Archives of Rehabilitation Research and Clinical Translation (2023) 000, 100315

ACRM
AMERICAN CONGRESS OF REHABILITATION MEDICINE

Archives of Rehabilitation Research and Clinical Translation

Archives of Rehabilitation Research and Clinical Translation 2023;000:100315

Available online at www.sciencedirect.com

Original Research

External Validation of a Dynamic Prediction Model for Upper Limb Function After Stroke

Iris C. Brunner, PhD [a,b], Eleni-Rosalina Andrinopoulou, PhD [c,d], Ruud Selles, PhD [e,f], Camilla Biering Lundquist, PhD [a], Asger Roer Pedersen, PhD [g]

[a] *Hammel Neurorehabilitation Centre and University Research Clinic, Hammel, Denmark*
[b] *Department of Clinical Medicine, Aarhus University, Hammel, Denmark*
[c] *Department of Biostatistics, University Medical Center Rotterdam, Netherlands*
[d] *Department of Epidemiology of the Erasmus MC, University Medical Center Rotterdam, Netherlands*
[e] *Department of Rehabilitation Medicine, Erasmus MC, University Medical Center Rotterdam, Netherlands*
[f] *Department of Plastic and Reconstructive Surgery of the Erasmus MC, University Medical Center Rotterdam, Netherlands*
[g] *University Research Clinic for Innovative Patient Pathways, Diagnostic Centre, Silkeborg Regional Hospital, Denmark*

KEYWORDS
Stroke;
Upper limb;
Neurological rehabilitation;
Algorithms

Abstract *Objective:* To externally validate the dynamic prediction model for prediction of upper limb (UL) function 6 months after stroke. The dynamic prediction model has been developed and cross-validated on data from 4 Dutch studies.
Design: Data from a prospective Danish cohort study were used to assess prediction accuracy.
Setting: A Danish neurorehabilitation hospital.
Participants: In this external validation study, follow-up data for 80 patients in the subacute phase after stroke (N=80), mean age 64 (SD11), 43% women, could be obtained. They were assessed at 2 weeks, 3 months, and 6 months after stroke with the Action Research Arm Test (ARAT), Fugl-Meyer Motor Assessment upper limb (FMA), and Shoulder Abduction (SA) Finger Extension (FE), (SAFE) test.
Intervention: Not applicable.
Main Outcome Measures: Prediction accuracy at 6 months was examined for 3 categories of ARAT (0-57 points): mild (48-57), moderate (23-47), and severe (0-22). Two individual predictions of ARAT scores at ±6 months post-stroke were computed based on, respectively, baseline (2 weeks) and 3 months ARAT, FE, SA values. The absolute individual differences between observed and predicted ARAT scores were summarized.

List of abbreviations: ARAT, Action Research Arm Test; FE, Finger extension; FMA, Fugl-Meyer Motor Assessment upper limb; SA, shoulder abduction; UL, upper limb.
Disclosures: There are no conflicts of interest.
Cite this article as: Arch Rehabil Res Clin Transl. 2023;000:100315

https://doi.org/10.1016/j.arrct.2023.100315

Results: The prediction model performed best for patients with relatively good UL motor function, with an absolute error median (IQR) of 3 (2-9), and worst for patients with severe UL impairment, with a median (IQR) of 30 (3-39) at baseline. In general, prediction accuracy substantially improved when data obtained 3 months after stroke was included compared with baseline at 2 weeks after stroke.

Conclusion: We found limited clinical usability due to the lack of prediction accuracy 2 weeks after stroke and for patients with severe UL impairments. The dynamic prediction model could probably be refined with data from biomarkers.

According to the World Stroke Organization, stroke has reached endemic proportions, as 1 in 4 adults over 25 will suffer from a stroke in their lifetime.[1] Because of aging populations in most Western and many developing countries, stroke incidence and consequently the need for effective rehabilitation will increase.[2] Personalized rehabilitation, based on the individual's recovery potential, life situation, and preferences, could help to find appropriate treatment strategies and thereby effectively use limited resources. Ideally, a prediction model should be able to incorporate relevant candidate predictors whenever they are available and for as long as the prediction is relevant. Predictor variables should be limited to the most necessary to facilitate application in clinical practice. Therapists can use prediction models to plan and implement targeted interventions. The prediction model should be stable enough to incorporate relevant predictor variables based on the patient status at the time of assessment, even if the assessment does not conform to set time points, such as day 3 or 3 months or 6 months post-stroke as used in formal model validation studies.

