Predicting specific abilities after disabling stroke

International Journal of Stroke WSO

Predicting specific abilities after disabling stroke: Development and validation of prognostic models

International Journal of Stroke
0(0) 1–9

DOI: 10.1177/1747493020982873
journals.sagepub.com/home/wso

Akila Visvanathan[1], Catriona Graham[2], Martin Dennis[1], Julia Lawton[3], Fergus Doubal[1], Gillian Mead[1] and William Whiteley[1]

Abstract

Background: Predicting specific abilities (e.g. walk and talk) to provide a functional profile six months after disabling stroke could help patients/families prepare for the consequences of stroke and facilitate involvement in treatment decision-making.

Aim: To develop new statistical models to predict specific abilities six months after stroke and test their performance in an independent cohort of patients with disabling stroke.

Methods: We developed models to predict six specific abilities (to be independent, walk, talk, eat normally, live without major anxiety/depression, and to live at home) using data from seven large multicenter stroke trials with multivariable logistic regression. We included 13,117 participants recruited within three days of hospital admission. We assessed model discrimination and derived optimal cut-off values using four statistical methods. We validated the models in an independent single-center cohort of patients ($n = 403$) with disabling stroke. We assessed model discrimination and calibration and reported the performance of our models at the statistically derived cut-off values.

Results: All six models had good discrimination in external validation (AUC 0.78–0.84). Four models (predicting to walk, eat normally, live without major anxiety/depression, live at home) calibrated well. Models had sensitivities between 45.0 and 97.9% and specificities between 21.6 and 96.5%.

Conclusions: We have developed statistical models to predict specific abilities and demonstrated that these models perform reasonably well in an independent cohort of disabling stroke patients. To aid decision-making regarding treatments, further evaluation of our models is required.

Keywords

Prognostic models, disability, stroke, prediction, prognosis, decision-making

Received: 30 June 2020; accepted: 11 November 2020

Introduction

Clinicians may estimate future recovery of patients who have had a disabling stroke based on their clinical experience. They may also incorporate information from prognostic models based on patient outcomes probability of good (survival with independence) and poor (death or dependency) outcome based on functional scales such as the modified Rankin scale (mRs).[6] However, two people who are in the same mRs category after stroke may vary with respect to

their "specific abilities" (e.g. to walk and to talk) and reported quality of life.[7] While some models have predicted recovery of one function, such as walking and arm function,[8] none have predicted multiple abilities to provide a functional profile for the patient and many have not been externally validated.[3]

Second, to influence decision-making regarding treatments, predictions need to be made early especially since treatments after stroke may have different consequences. For example, intermittent pneumatic compression[9] and enteral tube feeding[10] increase the probability of survival but do not improve functional recovery, therefore increasing the probability of survival with severe disability. However, the trajectory of acute stroke is difficult to predict, and therefore, decisions are made at a time of uncertainty.[4]

To facilitate shared decision-making about treatments in acute disabling stroke, patients and families need to understand possible outcomes, the uncertainty of prognosis, and possible effects of treatments. Therefore, we propose communicating prognosis with respect to "specific abilities" to provide a functional profile and in terms which are quantifiable in the early period after disabling stroke; to do so, we need adequately validated statistical models which provide predictions of "specific abilities."

Aims

a. Develop new models building on the existing six simple variable (SSV) models to predict the probability of a patient with stroke having "specific abilities" (being independent, able to walk, to talk, to eat normally, to live without major anxiety or depression, and to live at home) by six months.
b. Test the performance of these models in an independent cohort of patients with disabling stroke.

Methods

We report our methodology and results based on the TRIPOD (transparent reporting of a multivariable prognostic model for Individual Prognosis or Diagnosis) statement.[11]

therefore easily available to us. The "specific abilities" (to be independent, walk, talk, eat normally, live without major anxiety or depression, and live at home) could be derived from the outcomes these trials collected at six months.

The FOOD trial evaluated feeding policies in patients admitted to hospital with a recent stroke.[10] The three trials enrolled 5033 patients (November 1996–July 2003).

CLOTS tested external compression devices for prevention of deep venous thrombosis in acute stroke patients.[9] The three trials enrolled 8228 patients (March 2001–September 2012).

