

Beyond proportional recovery in wake-up stroke: unsupervised recovery clusters based on the NIHSS

Andrea Zanola^{1,2*}†, Antonio Luigi Bisogno^{1,2,3†},
Veronika Vadinova^{1,3}, Götz Thomalla⁴, Bastian Cheng⁴,
Manfredo Atzori^{1,2,5}, Maurizio Corbetta^{1,2,3}

¹*Department of Neuroscience, University of Padua, Padua, 35128, Italy.

²*Padua Neuroscience Center, University of Padua, Padua, 35128, Italy.

³Veneto Institute of Molecular Medicine, Padua, 35128, Italy.

⁴Department of Neurology, University Medical Center

Hamburg-Eppendorf, Hamburg, 20246, Germany.

⁵Information Systems Institute, University of Applied Sciences Western
Switzerland (HES-SO Valais), Sierre, 3960, Switzerland.

*Corresponding author(s). E-mail(s): andrea.zanola@studenti.unipd.it;

Contributing authors: antonioluigi.bisogno@studenti.unipd.it;

veronika.vadinova@unipd.it; thomalla@uke.de; b.cheng@uke.de;

manfredo.atzori@unipd.it; maurizio.corbetta@unipd.it;

†These authors contributed equally to this work.

Abstract

Post-stroke rehabilitation is a complex process influenced by several neurophysiological factors. The recovery is traditionally predicted based on initial impairment using linear models. The Proportional Recovery Rule (PRR), developed on the Fugl-Meyer scale, has even been proposed as a therapeutic target. In this framework, patients are classified as “fitters” or “non-fitters”, though this distinction depends on the methodology used. Additionally, issues like mathematical coupling and ceiling effects on clinical scales could raise concerns about the validity of these models. To overcome these issues, Repeated Spectral Clustering (RSC) was used to identify recovery patterns based on NIHSS. We selected 201 patients from the WAKE-UP trial, all moderately impaired at onset and still impaired at 22/36 hours. Clustering was performed using a similarity matrix based on pairwise absolute differences between recovery ratios, calculated from 22/36 hours to 90 days post-stroke. Cluster differences were tested with prognostic factors, including lesion volume, side, treatment, and the Heidelberg scale. The PRR was

fit to the cohort for comparison with clustering results. The linear fit reproduced findings consistent with the literature, such as a correlation of $\rho(x, \Delta) = .73$ and an average recovery ratio of 70% for the “fitters”. RSC grouped patients into six recovery clusters: C_0 (full recovery), C_1 (above average), C_2 and C_3 (average, PRR-aligned), C_4 (below average), and C_5 (deterioration). NIHSS scores in most patients declined non-proportionally. Lesion volume was not significantly different across clusters, while left-sided strokes were higher in low recovery clusters. Patients with a recovery ratio ≥ 0.3 within two weeks mostly fell into favorable clusters (C_0 – C_3), covering $\approx 90\%$ of such cases. The identified clusters provide a refined view of stroke recovery following wake-up stroke. Clustering better captures patient similarities, enabling the assessment of neurophysiological differences between groups and supporting tailored interventions.

Keywords: Stroke, clustering, longitudinal clustering, NIHSS, recovery ratio

1 Introduction

Recovery after stroke is a complex and multifaceted process that involves the interplay of neurological, physiological, and environmental factors. Longitudinal datasets offer a valuable opportunity to investigate recovery trajectories over time to identify prognostic factors and therapeutic targets. Traditional approaches to studying stroke recovery have relied heavily on evaluating outcomes in acute versus chronic phases. Although these methods yield important insights, they often fall short in capturing the heterogeneity of recovery patterns between individuals. A key concept in this domain is the categorization of patients into “fitters” vs. “non-fitters”, based on how closely their recovery trajectories align with established mathematical models [1, 2].

However, this dichotomy and related framework have notable limitations. First, from a methodological perspective, some authors [3, 4], have raised concerns about the issue of mathematical coupling, since proportional recovery is calculated based on acute vs. acute minus chronic scores, which inherently links the two variables. Second, widely described in recent literature [5], the ceiling effect is present in commonly used clinical scales. Essentially, this reflects a reduced sensitivity to the variability currently observed, implying that there may be additional variability that remains undetected. This leads to the possibility that unexplained variability is underestimated [3]. Another major concern, from a physiological standpoint, is that this theoretical framework has been interpreted to suggest that normal recovery entails regaining approximately 70% of the initial deficit. Such an interpretation risks oversimplifying the complex and individualized mechanisms of stroke recovery. Most importantly, this proportional recovery has often been interpreted as ground truth and, in certain cases, can be considered a valuable therapeutic target to obtain [6]. Unfortunately, strong neurophysiological evidence in this regard is still lacking, and most authors have focused on specific deficits and clinical scales (i.e. Fugl Meyer [7] and motor deficits). The literature reflects a lively debate, with some studies arguing for the utility of the “fitters” vs. “non-fitters” categorization and others highlighting its constraints and potential biases [4, 8–10].

In light of these challenges, the present study aims to provide a complementary approach to describe the recovery of stroke patients by using unsupervised techniques. Classical methodologies are focused on the statistical evaluation of the goodness-of-fit of a model for the whole population, assessing how one or more independent variables describe one or more dependent ones (e.g., how acute score describes the chronic one). Clustering, one of the main unsupervised techniques, is particularly effective in finding groups of subjects, i.e. phenotypes, within the analyzed population. Specifically, while statistical techniques focus primarily on correlation between variables, clustering instead focuses on grouping subjects based on their similarity to one another. From this perspective, clustering can be seen as an analysis of the “subject correlation” representing a complementary, but opposite, approach to the conventional analysis of feature correlation [11, 12]. As mentioned in Yang *et al.* [13], clustering is one of the most useful methods for analyzing similarities between patients for precision medicine. It groups patients into subsets that are clinically meaningful and can be used for a variety of tasks, such as policy making and personalized therapies. Kim *et al.* [14] stresses that finding similar clusters among stroke patients can be helpful from a medical perspective, as it can lead to the discovery of new patterns and ways to manage stroke more effectively.

In this work, we apply clustering algorithms tailored for biomedical data analysis [11] to the recovery ratio (RR) [15, 16] of patients, based on the National Institute of Health Stroke Scale (NIHSS) [17], in a large longitudinal sample of wake-up [18] stroke patients. The recovery ratio used for clustering is calculated using the total NIHSS scores measured in the subacute phase (22/36 hours) to 90 days. This approach presents several advantages over the existing methods. Clustering allows for the identification of diverse recovery patterns without presupposing specific models, in a data-driven manner. Moreover, it helps mitigate the issue of mathematical coupling inherent in traditional proportional recovery calculations by avoiding any model fitting in the statistical sense.

2 Methods

2.1 Dataset

WAKE-UP [18] is an investigator-initiated, multicenter, randomized, double-blind, placebo-controlled clinical trial involving stroke patients with an unknown time of onset. Patients were randomly assigned to receive intravenous alteplase (rtPA) or placebo. The study was registered under the identifier NCT01525290. The dataset includes 503 patients, with NIHSS [17] scores measured at four time points: onset, 22/36 hours, 7 days, and 90 days.

For the purposes of this study, patients with a Level-of-Consciousness Vigilance (LOC-V) score of zero and complete NIHSS data at all four time points (no missing values) were selected. Patients were further selected based on stroke severity, as measured by the total NIHSS score at onset. Only patients with a total NIHSS score greater than 4 (moderate stroke severity) were included; patients with a total NIHSS score of zero at 22/36 hours after the event were further excluded from the analysis. These selection criteria were designed to reduce the inherent ceiling effect of the

118 NIHSS and focus on patients with persistent deficits. As a result, the final analysis
119 focused on 201 patients who were still impaired at 22/36 hours. In the supplementary
120 material Section 1 a summary schema of the inclusion criteria and the consequent
121 effect on the size of the population is reported.

122 For each subject, general demographic information (age, sex and treatment group),
123 lesion characteristics (lesion side, lesion volume, large vessel occlusion and lacunes),
124 the 90-day modified Rankin Scale (mRS) [19], the Heidelberg hemorrhage classifica-
125 tion [20], and the NIHSS with all 15 sub-items are available. The mRS scale assesses
126 disability and dependence after a stroke, while the Heidelberg classification categorizes
127 intracranial hemorrhages following ischemic stroke. The demographic characteristics
128 of the selected patients are summarized in Table 1. A χ^2 independence test [21] did not
129 reveal statistically significant differences in the distribution of male and female par-
130 ticipants between the treated and non-treated groups ($\chi^2(1, 201) = 0.44$, $p = 0.51$).
131 Moreover, there are no significant differences in terms of median age between the four
132 groups, as assessed by the Kruskal-Willis test [22] ($H(3) = 3.53$, $p = .32$). For this
work, a significance level of $\alpha = 0.01$ was chosen.