Several models have been suggested to predict upper limb (UL) function after stroke, but they suffer from substantial shortcomings. The proportional recovery rule, first published by Prabhakaran et al in 2008, claims that people with stroke will recover 70% of their maximal possible improvement based on their initial Fugl-Meyer Motor Assessment Upper Limb (FMA) score.[3] Yet, this rule is weakened by a substantial part of "nonfitters" with severe paresis and appears to only apply to patients with mild to moderate paresis at a group level. Moreover, its fundamental assumptions have been challenged and debated.[4-6]

A simple bedside test based on shoulder abduction and finger extension (FE) 72 h, 5, and 9 days after stroke is the Early Prediction of Functional Outcome after Stroke (EPOS model),[7] which was recently externally validated.[8] While the EPOS model performed well in discriminating patients with no and some dexterity 3 and 6 months after stroke, it is still too coarse for clinical application because the outcome categories are too broad to be meaningful at an individual level in the clinical setting.

On an individual level, the Predict Recovery Potential (PREP) algorithm appears to be the only suitable model.[9,10] It comprises clinical assessments, and for patients with severe paresis or paralysis, an examination with transcranial magnetic stimulation (TMS) at predefined times after stroke. These predefined times and the need for specific equipment for a subgroup of patients are limiting factors of its application in clinical practice. When applied at a later point in time, the prediction accuracy of PREP expressed as Correct Classification Rate in 4 categories decreased from 75% to 60%, although the prediction accuracy for poor motor function remained high at 78%.[11]

Most prediction models are based on either linear or logistic regression, or Classification and Regression Tree Analysis, predicting an outcome at a defined endpoint, such as at 3 months or 6 months, based on data obtained at a specific baseline, such as day 3 post stroke. Consequently, they do not reflect the nonlinear course of recovery, with most spontaneous biological recovery occurring during the first weeks.[12]

Moreover, they depend on specific days after stroke for predictor measurements with only baseline values as a predictor and do not track interim progress.

Recently, a dynamic model of prediction was developed by Selles et al to address the limitations of other models.[13] This online-based mixed-effects prediction aims to provide patient-specific prediction independent of specific time points for data collection. The development of this model was based on data from 4 Dutch cohort studies, including a total of 450 patients.[7,14]

This model adds the substantial advantage of time-independent measurements, that is, not dependent on specific time points after stroke as suggested in other models.[8,15] In clinical routine, patients may not always be available for assessments on specific days because of other examinations and treatments. An online prediction visualization is available, including 68% and 98% prediction intervals reflecting prediction uncertainty. Moreover, the model has been cross-validated.

While conducting cross-validation with the same dataset is an essential step to evaluate a model's performance, external validation is necessary to assess its generalizability to other datasets. To address this need, we applied the model to a sample from a Danish cohort study[11] to perform an in-depth analysis of its clinical applicability for different levels of UL function after stroke. We hypothesized that the model would perform equally well in the validation cohort.

Methods

Design

The data set for external validation was obtained from a prospective cohort study where the prediction accuracy of

the components of the PREP2 algorithm was assessed when applied 2 weeks after stroke.[11] Based on the same cohort, 3- and 6-month follow-up data were obtained which have been published elsewhere[16,17] and were here used for the purpose of external validation.

Patients

All patients were consecutively recruited at a Danish Neurorehabilitation hospital from June 2018 to October 2019. The hospital has 110 beds for patients with acquired brain injuries. Patients ≥18 years were included if they suffered from impaired UL function (Shoulder Abduction Finger Extension [SAFE] score <10) after a first or recurrent stroke without any residual motor deficits. They had to be admitted to rehabilitation within 2 weeks after stroke and be able to comply with assessment procedures and provide informed consent. The included patients received standard rehabilitation for UL and did not participate in specific UL treatment studies. Ethical approval was obtained from the Ethical Committee of the Central Jutland Region, approval number 628213.