The IST3 assessed the benefits and harms of intravenous thrombolysis within six hours of acute ischemic stroke.[12] The trial enrolled 3035 patients (May 2000–July 2011).

Since we aimed to make predictions which might influence early decisions about treatment, we included only the 13,117 (79.1%) participants from these trials recruited by day 3 of hospital admission who had complete baseline data (Supplementary Table 1).

Validation cohort

We recruited an independent cohort of adults (>18 years) after disabling stroke from the UK teaching hospital (10 May 2017–25 May 2018) and followed them up for about six months. Due to challenges recruiting patients within three days after major stroke (patients were medically unwell, families needed more time to consider participation), the recruitment window was extended to 10 days. Our sample size was calculated based on a locally collected dataset to achieve an event per variable rate of at least 10.[13] We recruited patients who we defined as having had a disabling stroke; i.e. (a) mRs 3 or above at baseline or (b) mRs 0–2 but with two or more abilities (to walk, talk, and eat normally) affected by the stroke because our aim was to identify models which may be used in the prognostication of such patients. Patients or proxies (where the patient lacked capacity) provided informed consent.[7]

Definition of outcomes

stroke) may confound results, we chose to predict outcomes at six months. We had data on four of these "specific abilities" (to be independent, to walk, to live without major anxiety or depression, and to live at home) from all trials. Ability to talk was only available from IST-3 and ability to eat normally from the FOOD trials.

Selection of predictor variables

We used the SSVs (age, independent before stroke, living alone before stroke, being able to lift arms after stroke, being able to walk after stroke, and normal verbal score of the Glasgow Coma Scale after stroke) based on recommendations to build on existing validated models.[3,6] These models which predict survival and functional independence at six months have been validated for use in the acute setting[14] and perform as well as models including more variables.[6,15] The variables are clinically relevant and have good inter-rater reliability.[6] The SSVs were also common to both our development and validation datasets.

We had initially explored if adding some variables (sex, overweight, diabetes) to SSVs improved discrimination of our models. For two outcomes picked at random (to be independent and to talk), we also tested if model discrimination was improved if we assumed that a specific ability at baseline would be retained at six months. These did not improve the models and were therefore not included in our final models (Supplementary file 2).

Statistical methods

We developed the models using multivariable logistic regression.

We assessed discrimination (the ability of the model to separate individuals who develop the outcome of interest from those who do not) by calculating the area under the receiver operating characteristic (ROC) curve (AUC) of sensitivity versus 1 minus specificity. An area of 1 implies a test with perfect discrimination, whilst an area of 0.5 implies that the model's predictions are no better than chance.[16] We reported 95% confidence intervals of the AUCs. We determined optimal cut-off values in our development cohort from the ROC plots. We used four statistical methods to deter-

In our validation cohort, we assessed discrimination and calibration. The latter is an assessment of whether predicted probabilities of specific abilities of patients were higher or lower than those actually observed.[16] We plotted calibration curves based on tenths of patients. A model is well-calibrated if the predicted and observed probabilities are similar. We reported model performance (sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV)) at the statistically derived optimal cut-off values. Where the values were different using different methods, we reported all the solutions. We also calculated the Matthews correlation coefficient (MCC) which provides complementary information and may be less dependent on class imbalance:

$$\begin{aligned} \mathrm{MCC} = {} & (\text{True positive} \times \text{True negative}) \\ & - (\text{False positive} \times \text{False negative}) \\ & \sqrt{(\text{True positive} + \text{False Positive})} \\ & (\text{True positive} + \text{False negative}) \\ & (\text{True negative} + \text{False positive}) \\ & (\text{True negative} + \text{False negative}) \end{aligned}$$

An MCC value can be between −1 and +1; +1 describes a perfect prediction, 0 is no better than chance, and −1 describes inconsistency between prediction and observation.

We used SAS 9.4 (SAS institute, 2013) to develop our models and Stata 15 (StataCorp, College Station, TX, USA) for external validation.

Ethics

Ethical approval was not required to use anonymized data from the trials for the development cohort. The recruitment of our validation cohort was approved by the Scotland A Research Ethics Committee (Ref: 17/SS/0029).