Group \ Sex	Female	Male	
rtPA treated	36 (63.6 $\pm$ 12.1 years)	58 (64.8 $\pm$ 10.5 years)	94
non-treated	47 (67.5 $\pm$ 10.5 years)	60 (65.4 $\pm$ 11.6 years)	107
	83	118	201

Table 1 Demographical variables. The table reports the number of subjects divided by the two attributes, group (rtPA treated vs. non-treated) and sex (female vs. male). It also reports the mean ages for each group, with the standard deviation in years.

133
134 Regarding the lesion side, among the 201 subjects, 121 (60%) had left side lesions,
135 75 (37%) had right side lesions, and 5 (3%) had bilateral lesions. The average lesion
136 volume, measured at hospital admission using apparent diffusion coefficient MRI maps,
137 was 8.3 ± 12.1 mL (mean $\pm$ standard deviation). The distribution of cardiovascular
risk factors for the 201 selected subjects is reported in Table 2.

Cardiovascular Risk Factor	Non-Present	Present	Missing Value
Arterial Hypertension	102 (51%)	98 (49%)	1
Atrial Fibrillation	168 (84%)	30 (15%)	3
Transient Ischemic Attack	192 (96%)	5 (2%)	4
Previous Ischemic Stroke	176 (88%)	25 (12%)	0
Hypercholesterolemia	121 (62%)	74 (38%)	6
Diabetes mellitus type II	162 (82%)	36 (18%)	3

Table 2 The cardiovascular risk factors of the 201 subjects selected for this study are presented. Values represent the number of subjects.

2.2 Fitters versus non-fitters calculation

The chronic and sub-acute scores can be plotted against each other, as shown in Figure 1, Panel A. However, the most commonly used plot in the literature is shown in Panel B, which highlights the ceiling effect [4], indicated by the green line. The ceiling effect occurs when data points reach the maximum limit possible on the measurement scale, i.e. here 100% recovery. The correlation between the two variables, measured using the Spearman correlation, is strong and highly significant ($\rho_s(201) = .73$, $p < .001$). This value is consistent with similar findings reported in the literature [23]. A linear model can be used to fit the change in scores (sub-acute score minus chronic score) as the dependent variable, with the sub-acute score as the independent variable. In this work, the linear model is fitted without the intercept; the medical and statistical reasons behind can be found in the supplementary material Section 2; the proportional recovery rule (PRR) is shown in Figure 1, Panel B, with a dashed black line. In both

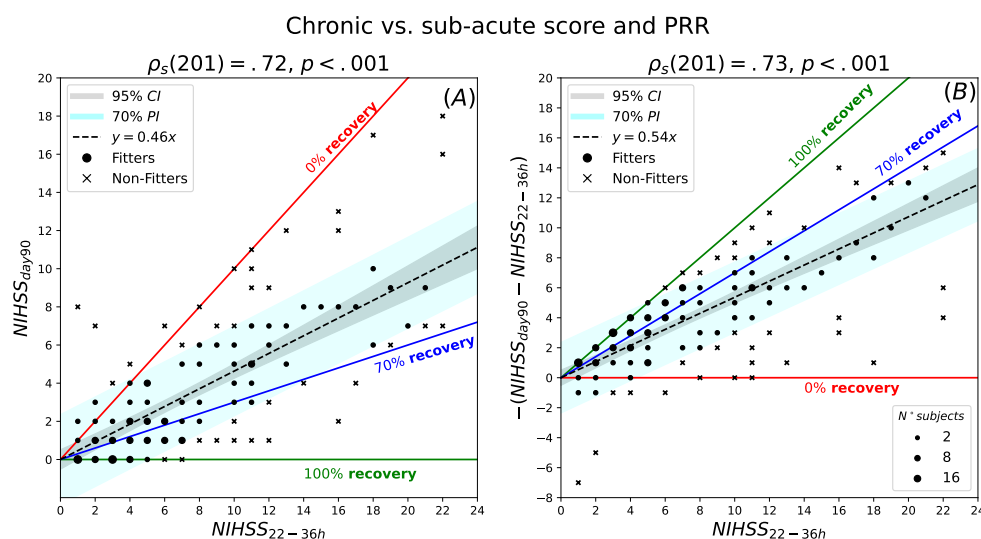


Fig. 1 Chronic vs. sub-acute score and PRR. Panel A, shows the total NIHSS at 90 days vs. total NIHSS at 22/36 hours (chronic vs. sub-acute). Panel B, shows the PRR i.e. the total NIHSS change vs. total NIHSS at 22/36 hours (change vs. sub-acute). In both panels, the title reports the Spearman's correlation. In both panels the red, blue and green lines indicate 0%, 70% and 100% recovery respectively. The dashed line represents the fitted linear model (PRR), in grey the 95% confidence interval (CI) and in light blue the 70% ($\approx 1\sigma$) prediction interval (PI). Subjects in both panels are marked with a circle (o) and identified as "fitters" if they stay within the prediction interval (light blue band), and with a cross (x) otherwise ("non-fitters"). The size of the markers, represent the number of subjects having that combination of values, as indicated by the corresponding legend.

panels of Figure 1, the 95% confidence interval (CI), representing the uncertainty of the fitted model, is shown in gray; the 70% prediction interval (PI) instead, which estimates where future observations are expected to fall, is shown in light blue. Both intervals were calculated following the approach described by Howell [24], 2007; for

further details, see the supplementary material Section 3. The prediction interval can be used to categorize the population into “fitters” and “non-fitters”. In both panels A and B of Figure 1, “fitters” are marked with a circle (o), while “non-fitters” with a cross (x). The prediction interval contains 70% of the subjects, who show an average recovery ratio of $\bar{R}R_{\text{fitters}} = 0.69 \pm 0.27$. This is in line with the well-established literature on the proportional recovery rule in stroke [23].

However, the distinction between “fitters” vs. “non-fitters” (or outliers) is far from being unique. Different techniques dichotomize the population in different ways, highlighting different aspects of the data. An notable effort has been made to compare different techniques in Kundert *et al.* [5], 2019. In order to briefly review and compare these methods, they have been applied to the population under study. Seven methods to define “non-fitters” or outliers have been used and reported in Figure 2; methodological details can be found in the supplementary material Section 4. The distribution of the recovery ratio for the entire population is also reported in the same figure. Interestingly, this distribution does not follow a normal distribution centered around 70%.

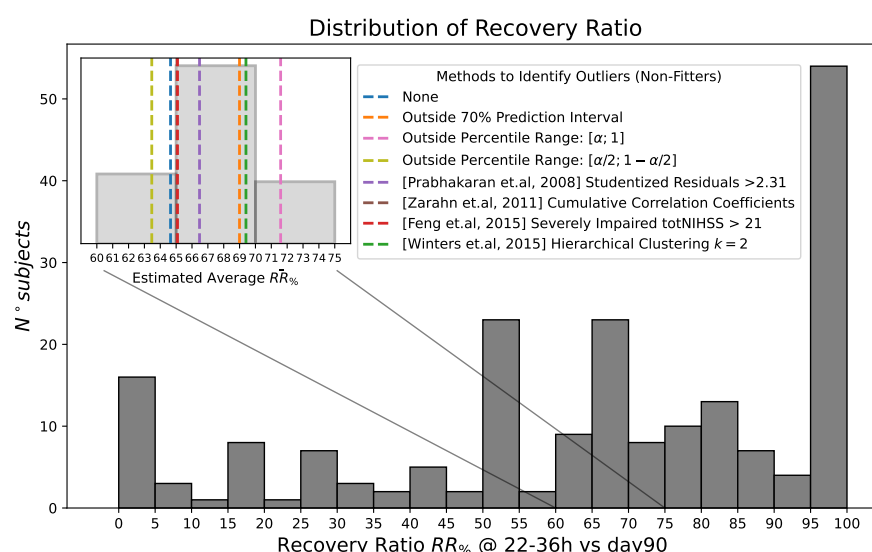


Fig. 2 Recovery ratio distribution. The main histogram depicts, in gray, the distribution of recovery ratios for the entire population. The inset panel provides a magnified view focusing on the recovery ratio range between 60% and 75%, highlighting the average recovery ratio estimates obtained from various methods. Different colors correspond to different estimation approaches, as indicated in the legend. For methodological details, refer to the supplementary material Section 4.