Assessments

Patients were assessed at 2 weeks, henceforth referred to as baseline, 3 months, and 6 months after stroke with the following measurements:

- SAFE score, based on shoulder abduction and FE, scored according to Medical Research Council (MRC) guidelines from 0 to 5 each, allowing for a maximum score of 10 (best).[18]
- Action Research Arm Test (ARAT), assessing UL function on a scale from 0 to 57 (best), divided into subdomains of gross motor function, grip, grasp, and fine motor skills.[19]
- Fugl-Meyer Motor Assessment Upper Limb (FMA), assessing UL impairment on a scale of 0-66 (best).[20]

Not all assessments were available at all time points; some had to be calculated based on other data described below. However, ARAT data were available for most patients at baseline, 3 months, and 6 months post stroke.

All assessments were performed by trained therapists, blinded for baseline scores at 3- and 6-month assessments.

Outcome

The main outcome was prediction accuracy for ARAT scores at 6 months post-stroke. For external validation, the prediction accuracy was compared with the accuracy of the development cohort.

The development cohort of the dynamic prediction model

The original dynamic prediction model[13] was first studied with data from 4 Dutch cohort studies.[7,14,21] Characteristics of the Dutch and the Danish sample are displayed in table 1. The model development and validation were extensively described elsewhere.[13] In short, during the development and evaluation of the model, 5 different model structures with different co-variates were considered. An internal validation using a cross-validation technique was performed. For the model development, a broad array of demographic and clinical co-variates was included, such as age, sex, body-side affected, stroke classification according to the Bamford Scale, thrombolysis, level of impairment according to the National Institute of Health Stroke Scale (NIHSS), and neglect.

The following time-dependent UL motor function variables were included in the dynamic prediction model: days of measurement post stroke, ARAT score, SAFE. Shoulder

Table 1 Patient demographics and medical data

	Original Cohort	Validation Cohort
Number of patients	450	80
Inclusion period	n/a	June 2018 to October 2019
Data sources	Data from 4 different studies in The Netherlands	Data from 1 longitudinal cohort study in Denmark
Time points for assessments post stroke	EPOS study: 3, 5, 6 days and 6 months EXPLORE and, 4D-EEG study: 5, 12, and 26 weeks EXPLICIT study: 1, 2, 3, 5, 12, and 26 weeks	14 days, 3 and 6 months post stroke
Age, years mean ± SD	65 (14)	64 (11)
Sex, women (%)	48	43
Type of stroke: ischemic/ hemorrhagic	450/0	60/20
NIHSS score mean ± SD	8 (5)	9.4 (5.0)
Affected body side (right) (%)	39	36
ARAT baseline score	Early post-stroke (within 1 week)	2 weeks post-stroke
Mean ± SD	14 (19)	22 (19)
FMA mean ± SD	25 (22)	34 (22)
Neglect (present) (%)	166 (36.9)	19 (24.1)

Abbreviations: ARAT, Action Research Arm Test; EPOS, Early Prediction of Functional Outcome after Stroke; EXPLICIT, EXplaining PLastICITy after stroke; FMA, Fugl-Meyer Motor Assessment upper limb; NIHSS, National Institute of Health Stroke Scale.

abduction (SA) was derived from Motricity Index (0-33), which contains 6 categories which are congruent with the SA score from the SAFE score (0= no active movement; 1= active movement palpable, not visible; 2= active movement, but against gravity; 3= active movement, but not against resistance; 4= active movement against resistance, but weaker than other side; 5= normal strength has compared with other side). The FE score was derived from FMA item 25 for FE (0= none, 1= partial, 2= full). The development included 5 different models containing different fixed-effect structures from the most complex to the simplest. This led to a final model based on the SAFE score only.[12] In general, the prediction error decreased as the number of measurements increased. In particular, a median error of 8.4 (Q1−Q3:1.7-28.1) was observed when 1 measurement early poststroke was used and a median error of 2.3 (Q1−Q3:1-7.2) was observed when 7 measurements were used. Furthermore, an increased error was observed when the baseline ARAT was low (between 0 and 22). In the development cohort, the predictive performance measure of the dynamic prediction model was investigated in 3 baseline ARAT categories (0-22, 23-47, 48-57 points). Both the predicted ARAT score at 6 months and the uncertainty were displayed.

The validation cohort

The external validation cohort consisted of the above-described patients from a longitudinal cohort study. The validation data set contained 103 patients at baseline. Complete datasets including both follow-up assessments at 3 and 6 months could be obtained for 80 patients and only those were used to test the model. Model predictions were derived twice (at baseline and at 3 months), leading to 80 predictions at each of the 2 time points.