Results

The characteristics and specific abilities of the patients in our development ($n = 13{,}117$) and validation ($n = 403$) cohorts are shown in Table 1. The mean dif-

Table 1. Baseline characteristics and specific abilities at six months

Variables	Development cohort $n = 13,117$ (n (%))	Validation cohort ($n = 403$) (n (%))	Mean difference in proportions (95% CI, p)
SSVs[6]			
Age; mean (standard deviation)	74.7 (12.3)	77.5 (11.8)	−2.8 (−3.97 – (−1.63), $p < 0.0001$
Independent before stroke	12,149 (92.6)	308 (76.4)	16.2 (12.0–20.4), $p < 0.0001$
Living alone before stroke	4456 (34.0)	158 (39.2)	−5.2 (−10.0 – (−0.37)), $p = 0.03$
Lift arms after stroke	5445 (41.5)	152 (37.8)	3.8 (−1.0–8.6), $p = 0.13$
Able to walk after stroke	864 (6.6)	28 (6.9)	−0.3 (−2.8–2.2), $p = 0.81$
Normal verbal score of Glasgow	8269 (63.0)	248 (61.5)	1.5 (−3.3–6.3), $p = 0.54$
Coma scale			
Sex			
Male	6466 (49.3)	179 (44.4)	4.9 (−0.0–9.8), $p = 0.05$
Female	6651 (50.7)	224 (55.6)	
Outcome at six months (scale/measure)			
Disability	Oxford Handicap Scale (OHS)	mRs	$\chi^2 = 128.6$, $p < 0.0001$, $df = 6$
0 (no symptoms)	692 (5.3)	8 (2.0)	
1 (no significant disability)	1323 (10.1)	45 (11.2)	
2 (slight disability)	1792 (13.7)	7 (1.7)	
3 (moderate disability)	2483 (18.9)	149 (37.0)	
4 (moderately severe disability)	1394 (10.6)	46 (11.4)	
5 (severe disability)	2090 (15.9)	36 (8.9)	
6 (dead)	3099 (23.6)	111 (27.5)	
Missing	24 (1.9)	1 (0.3)	
"Specific abilities" at six months			
To be independent			
mRs/OHS 0–2	3807 (29.0)	60 (14.9)	14.6 (11.1–18.2), $p < 0.0001$

Table 1. Continued

Variables	Development cohort $n = 13,117$ (n (%))	Validation cohort ($n = 403$) (n (%))	Mean difference in proportions (95% CI, p)
Dead	3099 (23.6)	111 (27.5)	
Missing	460 (3.5)	1 (0.3)	
To talk	$N = 3035$ (IST3 ONLY)		
No major problems (no dysphasia/mild-to-moderate dysphasia)	1332 (43.9)	278 (69.0)	−18.4 (−23.3– (−13.5)), $p < 0.0001$
Major problems (severe dysphasia/mute)	474 (15.6)	13 (3.3)	
Dead	815 (26.9)	111 (27.5)	
Missing	414 (13.6)	1 (0.3)	
To eat normally	$N = 1854$ (FOOD ONLY)		
Normal/oral modified	1409 (76.0)	286 (71.0)	5.1 (0.2–9.9), $p = 0.03$
Tube (side/nose/percutaneous)	51 (2.8)	4 (1.0)	
Dead	384 (20.7)	111 (27.5)	
Missing	10 (0.5)	2 (0.6)	
To live without major anxiety or depression	$N = 13,117$ (ALL)		
None/some (slight/moderate)	8680 (66.2)	252 (62.5)	6.0 (1.2–10.8), $p = 0.01$
Severe/extreme	853 (6.5)	39 (9.6)	
Dead	3099 (23.6)	111 (27.5)	
Missing	485 (3.7)	1 (0.3)	
To live at home			
Own home or relatives home	7777 (59.3)	218 (54.1)	7.0 (2.1–11.9), $p = 0.005$
Hospital/care home/residential	1826 (13.9)	72 (17.9)	
Dead	3099 (23.6)	111 (27.5)	
Unknown/other uncategorized[a]	0 (0.0)	1 (0.3)	
Missing	415 (3.2)	1 (0.3)	

[a]Discharged to prison

Figure 1. External validation: calibration curves for specific abilities that were well calibrated. AUCs are shown within each curve: (a) to walk; (b) to eat normally; (c) to live without major anxiety or depression (d) to live at home; (e) to be independent; and (f) to talk.

