Transl. Stroke Res. (2016) 7:388–394
DOI 10.1007/s12975-016-0488-0

SI: CHALLENGES AND CONTROVERSIES IN TRANSLATIONAL STROKE RESEARCH

How to Measure Recovery? Revisiting Concepts and Methods for Stroke Studies

Marc Hommel[1,2] · **Olivier Detante**[3] · **Isabelle Favre**[3] · **Emmanuel Touzé**[4] · **Assia Jaillard**[1,2]

Received: 11 February 2016 / Revised: 3 July 2016 / Accepted: 21 July 2016 / Published online: 8 August 2016

Abstract In clinical trials, assessing efficacy is based on validated scales, and the primary endpoint is usually based on a single scale. The aim of the review is to revisit the concepts and methods to design and analyze studies focused on restoration, recovery and or compensation. These studies are becoming more frequent with the development of restorative medicine. After discussing the definitions of recovery, we address the concept of recovery as the regain of lost capabilities, when the patient reaches a new equilibrium. Recovery is a dynamic process which assessment includes information from initial and final status, their difference, the difference between the final status of the patient and normality, and the speed of restoration. Finally, recovery can be assessed either for a specific function (focal restoration) or for a more global restoration. A single scale is not able to assess all the facets of a skill or a function, therefore complementary information should be collected and analyzed simultaneously to be tested in a single analysis. We are suggesting that recovery should be considered as a latent variable and therefore cannot be measured in pure form. We are also suggesting to customize the data collection and analysis according to the characteristics of the subjects, the mechanisms of action and consequences of the intervention. Moreover, recovery trials should benefit from latent variable analysis methods. Structural equation modeling is likely the best candidate for this approach applicable in preclinical and clinical studies.

Keywords Methodology · Intervention evaluation · Study design · Clinical scale · Latent variable · Modeling

✉ Marc Hommel
marc.hommel@ujf-grenoble.fr

[1] University Grenoble Alpes, AGEIS EA 7407, Grenoble, France

[2] CHU Grenoble, Pôle Recherche, Grenoble, France

[3] CHU Grenoble, Pôle Psychiatrie Neurologie, Stroke unit, Grenoble, France

[4] INSERM U919, Stroke unit, CHU, University of Caen Basse Normandie, Caen, France

Introduction

Many definitions of recovery have been proposed, from neuronal rewiring and anatomical restitution to improvement of performance or coping and social integration.

From the perspective of the conceptual framework of the WHO International Classification of Functioning (ICF) [1, 2], recovery refers to health condition, at different levels of body function and structure, activities, and participation. In order to determine the most important indicators of successful rehabilitation outcome, an attempt to classify instruments according to ICF was proposed such as the Fugl-Meyer scale [3] to assess impairment at the body structure level, the Barthel Index [4] and the mRankin scale [5] at the activities level, and the Euroqol-5D [6] at the participation level [7–9].

Furthermore, the term "true recovery" has been used to define the complete restitution of neuronal networks [10, 11]. However, there is still some ambiguity in defining true recovery, since this term is used to refer to both improved performances and return to normal clinical patterns [12]. Focusing on motor system, some motor tests based on video assessment have been developed to quantify recovery [10], assuming restitution (true recovery) is present if recovered movements have the same quality (patterns) as normal movements [10, 12, 13]. In contrast, impaired performance would reflect incomplete recovery that may be due to compensation [10, 11].

Functional neuroimaging (fMRI) that gives us access to neuronal network mapping has suggested that motor recovery was consistent with typical motor network pattern [14]. However, recovery of normal motor performance can be associated with vicarious processes that reflect appropriate reorganization of motor networks [15]. In that case, normal clinical motor patterns that are concomitant with reorganized brain network would not be defined as 'true recovery' although it has been claimed earlier [12].

Evidently, there are some heterogeneity between the definition of motor recovery, the evidence of its anatomical and functional correlates and the different scales used as endpoints in clinical trials. Moreover, if this hypothesis of normal patterns of functioning may be relevant for motor rehabilitation evaluation [10–12, 16], their assessment may hardly be generalized to all functions (i.e., cognition) and activity (i.e., social functioning). Nevertheless, the clinical scales used for stroke trials and measuring performance collapse recovery and compensation and are applicable to body function, activities and participation.