Because both cohorts did not have all variables in common, some predictor variables of the validation cohort had to be converted to be entered into the model. More specifically, the model by Selles et al requires measurements of ARAT and the SA item of the SAFE score at the same time point. These were not available for all patients in the current cohort, but we derived a proxy from the FMA item 15 (shoulder abduction to 90 degrees with forearm pronated and elbow extended, for the assessments at 3 and 6 months post-stroke. To do so, the FMA scores 0 (no movement) were replaced by SA score 1 (visible contraction without movement of the limb), FMA score 1 (partial movement) was replaced by SA score 2 (movement of the limb- but not against gravity), and the FMA score 2 (full movement) was replaced by SA score 3 (movement across the total movement range).

Statistical analyses

Individual predictions of ARAT scores at 6 months post stroke were computed based on the baseline values obtained around 2 weeks post-stroke of ARAT, FE, and the SA proxy as well as based on these values and the corresponding values obtained around 3 months post-stroke. The absolute individual differences between observed and predicted ARAT scores were summarized in box plots. In this comparison of observed and predicted ARAT values, predictions were truncated at ARAT values 0 (for 7 patients), and 57 (for 18 patients), to avoid predictions outside the ARAT scale, although truncation violates the assumptions behind the 95%-prediction intervals provided by the model. Individual prediction trajectories were investigated to see if they were within the ARAT range and if they were monotonic, that is, following a nonoscillating time course. Further analyses were performed to identify categories of prediction trajectories.

Results

The validation cohort consisted of complete datasets at 3 and 6 months post stroke for 80 patients. Patients in the original cohort were comparable in age and sex but were on average more severely affected and only patients with ischemic stroke were included, table 1.

The model performance illustrated with 6 cases

In figure 1, we present examples of model performance for different levels of impairment. In panels A1 and A2, 2 patients with severe impairment (paralysis) are depicted. The model performs well for A1 but underestimates the recovery of A2 from baseline to 3 months post-stroke. In 2 patients at the lower end of moderate impairment (ARAT 23-47) at baseline, the model predicts reasonably for B1 but again underestimates for B2 from baseline to 3 months. In some cases, as shown in panel C, the 95% prediction uncertainty of the model exceeds the upper limit of 57 on ARAT.

The overall model performance

We compared the prediction error to the error in the development cohort by setting up figure 2 directly comparable with a similar figure in Selles et al.[13] The absolute errors for the 3 ARAT categories of UL impairment, 0-22, 23-47, and 48-57, were similar to the original validation cohort. In figure 2, the model validation for the Danish cohort is depicted according to 3 ARAT categories applied in the publication by Selles et al.[13] The absolute error for ARAT scores 6 months post-stroke is largest at baseline (day 11-18) and for patients with severely impaired UL function (ARAT 0-22) with a median (min; IQR; max) 30 (1; 3 - 39; 56), at 3 months post-stroke 7 (0; 1 - 14; 21). For patients in the midrange (ARAT 23-47) of impairment level the absolute error is 11 (0; 7 - 16; 31) at baseline, diminished to 2 (0; 1 - 4; 11) 3 months after. In patients with relatively good baseline UL function (ARAT 48-57) the absolute error is small both at baseline, 3 (0; 2 - 5; 17), and 3 months, 0 (0; 0 - 3; 7), after stroke. These results are consistent with those obtained in the internal validation.

Individual prediction trajectories showed a large variety of recovery patterns, figure 3. In general, the model did not perform satisfactorily enough to be clinically relevant at baseline after stroke but improved over time. Thus, its clinical applicability is limited.

