Japanese Journal of Comprehensive Rehabilitation Science (2019)

Original Article

Comparison of prediction accuracy of the total score of FIM motor items at discharge in post-stroke patients in a Kaifukuki rehabilitation ward

Yoshitaka Wada, MD,[1] Shigeru Sonoda, MD, PhD,[1] Sayaka Okamoto, MD, PhD,[1] Makoto Watanabe, OTR, MS,[1] Hideto Okazaki, MD, PhD,[1] Yuko Okuyama, RPT[1]

[1]Fujita Health University Nanakuri Memorial Hospital, Tsu, Mie, Japan

ABSTRACT
Wada Y, Sonoda S, Okamoto S, Watanabe M, Okazaki H, Okuyama Y. Comparison of prediction accuracy of the total score of FIM motor items at discharge in post-stroke patients in a Kaifukuki rehabilitation ward. Jpn J Compr Rehabil Sci 2019; 10: 71−76.
Objective: We compared the accuracy of formulas for predicting ADL outcome constructed by multiple regression analysis in post-stroke patients admitted to a Kaifukuki rehabilitation ward.
Methods: We divided 1,502 post-stroke patients into a construction group used to generate prediction formulas, and a validation group used to confirm the prediction accuracy. Prediction formula S was constructed by conventional multiple regression analysis using Functional Independence Measure–motor score (mFIM) at discharge as the dependent variable. Prediction formula R was constructed by reciprocal multiple regression analysis. Prediction equation E was constructed by calculating mFIM at discharge via mFIM effectiveness. In the validation group, predicted mFIM at discharge was calculated, and intraclass correlation coefficient and absolute value of residual were compared.
Results: Intraclass correlation coefficients were 0.86 using prediction formula S, 0.90 using prediction formula R, and 0.89 using prediction formula E. Absolute values of residual were 9.38±6.62 using prediction formula S, 7.30±6.56 using prediction formula R, and 7.56±6.45 using prediction formula E. The Steel-Dwass test detected a significant difference between prediction formulas S and R, and between prediction formulas S and E (both $p<0.05$).
Conclusion: The prediction accuracy of formulas for predicting ADL outcome constructed by multiple regression analysis is improved by adding a transformation that brings the model toward linearity.
Key words: stroke, rehabilitation, multiple regression analysis, outcome prediction, Functional Independence Measure (FIM)

Correspondence: Yoshitaka Wada, MD
Fujita Health University Nanakuri Memorial Hospital,
424−1, Odori-cho, Tsu, Mie 514−1295, Japan.
E-mail: yoshi1201@med.showa-u.ac.jp
Accepted: July 17, 2019.
No conflicts of interest or research funds are provided in this study.

Introduction

Outcome prediction is important for the design and implementation of treatment plans in rehabilitation. Regarding predicting the outcome of post-stroke patients hospitalized in Kaifukuki rehabilitation wards, many reports of prediction formulas constructed by multiple regression analysis using the Barthel Index and Functional Independence Measure (FIM) score [1] have been published [2, 3].

To improve prediction accuracy, in addition to the selection of variables to be included in regression models [3], various methods have been proposed. These methods include prior transformation of the variables used for prediction [4], using predicted FIM effectiveness, which is the ratio of the actual amount of improvement achieved to the maximum amount that can be improved [5, 6], and construction of multiple prediction formulas within the same study population [7−9]. Although previous studies constructed prediction formulas by multiple regression analysis and compared their accuracy, the dependent variables and other conditions of analysis differed among the studies. Therefore it has not been possible to compare the accuracy of those prediction formulas under uniform conditions.

In the present research, under the same conditions, we constructed three prediction formulas according to previous studies and compared their prediction accuracy.

doi.org/10.11336/jjcrs.10.71

Methods

1. Subjects

Patients who had their first supratentorial cerebral hemorrhage or cerebral infarction and were hospitalized in the Kaifukuki rehabilitation ward of our hospital between February 2004 and March 2017 were eligible for the study. Among them, 1,502 patients who were aged 60 years or above, with a duration from stroke onset to admission of 7 to 60 days, no serious comorbidities (comorbidity index [10] 4 or above) that impeded training, and no acute exacerbation during hospitalization were included as subjects. All the subjects underwent the full-time integrated treatment program with training provided 7 days a week [11].

