Transforming self-reported outcomes from a stroke register to the modified Rankin Scale: a cross-sectional, explorative study

Tamar Abzhandadze¹,², Malin Reinholdsson³,², Annie Palstam¹, Marie Eriksson³ & Katharina S. Sunnerhagen¹

The aim was to create an algorithm to transform self-reported outcomes from a stroke register to the modified Rankin Scale (mRS). Two stroke registers were used: the Väststroke, a local register in Gothenburg, Sweden, and the Riksstroke, a Swedish national register. The reference variable, mRS (from Väststroke), was mapped with seven self-reported questions from Riksstroke. The transformation algorithm was created as a result of manual mapping performed by healthcare professionals. A supervised machine learning method—decision tree—was used to further evaluate the transformation algorithm. Of 1145 patients, 54% were male, the mean age was 71 y. The mRS grades 0, 1 and 2 could not be distinguished as a result of manual mapping or by using the decision tree analysis. Thus, these grades were merged. With manual mapping, 78% of the patients were correctly classified, and the level of agreement was almost perfect, weighted Kappa (Kw) was 0.81. With the decision tree, 80% of the patients were correctly classified, and substantial agreement was achieved, Kw = 0.67. The self-reported outcomes from a stroke register can be transformed to the mRS. A mRS algorithm based on manual mapping might be useful for researchers using self-reported questionnaire data.

Stroke registers are valuable data sources for scientific studies and for understanding the consequences of stroke. For stroke quality registers, data are gathered by staff or reported by patients. Self-reported information might have validity and reliability issues, but incorporating standardized assessment tools into quality registers can lead to an increased number of questions that might be difficult for the patients to answer.

The modified Rankin Scale (mRS) is one of the frequently used assessment instruments in stroke-related studies, but the administration of this instrument is staff dependent; the assessments are to be performed by trained personnel, either in person or by telephone interviews. Therefore, it can be difficult to obtain information from large geographical areas and at several time points. Eriksson et al. transformed five self-reported outcomes from the Swedish national stroke register, Riksstroke, to the mRS. The manual mapping method was used. Eriksson et al.'s algorithm allowed Riksstroke-based research to be compared across studies in which the mRS was used. Since the first transformation algorithm was developed, the questions in Riksstroke have been changed. Moreover, the previous algorithm could not distinguish mRS grades 0, 1 and 2 from each other. The supervised machine learning method could potentially be used to transform the self-reported data with more accuracy. The machine learning algorithms help label the input data and group the output into several classes.

The aims of this study were to create a new transformation algorithm for the modified Rankin Scale-Riksstroke (mRS-RS) based on self-reported outcomes from Riksstroke and to distinguish mRS-RS grades 0, 1 and 2 from each other. These aims were addressed by using a combination of manual mapping and a supervised machine learning method.

¹Institute of Neuroscience and Physiology, Rehabilitation Medicine, University of Gothenburg, Per Dubbsgatan 14, fl. 3, 413 45 Gothenburg, Sweden. ²Department of Occupational Therapy and Physiotherapy, Sahlgrenska University Hospital, Gothenburg, Sweden. ³Department of Statistics, USBE, Umeå University, Umeå, Sweden. *email: tamar.abzhandadze@gu.se
Methods

Study sample and procedure. This is a cross-sectional, explorative, register-based study, part of the Physical Activity Pre-Stoke In GOThenburg project. Two quality registers were used: the Väststroke, a local stroke register in Gothenburg, Sweden, and the Riksstroke, a Swedish national register for stroke. Acute care and three-month follow-up data were collected from January 1, 2015 to August 31, 2018 were extracted. Data for acute care were registered by hospital staff. The 3-month follow-up data were collected using postal, self-administered questionnaires as well as telephone interviews by trained nurses. The Väststroke and Riksstroke registers were linked by a statistician at Riksstroke by using the patients’ unique personal identification numbers.