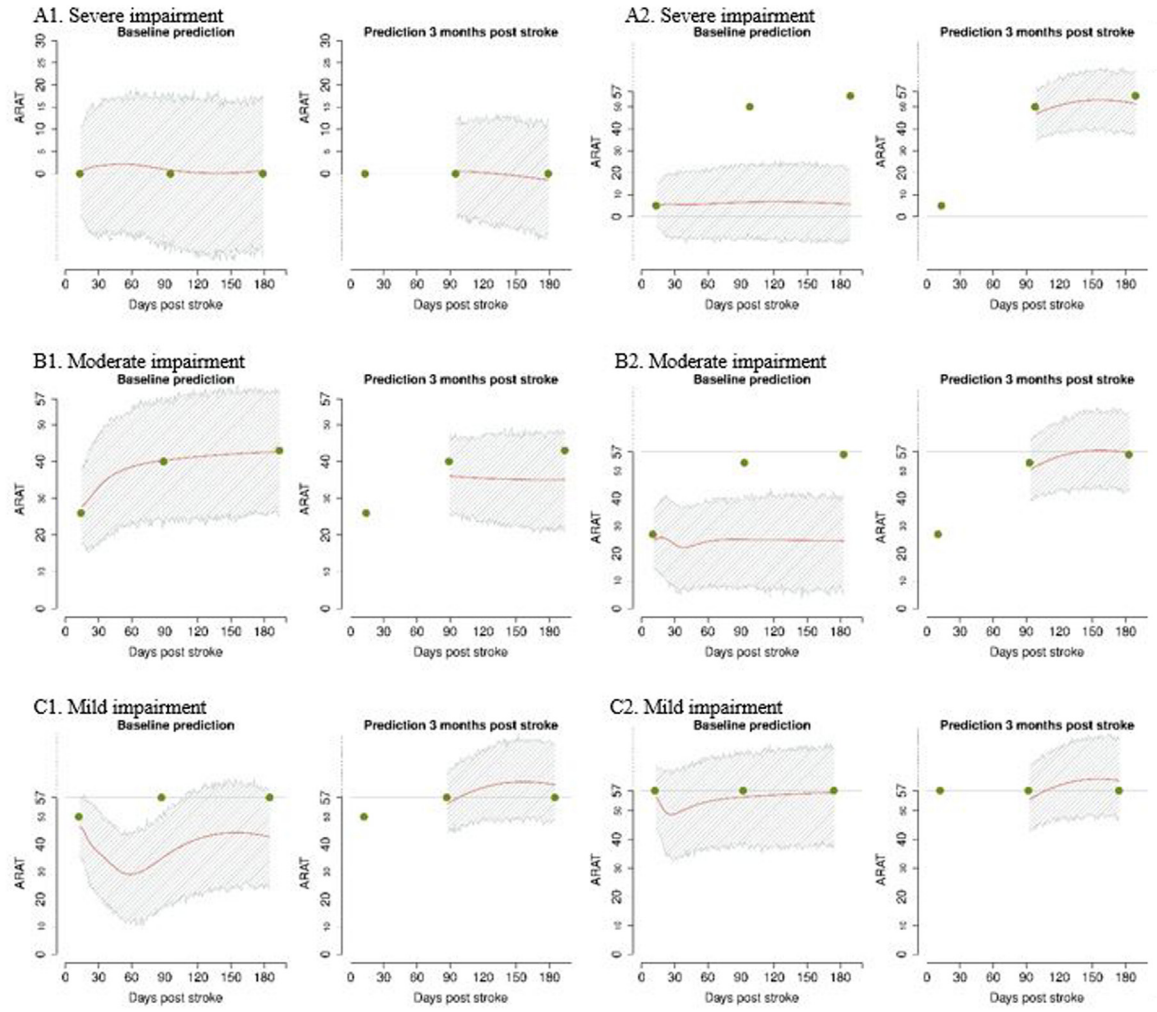

Fig 1 Examples of predictions for different levels of UL impairment. A1: the prediction interval (PI) is below 0; A2: The recovery potential is not detected by the model; B1: The model predicts well; B2: The model does not detect the recovery potential; C1: The model predicts deterioration and improvement; C2: The PI exceeds beyond the maximum score of 57.

