The distribution of the modified Rankin scale scores change according to eligibility criteria in acute ischemic stroke trials: A consideration for sample size calculations when using ordinal regression analysis

Verónica V. Olavarría a, Alejandro Brunser b, Norberto Cabral c, Sheila Martins d, Paula Muñoz-Venturelli b,e, Gabriel Cavada f, Pablo M. Lavados g,*

a Unidad de Neurología Vascular, Servicio de Neurología, Departamento de Neurología y Psiquiatría y Departamento de Paciente Crítico, Clínica Alemana de Santiago, Facultad de Medicina Clínica Alemana Universidad del Desarrollo, Santiago, Chile
b Unidad de Neurología Vascular, Servicio de Neurología, Departamento de Medicina Interna, Clínica Alemana de Santiago, Facultad de Medicina Clínica Alemana Universidad del Desarrollo, Santiago, Chile
c Universidade de Joinville (Unijove), Registro de AVC de Joinville, Clínica Neurologica de Joinville, Joinville, Brazil
d The George Institute for Global Health, The University of Sydney, Sydney, Australia
e Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil
f Unidad de Neurología Vascular, Servicio de Neurología, Departamento de Neurología y Psiquiatría, Clínica Alemana de Santiago, Facultad de Medicina Clínica Alemana Universidad del Desarrollo, Santiago, Chile
g Universidade de Joinville, Servicio de Neurología, Departamento de Ciencias Clínicas y Ensayos Clínicos, Departamento Científico Docente, Clínica Alemana de Santiago, Santiago, Chile

A R T I C L E   I N F O
Article history:
Received 4 August 2016
Received in revised form
19 December 2016
Accepted 10 January 2017
Available online 5 February 2017

Keywords:
Stroke
Sample size
Clinical trials
Ordinal logistic regression

A B S T R A C T

Background: Distribution shifts of the modified Rankin scale (mRs) scores change according to acute ischemic stroke (AIS) randomized controlled trials (RCT). Distribution across mRs strata is relevant for sample size calculations and may be affected by eligibility criteria. Aim: We aimed to assess the distribution of mRs scores across its different strata in AIS according to usual eligibility criteria. Methods: We computed follow-up mRs strata distribution between an unselected cohort and samples with (a) time from symptom onset < 6 h, (b) National Institutes of Health Stroke Scale (NIHSS) scores > 3 and < 25, and (c) both criteria combined. We compared distributions with Mann-Whitney U Test and calculated sample sizes for each distribution. Results: We included 5849 AIS patients. The unselected sample had a non-normal distribution with a median of 2. All selection criteria yielded significantly different distributions of mRs scores (p = 0.004, 0.02 and 0.02 respectively). This resulted in a significant variation in the calculated sample size when applying different selection criteria, with smaller numbers when RCT selection criteria are used (3616 versus 1553). Conclusions: The use of usual RCT eligibility criteria result in significant differences in mRs distribution and smaller sample sizes compared to unselected AIS samples.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The modified Rankin scale (mRs) is an ordinal disability score of 7 categories (0 = no symptoms to 6 = dead) and is currently the preferred method of evaluating outcome in patients with acute ischemic stroke (AIS) after intervention trials, particularly if the effect size is modest and distributed across all strata [1–3]. Thus, sample size calculations are increasingly determined by the distribution of the scores in the mRs at 90 days after symptom onset in modern AIS randomized clinical trials (RCT) [4–6].

http://dx.doi.org/10.1016/j.conctc.2017.01.008
2451-8654/© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
The data are analyzed using the t-test for difference of means or the Mann-Whitney U test or its variants - the Robust Rank Test and the Ordinal Logistic Regression (OLR) [7,8]. The latter also known as Whitehead’s Odds Ratio derives from the proportional odds model of McCullagh [9], and has been shown to produce the lowest set of criteria compared to the unselected sample. We then applied the method proposed by S. Simon to calculate sample sizes for an acute trial using the distribution of mRs according to the different sets of eligibility criteria [13]. We grouped categories mRs 5 and 6 into one, as is usual in the analysis of AIS RCTs. All sample sizes were calculated with a modest effect size of 2% between categories of mRs in each group, alfa was set at 0.05 and power at 80% and all calculation were 2 tailed.

2. Methods

In this retrospective cohort study, patients were selected from two prospective hospital stroke registries and one population-based registry (Joinville): two in Brazil (Porto Alegre and Joinville) and one in Chile (Santiago).

For the unselected cohort, we included all adult AIS patients (i.e. ≥ 18 years old) with clinical and imaging diagnosis of AIS, data available on NIHSS at admission, time from symptom onset to admission, mRs at follow up and informed consent given. Patients with Transient Ischemic Attacks and with incomplete data were excluded.

