2.0]	0.131
MRC_E	2.0 [3.0]	1.0 [3.0]	2.0 [3.0]	0.147
MRC_W	1.0 [3.0]	1.0 [3.0]	1.0 [3.0]	0.345
PCS	26.5 [9.7]	26.5 [9.6]	26.8 [10.4]	0.890
MCS	39.8 [17.0]	38.6 [15.0]	41.8 [17.9]	0.264
MBI	24.0 [37.0]	21.0 [42.0]	26.0 [29.0]	0.543
FAT	0.0 [2.0]	0.0 [2.0]	0.0 [2.0]	0.200
MI	30.0 [58.0]	29.0 [60.0]	30.0 [48.0]	0.295
B				
FMA	29.5 [35.0]	31.0 [34.0]	29.0 [36.0]	0.153
ARAT	15.0 [31.0]	13.0 [30.0]	15.0 [35.0]	0.246

Independent variables used as predictors are reported in Table 1A, whilst targeted outcomes are reported in Table 1B. For continuous variables, median and interquartile range (in brackets) are shown, while for categorical variables, count and percentages (in parentheses) are reported

ARAT: Action Research Arm Test; FAT: Frenchay Arm Test; FMA: Fugl-Meyer Assessment; IQR: Interquartile Range; MAS_E: Modified Ashworth Scale – Elbow; MAS_SA: Modified Ashworth Scale – Shoulder Abduction; MAS_SI: Modified Ashworth Scale – Shoulder Intra-rotation; MAS_W: Modified Ashworth Scale – Wrist; MRC_E: Medical Research Council – Elbow; MBI: Modified Barthel Index; MCS: Mental Composite Score; MI: Motricity Index; MRC_S: Medical Research Council – Shoulder; MRC_W: Medical Research Council – Wrist; NRS: Numerical Rating Scale; PCS: Physical Composite Score; TPO: Time Post- Onset

Statistically significant p-values are presented in bold

Results

Of the 224 participants enrolled in the RCT, 34 individuals (20 from the robotic group and 14 from the conventional group) did not complete the scheduled sessions due to clinical deterioration and were therefore considered dropouts. Consequently, the analyses included 190 participants: 99 assigned to the conventional therapy group and 91 to the robotic group. Missing data among the predictor variables accounted for only 0.16% of the dataset.

A summary of the overall sample and the two treatment groups is provided in Table 1. The median age [interquartile range, IQR] was 70.0 [19.0] years in the conventional group and 71.0 [19.0] years in the robotic group. Female participants represented 44.4% and 42.9% of the conventional and robotic groups, respectively. Baseline comparisons revealed no statistically significant differences between groups, except for the presence of language impairment (Table 1).

Preliminary univariate analyses revealed statistically significant associations between predictors and outcomes. Specifically, age showed a negative association with the FMA score in the conventional group, while no such associations were found for either the FMA or ARAT scores in the robotic group. In addition, a statistically significant difference by sex was observed for both outcomes in the conventional group, with males presenting higher outcome scores. No significant sex-based differences were found in the robotic group (Table 2).

Regarding the predictive models, the best test set results (validation results presented in Table 3) for ARAT prediction, expressed in MAE [IQR], were 3.0 [6.0] points for the conventional group and 5.0 [6.0] points for the robotic group. For FMA prediction, the best-performing models achieved MAEs of 5.0 [7.0] points and 5.0 [5.0] points, for the conventional and robotic groups, respectively. To evaluate the performance of the selected models at the individual level, we plotted predicted versus observed values for each subject. The scatter plots show a close alignment of predictions with actual outcomes, indicating that the models capture the variability across patients (Fig. 1). The diagonal line represents perfect prediction, and deviations from this line reflect estimation errors. Overall, the plots highlight both the accuracy of the predictions and the absence of systematic bias, with

Table 2 Univariate analyses on the control and experimental group, separately, considering the FMA and ARAT score at T1 as outcomes