In clinical studies, recovery describes progress from baseline to final evaluation, whether outcome generally refers to a well-defined status (e.g., death, dependence rate, clinical score). Reducing stroke occurrence or recurrence and vascular morbidity and mortality are the goals of prevention studies. In acute stroke, evaluating a therapy such as recanalization is based on correlated outcome instruments [17] focused, either on the neurological deficit using the National Institute of Health Stroke Scale (NIHSS) [18], or on independence using the Activities of Daily Living (ADL) [19], the Barthel index [4], Functional Independence Measure (FIM) [20], or on survival and handicap using the multi-dimension modified Rankin scale (mRankin) [5], or the Glasgow Outcome Scale (GOS) [21, 22]. These measures are used either alone [23, 24] or included in a global statistical test as the primary endpoint [25, 26]. Furthermore, cutoff values strategies are sometimes preferred to separate poor and good outcomes. However, choosing the right test and cutoff points as the primary endpoint may be challenging [27–29]. Considerable metrological efforts have been put on variance reduction [30], reliability, and measurements errors [31–36] and on other statistical approaches [37–42]. However, at the end of the trial, one has to confess that only a tiny amount of all the information collected is used for the primary endpoint assessment [43].

In this opinion paper, we revisit the concept of recovery from epistemological, methodological and statistical perspectives. Our purpose is to present a flexible approach reflecting the most likely manifestations of an intervention, far from standard single scales or tests. We also assume that using more appropriate information would improve our capability to demonstrate efficacy of intervention, by reducing type 1 and 2 errors. This customized approach should be relevant for studies on brain plasticity in restorative medicine. The scales in terms of metrological properties [7–9] and other methodological aspects are beyond the scope of this paper.

The Concept of Recovery

From a historical perspective of "the normal and the pathological", the epistemologist and medical doctor G. Canguilhem stressed that illness is defined by the "subject himself" [44]. In Canguilhem's perspective, recovery can never restore the initial state of "*biological innocence*." Canguilhem defined "Cure," i.e., recovery, as a new stable status that is reached when the *subject gives himself new norms* [44].

Nowadays, when referring to recovery, patients and medical doctors intuitively gather information related to several dimensions that are listed below.

- Recovery is a dynamic process. The term recovery refers to progress observed between time points, e.g., stroke onset or therapy onset and outcome evaluation. It means that recovery contains information from initial severity, from residual severity, and from the size of their difference. Speed of recovery also contributes to the description of this dynamic process.
- Recovery is related to *new norms* acquisition. One may consider that the closer from normal status these *new norms* are, the better recovery is. However, the term *new norms* may be ambiguous because it covers several dimensions from body structure to psychological and societal dimensions.
- Recovery can be local or global. A large body of evidence has shown a modular organization of brain functions. Therefore, recovery of single modules or functions can be assessed. For example, local recovery may represent the effects of a specific training of hand grip precision tasks for a patient suffering from hand paresis. This local recovery can be applied to every domain e.g., swallowing, speech. Global recovery may represent several local dimensions reunited, in a multiple modules scale (e.g., NIHSS, mRankin). A study can be aimed on local recovery only or combine local and global recovery assessments.
- No test is entirely selective to its targeted domain. For example, memory tests require attention and language process for performance. Therefore, the assessment of a single function should include several complementary tests. Whereas, there is no test designed for assessing the whole range of a function, it is impossible to structure a test that would consistently avoid "floor" and "ceiling" effects, [45]. For all the above listed reasons, it is unrealistic to expect that recovery can be assessed with one test or scale.

From Concept to Methodology

In a pragmatic approach, we are posing five specifications recovery assessment should fulfill.

1. A flexible representation of recovery in terms of modularity (i.e., local recovery, global recovery). It implies that multiple dimensions and biomarkers should be assessed.
2. Being appropriate for the intervention under study. For example, if a specific motor rehabilitation program is considered, a set of tests focused on this target should be included.
3. Presenting the contribution of each variables constituting recovery. It is indeed important that the effect size, its model, and its contribution to each dimension of recovery are documented.
4. The cornerstone of therapeutic research is the comparison of the effect of intervention between two groups of patients using a single test integrating all the dimensions of recovery.
5. Providing interpretation of gains. When explaining the expected effects of an intervention, it is important to quantify the improvement attributable to that intervention for each dimension.

