Transformative Medicine (T-Med)

Volume 1 | Number 2 Article 1

June 2022

The Modified Rankin Scale Can Accurately Be Derived From the Electronic Medical Record

Derek Isenberg
Temple University

Nelson Prus
Temple University

Frederick Ramsey
Temple Univeristy

Nina T. Gentile
Temple Univeristy

Recommended Citation
Isenberg D, Prus N, Ramsey F, Gentile NT. The Modified Rankin Scale Can Accurately Be Derived From the Electronic Medical Record. *Transformative Medicine (T-Med)*. 2022; 1(2):31-35. doi: https://doi.org/10.54299/tmed/jidk4591.

Original Reseach

The Modified Rankin Scale Can Accurately be Derived from the Electronic Medical Record

Derek Isenberg[1], Nelson Prus[1],
Frederick Ramsey[1], Nina T. Gentile[1]

1. Lewis Katz School of Medicine at Temple University

Published June 2022

ABSTRACT

INTRODUCTION: The modified Rankin Scale (mRS) is used to assess functional outcomes after a stroke and is the primary outcome in many stroke trials. For retrospective stroke research or stroke research in which patients are lost to study follow up, review of the electronic medical record (EMR) may be the sole way to estimate a patient's functional outcome. The purpose of this study is to determine if a mRS can be accurately derived from the electronic medical record EMR.

METHODS: This study used data from completed stroke studies in which in-person 90-day mRS were collected as part of the study protocol. These scores served as the reference standard. The EMR was searched to find a clinician note from the corresponding time as the 90-day post stroke assessment. These notes were given to three reviewers (an undergraduate research assistant, a medical student, and a neurology resident) to determine a mRS. Their scores were then compared to the in-person assessment and a kappa statistic was calculated.

RESULTS: 98 records were reviewed of which 60 met inclusion criteria. Comparing against the in-person mRS: the resident had a weighted kappa (kw) of 0.72, the medical student 0.71, and the research assistant 0.43. Aggregating the mRS into good outcome (mRS 0-2) vs poor outcome (mRS 3-5): the resident had a kw of 0.71, the medical student 0.78, and the research assistant 0.48.

DISCUSSION: This study demonstrates that both an absolute mRS and dichotomized mRS can be extracted from the EMR with good agreement by a medical student and neurology resident, but not by a research coordinator (with no formal medical education). Retrospective determination of a dichotomized mRS may be slightly more accurate than an absolute mRS. Researchers may use the EMR to estimate functional outcomes after stroke when in person assessment is not available.

Correspondence to Derek Isenberg at
Derek.isenberg@tuhs.temple.edu

Disclosure Statement: Nina Gentile received network support from the National Institute of Neurological Disorders and Stroke and the National Heart, Lung, and Blood Institute. Derek Isenberg received a grant for startup of OPUS-REACH from the American Heart Association. The remaining authors have no conflicts of interest to declare.

INTRODUCTION

The modified Rankin Scale (mRS) is commonly used to assess functional activity after a stroke.[1] The mRS grades the global disability of a stroke patient on a 0-6 scale with 0 being no symptoms and 6 being death. The mRS is generally obtained by in-person or telephone interview.[2-4] However, patients in acute stroke trials are sometimes lost to study follow up after discharge from the hospital. Although patients may not participate in the study follow up, they often have other contact with the healthcare system for routine or emergency care. Given that stroke studies need to examine long term outcomes, it would be helpful to be able to determine a mRS via chart review for patients lost to study follow up.

A 2008 study showed poor agreement with in-person determination compared to case-record appraisers, with a kappa of 0.34.[5] That study determined that, given the poor agreement, an "accurate mRS cannot be derived from standard hospital records." Given the advent of the electronic medical record (EMR), we hypothesized that electronic charts would have increased legibility over written records and provide a clearer view of the patient's functional status as compared to paper charts.

