KU KONKUK UNIVERSITY

More Accurate Methods for Predicting the Upper Extremity Motor Recovery After Ischemic Stroke: Simulation and Validation

Hyo Eun Kwon, MD.[1], Min Woo Woo, MD.[1], Hyun Haeng Lee, MD., MS.[1]

[1]Department of Rehabilitation Medicine, Konkuk University School of Medicine, Seoul, Republic of Korea

Introduction

To date, the prediction model for upper limb motor recovery following stroke has relied on the raw total score of the Fugl-Meyer Assessment upper extremity score (FMAUE). However, the FMAUE assumes multidimensionality, with scores of each item lacking homogeneous physiological meaning and correlating with each other. As a result, using the total score of FMAUE to construct motor recovery prediction model may result in biased estimates. To extract useful information with physiological meaning from each item of FMAUE while accounting for its multidimensionality and intercorrelated items, appropriate latent variables need to be extracted. Confirmatory factor analysis (CFA) is a standardized tool to assess the construct validity of an assessment tool, which is also recognized by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN). Previous studies have confirmed the construct validity of FMAUE using CFA, but the methods used in these studies have the following limitations. 1) There is no consensus on what constructs FMAUE consists of. 2) Failure to consider the possibility of not being able to identify structure with construct validity when a large number of patients with high FMAUE scores are included. In this study, we aimed to confirm the conditions not suitable for extracting valid latent variables from FMAUE through simulation studies. Furthermore, we seek to examine whether the extracted latent variables by a novel hybrid approach can improve the accuracy of the motor recovery prediction model.

Methods

First, we showed that when using CFA to extract latent variables, the appropriate latent variables are not extracted in the following cases; 1) In cases where there is no information on what combination of constructs make up the FMAUE; 2) In cases where the majority of patients have high Fugl-Meyer Assessment of the Upper Extremity (FMAUE) scores.

The procedure for simulation study is as follows; 1) The assumption for the combination of constructs in FMAUE is volition, coordination, and reflex; 2) Assume the motor recovery pattern over time after stroke separately for each construct and extract the values of the latent variable according to the assumed pattern; 3) The motor recovery pattern is assumed to follow an exponential function, and the parameters are differentiated for each construct. Random error was allowed in the extraction of latent variables for each construct. An example of the latent variables derived through the above process is shown in Fig. 1; 4) The values of the latent variable for each construct are used to derive the value of each item through an arbitrary cut-off value. The order of item difficulty is assumed to be constant; 5) We demonstrated that by using the derived item values in the two aforementioned cases (inaccurate information about the FMAUE construct combination and inclusion of patients with high FMAUE scores), an appropriate latent variable distribution is not obtained.

Considering the aforementioned situations where using CFA to extract latent variables is not appropriate, a novel hybrid approach was employed to extract latent variables while taking these circumstances into account. It is assumed that the FMAUE is a valid assessment tool with established construct validity when the aforementioned situations are excluded. 1) We utilized CFA to extract valid constructs from the entire dataset. The criteria for evaluating construct validity in this process are as follows Criteria for convergent validity: squared multiple correlation ≥0.5, composite reliability ≥0.7, average variance extracted (AVE) ≥0.5, Cronbach's alpha ≥0.7; Criteria for discriminant validity: AVE >r2 (r: correlation coefficient between each construct); The valid structures are extracted by dividing them into two groups according to the FMAUE level. The criteria for FMAUE levels are 60, 58, 56, 54, 52, and 50, and the criteria for evaluating construct validity are as described above; 3) Perform the above process for two construct combinations (case 1, volition, coordination, and reflex; case 2, volition of proximal part, volition of wrist, volition of hand, coordination, and reflex).

We call the above process 'a hybrid approach' because we set the potential combination of constructs of FMAUE through hypothesis-driven approach and simultaneously utilize an exploratory method to find a plausible structure by dividing subgroups according to the level of FMAUE in each case.

The real data used in this study is clinical data from those who were admitted to Konkuk University Hospital with a diagnosis of ischemic stroke from 2005 to 2022. We excluded the patients 1) with ischemic lesion in bilateral hemisphere or history of craniectomy, 2) missing values in initial (at 2 weeks (window period: 3 days)) and follow-up (after 3 months after stroke) motor impairment, 3) no initial motor impairment in the upper extremity, 4) no hand movement, and 5) report of hemorrhagic transformation, progression and recurrence of stroke up to 3 months.

To utilize the extracted latent variable based on the valid structure derived through the above process in motor recovery prediction model, we used the method as follows; 1) We extracted latent variables for each construct from the FMAUE item scores of patients in the subgroup for which valid constructs were derived (patients with a specific FMAUE score at both initial and follow-up assessment); 2) We divided the dataset corresponding to that subgroup into training dataset and test dataset; 3) We fit a linear regression model using the training dataset to determine the relationship between initial latent variables and latent variables at follow-up (explanatory variables: initial latent variable for each construct, age at stroke onset, initial impairment, and sex); 4) Using the test dataset, we derived the predicted value of each item using the linear regression model derived in step 3) and the factor loading estimates extracted through CFA, and check the difference between the sum of the predicted value of each item and the actual FMAUE value, which is root mean squared error (RMSE); 5) We calculated the RMSE of linear regression modeling without using latent variables and compared it with the above values; 6) Repeat steps 2) through 5) with a 5-fold cross validation method to determine the average value of the RMSE difference.