Variables	Conventional group (*N* = 99)				Robotic group (*N* = 91)			
	FMA		ARAT		FMA		ARAT	
	Metrics	*p*-value	Metrics	*p*-value	Metrics	*p*-value	Metrics	*p*-value
Age	-0.230	**0.022**	-0.195	0.053	-0.057	0.593	-0.066	0.536
Sex	Male: 34.0 [30.0] Female: 17.5 [33.0]	0.067	Male: 17.0 [30.0] Female: 0.0 [24.0]	0.015	Male: 38.5 [40.0] Female: 26.0 [28.0]	0.273	Male: 16.0 [43.0] Female: 15.0 [27.0]	0.472
TPO	0.028	0.782	-0.031	0.758	-0.109	0.304	-0.141	0.183
Aetiology	Isch.: 30.5 [34.0] Haem.: 31.0 [36.0]	0.222	Isch: 12.5 [29.0] Haem.: 17.0 [40.0]	0.158	Isch.: 29.5 [37.0] Haem.: 26.0 [36.0]	0.431	Isch.: 15.0 [38.0] Haem.: 15.0 [26.0]	0.867
Side of the lesion	Right: 34.0 [40.0] Left: 22.0 [32.0]	0.400	Right: 15.0 [34.0] Left: 3.0 [25.0]	0.109	Right: 36.5 [39.0] Left: 25.0 [35.0]	0.362	Right: 17.5 [41.0] Left: 15.0 [29.0]	0.316
Language impairment	Pres.: 29.0 [28.0] Abs.: 33.0 [35.0]	0.145	Pres.: 1.0 [23.0] Abs.: 14.5 [34.0]	0.177	Pres.: 19.0 [32.0] Abs.: 31.0 [36.0]	0.082	Pres.: 7.0 [13.0] Abs.: 15.0 [36.0]	0.082
Neglect	Pres.: 13.5 [25.0] Abs.: 34.0 [35.0]	**0.006**	Pres.: 0.0 [8.0] Abs.: 16.0 [32.0]	**0.007**	Pres.: 18.0 [11.0] Abs.: 36.5 [36.0]	**0.017**	Pres.: 0.0 [14.0] Abs.: 16.0 [35.0]	**0.003**
FMA	0.882	**< 0.001**	0.827	**< 0.001**	0.913	**< 0.001**	0.850	**< 0.001**
ARAT	0.782	**< 0.001**	0.845	**< 0.001**	0.800	**< 0.001**	0.847	**< 0.001**
NRS	-0.287	**0.004**	-0.256	**0.011**	-0.262	**0.012**	-0.277	**0.008**
MAS_SA	-0.162	0.109	-0.124	0.221	-0.127	0.230	-0.056	0.598
MAS_SI	-0.101	0.319	-0.091	0.369	-0.145	0.169	-0.085	0.426
MAS_E	0.033	0.744	0.020	0.844	-0.147	0.165	-0.074	0.486
MAS_W	-0.080	0.433	-0.111	0.275	-0.194	0.066	-0.146	0.166
MRC_S	0.792	**< 0.001**	0.785	**< 0.001**	0.773	**< 0.001**	0.738	**< 0.001**
MRC_E	0.803	**< 0.001**	0.789	**< 0.001**	0.823	**< 0.001**	0.778	**< 0.001**
MRC_W	0.766	**< 0.001**	0.769	**< 0.001**	0.847	**< 0.001**	0.783	**< 0.001**
PCS	0.356	**< 0.001**	0.356	**< 0.001**	0.239	**0.023**	0.228	**0.030**
MCS	0.135	0.198	0.103	0.326	0.154	0.148	0.204	0.054
MBI	0.571	**< 0.001**	0.497	**< 0.001**	0.514	**< 0.001**	0.566	**< 0.001**
FAT	0.707	**< 0.001**	0.700	**< 0.001**	0.755	**< 0.001**	0.786	**< 0.001**
MI	0.839	**< 0.001**	0.796	**< 0.001**	0.850	**< 0.001**	0.817	**< 0.001**

Metrics represent R or median [IQR] on the groups, in the presence of continuous variables or categorical ones, respectively. Values in bold represent statistically significant associations

Abs.: Absent; ARAT: Action Research Arm Test; FAT: Frenchay Arm Test; FMA: Fugl-Meyer Assessment; Haem.: Haemorrhagic; Isch.: Ischemic; IQR: Interquartile Range; MAS_E: Modified Ashworth Scale – Elbow; MAS_SA: Modified Ashworth Scale – Shoulder Abduction; MAS_SI: Modified Ashworth Scale – Shoulder Intra-rotation; MAS_W: Modified Ashworth Scale – Wrist; MRC_E: Medical Research Council – Elbow; MBI: Modified Barthel Index; MCS: Mental Composite Score; MI: Motricity Index; MRC_S: Medical Research Council – Shoulder; MRC_W: Medical Research Council – Wrist; NRS: Numerical Rating Scale; PCS: Physical Composite Score; Pres.: Present; TPO: Time Post- Onset

Statistically significant p-values are presented in bold

points scattered evenly around the diagonal. This visualization complements the aggregate metrics (MAE and IQR) by providing a clear view of model performance for individual participants. Comparison with the clinical benchmarks showed that the *ARAT_C* models consistently produced median MAEs below the MCID of 5.7, defined for the chronic stroke population. In contrast, some *ARAT_R* model predictions slightly exceeded this threshold. For the FMA outcome, *FMA_R* models and only *FMA_C* trained using SVR yielded median MAEs below the MCID of 5.25, which pertains to the upper extremity motor FMA validated in stroke populations (Table 4).

Further interpretability analyses were conducted using the SHAP technique on the best-performing algorithms (Fig. 2). These analyses confirmed the strong contribution of baseline FMA and ARAT scores for the corresponding models *FMA_R/FMA_C* and *ARAT_R*/*ARAT_C*) respectively. Notably, the FMA score emerged as the most influential predictor across all models. Consistent with statistical results, age was not a major contributing factor in the robotic group models. In contrast, the presence of neglect emerged as a key predictor in models developed for the experimental (robotic) group only. Additionally, wrist impairment (MRC_W) was found to contribute substantially to predictions in the robotic group models.