METHODS

This study utilized patient data from stroke studies performed at Temple University in which in-person 90-day mRS were collected as part of the study protocol. Patients were included from the following stroke trials: Stroke Hyperglycemia Insulin Network Effort (SHINE), Platelet-Oriented Inhibition in New TIA and minor ischemic stroke (POINT), Acute Stroke or Transient Ischemic Attack Treated with Aspirin or Ticagrelor and Patient Outcomes (SOCRATES).[6-8] These scores served as the reference mRS scores. We searched our EMR (Epic Systems Corporation, Verona, WI) to find a corresponding patient visit note within twenty days of the in-person mRS assessment. Patient notes were then given to three reviewers (an undergraduate research assistant, a second-year medical student, and a neurology resident) to determine a mRS. The reviewers were selected as a convenience sample. Reviewers were trained to derive the mRS by watching a web-based 16-minute video and completing the certification test at (https://webdcu.musc.edu/campus/). The performance of the abstractors was monitored through the first three patient notes by the primary investigator (DI). The abstractors could ask questions during that time. Neither the medical student nor the undergraduate research assistant had experience assessing stroke patients prior to this study.

Inclusion criteria included stroke patients with an in-person evaluation of 90-day mRS, a mRS between 0-5, and an electronic patient note within twenty days of the in-person mRS determination. The patient note could be from a neurology clinic, neurosurgery clinic, or other medical provider (e.g. physical therapist). However, all notes were written by at least a provider with doctoral level training. Exclusion criteria included a mRS of 6 (which indicates patient death) or the inability to locate a patient visit in the EMR within 20 days of the in-person mRS.

Once we identified an appropriate note in the EMR, we redacted all protected health information. We also redacted any mention of mRS or National Institutes of Health Stroke Score (NIHSS). We then gave each note to the three reviewers with instructions to make their best estimation of the patient's mRS. Reviewers were blinded to each other's evaluations. The only measured value was the modified Rankin Score which was determined by our reviewers. In a separate predetermined analysis, mRS were dichotomized to scores of 0 to 2 as a good functional outcome and 3 to 5 as a poor functional outcome.[9]

We tested the strength of agreement of the various raters against the in-person mRS with the Kappa statistic. The weighted kappa (kw) accounts for the extend of

Table 1. Demographics

Average Age (SD)	59.8 (+/- 10.03)	
Sex		
Female	36	(60%)
Male	24	(40%)
In-Person Evaluation mRS		
0	27	(45.0%)
1	14	(23.3%)
2	5	(8.33%)
3	9	(15.0%)
4	5	(8.3%)
5	0	(0%)
Dichotomous Comparison		
Good Outcome (mRS 0-2)	46	(76.7%)
Poor Outcome (mRS 3-5)	14	(23.3%)
Visit Type		
Neurology		
52 (86.7%)		
Hematology	1	(1.7%)
Internal Medicine	3	(5.0%)
Neurosurgery	3	(5.0%)
Physical Therapy	1	(1.7%)
Days after stroke, mean (SD)	89	(11)
Study		
POINT	38	(63.3%)
SOCRATES	13	(21.7%)
SHINE	9	(15%)

https://scholarcommons.towerhealth.org/t-med/vol1/iss2/1
DOI: https://doi.org/10.54299/tmed/jidk4591

Table 2. Comparison of Derived mRS versus In-Person mRS (n=60)

		Actual mRS				
Derived mRS		**0**	**1**	**2**	**3**	**4**
0	Neurology resident	22	1	0	0	0
	Medical Student	22	2	0	0	0
	Research Associate	14	3	2	0	0
1	Neurology resident	3	12	5	0	0
	Medical Student	3	7	2	0	0
	Research Associate	10	9	3	2	1
2	Neurology resident	0	1	0	3	1
	Medical Student	1	3	2	1	0
	Research Assistant	1	1	0	2	2
3	Neurology resident	2	0	0	3	0
	Medical Student	1	2	1	4	1
	Research Assistant	1	0	0	4	0
4	Neurology resident	0	0	0	3	4
	Medical Student	0	0	0	4	4
	Research Assistant	1	1	0	1	2

		In-Person mRS (Dichotomized)	
Derived mRS		**0-2**	**3-6**
0-2	Neurology resident	44	4
	Medical Student	42	1
	Research Associate	43	7
		0-2	**3-6**
3-6	Neurology resident	2	10
	Medical Student	4	13
	Research Associate	3	7

	Agreements (k)	
Reviewer	Chart Derived mRS vs. In-person Assessment kw (95% CI)	Chart Derived mRS vs. In-person Assessment (dichotomized) k (95% CI)
Resident	0.72 (0.60, 0.84)	0.71 (0.49, 0.93)
Medical Student	0.72 (0.60, 0.83)	0.78 (0.60, 0.96)
Research Assistant	0.43 (0.25, 0.60)	0.48 (0.21, 0.75)

disagreements on ordinal scales; that is, by how much the observations disagreed. Kw was used for the primary analysis. The simple kappa (k) is a measure of agreement of paired comparisons (i.e., two raters rating the same item) for categorical ordinal measures. We utilized the k statistic to analyze dichotomized mRS.