The variability in the average recovery ratios arises from the different criteria and algorithms applied between methods. Classifying the same patient as a “fitter” or a “non-fitter” depending on the method used, highlights the inherent limitations of an arbitrary dichotomization. In addition, Figure 2 clearly shows that the proportional recovery rule does not represent the typical amount of recovery (i.e. the recovery ratio

is not normally distributed). This work aims to move beyond this classification and advocates for a more nuanced, data-driven approach: clustering.

2.3 Recovery Ratio

A recovery ratio index (RR) of the form (acute score - chronic score)/acute score, was used to study the degree of recovery, as described in previous work [15], [16]. Let's denote as S , a score or a measure of choice, aimed at quantifying the impairment; then the recovery ratio index is defined as

$$RR = \frac{S^{t_1} - S^{t_0}}{\bar{S}_{CTL} - S^{t_0}} = 1 - \frac{S^{t_1}}{\bar{S}_{CTL} - S^{t_0}} \quad (1)$$

In Equation 1, S^{t_0} represents the sub-acute score measured at an initial time point t_0 , while S^{t_1} represents the chronic score measured at a later time point t_1 . The term $\bar{S}_{CTL}$ denotes the average score for the control population. When considering the total NIHSS as the score S , $\bar{S}_{CTL}$ equals zero, since healthy individuals have a total NIHSS of zero, which simplifies the formula. This formula applies only when the sub-acute score is greater than zero, i.e., only for impaired patients. Since some patients exhibit a worse condition at the chronic time point than at the sub-acute stage (deterioration), the recovery ratio can be negative in these cases. For the selected sample of 201 subjects, the average recovery ratio is $\bar{RR}_{\text{negative}} = 0.58 \pm 0.67$. Since the recovery ratio is intended to range between 0 and 1, the negative values are set to zero. With this convention, the average recovery ratio becomes $\bar{RR}_{\text{all}} = 0.65 \pm 0.31$.

2.4 Clustering

Clustering is a popular unsupervised technique, particularly useful in the identification of phenotypes [14], which can be effectively used to find groups of patients with common recovery. The following procedure was applied to cluster the data according to the recovery ratio.

1. The total NIHSS score at 22/36 hours (sub-acute score) and the total NIHSS score at 90 days (chronic score) are used to calculate the recovery ratio (RR) for each subject, as defined in Equation 1.
2. The distance between subjects is calculated using the Manhattan distance on their recovery ratios. Between two subjects i and j , the distance is $d_{ij} = |RR_i - RR_j|$. Distances are divided with respect to the maximum value found in the dataset.
3. The matrix $D_{(ij)} = d_{ij}$ is obtained from the pairwise distances between patients. After that, the similarity matrix ($W = \mathbb{I} - D$) is calculated and used as input for the clustering algorithm.
4. The application of the Repeated Spectral Clustering (RSC) [11] algorithm to the similarity matrix W allows the identification of groups of patients with similar recovery ratios.

The RSC algorithm is a blend of spectral clustering [25], known to work well with similarities matrices representing graphs, and consensus clustering [26], known to be

powerful since it allows to deal with the variability of results that stem from random initializations. In brief, RSC takes the matrix W as input to the spectral clustering procedure. Instead of running k -means just once, it is performed N times to account for randomness in centroid initialization. Across these N iterations, the number of times two patients are assigned to the same cluster is recorded. The more frequently two subjects are clustered together, the stronger the evidence that they should remain in the same cluster in the final assignment. The final cluster assignment is based on the co-occurrence matrix C , where each element C_{ij} represents the number of times subjects i and j were clustered together across all runs. For a complete description of the algorithm, including technical details and its application to NIHSS scores, refer to Tshimanga *et al.* [11], 2025.

The optimal number of clusters is determined by calculating the maximum spectral gap for different cluster configurations (i.e. different k) [11]. The elbow criterion is a simple and effective way to find the optimal number of clusters. In particular, for $k = 6$, the maximum spectral gap is at the end of a high plateau, indicating well-separated groups, while for $k = 7$, it decreases sharply, indicating subjects that are worst separated in groups. Thus, $k = 6$ is chosen as the optimal number of clusters. For more details, see the supplementary material Section 5.

The clustering algorithm uses the recovery ratio, which is derived from the total NIHSS at two distinct time points. The total NIHSS constitutes the aggregate of the 15 individual items that comprise the scale. All other features, including general subject information, lesion information, 90-day mRS, and the Heidelberg hemorrhage classification, were excluded from the clustering process. Instead, these features were used as validation features to characterize the resulting clusters and assess differences between them. The numerical and ordinal features between the six groups have been compared with the Kruskal-Wallis [22] test as omnibus, and then the Mann-Whitney U [27] as post hoc for pairwise comparisons. Boolean or categorical features have been compared between the six groups with the χ^2 independence test as omnibus and post hoc for pairwise comparisons. The p-values of the omnibus tests, as well as those of the post hoc test, were corrected for multiple comparisons using the FDR Benjamini-Hochberg correction [28]. The statistics used in this work to evaluate the effect size of the aforementioned tests are reported in the supplementary material, Section 6.

2.5 Chronic disability and early recovery

The early recovery ratio, calculated between 22/36 hours and 7 days as described in Equation 1, was also considered in this study. The 90-day mRS was dichotomized into two groups: scores ≤ 2 , and scores greater than 2. This threshold is clinically relevant, as an mRS score of 2 indicates a mild disability that does not interfere with daily activities. To examine the relationship between early recovery and long-term outcomes, we compared the number of subjects with mRS scores greater than 2 versus those with scores of 2 or less across different recovery intervals. The 95% confidence intervals for these proportions were estimated using binomial statistics and are shown in Figure 5, Panel A. Additionally, the percentage of subjects in each mRS group across recovery intervals is presented in Figure 5, Panel B.

3 Results

3.1 Clustering Recovery Ratio

The repeated spectral clustering (RSC) algorithm, as described in [section 2](#), uses exclusively the recovery ratio defined in [Equation 1](#). This variable is reported against the total NIHSS at 22/36 hours (sub-acute) in Panel A of [Figure 3](#). Although clustering is performed on the similarity matrix W , as described in [subsection 2.4](#), the visible effect of the algorithm is horizontal separation between subjects. This allows to have clusters with a defined recovery ratio with low variation. For example, cluster C_0 , which contains all subjects with a perfect recovery ratio of 100%, has a null standard deviation; C_5 instead, is the cluster with the worst mean recovery ratio which is close to 1%. Within this group, there are all subjects with a chronic score worse than sub-acute one. At the population level, the sub-acute score is significantly but moderately associated with the recovery ratio ($\rho_s(201) = -.34$, $p < .001$).

The bottom part of [Figure 3](#) illustrates the difference between how patients are grouped using a clustering approach (Panel B) versus how they are divided in “fitters” and “non-fitters” using the traditional statistical method (Panel C). Interestingly, cluster C_3 closely follows the fitted linear model, shown by the black dashed line.

Panel D of [Figure 3](#) shows the proportional recovery rules obtained fitting the subjects for each cluster. As reported in the legend, there is a strict correspondence between the average RR of each cluster and the angular coefficients of these lines. The 70% prediction interval is reported for every fitted line. Looking at the prediction intervals of each cluster, these models discriminate subjects at one standard deviation, when the sub-acute NIHSS score is greater than six. The cluster C_5 (in brown) overlaps with the cluster C_4 (in violet), probably due to the few subjects assigned to the first (19), and consequently a broad prediction interval. Moreover, cluster C_5 is clearly associated with people who actually become more impaired at 90 days compared to 22/36 hours (the intercept is -1.9). Only for the cluster C_5 , a linear model with the intercept is significantly better than one without ($F(1, 17) = 9.08$, $p < .01$, $f^2 = 0.53$).

In conclusions, while the PRR gives a methodological way to identify “fitters” vs. “non-fitters”, the clustering-based procedure gives rise to a data-driven family of groups, sharing similar recovery. As a consequence, subjects are not distinguished anymore between “fitters” vs. “non-fitters”, but rather between levels of recovery. In summary, C_0 (blue) has perfect recovery, C_1 (orange) has above average recovery, C_2 (green) and C_3 (red) have average recovery (in line with the expectations of the PRR), C_4 (violet) has below average recovery, and finally C_5 (brown) has no recovery.