We used the R programming language (version 4.0.5; https://www.r-project.org/) for statistical analysis and visualization.

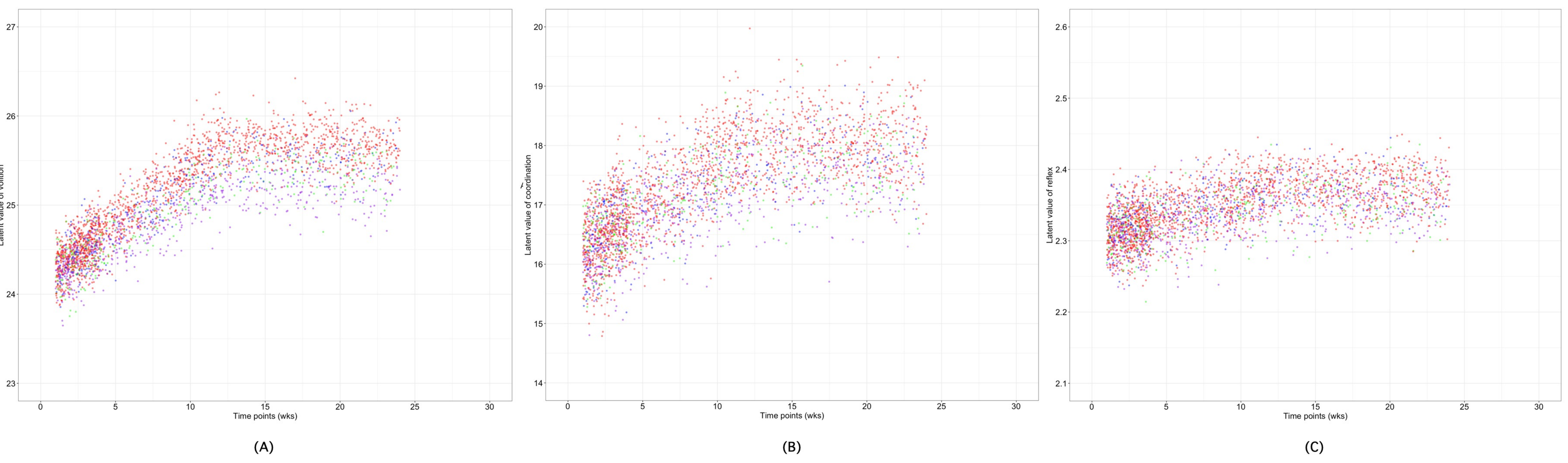

Figure 1. Simulated samples of latent variable for the respective construct; (A) latent variables for volition; (B) latent variables for coordination; (C) latent variables for reflex

Results

In our simulation study, we showed that the original latent variable was not extracted from the simulated data of the FMAUE item when we made incorrect assumptions about the construct and when there were considerable amount of patients with high FMAUE in the sample. As shown in Fig. 2, we obtained real data from 676 patients, with median (IQR) age of 68.5 (17), initial FMAUE of 57 (14) and initial assessment time of 2.2 (0.4) weeks, and follow-up FMAUE of 61 (12) and follow-up assessment time of 17.8 (7.1) weeks. The indicators of convergence and discriminative validity for the extracted structures for each dataset using real data and the model accuracy improved by the latent variable obtained using the structure are presented in Table 1.