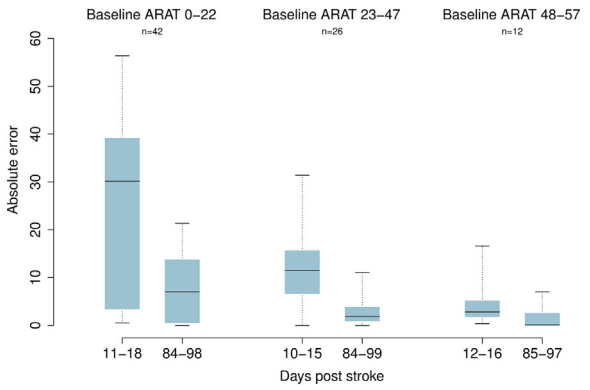

Fig 2 Model validation of the Danish data within the same ARAT categories used by Selles et al[13] (baseline ARAT score 0-22, 23-47, 48-57). Within categories, the absolute difference between observed and predicted ARAT score (absolute error), at approximately 180 days post stroke, is summarized by box and whiskers plots for predictions computed from, respectively, the baseline ARAT score (approximately 14 days post stroke) and both the baseline score and the score approximately 90 days post stroke.

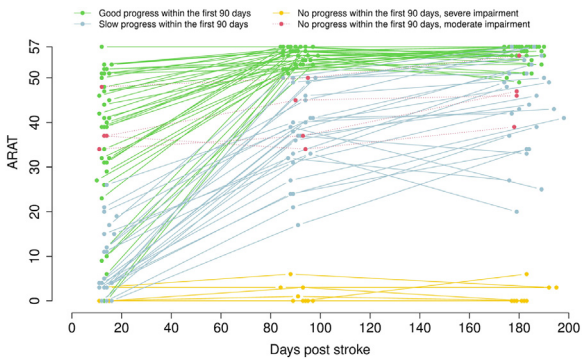

Fig 3 The large variation in actual recovery profiles of all included 80 patients in the validation cohort.

Discussion

In this study, we externally validated the computerized prediction model developed by Selles et al.[13] In accordance with the original cohort, the prediction model performed best for patients with relatively good UL motor function and worst for patients with severe UL impairment, especially at the first prediction, 2 weeks after stroke. In general, prediction accuracy substantially improved when obtained 3 months after stroke as compared with baseline. Consequently, the model performs best when there is less room for improvement, either as a factor of time or an already favorable initial UL function.

The dynamic prediction model developed by Selles et al provides some major improvements over other prediction models. First, it is independent of assessments being obtained at fixed time points after stroke and provides a prediction on an individual level which improves according to the number of measurements. Second, the easily available online tool can be connected to electronic patient records which adds to its clinical usefulness. Comparable with the original study by Selles et al, prediction errors at 6 months after stroke decreased in the present study, as the number of assessments and time points increased. However, in our cohort, the prediction uncertainty was still substantial when it was needed most, at baseline and for patients with severe UL impairment. Thus, while the model is independent of defined time points of measurement after stroke, its accuracy is not. Initial patient status affects accuracy of predictions as does time since stroke, regardless of using a dynamic model with flexible or specific days after stroke for measurements. Nevertheless, more measurements during the early post stroke phase are desirable. Unfortunately, our study did not incorporate additional measurements between the 2-week and 3-month post-stroke periods. Including more measurements during this timeframe would have likely mitigated prediction uncertainty.

It is common that a model performs worse in another cohort than the original development cohort, which could partly be attributed to differences between the cohorts.[22,23] While they were similar in age and sex, they differed in several other respects, such as type of stroke. We included both ischemic and hemorrhagic and recurrent strokes, while the development cohort only comprised ischemic anterior circulation first-ever strokes. Patients with hemorrhagic stroke may show a different recovery profile. In a cohort study from Sweden, it was demonstrated that patients with hemorrhagic stroke had a larger improvement during the first 3 months, though there was no difference at 1-year post-stroke.[24] In our cohort, only 3 assessments were performed contrary to the development cohort with a median of 6 assessments. Taking into account the nonlinear nature of recovery and clinical needs more assessments early, that is, during the first days and weeks after stroke are needed to improve prediction accuracy. Accordingly, the more frequent early measurements in the studies used in the development cohort provide increased prediction accuracy.[13]

At 3 months post-stroke, most patients have reached a plateau and completed the most intensive rehabilitation.[25] Thus, prediction is less meaningful both to the patient and health professionals. Prediction at an earlier time point would provide the opportunity to adapt the treatment plan to the individual recovery potential and focus on areas where improvement can be expected and focus on compensatory strategies where it cannot. Ideally, repeated measurements should be performed weekly during the first 4 weeks as part of clinical routine.