3.2 Differences between Clusters

The clustering algorithm uses only the recovery ratio, which is an efficient way to group patients who recover similarly together. Concerning the other features that have not been used by the algorithm, [Figure 4](#) gives a useful insight.

Panel A of [Figure 4](#) shows the linear relation between the lesion volume (expressed in mL) and the sub-acute score; points are colored by cluster assignment. The two variables are significantly correlated ($\rho_s(201) = .40$, $p < .001$), but there is

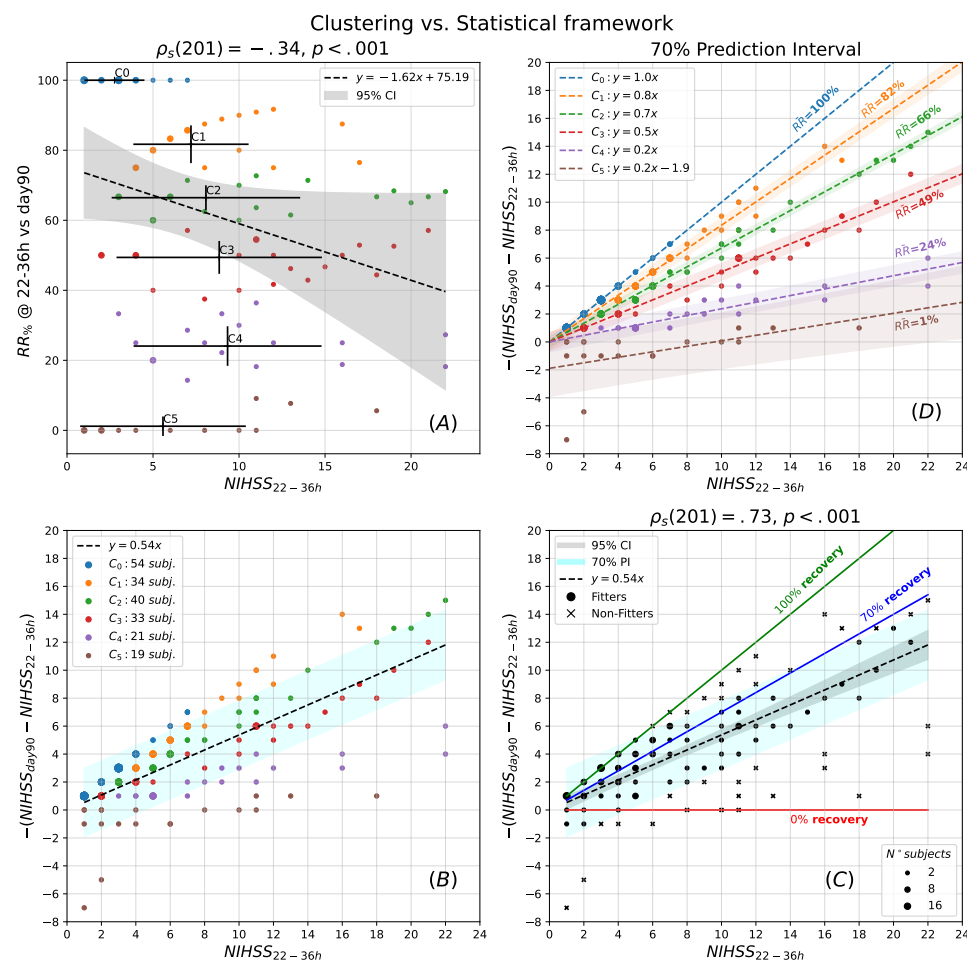


Fig. 3 Clustering results. Panel A, shows the recovery ratio (RR) vs. the total NIHSS at 22-36h, with colors indicating different clusters. Error-bars are ported for both variables and for each cluster. In gray is reported the 95% confidence interval (CI). Panel B and C are shown in order to compare how patients are colored based on cluster results vs. how they are labeled as “fitters” and “non-fitters”. Panel C is equal to Panel B in Figure 1, and visualized here for convenience. Finally Panel D, shows the six fitted linear models (proportional recovery rules) for the subjects contained in each cluster. The median recovery ratio for each cluster and the 70% prediction interval (PI) for every model is reported as well. The size of the markers in each of the panels represents the number of subjects having that combination of values, as indicated by the corresponding legend in Panel C.

no significant association between lesion volumes and clusters (Kruskal-Willis test, $H(5) = 7$, $p = .22$). In other words, there is no significant relation between lesion volume at admission and recovery measured at 90 days.

Panel B of Figure 4 represents the distribution of the lesion side (left, right, bilateral) in the obtained clusters. The clusters obtained with RSC show significant differences in lesion laterality ($\chi^2(5, 201) = 25.6$, $p = .004$, $V = .25$). In particular C_0 includes more right lesions than expected, while C_3 more left lesions (pairwise

post hoc test, $\chi^2(1,87) = 14.8$, $p = .009$, $V = .41$). In other words, the sides of the lesions are not randomly divided by clusters. Subjects with perfect recovery (C_0) have significantly more right than left lesions; conversely, patients in clusters with average recovery (C_3) have significantly more left than right lesions.

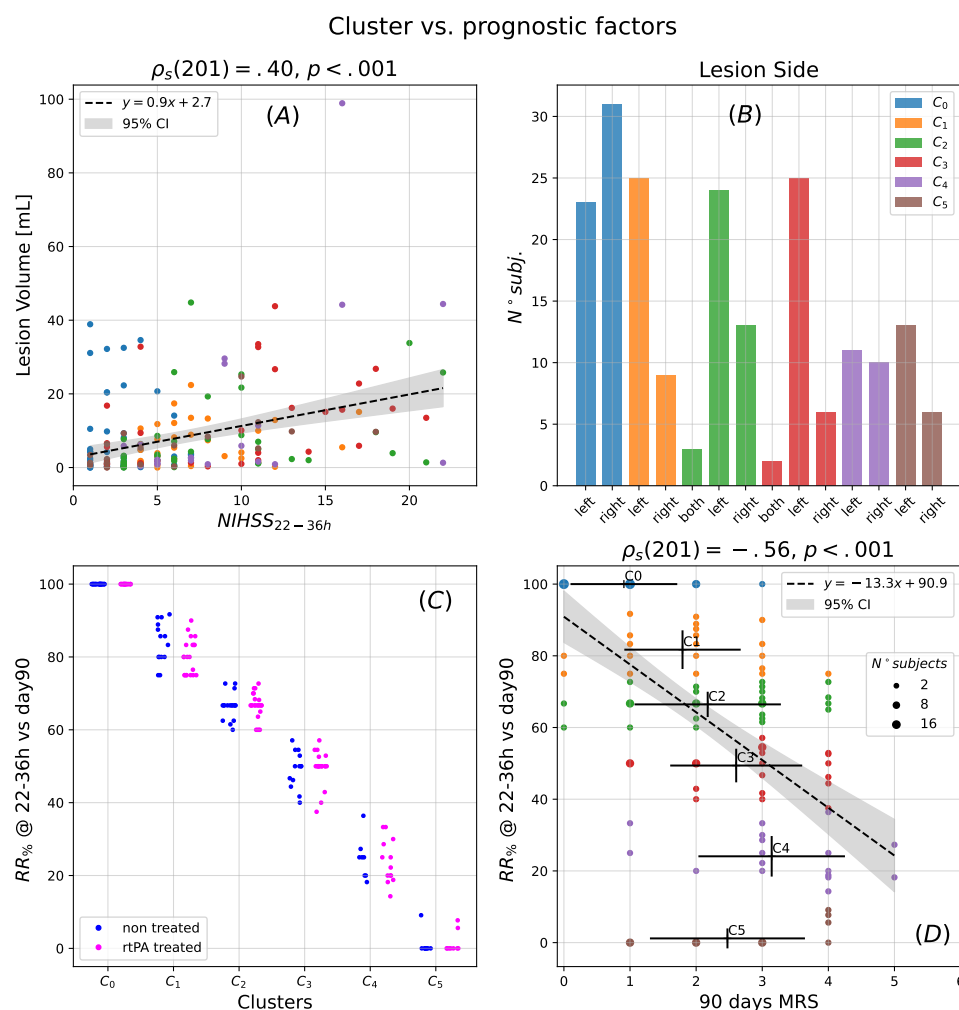


Fig. 4 Cluster's features. Panel A investigates the relation between lesion volume and total NIHSS at 22/36 hours. In both panels A and B, subjects (dots) are colored by cluster assignment, the dashed lines are the linear fitted-models, while the 95% confidence intervals (CI) are colored in gray. The lateralization distribution of the lesions (i.e. left, right or both) for each cluster is shown in panel B. Panel C shows the RR for the patients of each cluster, divided between rtPA treated group (magenta) and non-treated group (blue). Finally, Panel D shows instead the relation between the recovery ratio (RR) vs. 90 day MRS (modified rankin scale). Scatter size, represent the number of subjects having that value, as indicated by the corresponding legend in Panel D.