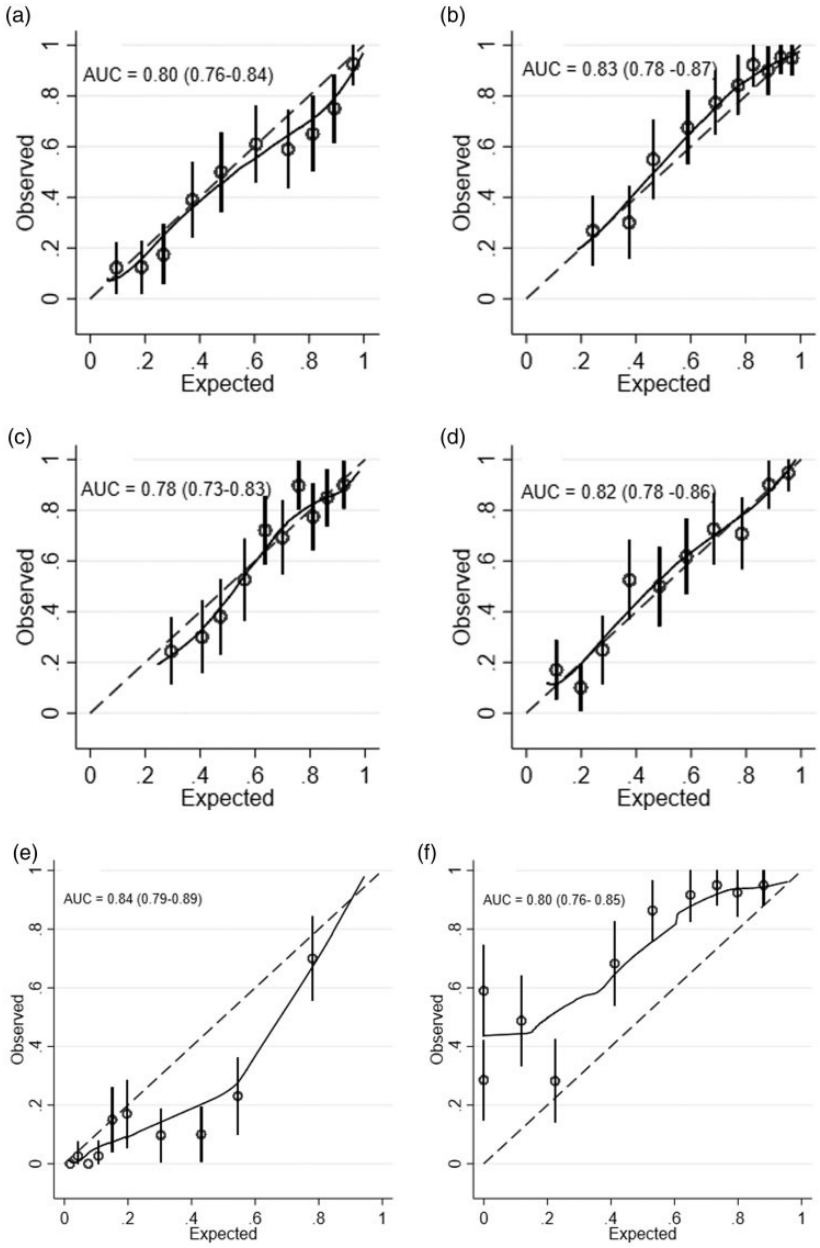

Table 2. Discrimination of models

Model for specific ability at six months	AUC (95% CI)	
	In development cohort	In validation cohort
To be independent	0.79 (0.78–0.80)	0.84 (0.79–0.89)
To walk	0.81 (0.80–0.81)	0.80 (0.76–0.84)
To talk	0.79 (0.77–0.81)	0.80 (0.76–0.85)
To eat normally	0.81 (0.78–0.83)	0.83 (0.79–0.87)
To live without major anxiety or depression	0.72 (0.71–0.73)	0.78 (0.73–0.83)
To live at home	0.80 (0.79–0.81)	0.82 (0.78–0.86)

sensitivities between 45.0 and 97.9%, specificities between 21.6 and 96.5%, and MCC between 0.3 and 0.5 (depending on the "specific ability").