The following eligibility criteria have been used in many acute stroke RCTs and were applied to the cohort to produce 3 samples: (a) time from symptom onset < 6 h (b) National Institutes of Health Stroke Scale (NIHSS) scores > 3 and < 25, and (c) both criteria combined.

The corresponding institutional review board and ethics committee approved the 3 registries were these data were derived from: The RECCA registry in Clínica Alemana, the Joinville Stroke Registry and Hospital das Clínicas, Porto Alegre Stroke Registry.

2.1. Analysis

We computed skewness, kurtosis, median values and used Mann-Whitney U Test to compare mRs strata distribution for each set of criteria compared to the unselected sample. We then applied the method proposed by S. Simon to calculate sample sizes for an acute trial using the distribution of mRs according to the different sets of eligibility criteria [13]. We grouped categories mRs 5 and 6 into one, as is usual in the analysis of AIS RCTs. All sample sizes were calculated with a modest effect size of 2% between categories of mRs in each group, alfa was set at 0.05 and power at 80% and all calculation were 2 tailed.

3. Results

The total unselected cohort consisted of 5849 consecutive patients, the demographic and clinical characteristics of which are described for each individual sample in Table 1. The selection criteria yielded 3275 (56.0%) patients for NIHSS alone, 2636 (45.1%) for time alone and 1134 (19.4%) for both criteria combined (Table 2).

The unselected sample had a non-normal distribution of mRs,

---

**Table 1**

| Variables                      | Clínica Alemana, Santiago | Hospital das Clínicas, Porto Alegre | Joinville |
|--------------------------------|---------------------------|------------------------------------|-----------|
|                                | N = 1078                  | N = 1830                           | N = 2941  |
| Mean age (SD)                  | 70 (16)                   | 66 (14)                            | 66 (14)   |
| Female (%)                     | 539 (50)                  | 934 (51)                           | 1375 (48) |
| Hypertension (%)               | 722 (67)                  | 1492 (82)                          | 2170 (74) |
| Diabetes Mellitus (%)          | 226 (21)                  | 510 (28)                           | 957 (33)  |
| Hypercholesterolemia (%)       | 259 (24)                  | 829 (45)                           | 887 (30)  |
| Current smoking (%)            | 205 (19)                  | 373 (20)                           | 602 (20)  |
| Atrial fibrillation (%)        | 172 (16)                  | 338 (19)                           | 394 (10)  |
| Ischemic heart disease (%)     | 197 (9)                   | 418 (23)                           | 170 (6)   |
| Previous stroke/TIA (%)        | 259 (24)                  | 653 (36)                           | 988 (34)  |
| Cardiogenic (%)                | 366 (34)                  | 490 (27)                           | 844 (29)  |
| Atherothrombotic (%)           | 140 (13)                  | 423 (23)                           | 706 (24)  |
| Lacunar (%)                    | 119 (11)                  | 377 (21)                           | 625 (21)  |
| Undetermined (%)               | 420 (39)                  | 347 (19)                           | 597 (20)  |
| Other determined (%)           | 32 (3)                    | 193 (10)                           | 169 (6)   |
| NIHSS Median (IQR)             | 5 (2–10)                  | 6 (3–13)                           | 4 (2–11)  |

---

**Table 2**

| mRs scores | Unselected cohort | NIHSS >3 and < 25 | Time from onset <6 h | Both criteria |
|------------|-------------------|-------------------|----------------------|---------------|
|            | N (%)             | N (%)             | N (%)                | N (%)         |
| 0          | 1211 (20.1)       | 391 (11.9)        | 531 (20.1)           | 132 (11.6)    |
| 1          | 1664 (28.4)       | 694 (21.2)        | 656 (24.9)           | 258 (22.6)    |
| 2          | 695 (11.8)        | 447 (13.6)        | 283 (10.7)           | 136 (11.9)    |
| 3          | 620 (10.6)        | 486 (14.8)        | 278 (10.5)           | 142 (12.5)    |
| 4          | 628 (10.7)        | 539 (16.5)        | 318 (12.1)           | 167 (14.7)    |
| 5–6        | 1031 (17.6)       | 2718 (21.9)       | 570 (21.6)           | 299 (26.3)    |
| Total      | 5849              | 3275              | 2636                 | 1134          |
skewed to the left, with a median of 2 (less disability). When admission NIHSS was used, the mRs distribution was significantly less skewed, with a median of 3 ($p = 0.04$). When only time from symptom onset was used, the mRs distribution more skewed to lower values but nevertheless different from the unselected sample ($p = 0.02$). When both criteria were used, the distribution was less skewed with a median mRs of 3 and also significantly different form the whole sample cohort ($p = 0.02$) (Fig. 1).