In Figure 4 Panel C, the recovery ratio distribution is reported for each group, where patients have been separated between treated with rtPA (magenta) and non-treated (blue). As can be seen in Figure 4, there are no significant differences between the distributions of the two groups for each cluster, and there is no significant association between treatment and clusters ($\chi^2(5, 201) = 5.7, p = .33$). Therefore, while rtPA improves NIHSS and mRS scores acutely and at 3 months (see supplementary material Section 7, Figure 3), data show that patients treated with rtPA have the same recovery trajectories than patients non treated with rtPA after the first 24 hours from the event. Finally, no significant differences were found between the clusters with respect to the Heidelberg hemorrhagic classification.

The recovery ratio and the 90 days mRS are significantly correlated ($\rho_s(201) = .56, p < .001$), and there is an association between the RSC clusters and motor abilities at 90 days (Kruskal-Willis' test, $H(5) = 74.3, p < .001, \eta^2 = .36$). Post hoc pairwise comparisons, conducted using the Mann-Whitney U test, reveal significant differences between C_0 and all other clusters, C_1 versus C_3 and C_4 , and C_2 versus C_4 . Finally, no cardiovascular risk factor is associated with a specific cluster, and we found no significant differences between clusters, both at $\alpha = 0.01$.

3.3 Prediction of long-term recovery by early recovery ratio

In addition, we investigated whether cluster identity, as well as overall outcome of each patient was related to initial recovery. This analysis is particularly interesting from a clinical perspective and provides valuable prognostic insights for the patient and caregivers during hospitalization. We tested the relation between disability at 90 days and the early recovery ratio measured between 22/36 hours and 7 days, in relation to the identified clusters. The results are reported in Figure 5. In Panel A, it can be observed that a recovery ratio between -0.1 and 0.3 does not provide sufficient information to predict whether chronic disability will be severe or not (random chance level). In contrast, an early recovery ratio greater than 0.3 substantially increases the probability of achieving a mild disability. Panel B, shows the same result from a clustering point of view. A recovery ratio between -0.1 and 0.3 does not provide sufficient information to determine which recovery cluster the patient will belong to. The probability of being assigned to a group characterized by above average recovery (C_0, C_1 or C_2) is 56% and 67% for the second and third ranges, and increases to 72% and 97% for the last two.

4 Discussion

In this work, we applied an unsupervised learning technique to the study of stroke recovery. Repeated spectral clustering (RSC), a recently developed procedure to tailor biomedical data heterogeneity [11], was applied to the recovery ratio based on total NIHSS in a large longitudinal sample of stroke patients. This approach grouped patients into distinct recovery clusters providing several clinical implications.

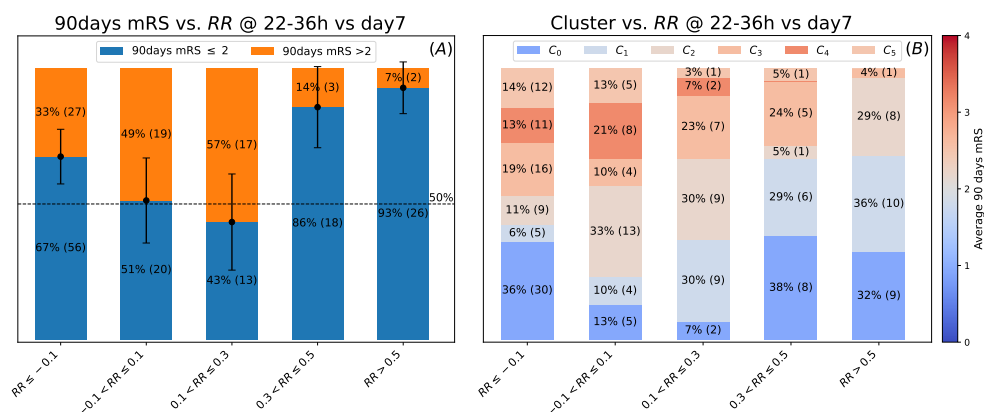


Fig. 5 90days mRS relation with early recovery. Panel A shows the amount of patients for each range of recovery ratio, for each of the two range of the mRS considered. In detail in blue when $mRS \leq 2$, while in orange > 2 . 95% confidence interval are associated to each counts (errorbars), using the binomial distribution. The dashed line indicates the random chance of 50%. Panel B instead shows the number of subjects in each range of recovery ratio for each cluster. Each block corresponds to a cluster in descending order (C_0 is always present on the bottom, in light blue). Color bar indicates the average 90 day mRS score for each cluster.

4.1 Recovery ratio and clustering

One of the main criticisms of the PRR is mathematical coupling, i.e. the fact that recovery is inherently dependent on the initial impairment. Essentially, from a mathematical point of view, fitting a linear model means establishing a linear relationship between an independent variable (the sub-acute score) and a dependent variable, the change (sub-acute score minus chronic score), which inherently includes the independent variable. This has caused much debate in the medical community, as evidenced by the extensive literature on the subject.

This methodology has been theoretically criticized by some and supported by others. The amount of recovery can be quantified as an absolute change, as previously mentioned, or in relative terms such as the recovery ratio index (RR). The definition of recovery ratio, in the framework of PRR, is induced by the variables chosen on the horizontal and vertical axes and by the linearity of the model. Instead, with clustering, the definition of recovery ratio could be carefully chosen a priori and not forced to be of the shape of Equation 1; the quantities of interest to quantify the recovery can be even nonlinearly correlated. Although it is true that RR defined by the PRR framework and used for this work has in its definition the ratio between chronic and acute scores, it is not calculated from a correlation or from the fit of a model; it simply represents a patient-specific estimate of recovery. In addition, this choice establishes a strong link between a traditional statistical framework and the clustering results obtained in this work.

Once the linear model is fitted, the identification of “fitters” vs. “non-fitters” may prove to be a great source of controversy. As shown in Figure 2, depending on the way patients are selected, different average recovery ratios can be estimated because

different patients are included in the estimate. This shows how the same theoretical framework can give subjective outcomes. Clustering the recovery ratio instead is an effective way both to group patients with similar recovery and to incorporate time into a clustering algorithm. The six clusters found exhibit different levels of recovery, with patients described by their own PRR as indicated by Figure 3, Panel D. The prediction intervals of these models are narrow and well separate subjects into clusters. This clustering scheme may provide a novel methodology capable of fully exploiting longitudinal ordinal data, such as those presented in this paper.

4.2 Clinical implications

From a clinical perspective, the described dataset includes a large sample of stroke patients with an unknown time of onset. Wake-up strokes generally show poorer outcomes at three months than known onset strokes [29]. Although several studies have explored the relationship between this stroke category and revascularization therapies [30, 31], leading to progressively expanded treatment indications, research on general prognostic markers for recovery remains limited. In this context, we were able to provide a fine-grained description of this population over time. Specifically, in most patients, NIHSS scores progressively declined in a non-proportional manner (see from Panel D to Panel I, Figure 5 in the supplementary material, Section 8). In other words, most of the patients did not recover a fixed proportional amount of their initial deficit (see Figure 2). On the contrary, the identified clusters included patients with distinct trajectories, such as additive, proportional, or no improvement profiles, while still capturing a consistent amount of recovery during the first three months (i.e., the critical period for spontaneous recovery) [32]. This becomes particularly evident when further differentiating between severity categories for moderate and severe patient trajectories (see from Panel A to Panel C, Figure 5 in the supplementary material, Section 5).