A priori, we defined a kappa of >0.80 excellent agreement, 0.60-0.80 good agreement, 0.40-0.60 moderate agreement, and <0.40 as poor agreement.

This study protocol was reviewed and approved by the Temple University Institutional Review Board, approval number 25662.

RESULTS

A total of 98 patient records were reviewed of which 60 met inclusion criteria. 45% of patients had a mRS of zero and for 87% of patients, the visit note was from the neurology clinic. The demographics of the sample are shown in Table 1.

Retrospective Modified Rankin Scale Score versus In-person Evaluation

In Table 2 we compare the mRS derived by the resident, medical student, and research assistant against the in-person mRS. The resident correctly derived the mRS in 68% of cases. 17% of the time resident underestimated the mRS and overestimated the mRS 15% of the time. The medical student correctly derived the mRS 65% of the time. The medical student underestimated the mRS 10% of the time and overestimated the mRS 25% of the time. The research assistant correctly derived the mRS 48% of the time. The research assistant underestimated the mRS 25% of the time and overestimated the mRS 27% of the time.

Comparing against the in-person mRS: the resident had a kw of 0.72 (95% CI: 0.60, 0.84), the medical student had a kw of 0.72 (95% CI: 0.60, 0.83), and the research assistant had a kw of 0.43 (95% CI: 0.25, 0.60).

Comparing the medical student to the resident, the medical Student and resident derived the same mRS 68% of the time. The medical student derived a lower mRS compared with the neurology resident 7% of the time and a higher mRS 25% of the time. (kw=0.74) (95% CI: 0.618, 0.857). Comparing the research assistant to the resident, the research assistant and resident derived the same mRS 57% of the time. The research assistant derived a lower mRS 23% of the time, a higher mRS 20% of the time. (kw=0.59 (95% CI: 0.44, 0.74).

Modified Rankin Scale Score Comparisons: Dichotomized Good vs Poor Outcomes

In Table 2, we compare the mRS derived by the resident, medical student, and research assistant dichotomized into good functional outcome (mRS 0-2) vs poor functional outcome (mRS 3-5) against the in-person mRS. Comparing against the in-person mRS: the resident had a k of 0.71 (95% CI: 0.49, 0.93), the medical student had a k of 0.78 (95% CI: 0.60, 0.96), and the research assistant had a k of 0.48 (95% CI: 0.21, 0.75).

Power analyses were conducted for the k analyses of the aggregated good functional outcome vs poor functional outcome assessments by the resident, medical student, and research assistant. As executed, the study sample size of 60 subjects achieves 85%, 89%, and 83% power, respectively, to detect a true k of 0.70 in a test of H 0: Kappa = κ 0 vs. H 1: Kappa ≠ κ 0 (where κ 0 = 0.30) when there are two categories with the respective observed frequencies of occurrence. These power calculations are based on a significance level (α) of 0.05.

DISCUSSION

In comparing our retrospective modified Rankin Scores to the in-person derived scores, we found a good level of agreement with the resident and medical student compared to the in-person mRS, but only moderate agreement between the research assistant and the in person mRS. As there are many different levels of training on a research team (e.g. research coordinator, research nurse, and research physician), we sought to find the minimum training needed to extract a mRS from the EMR. The moderate agreement of the research assistant (k=0.43) illustrates that just utilizing the online tutorial to determine mRS is insufficient and at least some medical education is necessary for reliable retrospective determination of the mRS.

After aggregating the mRS into good functional outcome (mRS 0-2) and poor functional outcome (mRS 3-5), we found a slightly improved level of agreement compared to exact mRS. As many stroke studies use good versus poor outcomes as dichotomous endpoints, it would be reasonable to extract these endpoints from the chart.

There have been studies of other stroke scales such as the NIH Stroke Scale, Canadian Neurological Scale and Scandinavian Stroke Scale that confirmed they can be reliably derived from patient records.[10-12] These scales rely more on physical exam findings such as facial symmetry and limb motor function which makes determining them a more objective process, especially when reading through patient notes. The mRS, on the other hand, is a more subjective assessment that asks whether patients can "carry out previous activities" or "walk and attend to bodily needs without assistance."