The inclusion criteria applied in this study were as follows: patients with first-ever stroke, diagnosed according to the International Classification of Diseases codes (ICD-10), those with an age > 18 y, complete data on those patients with mRS data registered in the Väststroke register and responses to Riksstroke’s questions that were to be used in the algorithm, as shown in Supplementary Table S1.

Ethics and informed consent. The data file that was used in the study was anonymized and individual patients could not be identified. The study was approved by the regional ethical review board in Gothenburg (#346–16) and the Swedish Ethical Review Authority (the amendment #2019-01251/346-16). The Declaration of Helsinki was followed. Informed consent: according to the Swedish Data Protection Authority, the handling of data generated within the framework of quality registers represents an exception from the general rule requiring written informed consent from the patients. Furthermore, the Personal Data Act (Swedish law #1998:204, issued 29 April 1998) allows data from medical charts to be collected for clinical purposes and quality control without written informed consent.

Variables. Variables used for the development of the modified Rankin Scale-Riksstroke (mRS-RS) algorithm. The reference variable mRS was derived from Väststroke’s three-month follow-up data. The mRS data were collected by experienced physicians or nurses. The telephone interview guidelines described by Bruno et al. were followed. The mRS is used to determine a patient’s level of functional disability after stroke; it uses a 7-level ordinal scale, with 0 = “no symptoms” and 6 = “death”.

The reference variable mRS (Väststroke) was ordinal. There were both ordinal and nominal index variables, which were the seven questions from Riksstroke. Thus, the nonparametric method was used for building the decision trees. The dataset was divided into training (80%) and test sets (20%). Several measurements were used to evaluate the performance of the decision trees. The classification accuracy was used to identify the overall rate of correctly classified patients. Classification accuracy was further compared with the no information rate to determine the usability of the model. The quadratic $K_r$ was used to study the agreement between the true and predicted values of the mRS. For each classification tree, confusion matrices were created.
Three individual decision trees were built:

Tree 1. For the validation of the mRS-RS algorithm that was developed based on manual mapping, we used the mRS with grades 0–2, 3, 4 and 5 as a target variable.

Tree 2. To distinguish mRS grades 0, 1 and 2 from each other, the full-scale mRS was used as a target variable.

Tree 3. Due to clinical reasoning and manual mapping issues, the target variable was the mRS with grades 0–1, 2–3 and 4–5, which indicated no disability, moderate disability and severe disability, respectively.

Descriptive statistical analyses of the study sample and agreement analyses were performed in SPSS Statistics 26.0 (IBM SPSS Statistics for Windows, Armonk, NY: IBM Corp). Decision trees were built using the conditional inference trees (ctree) approach from the toolkit for recursive partitioning (partykit) package17,18 (R, version 3.6.2). All tests were two-sided and conducted at the 5% significance level.

Results

The study sample, baseline demographics and clinical characteristics. In total, 1145 of 3567 patients with a first-ever stroke met the inclusion criteria (Fig. 1). Patients (n = 2245) were excluded either due to death or missing data at 3 months. Compared with the included patients, the excluded patients were more male (p < 0.05), had more severe stroke (p < 0.001) and an older age (p < 0.001).

The mean (SD) age of the patients was 71 y (14.3 y), and 54% were male. Before the stroke, the majority of the patients lived in their own homes without community services (87%). They were independent in terms of mobility (92%), toilet visits (97%) and getting dressed (96%). At the onset of stroke, 79% of patients had a mild stroke (NIHSS ≤ 5p), 89% had cerebral infarction, and 17% of the patients with infarction had received reperfusion treatment. Furthermore, 55% of the patients showed impaired cognition (≤ 25p), as assessed with the MoCA, during the hospital stay. The majority of the patients (72%) were discharged to their own homes (Table 1). Three months after the stroke, 21% of the patients had no symptoms (mRS = 0), 86% of the patients were living in their own homes with or without community services, but only 29% had been able to return to the lifestyle and the activities they had performed before the stroke (Table 1).