The individual prediction for patients with severe paresis during the first days and weeks after stroke seems to pose the biggest challenge. Clinical measures are useful for a wide range of patients with mild to moderate impairments. In the current model's development, several mixed models were examined containing all potentially relevant medical and demographic features, such as age, body side affected, in addition to time after stroke. Still, the model's best performance was found when ARAT scores were a function of the SAFE score and their interaction with time and model. This suggests that simple clinical assessments such as the SAFE score may be sufficient for a broad range of patients, as earlier suggested by Nijland et al.[7]

Veerbeek et al validated the model based on the SAFE score only. However, they only differentiated between patients who will reach some dexterity, defined $\geq$10 ARAT, and those who will not.[8] The return of FE during the first days after stroke is generally regarded as promising for future UL function. If FE is not present, prediction is difficult. In their PREP2 algorithm, Stinear et al therefore included an examination with TMS for patients with severe impairment.[15] TMS seems to increase prediction accuracy substantially and is recommended for stratification in clinical trials.[9,10,26] However, while TMS is relatively inexpensive and easy to administer, it is not available at all rehabilitation facilities.

Other demographic and clinical factors such as neglect, handedness, site of lesion, and so on have been examined in several studies with some, but not substantial contributions.[11,27-29] Consequently, a focus on initial impairment in combination with biomarkers for the severely impaired seems to be justified.

The online calculation and linkage to medical records is a huge advantage of the model proposed by Selles et al[13] and increases clinical usability.[30] Future prediction models should include these possibilities.

Study limitations

Some limitations of this study should be mentioned. This external validation cohort differed from the original cohort in several respects, such as the number of assessments, the average severity of impairment, and clinical features. Yet, both cohorts comprised patients for whom a prediction model should be applicable.

A limitation of the dynamic model applied is that it exceeds 57 and deceeds 0 which does not make sense in a clinical application because it is outside the scale of the ARAT. Future prediction models should take this into account and find ways of truncating these values without violating the model assumptions. Most challenging appears the prediction of UL function for patients with severe impairment during the first days and weeks after stroke where the use of biomarkers may substantially improve prediction accuracy.[30]

Conclusions

An increased demand for rehabilitation services will meet even more limited resources in the nearer future. There is a need for reliable prediction tools for core functional areas to target rehabilitation. In this study, we found limited clinical usability of an UL prediction model due to the lack of prediction accuracy when obtained 2 weeks after stroke and for patients with severe UL impairments. Prediction accuracy was reasonable for patients with mild impairment and when obtained 3 months after stroke. The dynamic prediction model provides an important step toward individual prediction of UL function and could probably be refined with data from biomarkers.

Corresponding author

Iris C. Brunner, PhD, Hammel Neurorehabilitation Centre and University Research Clinic, Voldbyvej 15, 8450 Hammel, Denmark. *E-mail address:* Iris.Brunner@rm.dk.