Statistical Analysis

1. For recovery assessment, we are suggesting to use methods based on latent variables

A latent variable cannot be directly observed or assessed in pure form. For instance, it can combine information on health condition provided by various scales and modalities. Practically, this latent variable has to be inferred from observed variables such as clinical scales, self-reports, laboratory results, and imaging results. These "indicators" share a commonality supposed to represent the measured recovery [46]. It means that one has to make strong and formal assumptions on the model of information determining this latent variable. Shaping this information according to the goals of the study, and constructing the latent variable using mathematical modeling, allows its estimation.

2. Simplification is mandatory

Measurement of clinically relevant recovery needs simplification with dimensions reduction. The price to pay lies in loss of dimensions and loss of completeness of information. A one-dimensional example of recovery is shown in Fig. 1 based on the same scale (one dimension). Four variables carrying convergent information can be used.

3. How to estimate a latent variable

Many tools may be used for computing latent variables, such as hidden Markov models, Bayesian models, exploratory factor analysis (EFA), principal component analysis (PCA), independent components analysis (ICA), and structural equation modeling (SEM). Hidden Markov models and Bayesian approaches are related to definite states (categorical variables) and the relationships between the factors are expressed as state transition probabilities. As compared to regression coefficients, probabilities are less meaningfully directly transposed to clinical world. From a Bayesian perspective, estimation is less about deducting the values of population parameters and more about updating, refining our beliefs about the empirical world. Factor analysis (PCA) is also a very traditional approach. However, it is exploratory and does not refer to an explicit pre-established model. It tries to account for all the variance and covariance of the set of variables and not for the portion of covariance the variables share in common. When running PCA, we hope there will be a single dominant factor. Moreover, the meaning of each factor is not straightforward. ICA attempts to decompose multivariate data into independent sets of random variables and is more powerful than PCA to separate sources. This technique implies to exit from the correlation toolbox. It shares with PCA the difficulty to give a straightforward signification to the factor.

According to the specifications of the concept of recovery, structural equations modeling (SEM) and for estimating recovery [47], its measurement subset confirmatory factor analysis (CFA) may represent a more relevant choice. In CFA, the number of latent variables and their construct are defined. CFA assumes that the latent variable accounts what the variables share in common and isolates shared variance. Moreover, the results may become stronger by removing measurement error (i.e., residual variance).

4. Using structural equations modeling to estimate recovery

SEM represents a flexible family of models allowing both exploratory and confirmatory modeling. SEM is more general than regression because it is based on multiple equations systems and the directionality of the relations is pre-specified in the model. In SEM, measured variables can be discrete, ordinal, or dummies [48], and multivariate normality distribution hypothesis can be relaxed using adequate estimator. CFA [49] is the best approach to build recovery (i.e., the latent variable). In the example of recovery model presented in Fig. 2, the investigator specified recovery as a single factor, based on clinical scales only. This factor combines initial severity (NIHSS at day 0), motor recovery (Fugl-Meyer motor scale [3, 50], day 0–90), deficit recovery (NIHSS day 0–90), and the global status at the end of follow-up (mRankin). When a

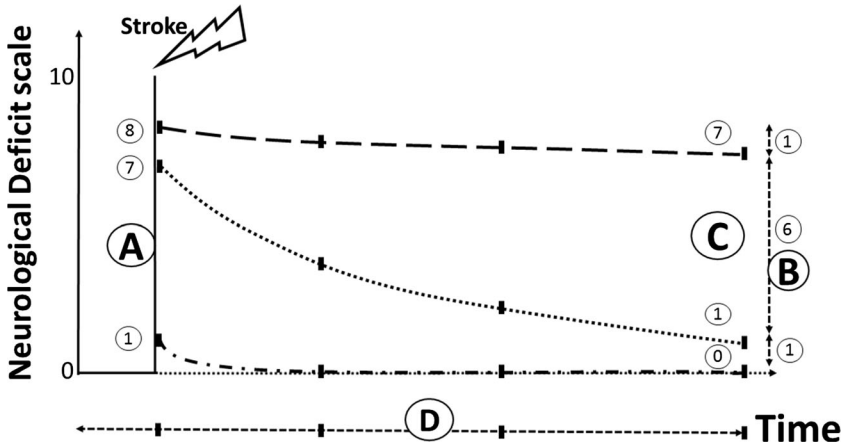

Fig. 1 Variables and time points contributing to the measure of recovery using a single scale. Three recovery paths of stroke patients are represented on a 10-point scale (0 is normal, 10 the maximum severity). The patient represented with a *dot-dash line* suffered a mild stroke and had a fast and complete recovery. The patient represented with a *dot line* suffered a moderate to severe stroke and had a sustained and important recovery. The patient represented with a *dash line* had a severe stroke and had a minor and poor recovery. Numbers in *A* represent initial severity. Numbers in *B* represent gains from onset (*A*–*C*). Numbers in **c** represent distance between the normal status and the recovered status, *D* with intermediate time points measurements informs on speed of recovery