Manual mapping-first transformation algorithm of mRS-RS. The results of the new manual mapping procedure were combined with a transformation algorithm that was previously developed by Eriksson et al., leading to the first version of the mRS-RS (Table 2). The distribution density of the mRS grades from Väststroke and Riksstroke’s 7 questions is presented in Supplementary Figure S1. The manual mapping results could not be used to distinguish between mRS grades 0, 1 and 2. Thus, this aim could not be fulfilled.
To develop a new transformation algorithm, mRS grades 0, 1 and 2 were merged (Table 2). The new algorithm achieved almost perfect agreement $K_w = 0.81$ ($p < 0.001$, 95% CI 0.77–0.85), Table 3. With the new transformation algorithm, 78% of the patients were correctly classified, 6% of the patients were given lower mRS grades than the reported grades, and 15% were given higher mRS grades (Table 3, section a). The number of correctly classified patients for each mRS grade as well as agreement on the individual mRS grades are presented in Supplementary Figure S2.

Decision trees. With the decision tree in which the mRS grades 0, 1 and 2 were merged, 80% of the patients were correctly classified. The level of agreement was substantial, with $K_w = 0.66$ and 0.67 for the training and testing sets, respectively (Table 4). The proportions of the patients that were correctly classified were similar for the decision tree (the training and testing sets) and manual mapping procedures; however, the decision tree could not be used to classify the patients with an mRS grade 4, which was a major difference between the two procedures (Table 3). The decision trees that were used to distinguish mRS grades 0, 1, 2 yielded correct classifications for only 46% of the patients (Supplementary Table S2—a). The overall $K_w$ indicated moderate agreement (0.58) (Table 4).

The decision trees are presented in Fig. 2.

### Table 1. Baseline demographics and clinical characteristics of the study sample (n = 1145). The sum may be different because of the missing values. Abbreviations: * SD standard deviation, † TIA a Transient Ischemic Attack, ‡ RLS the Reaction Level Scale (the range 1–8, where 1 means fully awake), § NIHSS the National Institute of Health Stroke Scale (the scores range from 0–42 points, a lower score indicates a less severe neurological status), ‖ (Q1–Q3) — the first quartile–the third quartile, ‖ MoCA—the Montreal Cognitive Assessment (the scores range from 0–30 points, a low score indicates more severe cognitive deficits). Variables with missing data n (%), presented in alphabetical order: diabetes 1 (≤ 1%), level of consciousness upon arrival at the hospital 17 (≤ 1%), lives alone 11 (≤ 1%), MoCA score 609 (53%), needed help prior to stroke 14 (≤ 1%), NIHSS 238 (21%), a history of TIA †, yes, n (%) 69 (6%), smoking, yes, n (%) 130 (13%).

| Baseline characteristics          | n (%)       |
|-----------------------------------|-------------|
| Female sex, n (%)                 | 523 (46%)   |
| Age in years, mean (SD*/range)    | 71 (14.3/19–100) |
| Lives in own home with/without community services, n (%) | 1105 (96%) |
| Lives in community facility or other, n (%) | 40 (4%) |
| Lives alone, yes/no, n (%)        | 501 (44%)/633 (56%) |
| Needed help prior to the stroke, yes/no, n (%) | 134 (14%)/977 (84%) |
| Diabetes, yes, n (%)              | 187 (16%)   |
| A history of TIA†, yes, n (%)     | 69 (6%)     |
| Smoking, yes, n (%)               | 130 (13%)   |

| Stroke-related features           | n (%)       |
|-----------------------------------|-------------|
| Stroke diagnosis, n (%)           | 127 (11%)   |
| Cerebral haemorrhage              | 1016 (89%)  |
| Cerebral infarctions              | 2 (0.2%)    |
| Reperfusion, n (%)                | 183 (17%)   |
| NIHSS§, median (Q1–Q3)‖/range     | 1 (0–5)/(0–26) |
| MoCA#, median (Q1–Q3)‖/range      | 25 (22–27)/(5–30) |
| Length of hospital stay in stroke units, days, median (Q1–Q3)‖/range | 7 (4–16)/(0–100) |