Our work has several strengths. We developed our models using the SSVs that are easy to collect with good inter-rater reliability.[2] Our development cohort included patients with a wide range of characteristics who were prospectively recruited using standardized definitions and methods of data collection with minimal losses to follow-up.[9,10,12] Our models are flexible; despite the difference in recruitment window and stroke severity between our development and validation cohorts, our models performed reasonably well.

However, our work also has limitations. Our development cohort was not designed for the purpose of predicting "specific abilities" after stroke. Therefore, certain baseline variables which might have improved predictions of certain "specific abilities" (e.g. anxiety or depression) were not collected. We were limited to predicting outcomes available in our development dataset at six months and therefore could not predict other specific abilities such as continence and cognition which would help complete the functional profile. Our validation cohort of 403 patients was of modest size, and hence our measures of model performance were relatively imprecise. The difference in prevalence of specific abilities in our cohorts may explain the poorer calibration of some models. For some "specific abilities," different outcome measures were used in the

We did not restrict our analysis to treatment or placebo groups because the effect sizes of the interventions studied by the trials are small compared to the effect of the predictive factors we included. Besides, in clinical practice, predictions are needed in patients who may or may not have had the trial interventions.

We reported the performance of the models in our validation dataset based on optimal cut-off values derived using purely statistical criteria. Whilst these each reflect a certain utility view for decision-making, they do not explicitly reflect the values of individual patients. It might be clinically more relevant to choose cut-offs based on judgments about the relative "cost" or importance to the patient of a false-positive or a false-negative classification. For instance, a false-positive prediction of having a specific ability may provide hope, leading to acceptance of treatments that prolonged survival. The outcome would be that the patient would survive with significant disability which they may judge to be worse than death. In contrast, a false-negative prediction may result in refusal of treatments. Therefore, patients may die before attaining the specific ability. This is challenging, and only individual patients and families are in a position to judge relative "costs" of predictions. An important step in using these models to support decision-making will be to establish peoples' preferences for living with or without specific abilities.

Whilst, for some specific abilities, the models might

Table 3. Optimal cut-off values and performance of models in validation dataset

Specific ability	Method	Cut-off	Sensitivity	Specificity	PPV	Negative predictive value	MCC
To be independent	Correct	0.46	45.0	96.5	69.2	90.9	0.50
	Distance	0.28	55.0	90.1	49.3	91.9	0.43
	Sens-Spec	0.28	55.0	90.1	49.3	91.9	0.43
	Youden	0.32	50.0	91.8	51.7	91.3	0.42
To walk	Correct	0.48	79.9	64.4	67.7	77.5	0.45
	Distance	0.58	72.2	75.0	72.9	74.3	0.47
	Sens-Spec	0.60	67.5	75.5	72.0	71.4	0.43
	Youden	0.58	72.2	75.0	72.9	74.3	0.47
To talk	Correct	0.50	81.7	47.6	77.7	53.6	0.30
	Distance	0.50	81.7	47.6	77.7	53.6	0.30
	Sens-Spec	0.50	81.7	47.6	77.7	53.6	0.30
	Youden	0.50	81.7	47.6	77.7	53.6	0.30
To eat normally	Correct	0.54	88.1	62.1	85.1	67.9	0.52
	Correct[2]	0.54	88.1	62.1	85.1	67.9	0.52
	Distance	0.26	97.9	21.6	75.5	80.7	0.33
	Sens-Spec	0.26	97.9	21.6	75.5	80.7	0.33
	Youden	0.26	97.9	21.6	75.5	80.7	0.33
To live without major anxiety/depression	Correct	0.48	86.5	56.7	77.0	71.4	0.46
	Distance	0.68	75.0	74.0	82.9	63.8	0.48
	Sens-Spec	0.68	74.6	74.0	82.8	63.4	0.47
	Youden	0.66	75.4	70.7	81.2	63.1	0.45
To live at home	Correct	0.53	80.7	70.7	76.5	75.6	0.52
	Distance	0.60	70.6	76.1	77.8	68.6	0.47
	Sens-Spec	0.60	70.6	76.1	77.8	68.6	0.47

understand what their future life might be like or even to help them make choices about treatments in the context of an acute disabling stroke.