model is specified, its fit should be carefully evaluated [49, 51]. To proceed, the model should be over-identified, i.e., the number of data points (variables) is greater than parameters to be estimated. Only over-identified models provide fit statistics as a mean of evaluating the fit of the overall model. In the model presented in Fig. 2, it is likely that the measurements errors of NIHSS at day 0 and the difference NIHSS day 0 minus day 90 are correlated. If the researcher wants to simplify further (e.g., remove Fugl-Meyer motor scale) this model becomes just identified as follows: the number of data points

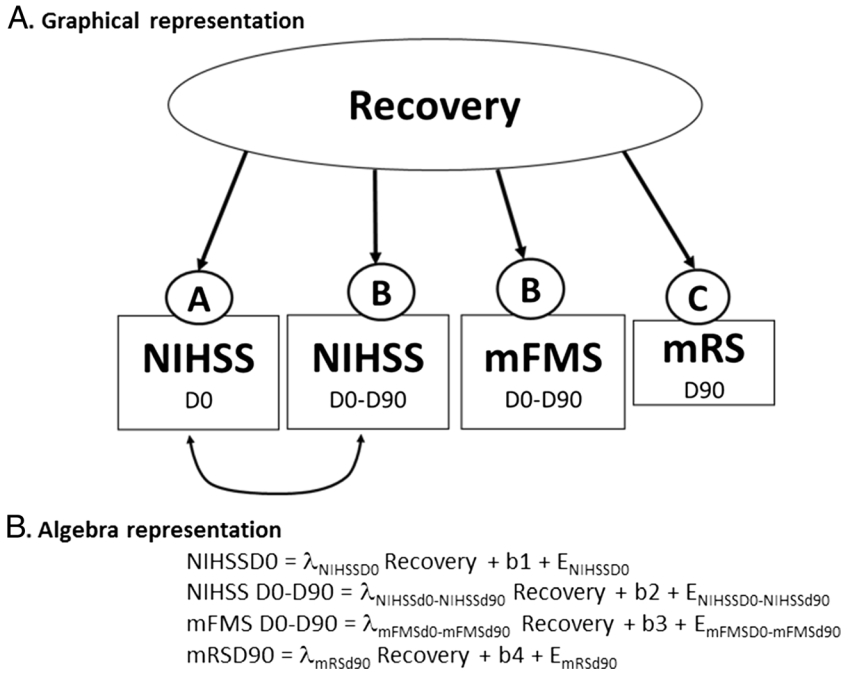

Fig. 2 Example of recovery measured using confirmatory factor analysis. **a** Graphical representation. As a convention in structural equation modeling, external measured variables are represented in squares (*rectangle*), directional arrow the direct effect and the latent variables in circle (*ellipse*). In this example of model, the confirmatory factor model (CFA) of recovery has four indicators, the National Institute of Health Stroke Scale (NIHSS) at the day of admission (D0) as the marker of initial severity (*A* in Fig. 1), the differences between day 0 and day 90 as a marker of the gain (*B* in Fig. 1) in NIHSS and in the modified Fugl-Meyer motricity scale, and the distance to normal status (*C* in Fig. 1) at day 90 for the modified Rankin scale (mRS). The correlation of the error terms of NIHSSd0 and NIHSSd0-d90 is represented with a curved path and this extra parameter is also estimated. **b** Algebra representation. Recovery is represented with a system of four equations; **λ** is the coefficient of the equation and E the measurement error (residual variance)