By eliminating patients with an mRS=6 (death), we likely skewed our results in the negative direction. The agreement of whether a patient was dead is likely to be high, thus increasing the k. However, we felt that this would artificially inflate the k statistic. We do not feel that including mRS=0 changed the kappa as there could be subjectivity between an mRS of 0 and 1.

There are several limitations to our study. As 87% of the notes were from the neurology clinic, this study is not generalizable to non-neurology patient notes. Patient notes were not recorded on the exact day that the in-person mRS was obtained. It is possible that the two scores varied because of the timing of the note, though we do not believe that mRS would vary in such a short time. Additional study limitations include the small sample size, both in terms of the number of raters, the level of training for the raters, and the single-center design. Future studies, with a broader range of note types and mRS, would be valuable in affirming the findings of this study.

We conclude that, it is feasible to derive either the exact mRS or a dichotomous outcome from the EMR. However, the accuracy was better in the dichotomized group suggesting this may be the preferred approach rather than estimating an exact mRS. Researchers may use the EMR to estimate a mRS after stroke when in person assessment is not available.

REFERENCES

1. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke. May 1988;19(5):604-7. https://doi.org/10.1161/01.STR.19.5.604
2. Savio K, Pietra GL, Oddone E, Reggiani M, Leone MA. Reliability of the modified Rankin Scale applied by telephone. Neurol Int. Feb 11 2013;5(1):e2. https://doi.org/10.4081/ni.2013.e2
3. Merino JG, Lattimore SU, Warach S. Telephone assessment of stroke outcome is reliable. Stroke. Feb 2005;36(2):232-3. https://doi.org/10.1161/01.STR.0000153055.43138.2f
4. Baggio JA, Santos-Pontelli TE, Cougo-Pinto PT, et al. Validation of a structured interview for telephone assessment of the modified Rankin Scale in Brazilian stroke patients. Cerebrovasc Dis. 2014;38(4):297-301. https://doi.org/10.1159/000367646
5. Quinn TJ, Ray G, Atula S, Walters MR, Dawson J, Lees KR. Deriving modifiedRankinscoresfrommedicalcase-records.Stroke.Dec2008;39(12):3421-3. https://doi.org/10.1161/STROKEAHA.108.519306
6. Johnston SC, Easton JD, Farrant M, et al. Clopidogrel and Aspirin in Acute Ischemic Stroke and High-Risk TIA. N Engl J Med. Jul 19 2018;379(3):215-225. https://doi.org/10.1056/NEJMoa1800410
7. Johnston SC, Amarenco P, Albers GW, et al. Ticagrelor versus Aspirin in Acute Stroke or Transient Ischemic Attack. N Engl J Med. Jul 2016;375(1):35-43. https://doi.org/10.1056/NEJMoa1603060
8. Johnston KC, Bruno A, Pauls Q, et al. Intensive vs Standard Treatment of Hyperglycemia and Functional Outcome in Patients With Acute Ischemic Stroke: The SHINE Randomized Clinical Trial. JAMA. Jul 23 2019;322(4):326-335. https://doi.org/10.1001/jama.2019.9346
9. Weisscher N, Vermeulen M, Roos YB, de Haan RJ. What should be defined as good outcome in stroke trials; a modified Rankin score of 0-1 or 0-2? J Neurol. Jun 2008;255(6):867-74. https://doi.org/10.1007/s00415-008-0796-8
10. Barber M, Fail M, Shields M, Stott DJ, Langhorne P. Validity and reliability of estimating the scandinavian stroke scale score from medical records. Cerebrovasc Dis. 2004;17(2-3):224-7. https://doi.org/10.1159/000075795
11. Kasner SE, Chalela JA, Luciano JM, et al. Reliability and validity of estimating the NIH stroke scale score from medical records. Stroke. Aug 1999;30(8):1534-7. https://doi.org/10.1161/01.STR.30.8.1534
12. Stavem K, Lossius M, Ronning OM. Reliability and validity of the Canadian Neurological Scale in retrospective assessment of initial stroke severity. Cerebrovasc Dis. 2003;16(3):286-91. https://doi.org/10.1159/000071129