| Discharge destination, n (%)      |           |
|-----------------------------------|-----------|
| Own home                          | 824 (72%) |
| Community facility                | 174 (15%) |
| Another acute clinic              | 13 (1%)   |
| Geriatric/rehab unit              | 132 (11%) |
| Other stroke units                | 2 (0.2%)  |

| Functional outcome 3 months after stroke (modified Rankin Scale), n (%) |           |
|------------------------------------------------------------------------|-----------|
| No symptoms at all                                                    | 242 (21%) |
| No significant disability despite symptoms                             | 314 (27%) |
| Slight disability, unable to carry out all previous activities         | 264 (23%) |
| Moderate disability, requiring some help, able to walk without assistance | 178 (15%) |
| Moderately severe disability, unable to walk without assistance        | 85 (7%)   |
| Severe disability, bedridden                                          | 62 (5%)   |

To develop a new transformation algorithm, mRS grades 0, 1 and 2 were merged (Table 2). The new algorithm achieved almost perfect agreement $K_w = 0.81$ ($p < 0.001$, 95% CI 0.77–0.85), Table 3. With the new transformation algorithm, 78% of the patients were correctly classified, 6% of the patients were given lower mRS grades than the reported grades, and 15% were given higher mRS grades (Table 3, section a). The number of correctly classified patients for each mRS grade as well as agreement on the individual mRS grades are presented in Supplementary Figure S2.
Table 2. The manual mapping results. The translation algorithm used to transform the self-reported questions from Riksstroke into modified Rankin Scale (mRS) grades. *The answers of the multiple-choice questions are given as stated in Riksstroke’s follow-up questionnaire.

| mRS grades | The answer choices to the self-reported questions from Riksstroke* | Answer codes as defined in Supplementary Table S1 |
|------------|---------------------------------------------------------------|-----------------------------------------------|
| 0–2        | Q1 Not or partly dependent on support or assistance from relatives/friends OR Q6 All problems have completely gone OR Q7 Can return to the life and activities I had prior to stroke Q2 Live in my own home, without community services Q3 Can get around both indoors and outdoors without the help of another person Q4 Can manage to visit the toilet by myself Q5 Can manage to get dressed and undressed by myself | 3 or 4 1 or 2 1 or 2 1 | |
| 3          | Q1 Completely dependent on a next of kin for help/support OR Q2 Live in own home with community support OR Q3 Can move around without help at least indoors Q6 All problems have completely gone | 2 2 1 or 2 2 |
| 4          | Q3 Cannot move around without help indoors OR Q4 Need help to go to the toilet OR Q5 Need help to get dressed and undressed Q6 All problems are completely resolved Q7 Cannot return to the life and activities I had prior to stroke | 3 2 2 2 3 |
| 5          | Q2 Not living in my own home Q3 Need another person’s help to move Q4 Need help to go to the toilet Q5 Need help to get dressed and undressed | 3 2 2 2 |

Table 3. The confusion matrices of the manual mapping algorithm and the decision tree for the modified Rankin Scale grades 0—2, 3, 4 and 5. The results are presented as the number of patients. n = 981 patients due to characteristics of the manual translation algorithm comprising OR and AND conditions.

| Predicted value | True value | a. Manual mapping (n = 981) | b. Decision tree, train data set (n = 911) | c. Decision tree, test data set (n = 234) |
|-----------------|------------|-----------------------------|---------------------------------------------|-------------------------------------------|
| mRS grades 0-2, 3, 4, 5 | 0 - 2 | 623 90 8 5 | 0 - 2 | 626 78 8 1 | 0 - 2 | 160 18 4 2 |
| 3 | 32 76 15 5 | 3 | 23 55 25 | 3 | 13 4 1 |
| 4 | 2 13 19 27 | 4 | 0 0 0 0 | 4 | 0 0 0 0 |
| 5 | 0 0 9 44 | 5 | 5 12 33 | 5 | 2 1 12 13 |