equals the number of parameters to be estimated. This latter model fits the data perfectly and thus is of little use. Therefore, when designing a study, enough dimensions (variables) should be collected to build a latent variable. The investigator may consider adding other relevant indicators (variables). For example, if intervention was focused on a motor rehabilitation program, he may add a 10-Meter Walk Test [52] difference between day 0 and 90, and for the upper limb the differences between day 0 and 90 for ARAT [53]. He also may be interested in the contribution of a more general evaluation of the patient's status at 6 months regarding the cognitive functioning using MoCA [54] and the social impact of stroke using the work and social adjustment scale (WSAS) [55]. Redundant information (variables with a correlation coefficient >0.8) should be removed. Variables representing complementary information on different facets are to be preferred. If interested, the investigator can build a latent variable for compensation and another one for true recovery and test their indicators and relations [10]. Alternatively, he can consider another factor related to biomarkers of recovery (e.g., imaging, biology) and may be interested in the relationship between clinical and biomarkers factors. Neuroimaging has become a privileged tool to explore stroke and brain plasticity. A multimodal approach aimed at anatomical connectivity, activation studies, and resting state generates many imaging biomarkers. These imaging biomarkers can constitute imaging latent variables that can be used into a multiple factors measurement model to evaluate their relations with clinical variables and intervention.

Eventually, the investigator has to keep in mind that other models may explain as well or even better the concept he attempts to analyze. Therefore, re-specification may be needed, and the choice between equivalent models or of the best model has to be documented [47, 49].

CFA provides for each indicator (measured variable; Fig. 2b) estimates of the standardized and unstandardized coefficient, its standard error, a *z* test, and confidence intervals. Standardized coefficients are interesting for measuring the strength of the associations. In unstandardized solutions, the metrics of both indicators and the latent variable are based on the original metrics of the indicator, and coefficients can be interpreted as unstandardized regression coefficients. Thus, effect size of each indicator can be estimated using the unstandardized coefficients and allowing to discuss the minimal clinically important difference. Percentage of variance explained by the model and for each indicator can be estimated.

A key issue in experimental designs is about the comparison between groups of intervention [56]. When comparing the means between treated and control groups, the latent variable is computed for each group, the coefficient, the *z* value, its significance, and confidence intervals are shown. This significance allows a clear conclusion on primary hypothesis of H0 rejection. If the model is composed of multiple factors, a multiple indicators and multiple causes (MIMIC) model can be used to evaluate intervention. These multivariate models groups' comparisons are a major advance beyond traditional bivariate comparison tests.

The introduction of auxiliary variables [57] will add information in the model without being directly included in the estimation of the factor. Such information can for example be appropriate for variables influencing recovery (e.g., age, depression) and give more relevance to the factor assessed.

Additionally, SEM can be used for presenting a model that conveys causal assumptions [58], or path analysis including multiple latent variables and measured variables. In longitudinal studies, multiple time points can be measured, and latent growth curves modeling can estimate the speed of recovery [59], like multilevel longitudinal models can do.

In summary, SEM fulfills all the above specifications and should be considered to estimate recovery. However, this is a complex, time consuming, and conceptually demanding approach. As in other statistical approaches, sensitivity analysis can be performed with different conditions (e.g., diagnostic classification, measurements of indicators, normality assumptions).

Practical Issues in Structural Equations Modeling

Many software are available, either free such as R with several packages dealing with SEM (e.g., "lavaan," OpenMx, and "sem"), the graphical interface for structural equation modeling Ωnyx, and the online software WebSEM, or they are commercial smartPLS, AMOS in SPSS, Stata, SAS, MPlus, LISREL, EQS, PLS-Graph, and WarpPLS.

How to Communicate?

Communicating results of a clinical trial is usually based on a contingency table via the relative risk reduction and the number needed to treat. It has to be stressed that based on the same four numbers, both ways of communication are referring for the first one to a multiplicative model and for the second one to an additive model. Both models are very unlikely fitting the genuine effect model of the intervention because additive and multiplicative model are rare and effect models usually fit curved models [60, 61]. It can be concluded that accurate communication is not easy with classical tools.

The CFA group comparison or the MIMIC model shows clearly the statistical significance of the study. A reasonable calculation of the effect size with the unstandardized solution can be done with the ratio of the difference between the means of the latent variable of the treated and control groups on their pooled standard deviations [62]. Thus, to report this effect on the indicators, the unstandardized solution gives for the clinical scales used, the difference in means between the groups.

This simultaneous translation on different clinical scales allows discussing the clinical utility of an intervention.

Limitations

As compared with well-established rules for analyzing clinical studies, SEM is clearly not so straightforward. To be convincing, quality rules have to be explicitly followed. It concerns traceable a priori assumptions on the model, assessment of model's fit, and changes from initial model explicated.