Discussion

This study aimed to develop ML-based predictive models for two cohorts of post-stroke patients undergoing either robot-assisted or conventional rehabilitation of the upper limb. Predictive performances were promising, with

Table 3 Validation median absolute errors for each fold, algorithm, and model

Algorithm	Fold	FMA_C	FMA_R	ARAT_C	ARAT_R
LASSO	1	6.107	4.102	5.111	4.433
	2	5.141	4.614	4.765	3.998
	3	4.567	4.255	4.444	4.230
	4	4.262	4.699	4.750	4.124
	5	4.742	3.430	3.878	4.557
	Median	*4.724*	*4.255*	*4.750*	*4.230*
SVR	1	4.047	4.066	3.18	2.155
	2	4.350	3.095	3.033	3.156
	3	4.548	3.611	3.726	3.548
	4	4.105	3.140	1.411	3.524
	5	3.836	2.816	1.556	3.486
	Median	*4.105*	*3.611*	*3.033*	*3.486*
kNN	1	4.829	3.642	3.636	1.517
	2	4.520	3.542	2.940	3.282
	3	4.482	3.714	2.717	2.900
	4	4.730	3.833	2.909	3.717
	5	3.600	3.513	1.538	3.575
	Median	*4.520*	*3.642*	*2.909*	*3.282*
RF	1	4.629	4.125	4.203	3.112
	2	4.696	3.861	3.372	3.004
	3	4.855	4.192	4.454	4.725
	4	4.892	3.382	3.802	4.963
	5	3.835	3.329	2.480	3.743
	Median	*4.696*	*3.861*	*3.802*	*3.743*

ARAT_C: Model with outcome Action Research Arm Test and trained on the Conventional group; ARAT_R: Model with outcome Action Research Arm Test and trained on the Robotic group; FMA_C: Model with outcome Fugl-Meyer Assessment and trained on the Conventional group; FMA_R: Model with outcome Fugl-Meyer Assessment and trained on the Robotic group; kNN: k-Nearest Neighbours; LASSO: Lasso regularised regression; RF: Random Forest; SVR: Support Vector Regression

ARAT predictions being 5.0 [6.0] in the robotic group and 3.0 [6.0] in the conventional group. For FMA prediction, results were 5.0 [5.0] and 5.0 [7.0], respectively. Beyond overall model performance, the relative contribution of baseline predictors varied meaningfully according to treatment type. Baseline motor impairment, as captured by FMA at admission, consistently emerged as a strong determinant of ARAT outcomes across both groups, underscoring its central role in upper-limb functional recovery. In contrast, predictors influencing FMA outcomes differed by intervention: the presence of neglect was more influential in the robotic group, whereas age played a prominent role in the conventional group for both ARAT and FMA predictions. Taken together, these findings highlight that the relevance of baseline clinical characteristics is not uniform across rehabilitation approaches, supporting the need for treatment-specific predictive frameworks. This study identifies key factors associated with differential responsiveness to rehabilitation, providing insights that may inform the future development of clinical decision support tools aiming at tailored therapeutic approaches.

The FMA, particularly its upper extremity motor component, is a well-validated tool [22] with excellent psychometric properties, including interrater reliability [38], internal consistency [39], and validity [40]. It assesses joint coordination, speed, and range of motion [41]. Conversely, the ARAT evaluates a patient's ability to handle objects of varying size, shape, and weight, making it a specific measure of activity limitation in the arm [40]. The ARAT also demonstrates excellent intra- and interrater reliability [42], validity [36], and internal consistency [43].

Given the robust validation of both measures, their respective MCID values were used as clinical benchmarks for evaluating model performance (Table 3). For the *ARAT_C* models, all median absolute error (MAE) values fell below the MCID of 5.7, established for chronic stroke patients. For the *ARAT_R* models, some predictions slightly exceeded this threshold. In the case of FMA predictions, *FMA_R* and only *FMA_C* trained on the SVR algorithm achieved MAEs below the MCID of 5.25, the established benchmark for upper extremity in stroke populations (Table 4). Notably, the IQRs for prediction errors were close to the MAEs in each case, reflecting a limitation linked to the relatively small sample size in our retrospective study. We expect that a larger sample would likely yield narrower IQRs and more precise estimates.

The primary focus of this study, however, was to explore the differential contributions of predictors across treatment types and outcomes. Therefore, results from the SHAP interpretability analysis were compared, and separate insights were derived for the two outcomes.

Regarding the FMA outcome, differences between robotic and conventional groups were particularly insightful. Age emerged as a significant predictor in the conventional group, consistent with earlier biostatistical analyses [21], but was not influential in the robotic group. One advantage of robotic therapy may lie in its capacity to engage patients across age groups. Robotic devices, especially those used in this study, often incorporate virtual reality scenarios that create appealing, task-oriented exercises, which may help sustain motivation and involvement regardless of age [21, 44].