Table 4. Summary statistics of the three decision trees based on different versions of the modified Rankin Scale (Väststroke) variables. CA—classification accuracy, the rate of correctly classified patients, Kω—quadratic weighted kappa, 95% CI—95% confidence interval for Kω, NIR—no information rate. *n = 981 patients due to characteristics of the manual translation algorithm comprising OR and AND conditions.

| mRS variables | Decision tree, training data (n = 911) | Decision tree, test data (n = 234) | Manual mapping (n = 981)* |
|---------------|----------------------------------------|-----------------------------------|---------------------------|
| mRS grades 0-2, 3, 4, 5 | CA | 0.79 | 0.61 - 0.70 | 0.71 | 0.80 | 0.58 - 0.76 | Kω | 0.77 - 0.85 |
| mRS grades 0-1, 2-3, 4-5 | CA | 0.46 | 0.53 - 0.60 | 0.26 | 0.45 | 0.51 - 0.65 | Kω | 0.32 |
| mRS grades 0-1, 2-3, 4-5 | CA | 0.70 | 0.52 - 0.61 | 0.48 | 0.67 | 0.47 - 0.65 | Kω | 0.51 |
Figure 2. Graphical representation of the decision trees. The classification of Riksstroke's seven questions into the modified Rankin Scale (mRS). The different colours indicate various mRS grades. Note ***p < 0.001; **p < 0.01; *p < 0.05.
Discussion

This study demonstrates a methodological approach for transforming self-reported functional outcomes from a stroke register into the mRS. Manual mapping and a supervised machine learning method, namely, decision trees, were used. The mRS grades 0, 1 and 2 could not be distinguished from each other with either of these methods; therefore, these grades were merged. Substantial classification accuracy and almost perfect agreement for manual mapping were obtained as the result of the implementation of both methodological approaches. However, the patients with mRS grade 4 could not be classified with the decision trees. Therefore, the results of the manual transformation algorithm can be used for comparative studies based on Riksstroke data, where the mRS score is an outcome variable.

One of the aims of this study was to distinguish mRS grades 0, 1 and 2. The aim was challenging, although the mapping was conducted by experienced health care professionals in stroke rehabilitation (T.A., M.R., K.S.S.). The aim could not be fulfilled after applying a machine learning method for classification purposes. There are several explanations for this result. First, although Riksstroke’s questions included in the algorithm have been validated, they lack specificity in the wording. Furthermore, multiple choice questions from Riksstroke register are binary or ordinal and grouped into three categories. This limitation in the response options can lead to difficulty in distinguishing “no symptoms”, “no significant disability” and “slight disability”. Second, the limited accuracy of self-reported information should be considered in interpreting the results of the study. Motivational and cognitive processes of the patients can bias their responses to the self-reported questions. Third, the data on the mRS (Väststroke) were obtained by several research nurses using the telephone interview guideline. Although the interview guideline has shown substantial interrater agreement, it showed a lower level of agreement between some grades of mRS.

The other aim of the study was to create a new transformation algorithm based on the Riksstroke questions. Although mRS grades 0, 1 and 2 were merged, misclassification could not be avoided; the results were overestimated for 15% of the patients and underestimated for 6%. The current study results are relatively similar to the classification characteristics of the first transformation algorithm developed by Eriksson et al. In that study, the results were overestimated for slightly fewer patients (14.2%) underestimated for more patients (9.5%). Furthermore, the proportion of correctly classified patients was substantial in this study, which was in agreement with the results of Eriksson et al. However, Eriksson et al. reported a lower kappa compared with that in our study. When a machine learning method was applied for the sensitive analyses, similar levels of classification accuracy were achieved for the training and testing datasets, but mRS grade 4 could not be classified. In conclusion, it is suggested that the manual algorithm for mRS-RS, where the mRS grades are coded as 0–2, 3, 4 and 5, can be useful for the classification of functional disability.