2. Changes from admission to discharge

Scatter plots of the total score of FIM motor items (mFIM) on admission versus mFIM at discharge were generated for all the subjects, and the regression equations were computed. Then, mFIM effectiveness was calculated as follows: mFIM gain was obtained by subtracting mFIM on admission from mFIM at discharge; next, mFIM on admission was subtracted from 91 (maximum mFIM) and the value obtained was used to divide mFIM gain [5]. Scatter plots of mFIM on admission versus mFIM effectiveness were also generated (Figure 2).

3. Three methods of multiple regression analysis

Based on the report by Tokunaga et al. [6], the independent variables used in this study were: age, duration from stroke onset to admission, mFIM on admission, total score of FIM cognitive items (cFIM) on admission, gender (male/female), and type of stroke (hemorrhage/cerebral infarction). Dummy variables were used for gender and type of stroke (male 0/female 1, cerebral hemorrhage 0/cerebral infarction 1).

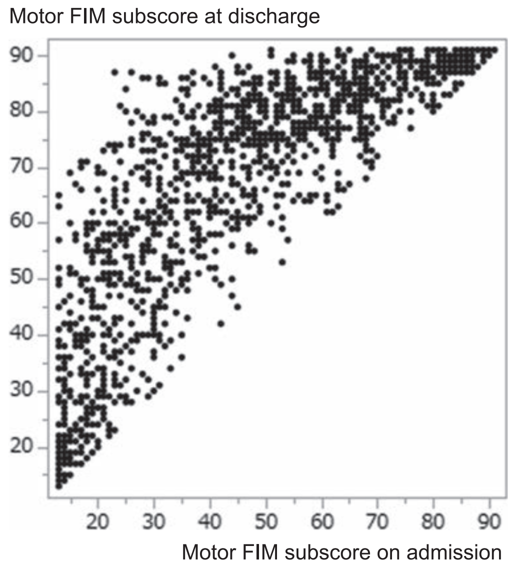

Figure 1. A scatter plot of mFIM on admission versus mFIM at discharge.

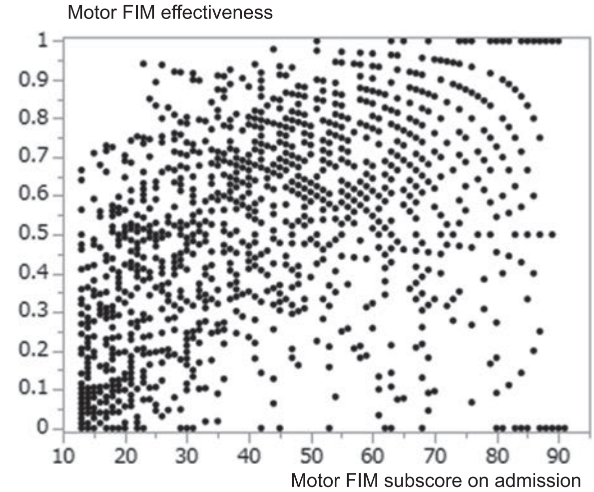

Figure 2. A scatter plot of mFIM on admission versus mFIM effectiveness.

Prediction formula S: This prediction formula was constructed by multiple regression analysis using mFIM at discharge as the dependent variable and the above-mentioned items as independent variables.

Prediction formula R: This prediction formula was constructed by multiple regression analysis using mFIM at discharge as the dependent variable and the above-mentioned items as independent variables, except that mFIM on admission was substituted by the reciprocal of mFIM on admission (1/mFIM on admission) according to Sonoda et al. [4].