The identified clusters are likely to reflect underlying neurophysiological differences that influence recovery trajectories. To evaluate this, we tested available prognostic factors (i.e. well-established post stroke recovery prognostic biomarkers) as validation features of the proposed approach. No significant differences in recovery ratio (RR) (i.e. from 22-36h to 90 days) were observed between patients who received rtPA and those who received a placebo. This finding is consistent with expectations, as RR calculation spans from 22–36 hours to 90 days post-stroke, whereas rtPA effects mainly affects the initial hours after the event [33]. Figures 3 and 4 in the supplementary material show improved early recovery (from onset to 22/36h) after acute revascularization treatment and no difference in the subsequent, spontaneous recovery from 22/36h to 90 days. A positive correlation between stroke volume and NIHSS at 22/36 hours supports the expected relationship between initial stroke severity and early impairment [34, 35]. However, stroke volume did not correlate with RR , indicating that lesion size alone does not predict long-term recovery patterns. This finding corroborates existing evidence that volume explains only a limited portion of outcomes, primarily related to initial impairment also following wake up stroke [36–38]. A negative correlation was observed between the mRS scores and both the RR and recovery groups, suggesting that patients with worse functional outcomes tend to exhibit poorer recovery trajectories. Worse recovery was also observed in patients with lesions of the left

hemisphere. While this is consistent with some previous studies reporting worse functional outcomes for left-sided strokes in known-onset stroke [39], several studies have shown right hemisphere lesions also significantly affect recovery, especially reducing patients' compliance to rehabilitation [40, 41]. Our results may, therefore, be explained by the NIHSS's well-established increased sensitivity to left-sided strokes and should be validated using additional clinical scales in specific behavioral domains [42, 43]. Interestingly, no cluster presented a significantly higher frequency of patients with hemorrhagic transformation following rtPA. This suggests a relatively small impact of this feature on long-term disability of wake-up strokes undergoing thrombolytic treatment. These results underscore the substantial benefits of rtPA over its risks in this typology of stroke. Future work should evaluate the role of additional prognostic features of wake-up strokes, including vessel occlusion (i.e. anterior versus posterior circulation), comorbidities (i.e. hypertension as well as other cardiovascular risk factors) and neurophysiology indices (i.e. presence/absence of motor evoked potentials) on final disability, using a similar framework. Finally, patients who achieved an RR of 0.3 or higher within two weeks post-hospitalization consistently clustered into favorable outcome groups, with approximately 90% falling within clusters C_0 , C_1 , C_2 , and C_3 . A similar pattern emerged when considering mRS scores dichotomized for good (mRS ≤ 2) and poor (mRS > 2) outcomes. These results provide a valuable stratification, readily available during hospitalization (i.e. the first week after the event) with several prognostic implications for the patient and caregivers. In addition, these findings highlight the potential clinical application of this clustering approach, demonstrating its scalability across large datasets. Specifically, this methodology draws a useful parallel to how the mRS is employed in clinical trials for rtPA and thrombectomy. Just as these trials aim to identify an increased proportion of patients achieving favorable mRS outcomes, future interventions on neuroplasticity and functional re-organization after stroke could focus on shifting more patients toward the most favorable recovery clusters.

4.3 Limitations and future directions

In regards to the limitations of this study, a potential caveat of this approach is that for NIHSS at 22/36 hours, less than 6, the six linear models found with their relative prediction interval do not discriminate accurately between patients. This could somewhat limit the analysis by questioning the validity of these clusters and the separability between them. This is related to the "ceiling" effect. In fact, in the statistical model shown in Figure 3 Panel C, the prediction interval (in light blue) for a NIHSS less than 6 hardly discriminates points between "fitters" and "non-fitters", questioning the validity of this subdivision for low-impaired subjects. Other clinical measures could be used to discriminate subjects even in this area of low impairment. Specifically, while the clustering approach was limited to only one variable (i.e. the recovery ratio), this was calculated from two variables at different times. Although the use of one variable could be criticized, it is worth underlining that clustering is not a dimensionality reduction technique. Future work could exploit the combination of additional variables for the purpose of phenotyping patients; nonetheless, clustering the recovery ratio can be seen as a clever strategy to embody time in a clustering algorithm. The full

deficit dynamics at additional time points could be explored, enriching current knowledge about stroke recovery. Additional neurophysiological differences across clusters deserve further exploration. These include neuroimaging features (i.e. lesion location, lesion induced disconnection) as well as electro-physiological signals (i.e. evoked potentials and resting state electroencephalography). Finally, it should be noted that these results are limited to the sample in analysis and only indirectly account for several well described prognostic factors (i.e. hypertension, vessel occlusion, comorbidities, corticospinal tract integrity, among others). Therefore, replication in independent cohorts and evaluations using alternative clinical evaluations will be essential to validate and expand these results.

5 Conclusion

This paper aims to pave the way for a new, data-driven approach to studying stroke recovery. By using unsupervised techniques such as Repeated Spectral Clustering, we offer a more precise sub-division of patients based on their recovery trajectories. Clustering the recovery ratio is an efficient way to embody time within the algorithm, while also accommodating the inherent variability in recovery across individuals. This flexible and scalable framework supports the analysis of larger, more heterogeneous datasets and enables continuous evaluation of neurophysiological differences between groups. Ultimately, it offers insights into the mechanisms of post-stroke recovery that extend beyond the Proportional Recovery Rule.

6 Acknowledgment

- Funding: This work was supported by the STARS@UNIPD funding program of the University of Padova, Italy, through the project: MEDMAX. This project has received funding from the European Union’s Horizon Europe research and innovation program under grant agreement no 101137074 - HEREDITARY. ALB, VV and MC were supported by HORIZON-ERC-SyG (Grant No.101071900) “Neurological Mechanisms of Injury And Sleep-Like Cellular Dynamics (NEMESIS)”. MC was supported by HORIZON- INFRA-2022 SERV (Grant No.101147319) “EBRAINS 2.0: A Research Infrastructure to Advance Neuroscience and Brain Health”.
- Conflict of interest: The authors have no competing interests to declare that are relevant to the content of this article.
- Code availability: code is available at [MedMaxLab/beyondPRR](https://github.com/MedMaxLab/beyondPRR).
- Authors’ contributions: AZ: Conceptualization, Data Curation, Methodology, Clustering, Statistical Analysis, Writing - Original Draft. ALB: Conceptualization, Methodology, Data Curation, Writing - Original Draft. VV: Methodology. MA: Project Administration, Supervision. MC: Conceptualization, Project Administration, Supervision. All authors reviewed and edited the manuscript.

References

- [1] Krakauer, J., Marshall, R.: The proportional recovery rule for stroke revisited. *Annals of Neurology* **78**(6), 845–847 (2015) <https://doi.org/10.1002/ana.24537>

- 501 [2] Prabhakaran, S., Zarah, E., Riley, C., Speizer, A., Chong, J.Y., Lazar,
502 R.M., Marshall, R.S., Krakauer, J.W.: Inter-individual variability in the capac-
503 ity for motor recovery after ischemic stroke. *Neurorehabilitation and Neural*
504 *Repair* **22**(1), 64–71 (2008) <https://doi.org/10.1177/1545968307305302> . PMID:
505 17687024
- 506 [3] Bowman, H., Bonkhoff, A., Hope, T., Grefkes, C., Price, C.: Inflated estimates of
507 proportional recovery from stroke. *Stroke* **52**(5), 1915–1920 (2021) [https://doi.](https://doi.org/10.1161/STROKEAHA.120.033031)
508 [org/10.1161/STROKEAHA.120.033031](https://doi.org/10.1161/STROKEAHA.120.033031)
- 509 [4] Hawe, R.L., Scott, S.H., Dukelow, S.P.: Response by hawe et al to letter regarding
510 article, “taking proportional out of stroke recovery”. *Stroke* **50**(5), 126–126 (2019)
511 <https://doi.org/10.1161/STROKEAHA.119.024794>
- 512 [5] Kundert, R., Goldsmith, J., Veerbeek, J.M., Krakauer, J.W., Luft, A.R.: What
513 the proportional recovery rule is (and is not): Methodological and statistical
514 considerations. *Neurorehabilitation and Neural Repair* **33**(11), 876–887 (2019)
515 <https://doi.org/10.1177/1545968319872996> . PMID: 31524062
- 516 [6] Bonkhoff, A.K., Hope, T., Bzdok, D., Guggisberg, A.G., Hawe, R.L., Dukelow,
517 S.P., Rehme, A.K., Fink, G.R., Grefkes, C., Bowman, H.: Bringing proportional
518 recovery into proportion: Bayesian modelling of post-stroke motor impairment.
519 *Brain* **143**(7), 2189–2206 (2020) <https://doi.org/10.1093/brain/awaa146>
- 520 [7] Fugl-Meyer, A., Jääskö, L., Leyman, I., Olsson, S., Steglind, S.: The post-
521 stroke hemiplegic patient. 1. a method for evaluation of physical performance.
522 *Scandinavian journal of rehabilitation medicine* **7**(1), 13–31 (1975)
- 523 [8] Hope, T.M.H., Friston, K., Price, C.J., Leff, A.P., Rotshtein, P., Bowman, H.:
524 Recovery after stroke: not so proportional after all? *Brain* **142**(1), 15–22 (2018)
525 <https://doi.org/10.1093/brain/awy302>
- 526 [9] Vliet, R., Selles, R.W., Andrinopoulou, E.-R., Nijland, R., Ribbers, G.M., Frens,
527 M.A., Meskers, C., Kwakkel, G.: Predicting upper limb motor impairment recov-
528 ery after stroke: A mixture model. *Annals of Neurology* **87**(3), 383–393 (2020)
529 <https://doi.org/10.1002/ana.25679>
- 530 [10] Winters, C., Wegen, E.E.H., Daffertshofer, A., Kwakkel, G.: Generalizability of
531 the proportional recovery model for the upper extremity after an ischemic stroke.
532 *Neurorehabilitation and Neural Repair* **29**(7), 614–622 (2015) [https://doi.org/10.](https://doi.org/10.1177/1545968314562115)
533 [1177/1545968314562115](https://doi.org/10.1177/1545968314562115) . PMID: 25505223
- 534 [11] Tshimanga, L.F., Zanola, A., Facchini, S., Bisogno, A.L., Pini, L., Atzori, M., Cor-
535 betta, M.: Behavioral clusters and lesion distributions in ischemic stroke, based
536 on nihss similarity network. *Journal of Healthcare Informatics Research* (2025)
537 <https://doi.org/10.1007/s41666-025-00197-6>