A significant variation in the calculated sample size was found when applying different selection criteria to the sample, with need of a smaller recruited number of patients when usual RCT selection criteria are used (Table 3).

4. Discussion

Our results show sample sizes for OLR are significantly affected by the different distributions of follow-up mRs produced when applying usual eligibility criteria for AIS RCTs, being smaller when the resulting distribution is less skewed. In this sample, the effect was greater when stroke severity, measured with NIHSS, was used as a selection criterion together with time from symptom onset. We also found that the 6-h time window does NOT influence the sample size.

Even though mRs is an ordinal scale, it has been traditionally and still is analyzed as a binary outcome measure with many possible cut-off points [14–17]. Only recently large stroke trials have used its distribution across all strata as primary end point and thus for sample size calculations [18–21].

An important advantage of using ordinal or continuous analysis of mRs as primary outcome is the need for smaller sample sizes in effective therapies with small effect size or effects that are distributed across all strata [22]. This advantage may be neglected if the impact of patient selection through eligibility criteria on the distribution of mRs strata is not accounted for. This same analysis could apply to other areas of neurology and medicine that are using ordinal scales as outcome measures. A similar approach to calculate the appropriate sample size, could also be used in case an ordinal scale is dichotomized and the distribution of outcome varies in the subgroups after using different selection criteria.

The main strength of this study is that the populations were large and unselected cohorts of patients with AIS presenting at large academic medical centers that may be eligible for any acute treatment or clinical trial enrolment.

The main weakness is that we assumed an even effect size for all categories in mRs, that is improving 2% for 0–2 and worsening 2% for 3–6. This is not the case in many trials in which there could be an improvement in 0–1 and worsening of 5–6 categories or no change in mortality (mRs = 6).

In conclusion, we found that the application of usual eligibility criteria produced a 50% reduction in the required potential sample sizes for AIS RCT analyzing the mRs scale with OLR. The distribution of follow-up mRs scores are significantly different from the unselected data sample. Sample size calculations should consider this distribution in AIS, if primary outcome is ordinal analysis of the distribution of mRs scores with ordinal logistic regression.

Table 3

| Criteria | Median | Skewness | Kurtosis | Sample size |
|----------|--------|----------|----------|-------------|
| None     | 2      | 0.56     | −0.72    | 3616        |
| Time from onset <6 h | 2 | 0.53 | −0.85 | 3364 |
| NIHSS > 3 and <25 | 3 | 0.09 | −0.9 | 1544 |
| Both criteria | 3 | 0.18 | −1.07 | 1553 |

Fig. 1. Percentage distribution of patients modified Rankin scale scores at discharge in the total cohort according to usual acute ischemic stroke trials selection criteria.

Table 3

Descriptive statistics of samples and resulting sample sizes according to the distribution of modified Rankin scale scores at follow up after applying different selection criteria to the cohort and resulting odd ratios.

Funding

This work was supported by unconditional grants form from Lundbeck Chile and Clínica Alemana de Santiago. It also received financial support from the Fundacao de Amparo a Pesquisa do Estado de Santa Catarina, University of the Region of Joinville and the Joinville Municipal Health Department. Grant support: Fundacao de Apoio a Pesquisa Cientifica do Estado de Santa Catarina-FAPESC, University of the Region of Joinville Unive. The funding sources had no involvement in study design; collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.
Acknowledgements

We would like to thank our colleagues and all the staff in the participating centers, the patients and their families for making this work possible.