In this study, it was challenging to distinguish mRS grades 2 from 3 and 4 from 5. The same issue was mentioned by Eriksson et al. The clinical model was created by merging mRS grades 0–1 (no functional disability/functional independence), 2–3 (moderate functional disability) and 4–5 (severe functional disability). Choosing mRS grades 0–1 for functional independence was suggested in different stroke trials and linked with independence in everyday life as measured by the Barthel Index. The decision tree model showed substantial classification accuracy and moderate agreement. Implementing this model could avoid misclassification of the patients, especially towards better outcomes that can have extensive consequences for patients’ everyday lives, because of the risk that they will not receive the necessary care and rehabilitation.

This study has several strengths and limitations. Cohen’s kappa is a robust statistical method feasible for studying interrater agreement, but there are several uncertainties with kappa. Gisev et al. argue that based on the formula of the kappa, it can be difficult to achieve perfect agreement; moreover, low kappa does not always correspond to low agreement. Hence, we have chosen to also present the classification accuracy (%). Weighted kappa represents an extension of Cohen’s kappa, and it is useful for rating items with more than two categories. Studying the agreement between major categories and identifying how they differ from each other. However, the Kappa coefficients can differ by the number of categories, which makes it difficult to compare the coefficients with each other. Furthermore, interrater agreement bias, as well as disagreement in the classification items, are difficult to avoid. These issues were not present in the Eriksson et al. study, where one experienced nurse gathered the data on the mRS as well as Riksstroke’s 3-month follow-up questionnaire.

This study was restricted to first strokes, and the results may differ in recurrent strokes. From a patient perspective, there can be a big difference from having no symptoms at all from a stroke (mRS 0) and having a stroke that leaves you dependent on others for certain activities (mRS 2). By combining these values, we have probably missed important information on patient recovery. The use of the full ordinal scale would be more efficient, prospective, there can be a big difference from having no symptoms at all from a stroke (mRS 0) and having a stroke that leaves you dependent on others for certain activities (mRS 2). By combining these values, we have probably missed important information on patient recovery. The use of the full ordinal scale would be more efficient, however, the results of the manual mapping of the Riksstroke questions to mRS outcome of good (mRS grades 0–2) and poor outcome (mRS grades 3, 4 and 5) is still useful when the aim is to classify patients.

Supervised machine learning methods can be used in register-based studies. By applying the decision tree method, a more sensitive method of classification was expected but only partly achieved. It is possible that the data were not balanced. Several physicians and research nurses collected the mRS data over 31 months. The physicians/nurses did not undergo training for calibration, and the interrater agreement between the nurses/physicians could not be assessed because of the high rate of employee turnover in the stroke units. Furthermore, decision trees tend to have high accuracy because of overfitting. This problem was addressed by introducing the rule of minimal split at n = 50.