- 538 [12] Cheng, B., Chen, J., Königsberg, A., Mayer, C., Rimmele, L., Patil, K.R., Gerloff,
539 C., Thomalla, G., Eickhoff, S.B.: Mapping the deficit dimension structure of the
540 national institutes of health stroke scale. *eBioMedicine* **87**, 104425 (2023) <https://doi.org/10.1016/j.ebiom.2022.104425>
541
- 542 [13] Yang, W.-C., Lai, J.-P., Liu, Y.-H., Lin, Y.-L., Hou, H.-P., Pai, P.-F.: Using
543 medical data and clustering techniques for a smart healthcare system. *Electronics*
544 **13**(1) (2024) <https://doi.org/10.3390/electronics13010140>
- 545 [14] Kim, J.-T., Kim, N.R., Choi, S.H., Oh, S., Park, M.-S., Lee, S.-H., Kim, B.C.,
546 Choi, J., Kim, M.S.: Neural network-based clustering model of ischemic stroke
547 patients with a maximally distinct distribution of 1-year vascular outcomes. *Sci-*
548 *entific Reports* **12**(1), 9420 (2022) <https://doi.org/10.1038/s41598-022-13636-w>
- 549 [15] Ramsey, L.E., Siegel, J.S., Lang, C.E., Strube, M., Shulman, G.L., Corbetta,
550 M.: Behavioural clusters and predictors of performance during recovery from
551 stroke. *Nature Human Behaviour* **1**(3), 0038 (2017) <https://doi.org/10.1038/s41562-016-0038>
552
- 553 [16] Lazar, R.M., Minzer, B., Antonello, D., Festa, J.R., Krakauer, J.W., Marshall,
554 R.S.: Improvement in aphasia scores after stroke is well predicted by initial sever-
555 ity. *Stroke* **41**(7), 1485–1488 (2010) <https://doi.org/10.1161/STROKEAHA.109.577338>
556
- 557 [17] Chalos, V., Ende, N.A.M., Lingsma, H.F., Mulder, M.J.H.L., Venema, E., Dijk-
558 land, S.A., Berkhemer, O.A., Yoo, A.J., Broderick, J.P., Palesch, Y.Y., Yeatts,
559 S.D., Roos, Y.B.W.E.M., Oostenbrugge, R.J., Zwam, W.H., Majoie, C.B.L.M.,
560 Lugt, A., Roozenbeek, B., Dippel, D.W.J., MR CLEAN Investigators*, Berkhe-
561 mer, O.A., Fransen, P.S.S., Beumer, D., Berg, L.A., Lingsma, H.F., Yoo, A.J.,
562 Schonewille, W.J., Vos, J.A., Nederkoorn, P.J., Wermer, M.J.H., Walderveen,
563 M.A.A., Staals, J., Hofmeijer, J., Oostayen, J.A., Nijeholt, G.J.L., Boiten, J.,
564 Brouwer, P.A., Emmer, B.J., Bruijn, S.F., Dijk, L.C., Kappelle, L.J., Lo, R.H.,
565 Dijk, E.J., Vries, J., Kort, P.L.M., Rooij, W.J.J., Berg, J.S.P., Hasselt, B.A.A.M.,
566 Aerden, L.A.M., Dallinga, R.J., Visser, M.C., Bot, J.C.J., Vroomen, P.C., Eshghi,
567 O., Schreuder, T.H.C.M.L., Heijboer, R.J.J., Keizer, K., Tielbeek, A.V., Hertog,
568 H.M., Gerrits, D.G., Berg-Vos, R.M., Karas, G.B., Steyerberg, E.W., Flach, H.Z.,
569 Marquering, H.A., Sprengers, M.E.S., Jenniskens, S.F.M., Beenen, L.F.M., Berg,
570 R., Koudstaal, P.J., Zwam, W.H., Roos, Y.B.W.E.M., Lugt, A., Oostenbrugge,
571 R.J., Majoie, C.B.L.M., Dippel, D.W.J.: National institutes of health stroke
572 scale. *Stroke* **51**(1), 282–290 (2020) <https://doi.org/10.1161/STROKEAHA.119.026791>
573
- 574 [18] Aschman, T.A.D., Audebert, H.J., Nitschmann, S.: Mrt-gesteuerte thrombolyse
575 bei schlaganfall. *Der Internist* **60**(4), 420–423 (2019) <https://doi.org/10.1007/s00108-018-0544-9>
576

- 577 [19] Rankin, J.: Cerebral vascular accidents in patients over the age of 60: II. prog-
578 nosis. *Scottish Medical Journal* **2**(5), 200–215 (1957) [https://doi.org/10.1177/](https://doi.org/10.1177/003693305700200504)
579 [003693305700200504](https://doi.org/10.1177/003693305700200504) . PMID: 13432835
- 580 [20] Kummer, R., Broderick, J.P., Campbell, B.C.V., Demchuk, A., Goyal, M., Hill,
581 M.D., Treurniet, K.M., Majoie, C.B.L.M., Marquering, H.A., Mazya, M.V.,
582 Román, L.S., Saver, J.L., Strbian, D., Whiteley, W., Hacke, W.: The heidelberg
583 bleeding classification. *Stroke* **46**(10), 2981–2986 (2015) [https://doi.org/10.1161/](https://doi.org/10.1161/STROKEAHA.115.010049)
584 [STROKEAHA.115.010049](https://doi.org/10.1161/STROKEAHA.115.010049)
- 585 [21] Agresti, A.: *An Introduction to Categorical Data Analysis*, 2nd edn. Wiley,
586 Hoboken, NJ (2007)
- 587 [22] Kruskal, W.H., Wallis, W.A.: Use of ranks in one-criterion variance analysis.
588 *Journal of the American Statistical Association* **47**(260), 583–621 (1952) <https://doi.org/10.1080/01621459.1952.10483441>
589 <https://doi.org/10.1080/01621459.1952.10483441>
- 590 [23] Chong, B., Wang, A., Stinear, C.M.: Proportional recovery after stroke: Address-
591 ing concerns regarding mathematical coupling and ceiling effects. *Neuroreha-*
592 *bilitation and Neural Repair* **37**(7), 488–498 (2023) [https://doi.org/10.1177/](https://doi.org/10.1177/15459683231177598)
593 [15459683231177598](https://doi.org/10.1177/15459683231177598) . PMID: 37269116
- 594 [24] Howell, D.C.: *Statistical Methods for Psychology*. Thomson Wadsworth, ???
595 (2007)
- 596 [25] Luxburg, U.: A tutorial on spectral clustering. *Statistics and Computing* **17**(4),
597 395–416 (2007) <https://doi.org/10.1007/s11222-007-9033-z>
- 598 [26] Fred, A.L.N., Jain, A.K.: Data clustering using evidence accumulation. In: 2002
599 International Conference on Pattern Recognition, vol. 4, pp. 276–2804 (2002).
600 <https://doi.org/10.1109/ICPR.2002.1047450>
- 601 [27] Mann, H.B., Whitney, D.R.: On a test of whether one of two random variables is
602 stochastically larger than the other. *The annals of mathematical statistics*, 50–60
603 (1947)
- 604 [28] Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: A practical and
605 powerful approach to multiple testing. *Journal of the Royal Statistical Society.*
606 *Series B (Methodological)* **57**(1), 289–300 (1995)
- 607 [29] Hervella, P., Alonso-Alonso, M.L., Pérez-Mato, M., Rodríguez-Yáñez, M., Arias-
608 Rivas, S., López-Dequidt, I., Pumar, J.M., Sobrino, T., Campos, F., Castillo, J.,
609 Iglesias-Rey, R.: Surrogate biomarkers of outcome for wake-up ischemic stroke.
610 *BMC Neurology* **22**(1), 215 (2022) <https://doi.org/10.1186/s12883-022-02740-z>
- 611 [30] Schlemm, L., Braemswig, T.B., Boutitie, F., Vynckier, J., Jensen, M., Galinovic,
612 I., Simonsen, C.Z., Cheng, B., Cho, T.-H., Fiehler, J., Puig, J., Thijs, V., Fiebach,