In terms of upper limb motor function, other key predictors were the Motricity Index (MI) in models for the conventional group (*FMA_C* and *ARAT_C*), and wrist muscle strength (MRC_W) in models for the robotic group (*FMA_R* and *ARAT_R*). This distinction suggests that while general arm strength predicts recovery under conventional therapy, wrist functionality plays a more central role in robot-assisted rehabilitation. One plausible explanation is the use of end-effector devices in the robotic group, where movement is initiated from a single

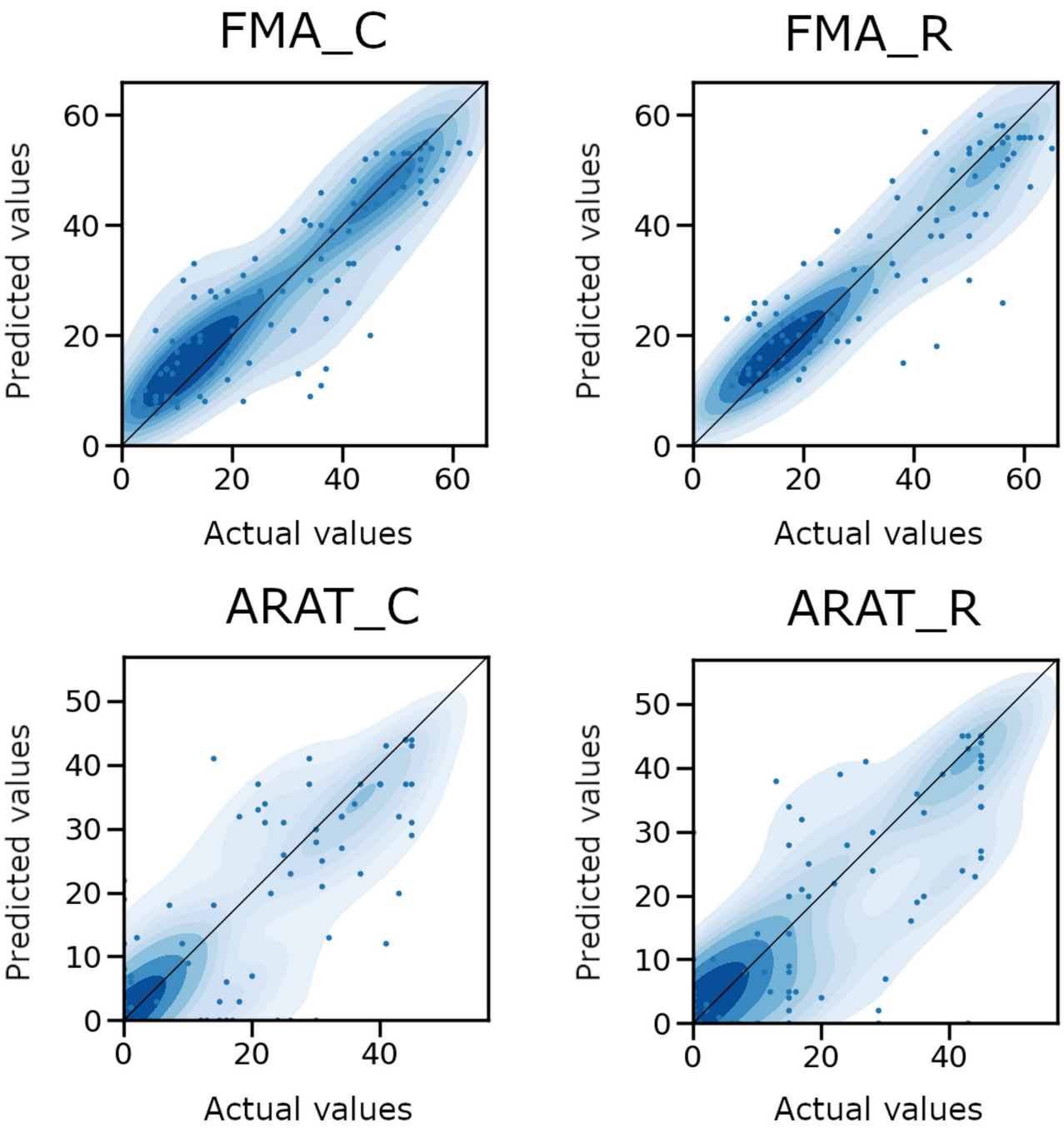


Fig. 1 Scatter plot of the predicted vs. actual values in the test set. Overall, the plots highlight both the accuracy of the predictions and the absence of systematic bias, with points scattered evenly around the diagonal

distal contact point at the hand. In such systems, residual wrist function may be critical to generating effective motion throughout the limb. Further studies are warranted to clarify these mechanisms.