Prediction formula E: First, according to the report by Tokunaga et al. [6], multiple regression analysis was performed using mFIM effectiveness as the dependent variable and the above-mentioned items as independent variables. Using the predicted mFIM effectiveness obtained, predicted mFIM at discharge was calculated using the equation: mFIM on admission + predicted mFIM effectiveness × (91 − mFIM on admission).

4. Validation methods

The 1,502 subjects were assigned randomly in two groups of 751 subjects each, using the formula for simple random sampling in JMP Pro14. One group was used to construct prediction formulas by multiple regression analysis (construction group). The other group was used to calculate predicted values using the constructed prediction formulas (validation group). In the construction group, prediction formula S, prediction formula R, and prediction formula E were constructed. In the validation group, the predicted values were determined using each of the prediction formulas, and scatter plots of measured mFIM at discharge versus predicted mFIM at discharge were generated. Then, the intraclass correlation coefficient between measured and predicted values and the residual (value obtained by subtracting the predicted value from the measured value) were calculated. Since a Bartlett test performed in advance found that the data of both groups did not follow a normal distribution, the median of the absolute value of residual was compared among the three prediction formulas using the Kruskal–Wallis test, followed by post-hoc analysis using the Steel–Dwass test. For all statistical analyses, the statistical level was set at less than 5% ($p < 0.05$).

In conducting this series of research, comprehensive consent for the use of data was obtained from patients at the time of admission, and consideration was given to protect personal information during analysis. The statistical software used was JMP Pro14.

Results

Patient characteristics are shown in Table 1. A scatter plot of mFIM on admission versus mFIM at discharge is shown in Figure 1, and a scatter plot of mFIM on admission versus mFIM effectiveness is shown in Figure 2. Scatter plots of measured mFIM versus predicted mFIM are shown in Figure 3. In the validation group, the intraclass correlation coefficients between predicted values and measured values were 0.86 using prediction formula S, 0.90 using prediction formula R, and 0.89 using prediction formula E. The residuals were −0.83±11.5 (median −0.02) using prediction formula S, 0.83±9.79 (median 0.95) using prediction formula R, and 1.03±9.89 (median 0.61) using prediction formula E. The absolute values of residual were 9.38±6.62 (median 7.90) using prediction formula S, 7.30±6.56 (median 5.67) using prediction formula R, and 7.56±6.45 (median 5.99) using prediction formula E. The Kruskal–Wallis test detected a significant difference in the absolute values of residual among the three groups ($p < 0.05$). The post-hoc Steel–Dwass test found a significant difference between prediction formula S and prediction formula R, and between prediction formula S and prediction formula E (both $p < 0.05$), and the absolute value of residual was greater for prediction formula S than for the other two formulas. No significant difference was found between prediction formula R and prediction formula E ($p = 0.52$) (Figure 4).

Discussion

The present study revealed that compared with conventional multiple regression analysis, the method using the reciprocal of mFIM on admission and the prediction method via predicting mFIM effectiveness both improved the accuracy of predicting mFIM at discharge. In the following sections, the methods of comparing outcome prediction and the characteristics

Table 1. Patient characteristics.

	Prediction group *N*=751	Validation group *N*=751
Sex (Male/Female)	432/319	423/328
Age	72.5±7.8 (72)	72.6±7.9 (72)
Number of days from onset to admission	32.6±12.1 (31)	32.5±11.9 (31)
Cerebral hemorrhage/Cerebral infarction	310/441	304/447
Motor FIM subscore on admission	45.0±22.4 (44)	43.5±22.0 (41)
Cognitive FIM subscore on admission	22.9±8.7 (24)	22.4±9.0 (23)
Motor FIM subscore at discharge	65.6±22.0 (73)	63.8±22.8 (71)

Data for this study are expressed as number of patients or mean ± standard deviation (median value).