Conclusions

To the best of our knowledge, this is the first study in which a combination of manual mapping and a machine learning method was applied to transform register-based self-reported functional outcomes to a standardized assessment instrument by using the data from two stroke registers. The method presented in this study can be
References
1. Lin, C. H. et al. Evaluation of machine learning methods to stroke outcome prediction using a nationwide disease registry. *Comput. Methods Prog. Biomed.* **190**, 105381. https://doi.org/10.1016/j.cmpb.2020.105381 (2020).
2. McDermid, I. et al. Home-time is a feasible and valid stroke outcome measure in national datasets. *Stroke* **50**, 1282–1285 (2019).
3. Rankin, J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. *Scott. Med. J.* **2**, 200–213 (1957).
4. Banks, J. L. & Marotta, C. A. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. *Stroke* **38**, 1091–1096 (2007).
5. Broderick, J. P., Adeoye, O. & Elmn, J. Evolution of the modified Rankin scale and its use in future stroke trials. *Stroke* **48**, 2007–2012 (2017).
6. Drozdowska, B. A., Singh, S. & Quinn, T. J. Thinking about the future: a review of prognostic scales used in acute stroke.
7. Bruno, A. Evolution of the modified Rankin scale and its use in future stroke trials. *Stroke* **48**, 2007–2012 (2017).
8. Lekander, I. June 10. Lekander, I.
9. Banks, J. L. & Marotta, C. A. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. *Stroke* **38**, 1091–1096 (2007).
10. McHugh, M. L. Interrater reliability: the kappa statistic. *Fam. Med.* **37**, 360–363 (2005).
11. Hothorn, T., Hornik, K. & Zeileis. A. Biased recursive partitioning: a conditional inference framework. *J. Comput. Graph. Stat.* **15**, 651–674 (2006).
12. Hothorn, T. & Zeileis. A. Partykit: a modular toolkit for recursive partitioning in *R*. *J. Mach. Learn. Res.* **10**, 276–282 (2012).
13. Gislov, N., Bell, J. S. & Chen, T. F. Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Res. Soc. Adm. Pharm.* **9**, 330–338 (2013).
14. Viera, A. J. & Garrett, J. M. Understanding interobserver agreement: the kappa statistic. *Eur. J. Epidemiol.* **22**, 276–282 (2007).
15. Nyhuis, J. C. & Chen, T. F. Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Res. Soc. Adm. Pharm.* **9**, 330–338 (2013).
16. Fleiss, J. L. & Cohen, J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educa. Psychol. Meas.* **33**, 613–619 (2016).
17. Brenner, H. & Kliebsch, U. Dependence of weighted kappa coefficients on the number of categories. *Epidemiology* **7**, 199–202 (1996).
18. Maclure, M. & Willett, W. C. Misinterpretation and misuse of the kappa statistic. *Am. J. Epidemiol.* **126**, 161–169 (1987).
19. Viera, A. J. & Garrett, J. M. Understanding interobserver agreement: the kappa statistic. *Eur. J. Epidemiol.* **22**, 276–282 (2007).
20. Vanbelle, S. A new interpretation of the weighted kappa coefficients. *Psychometrika* **81**, 399–410 (2016).
21. König, M. A. & Altman, N. Classification and regression trees. *Nat. Methods.* **14**, 757–758 (2017).

Acknowledgements

We thank all health care professionals who gathered and recorded the data in the Rikstroke and Väststroke registers. We also thank the patients and caregivers for answering the 3-month follow-up questionnaires.

Author contributions

T.A.—designed the study, analyzed and interpreted the data, and drafted the manuscript. M.R.—retrieved data from the registers, analyzed the data, and revised the manuscript for intellectual content. A.P.—revised the manuscript for intellectual content. M.E.—designed the study, interpreted the data, and revised the manuscript for intellectual content. K.S.S.—analyzed and interpreted the data and revised the manuscript for intellectual content. All authors approved the submitted version of the manuscript.

Funding

Open Access funding provided by Gothenburg University Library. The study was financed by grants from the Swedish state under an agreement between the Swedish government and the county councils, the ALF agreement (ALFGBG-718711, ALFGBG-877961), the Swedish National Stroke Association, the Local Research and Development Board for Gothenburg and Södra Bohuslän, Greta and Einar Askér’s Foundation, Rune and Greta Almöv’s Foundation for Neurological Research, Hjalmar Svensson’s Research Foundation, Gunn and Bertil Stohne’s Foundation, Herbert and Karin Jacobson’s foundation, and Doktor Felix Neubergh’s foundation.
Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-73082-4.

Correspondence and requests for materials should be addressed to T.A.

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

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020