Another variable of note was the presence of neglect, which proved influential in robotic but not conventional models. This may reflect the inability of standard robotic programs to adapt to cognitive impairments such as spatial neglect, whereas human therapists can modify treatment accordingly. These findings highlight the need to develop robot-assisted therapies that incorporate cognitive or attentional dimensions. Prior research has shown that when such impairments are addressed directly, robotic therapy can also enhance cognitive functions, including memory, visual attention, and numerical processing [45]. Unfortunately, the current dataset did not differentiate neglect subtypes, limiting more nuanced

Table 4 Median absolute errors and interquartile ranges for each algorithm in each model

Algorithm	FMA_C	FMA_R	ARAT_C	ARAT_R
LASSO	6.0 [8.0]	5.0 [5.5]	5.0 [8.5]	6.0 [6.0]
SVR	**5.0 [7.0]**	5.0 [5.5]	3.0 [10.0]	4.0 [7.5]
kNN	6.0 [7.0]	**5.0 [5.0]**	**3.0 [6.0]**	**5.0 [6.0]**
RF	6.0 [7.0]	5.0 [5.5]	3.0 [10.0]	4.0 [8.0]

The four models on the Fugl-Meyer Assessment and Action Research Arm Test scores on the conventional and experimental groups, FMA_C, FMA_R, ARAT_C, and ARAT_R, respectively, are presented

ARAT_C: Model with outcome Action Research Arm Test and trained on the Conventional group; ARAT_R: Model with outcome Action Research Arm Test and trained on the Robotic group; FMA_C: Model with outcome Fugl-Meyer Assessment and trained on the Conventional group; FMA_R: Model with outcome Fugl-Meyer Assessment and trained on the Robotic group; kNN: k-Nearest Neighbours; LASSO: Lasso regularised regression; RF: Random Forest; SVR: Support Vector Regression

Values in bold represent the results of the best-performing algorithm on the validation set for each model

analyses. Nonetheless, the results underscore the need for future investigation.

In summary, this study identified distinct predictor patterns for motor recovery in conventional versus robot-assisted rehabilitation of the upper limb in post-stroke patients. While our previous study on the same population has identified individual predictors of functional recovery after stroke, the present work extends these findings by evaluating multiple baseline characteristics simultaneously within a machine learning framework. This approach, based on nested cross-validation, provides a more robust estimate of predictive performance and highlights the relative contribution of different factors, thereby complementing traditional statistical analyses and supporting the development of individualized rehabilitation strategies.

Limitation

The present study has important limitations that should be acknowledged. First, the relatively small sample size substantially restricts the statistical power of our analyses and limits the generalizability of the findings. With such a sample, the robustness of the predictive models remains uncertain, and subtle but clinically relevant associations may have been overlooked. A larger dataset would not only enhance the generalizability of findings but also enable more detailed error analyses, allowing patterns in model misclassification to be systematically explored.

Another limitation is the absence of objective kinematic data from robotic devices [46, 47], which were only available for participants in the experimental group. The inclusion of baseline robotic assessments for all participants would allow for a more comprehensive, model-based integration of movement data.

It should also be emphasized that the present analysis is secondary and based on data from a randomized controlled trial originally designed to evaluate the efficacy of robotic rehabilitation, rather than to identify predictors of recovery. As a result, certain potentially important dimensions were not included in the trial protocol. Most notably, no cognitive assessments were performed, despite growing evidence that cognitive functioning plays a critical role in recovery trajectories and can substantially influence both rehabilitation engagement and outcomes. The absence of cognitive measures prevented us from accounting for this potentially moderating factor in our predictive models.

In terms of intervention, although multiple devices were employed in the present study, they were all confined to a predefined set of upper-limb robotic rehabilitation tools, which in the parent study were conceptualized and delivered as a coherent intervention strategy based on a battery of robotic devices rather than as independent or unrelated treatments. Nevertheless, this design choice inevitably introduces a degree of intervention heterogeneity, and the results should therefore be interpreted in light of this consideration.

At the same time, the use of a battery of upper-limb robotic devices offers important advantages, as it enables the intervention to be tailored to patients with different levels and profiles of motor impairment, thereby improving the representativeness of the clinical sample and reflecting real-world rehabilitation practice.

Finally, only short- to medium-term outcomes were analyzed, leaving open the question of whether the identified predictors retain their prognostic value over longer follow-up periods.

Despite its limitations, this preliminary study provides an initial foundation for identifying predictive factors in upper limb recovery and may contribute to the development of personalized rehabilitation strategies. Moreover, the results support the development of robotic therapy programs that integrate cognitive rehabilitation elements for patients with complex clinical profiles.

Conclusions

This study lays the groundwork for developing tools to guide rehabilitation treatment selection based on individual patient characteristics. Although further development is needed, our analyses highlighted baseline clinical features most predictive of motor recovery following conventional or robot-assisted therapy. Noteworthy findings include the divergent influence of age across treatment types, the consistent importance of the FMA score across all models, and the distinct role of neglect in robotic versus conventional rehabilitation outcomes.