References

[1] J. Rankin, Cerebral vascular accidents in patients over the age of 60. II. Prognosis, Scott. Med. J. 2 (5) (1957 May) 200–215.
[2] J.L. Saver, Novel end point analytic techniques and interpreting shifts across the entire range of outcome scales in acute stroke trials, Stroke 38 (11) (2007 Nov) 3055–3062.
[3] J.L. Saver, Optimal end points for acute stroke therapy trials: best ways to measure treatment effects of drugs and devices, Stroke 42 (8) (2011 Aug) 2356–2362.
[4] K.R. Lees, P.M.W. Bath, P.D. Schellinger, D.M. Kerr, R. Fulton, W. Hacke, et al., Contemporary outcome measures in acute stroke research: choice of primary outcome measure, Stroke 43 (4) (2012 Apr) 1163–1170.
[5] J.L. Saver, S. Starkman, M. Eckstein, S.J. Stratton, F.D. Pratt, S. Hamilton, et al., Prehospital use of magnesium sulfate as neuroprotection in acute stroke, N. Engl. J. Med. 372 (6) (2015 Feb 5) 528–536.
[6] ENOS Trial Investigators, P.M.W. Bath, L. Woodhouse, P. Scutt, K. Krishnan, J.M. Wardlaw, et al., Efficacy of nitric oxide, with or without continuing anti-hypertensive treatment, for management of high blood pressure in acute stroke (ENOS): a partial-factorial randomised controlled trial, Lancet 385 (9968) (2015 Feb 14) 617–628.
[7] M.J. Campbell, S.A. Julious, D.C. Altman, Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons, BMJ 311 (7013) (1995 Oct 28) 1145–1148.
[8] P.M.W. Bath, L.J. Gray, T. Collier, S. Pocock, J. Carpenter, Can we improve the statistical analysis of stroke trials? Statistical reanalysis of functional outcomes in stroke trials, Stroke 38 (6) (2007 Jun) 1911–1915.
[9] P. McCullagh, Regression models for ordinal data, J. R. Stat. Soc. Ser. B Methodol. 42 (2) (1980 Jan 1) 109–142.
[10] Optimising Analysis of Stroke Trials Collaboration, Calculation of sample size for stroke trials assessing functional outcome: comparison of binary and ordinal approaches, Int. J. Stroke 3 (2) (2008 May) 78–84.
[11] J. Whitehead, Sample size calculations for ordered categorical data, Stat. Med. 12 (24) (1993 Dec 30) 2257–2271.
[12] S.J. Walters, Sample size and power estimation for studies with health related quality of life outcomes: a comparison of four methods using the SF-36, Health Qual. Life Outcomes 2 (2004 May) 26.
[13] Stats: Sample Size for an Ordinal Outcome, September 22, 2004 [Internet]. [cited 2015 Jul 18]. Available from: http://www.pmean.com/04/OrdinalLogistic.html.
[14] P.W. Duncan, H.S. Jorgensen, D.T. Wade, Outcome measures in acute stroke trials: a systematic review and some recommendations to improve practice, Stroke 31 (6) (2000 Jun) 1429–1438.
[15] AVERT Trial Collaboration group, J. Bernhardt, P. Langhorne, R.I. Lindley, A.G. Thrift, F. Eley, et al., Efficacy and safety of very early mobilisation within 24 h of stroke onset (AVERT): a randomised controlled trial, Lancet 386 (9988) (2015 Jul 4) 46–55.
[16] IST-3 collaborative group, P. Sandercoc, J.M. Wardlaw, R.I. Lindley, M. Dennis, G. Cohen, et al., The benefits and harms of intravenous thrombolysis with recombinant tissue plasminogen activator within 6 h of acute ischaemic stroke (the third international stroke trial [IST-3]): a randomised controlled trial, Lancet 379 (9834) (2012 Jun 23) 2352–2363.
[17] C.S. Anderson, E. Heeley, Y. Huang, J. Wang, C. Stapf, C. Delcourt, R. Lindley, T. Robinson, P. Lavados, B. Neal, J. Hata, H. Arima, M. Parsons, Y. Li, J. Wang, S. Herrtger, Q. Li, M. Woodward, R.J. Simes, S.M. Davis, J. Chalmers, INTERACT2 Investigators, Rapid blood-pressure lowering in patients with acute intracerebral hemorrhage, N. Engl. J. Med. 368 (25) (2013 Jun 20) 2355–2365.
[18] M. Goyal, A.M. Demchuk, B.K. Monon, M. Eesa, J.I. Rempel, J. Thornton, et al., Randomized assessment of rapid endovascular treatment of ischemic stroke, N. Engl. J. Med. 372 (11) (2015 Mar 12) 1019–1030.
[19] J.L. Saver, M. Goyal, A. Bonahe, H.-C. Diener, E.I. Levy, V.M. Pereira, et al., Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke, N. Engl. J. Med. 372 (24) (2015 Jun 11) 2285–2295.
[20] T.G. Jovin, A. Chamorro, E. Cobo, M.A. de Miquel, C.A. Molina, A. Rozvi, et al., Thrombectomy within 8 hours after symptom onset in ischemic stroke, N. Engl. J. Med. 372 (24) (2015 Jun 11) 2296–2306.
[21] P. Muñoz-Venturelli, H. Arima, P. Lavados, A. Brunser, B. Peng, L. Cui, et al., Head Position in Stroke Trial (HeadPoST) – sitting-up vs lying-flat positioning of patients with acute stroke: study protocol for a cluster randomised controlled trial, Trials 16 (2015) 256.
[22] P.M.W. Bath, K.R. Lees, P.D. Schellinger, H. Altman, M. Bland, C. Hogg, et al., Statistical analysis of the primary outcome in acute stroke trials, Stroke 43 (4) (2012 Apr) 1171–1178.