Conclusions

Assessment needs to be appropriate to the goals [63]. In our review, we emphasize the potential gain of customizing the collection of information and data analysis focusing on the potential effects of the intervention. Computationally intensive structural equation modeling (SEM) approaches have been developed over much of the twentieth century. SEM tools are flexible and allow estimating latent variables shaped according to the purpose of the study. However, this approach still has to be evaluated in simulation and real life studies and confronted with the classical statistical methods for preclinical and clinical research.

Acknowledgments We thank Jean Luc Cracowski and Jacques Demongeot for their comments on earlier versions of the manuscript.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Disclosures None.

Funding Information PHRCI ISIS Dr Olivier Detante
PHRCI HERMES Dr Assia Jaillard
RESSTORE Horizon 2020 Dr Olivier Detante
ANR e-Swallhome Prof Emmanuel Touzé

References

1. WHO/EIP/GPE/CAS/01.3: Towards a common language for functioning, disability and health icf, WHO; 2002, 2015.
2. Barak S, Duncan PW. Issues in selecting outcome measures to assess functional recovery after stroke. NeuroRx: J Am Soc Exp NeuroTher. 2006;3:505–24.
3. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. Scand J Rehabil Med. 1975;7:13–31.
4. Mahoney FI, Barthel DW. Functional evaluation: the barthel index. Maryland State Med J. 1965;14:61–5.
5. Rankin J. Cerebral vascular accidents in patients over the age of 60: ii. Prognosis Scottish Med J. 1957;2:200–15.
6. Golicki D, Niewada M, Karlinska A, Buczek J, Kobayashi A, Janssen MF, Pickard AS. Comparing responsiveness of the eq-5d-5 l, eq-5d-3 l and eq vas in stroke patients. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2014
7. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J. Issues for selection of outcome measures in stroke rehabilitation: icf body functions. Disabil Rehabil. 2005a;27:191–207.
8. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: icf activity. Disabil Rehabil. 2005b;27:315–40.
9. Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: icf participation. Disabil Rehabil. 2005c;27:507–28.
10. Buma F, Kwakkel G, Ramsey N. Understanding upper limb recovery after stroke. Restor Neurol Neurosci. 2013;31:707–22.
11. Levin MF, Kleim JA, Wolf SL. What do motor "recovery" and "compensation" mean in patients following stroke? Neurorehabil Neural Repair. 2009;23:313–9.
12. Wolf SL, Catlin PA, Ellis M, Archer AL, Morgan B, Piacentino A. Assessing wolf motor function test as outcome measure for research in patients after stroke. Stroke. 2001;32:1635–9.
13. Whishaw IQ. Loss of the innate cortical engram for action patterns used in skilled reaching and the development of behavioral compensation following motor cortex lesions in the rat. Neuropharmacology. 2000;39:788–805.
14. Favre I, Zeffiro TA, Detante O, Krainik A, Hommel M, Jaillard A. Upper limb recovery after stroke is associated with ipsilesional primary motor cortical activity: a meta-analysis. Stroke. 2014;45:1077–83.
15. Jaillard A, Martin CD, Garambois K, Lebas JF, Hommel M. Vicarious function within the human primary motor cortex? A longitudinal fmri stroke study. Brain J Neurol. 2005;128:1122–38.
16. Moon SK, Alaverdashvili M, Cross AR, Whishaw IQ. Both compensation and recovery of skilled reaching following small photothrombotic stroke to motor cortex in the rat. Exp Neurol. 2009;218:145–53.
17. Goldie FC, Fulton RL, Frank B, Lees KR. Interdependence of stroke outcome scales: reliable estimates from the virtual international stroke trials archive (vista). Int J Stroke: Off J Int Stroke Soc. 2014;9:328–32.
18. Brott T, Adams Jr HP, Olinger CP, Marler JR, Barsan WG, Biller J, Spilker J, Holleran R, Eberle R, Hertzberg V, et al. Measurements of acute cerebral infarction: a clinical examination scale. Stroke. 1989;20:864–70.
19. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. The Gerontologist. 1969;9:179–86.
20. Hamilton BB, Laughlin JA, Fiedler RC, Granger CV. Interrater reliability of the 7-level functional independence measure (fim). Scand J Rehabil Med. 1994;26:115–9.
21. Jennett B, Snoek J, Bond MR, Brooks N. Disability after severe head injury: observations on the use of the Glasgow outcome scale. J Neurol Neurosurg Psychiatry. 1981;44:285–93.
22. McArthur K, Fan Y, Pei Z, Quinn T. Optimising outcome assessment to improve quality and efficiency of stroke trials. Exp Rev Pharmacoeconomics Outcome Res. 2014;14:101–11.
23. The Multicenter Acute Stroke Trial–Europe Study Group. Thrombolytic therapy with streptokinase in acute ischemic stroke. N Engl J Med. 1996;335:145–50.
24. Chollet F, Cramer SC, Stinear C, Kappelle LJ, Baron JC, Weiller C, Azouvi P, Hommel M, Sabatini U, Moulin T, Tardy J, Valenti M, Montgomery S, Adams Jr H. Pharmacological therapies in post stroke recovery: recommendations for future clinical trials. J Neurol. 2014;261:1461–8.

25. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. N Engl J Med. 1995;333:1581–7.
26. Tilley BC, Marler J, Geller NL, Lu M, Legler J, Brott T, Lyden P, Grotta J. Use of a global test for multiple outcomes in stroke trials with application to the national institute of neurological disorders and stroke t-pa stroke trial. Stroke. 1996;27:2136–42.
27. Hacke W, Kaste M, Fieschi C, Toni D, Lesaffre E, von Kummer R, Boysen G, Bluhmki E, Hoxter G, Mahagne MH, et al. Intravenous thrombolysis with recombinant tissue plasminogen activator for acute hemispheric stroke. The European Cooperative Acute Stroke Study (ECASS). JAMA. 1995;274:1017–25.
28. Hacke W, Kaste M, Fieschi C, von Kummer R, Davalos A, Meier D, Larrue V, Bluhmki E, Davis S, Donnan G, Schneider D, Diez-Tejedor E, Trouillas P. Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). Second European-Australasian Acute Stroke Study Investigators. Lancet. 1998;352:1245–51.
29. Hacke W, Kaste M, Bluhmki E, Brozman M, Davalos A, Guidetti D, Larrue V, Lees KR, Medeghri Z, Machnig T, Schneider D, von Kummer R, Wahlgren N, Toni D. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. N Engl J Med. 2008;359:1317–29.
30. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke. 1988;19:604–7.
31. Wilson JT, Hareendran A, Grant M, Baird T, Schulz UG, Muir KW, Bone I. Improving the assessment of outcomes in stroke: use of a structured interview to assign grades on the modified Rankin scale. Stroke. 2002;33:2243–6.
32. Wilson JT, Hareendran A, Hendry A, Potter J, Bone I, Muir KW. Reliability of the modified Rankin scale across multiple raters: benefits of a structured interview. Stroke. 2005;36:777–81.
33. Quinn TJ, Lees KR, Hardemark HG, Dawson J, Walters MR. Initial experience of a digital training resource for modified Rankin scale assessment in clinical trials. Stroke. 2007;38:2257–61.
34. Saver JL, Filip B, Hamilton S, Yanes A, Craig S, Cho M, Conwit R, Starkman S. Improving the reliability of stroke disability grading in clinical trials and clinical practice: the Rankin focused assessment (rfa). Stroke. 2010;41:992–5.
35. Quinn TJ, Dawson J, Walters MR, Lees KR. Reliability of the modified Rankin scale: a systematic review. Stroke. 2009;40:3393–5.
36. Banks JL, Marotta CA. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. Stroke. 2007;38:1091–6.
37. Saver JL. Novel end point analytic techniques and interpreting shifts across the entire range of outcome scales in acute stroke trials. Stroke. 2007;38:3055–62.
38. Saver JL, Gornbein J. Treatment effects for which shift or binary analyses are advantageous in acute stroke trials. Neurology. 2009;72:1310–5.
39. Bath PM, Gray LJ, Collier T, Pocock S, Carpenter J. Can we improve the statistical analysis of stroke trials? Statistical reanalysis of functional outcomes in stroke trials. Stroke. 2007;38:1911–5.
40. Savitz SI, Lew R, Bluhmki E, Hacke W, Fisher M. Shift analysis versus dichotomization of the modified Rankin scale outcome scores in the NINDS and ECASS-II trials. Stroke. 2007;38:3205–12.
41. Howard G, Waller JL, Voeks JH, Howard VJ, Jauch EC, Lees KR, Nichols FT, Rahlfs VW, Hess DC. A simple, assumption-free, and clinically interpretable approach for analysis of modified Rankin outcomes. Stroke. 2012;43:664–9.
42. Churilov L, Arnup S, Johns H, Leung T, Roberts S, Campbell BC, Davis SM, Donnan GA. An improved method for simple, assumption-free ordinal analysis of the modified Rankin scale using generalized odds ratios. Int J Stroke: Off J Int Stroke Soc. 2014;9:999–1005.