Acknowledgements

No other persons have made substantial contributions to this manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: W Whiteley is funded by the Chief Scientist Office. F Doubal is funded by the Stroke Association. A Visvanathan received a fellowship from the Chief Scientist Office (CAF/16/01).

ORCID iDs

Akila Visvanathan https://orcid.org/0000-0002-9713-8713
Gillian Mead https://orcid.org/0000-0001-7494-2023

Supplemental material

Supplemental material for this article is available online.

References

1. Counsell C and Dennis M. Systematic review of prognostic models in patients with acute stroke. *Cerebrovasc Dis* 2001; 12: 159–170.
2. Teale EA, Forster A, Munyombwe T and Young JB. A systematic review of case-mix adjustment models for stroke. *Clin Rehabil* 2012; 26: 771–786.
3. Fahey M, Crayton E, Wolfe C and Douiri A. Clinical prediction models for mortality and functional outcome following ischemic stroke: a systematic review and meta-analysis. *PLoS One* 2018; 13: 1–13.
4. Creutzfeldt CJ, Longstreth WT and Holloway RG. Predicting decline and survival in severe acute brain injury: the fourth trajectory. *BMJ* 2015; 351: 6–8.
5. Visvanathan A, Dennis M, Mead G, Whiteley WN, Lawton J and Doubal FN. Shared decision making after severe stroke – how can we improve patient and family involvement in treatment decisions? *Int J Stroke* 2017; 12: 920–922.
7. Visvanathan A, Whiteley W, Mead G, Lawton J, Doubal FN and Dennis M. Reporting "specific abilities" after major stroke to better describe prognosis. *J Stroke Cerebrovasc Dis* 2020; 29: 1–6.
8. Kwah LK and Herbert RD. Prediction of walking and arm recovery after stroke: a critical review. *Brain Sci* 2016; 6: 53.
9. Dennis M, Sandercock P, Graham C and Forbes J. The Clots in Legs or sTockings after Stroke (CLOTS) 3 trial: a randomised controlled trial to determine whether or not intermittent pneumatic compression reduces the risk of post-stroke deep vein thrombosis and to estimate its cost-effectiveness. *Health Technol Assess* 2015; 19: 1–90.
10. Dennis M, Lewis S, Cranswick G and Forbes J. FOOD: a multicentre randomised trial evaluating feeding policies in patients admitted to hospital with a recent stroke. *Health Technol Assess* 2006; 10: 1–120.
11. TRIPOD Checklist: Prediction Model Development. Julius Center, 2017. p.22, www.tripod-statement.org/TRIPOD/TRIPOD-Checklists/TRIPOD-Checklist-Prediction-Model-Development (accessed 6 May 2020).
12. Sandercock P, Wardlaw JM, Lindley RI, 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* 2012; 379: 2352–2363.
13. Harrell FEJ, Lee KL and Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; 15: 361–387.
14. Reid J, Gubitz G, Dai D, et al. External validation of a six simple variable model of stroke outcome and verification in hyper-acute stroke. *J Neurol Neurosurg Psychiatr* 2007; 78: 1390–1391.
15. Ayis SA, Coker B, Rudd AG, Dennis MS and Wolfe CDA. Predicting independent survival after stroke: a European study for the development and validation of standardised stroke scales and prediction models of outcome. *J Neurol Neurosurg Psychiatry* 2013; 84: 288–296.
16. Steyerberg EW, Moons KGM, et al. Prognosis research strategy (PROGRESS) 3: prognostic model research. *Plos Med* 2013; 10: e1001381.
17. Morrison V, Pollard B, Johnston M and MacWalter R. Anxiety and depression 3 years following stroke: demographic, clinical, and psychological predictors. *J Psychosom Res* 2005; 59: 209–213.