- 613 J., Muir, K., Nighoghossian, N., Ebinger, M., Pedraza, S., Thomalla, G., Gerloff,
614 C., Endres, M., Lemmens, R., Nolte, C.H., WAKE-UP Investigators: Cerebral
615 microbleeds and treatment effect of intravenous thrombolysis in acute stroke. *Neu-*
616 *rology* **98**(3), 302–314 (2022) <https://doi.org/10.1212/WNL.0000000000013055>
617 <https://www.neurology.org/doi/pdf/10.1212/WNL.0000000000013055>
- 618 [31] Thomalla, G., Simonsen, C.Z., Boutitie, F., Andersen, G., Berthezene, Y., Cheng,
619 B., Cheripelli, B., Cho, T.-H., Fazekas, F., Fiehler, J., Ford, I., Galinovic, I.,
620 Gellissen, S., Golsari, A., Gregori, J., Günther, M., Guibernau, J., Häusler, K.G.,
621 Hennerici, M., Kemmling, A., Marstrand, J., Modrau, B., Neeb, L., Ossa, N.P.,
622 Puig, J., Ringleb, P., Roy, P., Scheel, E., Schonewille, W., Serena, J., Sunaert,
623 S., Villringer, K., Wouters, A., Thijs, V., Ebinger, M., Endres, M., Fiebach, J.B.,
624 Lemmens, R., Muir, K.W., Nighoghossian, N., Pedraza, S., Gerloff, C.: Mri-guided
625 thrombolysis for stroke with unknown time of onset. *New England Journal of*
626 *Medicine* **379**(7), 611–622 (2018) <https://doi.org/10.1056/NEJMoa1804355>
- 627 [32] Wade, D.T., Wood, V.A., Hower, R.L.: Recovery after stroke—the first 3 months.
628 *Journal of Neurology, Neurosurgery & Psychiatry* **48**(1), 7–13 (1985) <https://doi.org/10.1136/jnnp.48.1.7>
629
- 630 [33] null: Tissue plasminogen activator for acute ischemic stroke. *New Eng-*
631 *land Journal of Medicine* **333**(24), 1581–1588 (1995) [https://doi.org/10.1056/](https://doi.org/10.1056/NEJM199512143332401)
632 [NEJM199512143332401](https://doi.org/10.1056/NEJM199512143332401)
- 633 [34] Ghoneem, A., Osborne, M.T., Abohashem, S., Naddaf, N., Patrich, T., Dar, T.,
634 Abdelbaky, A., Al-Quthami, A., Wasfy, J.H., Armstrong, K.A., Ay, H., Tawakol,
635 A.: Association of socioeconomic status and infarct volume with functional out-
636 come in patients with ischemic stroke. *JAMA Network Open* **5**(4), 229178–229178
637 (2022) <https://doi.org/10.1001/jamanetworkopen.2022.9178>
- 638 [35] Lee, K.-J., Kim, J.Y., Kang, J., Kim, B.J., Kim, S.-E., Oh, H., Park, H.-K., Cho,
639 Y.-J., Park, J.-M., Park, K.-Y., Lee, K.B., Lee, S.J., Park, T.H., Lee, J.S., Lee, J.,
640 Yang, K.H., Choi, A.R., Kang, M.Y., Saposnik, G., Bae, H.-J.: Hospital volume
641 and mortality in acute ischemic stroke patients: Effect of adjustment for stroke
642 severity. *Journal of Stroke and Cerebrovascular Diseases* **29**(5), 104753 (2020)
643 <https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.104753>
- 644 [36] Rosso, C., Samson, Y.: The ischemic penumbra: the location rather than the
645 volume of recovery determines outcome. *Current Opinion in Neurology* **27**(1)
646 (2014)
- 647 [37] Koch, P.J., Rudolf, L.F., Schramm, P., Frontzkowski, L., Marburg, M., Matthis,
648 C., Schacht, H., Fiehler, J., Thomalla, G., Hummel, F.C., Neumann, A., Münte,
649 T.F., Rojl, G., Machner, B., Schulz, R.: Preserved corticospinal tract revealed
650 by acute perfusion imaging relates to better outcome after thrombectomy in
651 stroke. *Stroke* **54**(12), 3081–3089 (2023) [https://doi.org/10.1161/STROKEAHA.](https://doi.org/10.1161/STROKEAHA.123.044221)
652 [123.044221](https://doi.org/10.1161/STROKEAHA.123.044221)

- 653 [38] Schlemm, E., Ingwersen, T., Königsberg, A., Boutitie, F., Ebinger, M., Endres,
654 M., Fiebach, J.B., Fiehler, J., Galinovic, I., Lemmens, R., Muir, K.W., Nighoghos-
655 sian, N., Pedraza, S., Puig, J., Simonsen, C.Z., Thijs, V., Wouters, A., Gerloff, C.,
656 Thomalla, G., Cheng, B.: Preserved structural connectivity mediates the clinical
657 effect of thrombolysis in patients with anterior-circulation stroke. *Nature Com-*
658 *munications* **12**(1), 2590 (2021) <https://doi.org/10.1038/s41467-021-22786-w>
- 659 [39] Hedna, V.S., Bodhit, A.N., Ansari, S., Falchook, A.D., Stead, L., Heilman, K.M.,
660 Waters, M.F.: Hemispheric differences in ischemic stroke: Is left-hemisphere stroke
661 more common? *J Clin Neurol* **9**(2), 97–102 (2013) [https://doi.org/10.3988/jcn.](https://doi.org/10.3988/jcn.2013.9.2.97)
662 [2013.9.2.97](https://doi.org/10.3988/jcn.2013.9.2.97)
- 663 [40] Aszalós, Z., Barsi, P., Vitrai, J., Nagy, Z.: Lateralization as a factor in the prog-
664 nosis of middle cerebral artery territorial infarct. *European Neurology* **48**(3),
665 141–145 (2002) <https://doi.org/10.1159/000065515>
- 666 [41] Ween, J.E., Alexander, M.P., D’Esposito, M., Roberts, M.: Factors predictive
667 of stroke outcome in a rehabilitation setting. *Neurology* **47**(2), 388–392 (1996)
668 <https://doi.org/10.1212/WNL.47.2.388>
- 669 [42] Woo, D., Broderick, J.P., Kothari, R.U., Lu, M., Brott, T., Lyden, P.D., Marler,
670 J.R., Grotta, J.C.: Does the national institutes of health stroke scale favor left
671 hemisphere strokes? *Stroke* **30**(11), 2355–2359 (1999) [https://doi.org/10.1161/](https://doi.org/10.1161/01.STR.30.11.2355)
672 [01.STR.30.11.2355](https://doi.org/10.1161/01.STR.30.11.2355)
- 673 [43] Fink, J.N., Selim, M.H., Kumar, S., Silver, B., Linfante, I., Caplan, L.R., Schlaug,
674 G.: Is the association of national institutes of health stroke scale scores and acute
675 magnetic resonance imaging stroke volume equal for patients with right- and
676 left-hemisphere ischemic stroke? *Stroke* **33**(4), 954–958 (2002) [https://doi.org/](https://doi.org/10.1161/01.STR.0000013069.24300.1D)
677 [10.1161/01.STR.0000013069.24300.1D](https://doi.org/10.1161/01.STR.0000013069.24300.1D)