43. Saver JL, Yafeh B. Confirmation of tpa treatment effect by baseline severity-adjusted end point reanalysis of the NINDS-tPA stroke trials. Stroke. 2007;38:414–6.
44. Canguilhem G. The normal and the pathological, ed 5th printing, 2007. New York, NY: Zone books, Urzone Inc; 1991.
45. Weintraub S. Neuropsychological assessment of mental state. In: Mesulam M, editor. Principles of behavioral and cognitive neurology, Oxford University Press; 2000, p. 121–73.
46. Bollen AK, Hoyle RH. Latent variable in structural equation modeling. In: Hoyle RH, editor. Handbook of structural equation modeling, The Guilford Press; 2012, p. 56–67.
47. Lei PW, Wu Q. Estimation in structural equation modeling. In: Hoyle RH, editor. Handbook of structural equation modeling, The Guilford Press; 2012, p 164–80.
48. Edwards MC, Wirth RJ, Houts CR, Xi N. Categorical data in the structural equation modeling framework. In Hoyle, RH, editor. Handbook of structural equation modeling, The Guilford Press, 2012, p. 195–208.
49. Brown TA. Confirmatory factor analysis for applied research. New York, NY: The Guilford Press; 2006.
50. Hsieh YW, Hsueh IP, Chou YT, Sheu CF, Hsieh CL, Kwakkel G. Development and validation of a short form of the Fugl-Meyer motor scale in patients with stroke. Stroke. 2007;38:3052–4.
51. West SG, Taylor AB, Wu W. Model fit and model selection in structural equation modeling. In: Hoyle, RH, editor. Handbook of structural equation modeling, The Guilford Press: 2012, p. 209–31.
52. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc. 2006;54:743–9.
53. Yozbatiran N, Der-Yeghiaian L, Cramer SC. A standardized approach to performing the action research arm test. Neurorehabil Neural Repair. 2008;22:78–90.
54. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. The Montreal cognitive assessment, Moca: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc. 2005;53:695–9.
55. Hommel M, Trabucco-Miguel S, Joray S, Naegele B, Gonnet N, Jaillard A. Social dysfunctioning after mild to moderate first-ever stroke at vocational age. J Neurol Neurosurg Psychiatry. 2009;80:371–5.
56. Green SB, Thompson MS. A flexible structural equation modeling approach for analyzing means. In: Hoyle RH, editor. Handbook of structural equation modeling, The Guilford Press: 2012, p. 393–416.
57. Heckman J. Instrumental variables—a study of implicit behavioral assumptions used in making program evaluations. J Hum Resour. 1997;32:441–62.
58. Pearl J. The causal foundations of structural equation modeling. In Hoyle RH, editor. Handbook of structural equation modeling, The Guilford Press; 2012, p. 68–91.
59. McArdle JJ: Latent curve modeling of longitudinal growth data. In: Hoyle RH, editor. Handbook of structural equation modeling, The Guilford Press; 2012, p. 547–70.
60. Boissel JP, Cucherat M, Nony P, Chabaud S, Gueyffier F, Wright JM, Lievre M, Leizorovicz A. New insights on the relation between untreated and treated outcomes for a given therapy effect model is not necessarily linear. J Clin Epidemiol. 2008;61:301–7.
61. Mishra NK, Lyden P, Grotta JC, Lees KR. Thrombolysis is associated with consistent functional improvement across baseline stroke severity: a comparison of outcomes in patients from the virtual international stroke trials archive (vista). Stroke. 2010;41:2612–7.
62. Acock AC. Discovering structural equation modeling using stata. College Station: Stata Press Publication; 2013.
63. Tilley BC. Contemporary outcome measures in acute stroke research: choice of primary outcome measure and statistical analysis of the primary outcome in acute stroke trials. Stroke. 2012;43:935–7.