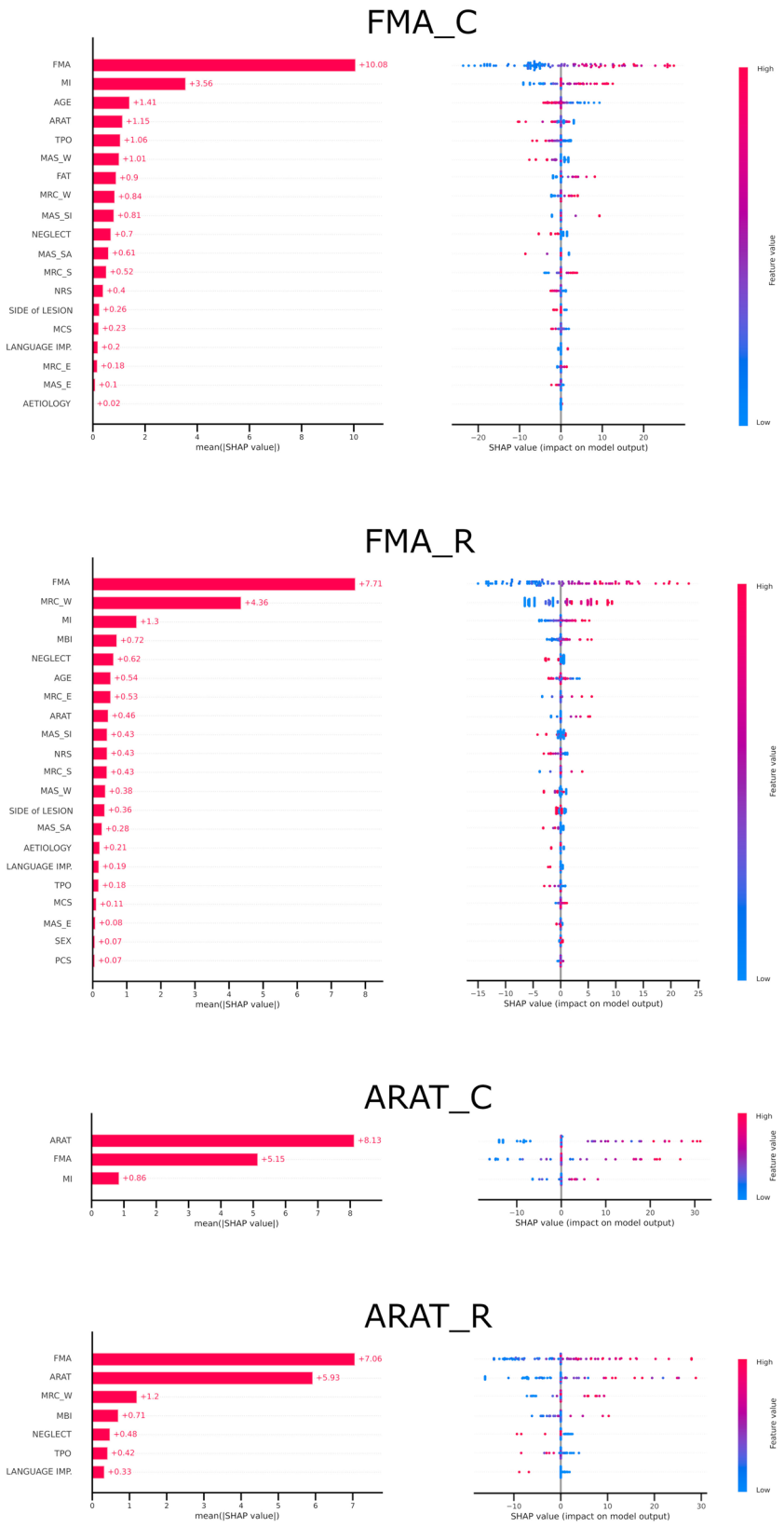


Fig. 2 Bar chart (on the left) and beeswarm plot (on the right) representing the contributions of the predictors for each model. The features never selected during the feature selection obtained null contributions and were thus excluded from the graphs. The FMA score emerged as the most influential predictor across all models. In contrast, the presence of neglect emerged as a key predictor in models developed for the experimental (robotic) group only. ARAT: Action Research Arm Test; FMA: Fugl-Meyer Assessment

Abbreviations

ARAT	Action Research Arm Test
CDSS	Clinical Decision Support Systems
FAT	Frenchay Arm Test
FMA	Fugl-Meyer Assessment
IQR	interquartile range
kNN	k-Nearest Neighbours
LASSO	Lasso regularised regression
MAE	Median Absolute Error
MBI	Modified Barthel Index
MCS	Mental Composite Score
MI	Motricity Index
ML	Machine Learning
MRC	Medical Research Council
NRS	Numerical Rating Scale
PCS	Physical Composite Score
RCT	Randomized Controlled Trial
RF	Random Forest
SHAP	Shapley Additive exPlanations
SVR	Support Vector Regression
TPO	Time Post-Onset

Author contributions

SC, MG, FC, IGA, and AM conceptualized the research. AP and AF collected the data. SC and PL analyzed the data. SC, MG, and AM prepared the initial draft. All authors edited and/or commented on the manuscript.