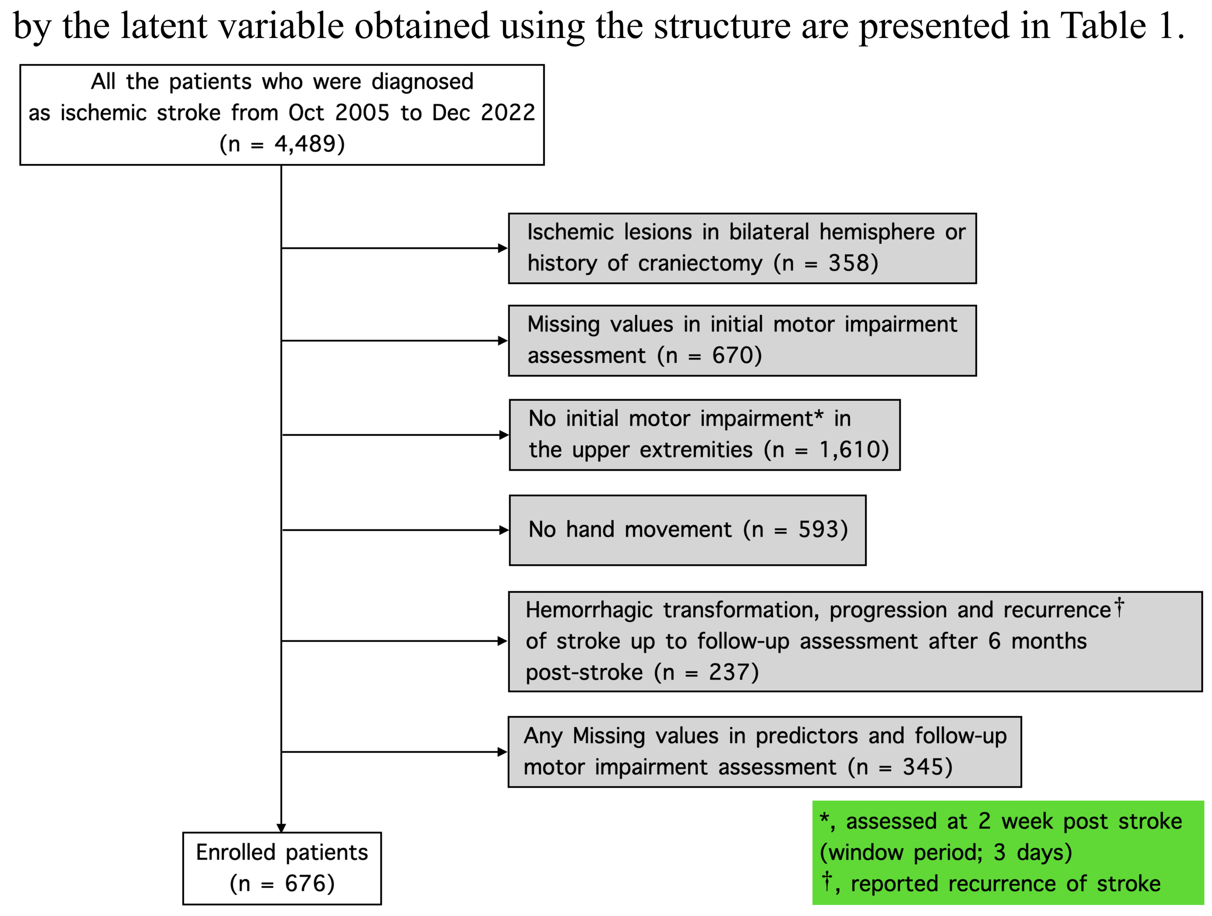

Figure 2. Flow of patients enrollment

Table 1. Convergent validity and discriminant validity of structure obtained from confirmatory factor analysis in the various dataset according to the combination of constructs and degree of Fugl-Meyer Assessment of the upper extremity score

Assumption of construct	Dataset	CFI	TLI	RMSEA	SRMR	Improved accuracy (%)
Volition, coordination, and reflex	Whole	0.6103441	0.5794037	0.2056866	0.089304885	14.4
	>60	1.0000000	1.0030243	0.0000000	0.004646706	20.2
	≤60	0.8205949	0.7813500	0.1960258	0.054631182	17.5
	>58	NA	NA	NA	NA	NA
	≤58	0.8118911	0.7707423	0.1976592	0.057898385	15.5
	>56	1.0000000	1.0048957	0.0000000	0.001937311	18.3
	≤56	0.8534556	0.8140013	0.1754110	0.051075524	16.0
	>54	NA	NA	NA	NA	NA
	≤54	0.8645589	0.8280940	0.1665563	0.048697988	17.3
	>52	0.9687725	0.9375451	0.1069333	0.028136822	25.5
	≤52	0.8552307	0.8162543	0.1682162	0.052803041	23.5
	>50	0.9330700	0.8884499	0.2359731	0.054488399	24.9
	≤50	0.8609409	0.8235019	0.1616207	0.052232702	24.1
Volition of proximal part, volition of wrist, volition of hand, coordination, and reflex	Whole	0.8275533	0.8105699	0.13803773	0.05849907	20.1
	>60	NA	NA	NA	NA	NA
	≤60	0.8269037	0.8069558	0.13553336	0.06202570	18.1
	>58	NA	NA	NA	NA	NA
	≤58	0.8270592	0.8032053	0.14626336	0.06959555	18.2
	>56	NA	NA	NA	NA	NA
	≤56	0.8805291	0.8571362	0.12827057	0.06097986	18.9
	>54	NA	NA	NA	NA	NA
	≤54	0.8851473	0.8626587	0.12278028	0.05493135	19.9
	>52	0.9598288	0.9423631	0.07570517	0.02876952	27.6
	≤52	0.8767151	0.8525753	0.12506780	0.05565878	23.1
	>50	0.9354151	0.9013286	0.13250032	0.03934270	24.1
	≤50	0.8837254	0.8609584	0.12002193	0.05880662	22.2

Conclusion

Through this study, we demonstrated the conditions which is not suitable for extracting latent variables from FMAUE by CFA. We also showed that hybrid approach for extracting latent variables is useful through the improved accuracy of motor recovery prediction model. In future, we will examine the effect of clinical characteristics such as age or ischemic lesion location on the order of factor loadings in CFA model and compare usefulness of this hybrid approach with other data-driven dimension reduction techniques.