References

1. Thayabaranathan T, Kim J, Cadilhac DA, et al. Global stroke statistics 2022. Int J Stroke 2022:17474930221123175.
2. Feigin VL, Norrving B, Mensah GA. Global burden of stroke. Circ Res 2017;120:439–48.
3. Prabhakaran S, Zarahn E, Riley C, et al. Inter-individual variability in the capacity for motor recovery after ischemic stroke. Neurorehabil Neural Repair 2008;22:64–71.
4. Hawe RL, Scott SH, Dukelow SP. Taking proportional out of stroke recovery. Stroke 2019;50:204–11.
5. Kundert R, Goldsmith J, Veerbeek JM, Krakauer JW, Luft AR. What the proportional recovery rule is (and is not): methodological and statistical considerations. Neurorehabil Neural Repair 2019;33:876–87.
6. Goldsmith J, Kitago T, Garcia de la Garza A, et al. Arguments for the biological and predictive relevance of the proportional recovery rule. Elife 2022: 11.
7. Nijland RH, van Wegen EE, Harmeling-van der Wel BC, Kwakkel G. Presence of finger extension and shoulder abduction within 72 hours after stroke predicts functional recovery: early prediction of functional outcome after stroke: the EPOS cohort study. Stroke 2010;41:745–50.
8. Veerbeek JM, Pohl J, Luft AR, Held JPO. External validation and extension of the Early Prediction of Functional Outcome after Stroke (EPOS) prediction model for upper limb outcome 3 months after stroke. PLoS One 2022;17:e0272777.
9. Stinear CM, Smith MC, Byblow WD. Prediction tools for stroke rehabilitation. Stroke 2019;50:3314–22.
10. Stinear CM, Byblow WD, Ackerley SJ, Smith MC, Borges VM, Barber PA. Proportional motor recovery after stroke: implications for trial design. Stroke 2017;48:795–8.
11. Lundquist CB, Nielsen JF, Arguissain FG, Brunner IC. Accuracy of the upper limb prediction algorithm PREP2 applied 2 weeks poststroke: a prospective longitudinal study. Neurorehabil Neural Repair 2021;35:68–78.
12. Krakauer JW. Arm function after stroke: from physiology to recovery. Semin Neurol 2005;25:384–95.
13. Selles RW, Andrinopoulou ER, Nijland RH, et al. Computerised patient-specific prediction of the recovery profile of upper limb capacity within stroke services: the next step. J Neurol Neurosurg Psychiatry 2021;92:574–81.
14. Kwakkel G, Winters C, van Wegen EE, et al. Effects of unilateral upper limb training in two distinct prognostic groups early after stroke: the EXPLICIT-stroke randomized clinical trial. Neurorehabil Neural Repair 2016;30:804–16.
15. Stinear CM, Byblow WD, Ackerley SJ, Barber PA, Smith MC. Predicting recovery potential for individual stroke patients increases rehabilitation efficiency. Stroke 2017;48:1011–9.
16. Lundquist CB, Nguyen BT, Hvidt TB, Stabel HH, Christensen JR, Brunner I. Changes in upper limb capacity and performance in the early and late subacute phase after stroke. J Stroke Cerebrovasc Dis 2022;31:106590.
17. Essers B, Biering Lundquist C, Verheyden G, Brunner IC. Determinants of different aspects of upper-limb activity after stroke. Sensors (Basel) 2022;22:2273.
18. Paternostro-Sluga T, Grim-Stieger M, Posch M, et al. Reliability and validity of the Medical Research Council (MRC) scale and a modified scale for testing muscle strength in patients with radial palsy. J Rehabil Med 2008;40:665–71.
19. Platz T, Pinkowski C, van Wijck F, Kim IH, di Bella P, Johnson G. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multicentre study. Clin Rehabil 2005;19:404–11.
20. Lundquist CB, Maribo T. The Fugl-Meyer assessment of the upper extremity: reliability, responsiveness and validity of the Danish version. Disabil Rehabil 2017;39:934–9.
21. Saes M, Meskers CGM, Daffertshofer A, van Wegen EEH, Kwakkel G. Are early measured resting-state EEG parameters predictive for upper limb motor impairment six months poststroke? Clin Neurophysiol 2021;132:56–62.
22. Siontis GC, Tzoulaki I, Castaldi PJ, Ioannidis JP. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. J Clin Epidemiol 2015;68:25–34.
23. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med 2013;10:e1001381.
24. Persson HC, Opheim A, Lundgren-Nilsson A, Alt Murphy M, Danielsson A, Sunnerhagen KS. Upper extremity recovery after ischaemic and haemorrhagic stroke: part of the SALGOT study. Eur Stroke J 2016;1:310–9.
25. Kitago T, Krakauer JW. Motor learning principles for neurorehabilitation. Handb Clin Neurol 2013;110:93–103.
26. Boyd LA, Hayward KS, Ward NS, et al. Biomarkers of stroke recovery: consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. Int J Stroke 2017;12:480–93.
27. Kwakkel G, Kollen BJ, van der Grond J, Prevo AJ. Probability of regaining dexterity in the flaccid upper limb: impact of severity of paresis and time since onset in acute stroke. Stroke 2003;34:2181–6.
28. Doron N, Rand D. Is unilateral spatial neglect associated with motor recovery of the affected upper extremity post-stroke? A systematic review. Neurorehabil Neural Repair 2019;33:179–87.
29. Coupar F, Pollock A, Rowe P, Weir C, Langhorne P. Predictors of upper limb recovery after stroke: a systematic review and meta-analysis. Clin Rehabil 2012;26:291–313.
30. Hayward KS, Schmidt J, Lohse KR, et al. Are we armed with the right data? Pooled individual data review of biomarkers in people with severe upper limb impairment after stroke. Neuroimage Clin 2017;13:310–9.