Funding

This work was funded by the Italian Ministry of University and Research, under the complementary actions to the NRRP "Fit4MedRob - Fit for Medical Robotics" Grant (# PNC0000007). Additionally, the study was funded by the Italian Ministry of Health under: the "Ricerca Corrente" program; the "Ricerca Finalizzata" program thorugh the project "Clinical validatiOn of an AI-based DEcision support system for Robotic upper limb rehabilitation in patients with Stroke: A CO-AIDER Study" (RF-2021-12375472) and under the frame of Transforming Health and Care Systems, THCS, (NeuroRehab4EU, GA N° 101095654 of the EU Horizon Europe Research and Innovation Programme)"

Data availability

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

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee "Comitato Etico Lazio 1" on May 6, 2021 (610/CE Lazio 1). All recruited participants provided written informed consent to participate in this study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]IRCCS Fondazione Don Carlo Gnocchi ETS, Via di Scandicci, 269, 50143 Firenze, Italy
[2]Department of Experimental and Clinical Medicine, University of Florence, Firenze, Italy

Received: 20 December 2024 / Accepted: 29 April 2026
Published online: 12 May 2026

References

1. Chang WH, Kim Y-H. Robot-assisted Therapy in Stroke Rehabilitation. J Stroke. 2013;15:174.
2. Poli P, Morone G, Rosati G, Masiero S. Robotic Technologies and Rehabilitation: New Tools for Stroke Patients' Therapy. Biomed Res Int. 2013;2013:1–8.
3. Weber LM, Stein J. The use of robots in stroke rehabilitation: a narrative review. NeuroRehabilitation. 2018;43:99–110.
4. Mehrholz J, Pollock A, Pohl M, Kugler J, Elsner B. Systematic review with network meta-analysis of randomized controlled trials of robotic-assisted arm training for improving activities of daily living and upper limb function after stroke. J Neuroeng Rehabil. 2020;17:83.
5. Pignolo L, Servidio R, Basta G, Carozzo S, Tonin P, Calabrò RS, et al. The Route of Motor Recovery in Stroke Patients Driven by Exoskeleton-Robot-Assisted Therapy: A Path-Analysis. Med Sci. 2021;9:64.
6. Franceschini M, Goffredo M, Pournajaf S, Paravati S, Agosti M, De Pisi F, et al. Predictors of activities of daily living outcomes after upper limb robot-assisted therapy in subacute stroke patients. PLoS ONE. 2018;13:e0193235.
7. Duret C, Pila O, Grosmaire A-G, Koeppel T. Can robot-based measurements improve prediction of motor performance after robot-assisted upper-limb rehabilitation in patients with moderate-to-severe sub-acute stroke? Restor Neurol Neurosci. 2019;37:119–29.
8. Leem MJ, Kim GS, Kim KH, Yi TI, Moon HI. Predictors of functional and motor outcomes following upper limb robot-assisted therapy after stroke. Int J Rehabil Res. 2019;42:223–8.
9. Lee JJ, Shin J-H. Predicting Clinically Significant Improvement After Robot-Assisted Upper Limb Rehabilitation in Subacute and Chronic Stroke. Front Neurol. 2021;12:668923.
10. Machine Learning for Healthcare Applications. Wiley. https://www.wiley.com/en-us/Machine+Learning+for+Healthcare+Applications-p-9781119791812. Accessed 28 Feb 2022.
11. Heyen NB, Salloch S. The ethics of machine learning-based clinical decision support: an analysis through the lens of professionalisation theory. BMC Med Ethics. 2021;22:112.
12. Mainali S, Darsie ME, Smetana KS. Machine learning in action: stroke diagnosis and outcome prediction. Front Neurol. 2021. https://doi.org/10.3389/fneur.2021.734345.
13. 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.
14. Deo RC. Machine Learning in Medicine. Circulation. 2015;132:1920–30.
15. Iwamoto Y, Imura T, Tanaka R, Mitsutake T, Jung H, Suzukawa T, et al. Clinical Prediction Rule for Identifying the Stroke Patients who will Obtain Clinically Important Improvement of Upper Limb Motor Function by Robot-Assisted Upper Limb. J Stroke Cerebrovasc Dis. 2022;31:106517.
16. Quattrocelli S, Russo E, Gatta M, Filoni S, Pellegrino R, Cangelmi L, et al. Integrating Machine Learning with Robotic Rehabilitation May Support Prediction of Recovery of the Upper Limb Motor Function in Stroke Survivors. Brain Sci. 2024;14:759.
17. Rahmani AM, Yousefpoor E, Yousefpoor MS, Mehmood Z, Haider A, Hosseinzadeh M, et al. Machine Learning (ML) in Medicine: Review, Applications, and Challenges. Mathematics. 2021;9:2970.
18. Shapley LS. 17. A value for n-person games. In: Kuhn HW, Tucker AW, editors. Contributions to the theory of games (AM-28), Vol. II. Princeton University Press; 1953. pp. 307–18. https://doi.org/10.1515/9781400881970-018.
19. Kalai E, Samet D. On weighted Shapley values. Int J Game Theory. 1987;16:205–22.
20. Aprile I, Germanotta M, Cruciani A, Loreti S, Pecchioli C, Cecchi F, et al. Upper Limb Robotic Rehabilitation After Stroke: A Multicenter, Randomized Clinical Trial. J Neurol Phys Ther. 2020;44:3–14.
21. Cecchi F, Germanotta M, Macchi C, Montesano A, Galeri S, Diverio M, et al. Age is negatively associated with upper limb recovery after conventional but not robotic rehabilitation in patients with stroke: a secondary analysis of a randomized-controlled trial. J Neurol. 2021;268:474–83.
22. Cecchi F, Carrabba C, Bertolucci F, Castagnoli C, Falsini C, Gnetti B, et al. Transcultural translation and validation of Fugl-Meyer assessment to Italian. Disabil Rehabil. 2021;43:3717–22.
23. Lyle RC. A performance test for assessment of upper limb function in physical rehabilitation treatment and research. Int J Rehabil Res. 1981;4:483–92.
24. World Health Organization. IFC: international classification of functioning, disability and health. 2001.
25. High AS, Macgregor AJ, Tomlinson GE, Salkouskis PM. A gnathodynanometer as an objective means of pain assessment following wisdom tooth removal. Br J Oral Maxillofac Surg. 1988;26:284–91.
26. Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther. 1987;67:206–7.
27. Paternostro-Sluga T, Grim-Stieger M, Posch M, Schuhfried O, Vacariu G, Mittermaier C, et al. Reliability and validity of the Medical Research Council (MRC)

scale and a modified scale for testing muscle strength in patients with radial palsy. J Rehabil Med. 2008;40:665–71.
28. Apolone G, Mosconi P. The Italian SF-36 Health Survey: translation, validation and norming. J Clin Epidemiol. 1998;51:1025–36.
29. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol. 1989;42:703–9.
30. Bohannon R. Motricity Index Scores are Valid Indicators of Paretic Upper Extremity Strength Following Stroke. J Phys Therapy Sci - J PHYS THER SCI. 1999;11:59–61.
31. Heller A, Wade DT, Wood VA, Sunderland A, Hewer RL, Ward E. Arm function after stroke: measurement and recovery over the first three months. J Neurol Neurosurg Psychiatry. 1987;50:714–9.
32. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, et al. Missing value estimation methods for DNA microarrays. Bioinformatics. 2001;17:520–5.
33. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O et al. Scikit-learn: machine learning in python. Mach Lear Python. 6.
34. Cawley GC, Talbot NLC. On over-fitting in model selection and subsequent selection bias in performance evaluation.
35. Raschka S, MLxtend. Providing machine learning and data science utilities and extensions to Python's scientific computing stack. J Open Source Softw. 2018;3:638.
36. van der Lee JH, Beckerman H, Lankhorst GJ, Bouter LM. The responsiveness of the Action Research Arm test and the Fugl-Meyer Assessment scale in chronic stroke patients. J Rehabil Med. 2001;33:110–3.
37. Page SJ, Fulk GD, Boyne P. Clinically Important Differences for the Upper-Extremity Fugl-Meyer Scale in People With Minimal to Moderate Impairment Due to Chronic Stroke. Phys Ther. 2012;92:791–8.
38. Duncan PW, Propst M, Nelson SG. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. Phys Ther. 1983;63:1606–10.
39. Wood-Dauphinee SL, Williams JI, Shapiro SH. Examining outcome measures in a clinical study of stroke. Stroke. 1990;21:731–9.
40. Platz T, Pinkowski C, van Wijck F, Kim I-H, di Bella P, Johnson G. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multicentre study. Clin Rehabil. 2005;19:404–11.
41. Fugl-Meyer AR, Jääskö 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.
42. Van der Lee JH, De Groot V, Beckerman H, Wagenaar RC, Lankhorst GJ, Bouter LM. The intra- and interrater reliability of the action research arm test: a practical test of upper extremity function in patients with stroke. Arch Phys Med Rehabil. 2001;82:14–9.
43. Nijland R, van Wegen E, Verbunt J, van Wijk R, van Kordelaar J, Kwakkel G. A comparison of two validated tests for upper limb function after stroke: The Wolf Motor Function Test and the Action Research Arm Test. J Rehabil Med. 2010;42:694–6.
44. Feingold-Polak R, Elishay A, Shahar Y, Stein M, Edan Y, Levy-Tzedek S. Differences between young and old users when interacting with a humanoid robot: a qualitative usability study. Paladyn J Behav Rob. 2018;9:183–92.
45. Aprile I, Guardati G, Cipollini V, Papadopoulou D, Mastrorosa A, Castelli L, et al. Robotic rehabilitation: an opportunity to improve cognitive functions in subjects with stroke: an explorative study. Front Neurol. 2020. https://doi.org/10.3389/fneur.2020.588285.
46. Germanotta M, Cruciani A, Pecchioli C, Loreti S, Spedicato A, Meotti M, et al. Reliability, validity and discriminant ability of the instrumental indices provided by a novel planar robotic device for upper limb rehabilitation. J Neuroeng Rehabil. 2018;15:39.
47. Germanotta M, Gower V, Papadopoulou D, Cruciani A, Pecchioli C, Mosca R, et al. Reliability, validity and discriminant ability of a robotic device for finger training in patients with subacute stroke. J Neuroeng Rehabil. 2020;17:1.

Publisher's Note

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