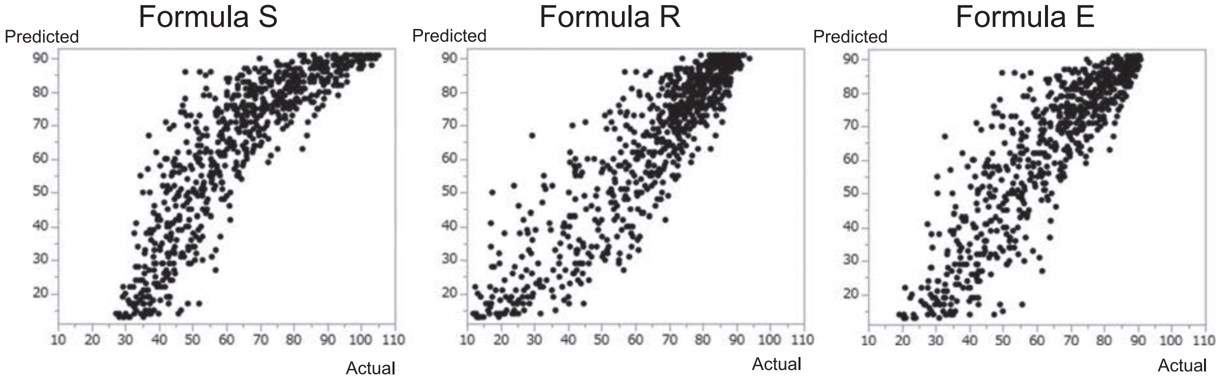

Figure 3. Scatter plots of measured mFIM versus predicted mFIM.

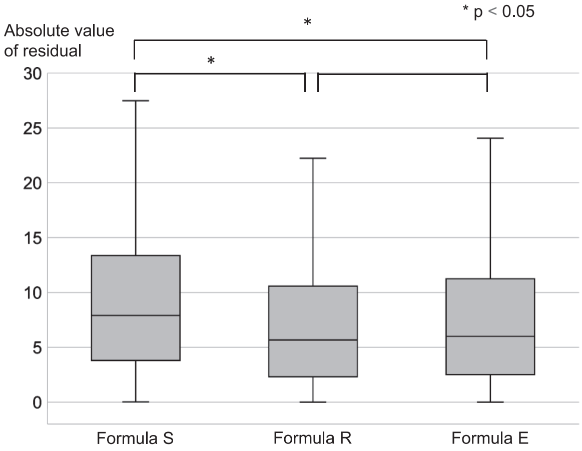

Figure 4. Comparison of absolute value of residual among the three groups.

of the prediction methods used in this study are discussed.

Methods of comparing outcome prediction

Several reports comparing the methods of outcome prediction in stroke patients have been published [2−4, 12]. Heinemann et al. [3] summarized 17 studies as a comparison between studies that conducted outcome prediction from the Uniform Data System for Medical Rehabilitation database, and revealed variations in the prediction rate and the level of contribution of variables. Meyer et al. [2] conducted a systematic review including only reports of multiple regression analysis predicting FIM and Barthel Index of stroke patients. They identified 27 studies and 63 multiple regression formulas, and their discussions centered on which variables made a significant contribution without examining the pros and cons of the prediction methods.

When comparing outcome prediction methods, it is important to confirm whether the conditions of comparison are uniform between studies. In meta-analyses and systematic reviews that pool the data of the same type of studies with different study populations, an advantage of being able to gather a large number of cases often coexists with difficulties in interpretation due to the differences between patient groups. Therefore, a comparison of outcome prediction methods using the same patient population for constructing prediction formulas and for confirming the accuracy of outcome prediction is useful, and is even better if the patient population contains a large number of homogeneous subjects. Although outcome comparison studies using the data of a large number of cases from the Japan Association of Rehabilitation Database have been conducted, there is a limitation that the data are submitted voluntarily from diverse rehabilitation environments [8, 13−17]. In the present study, all the subjects participated in the full-time integrated treatment program incorporating motor learning. Hence, we were able to utilize a relatively large patient population in the Kaifukuki rehabilitation ward who underwent the same rehabilitation program.

When comparing different statistical methods, it is important to use the same independent variable and the same amount of information provided. In this research, we planned the comparisons using the same independent variable, and we consider that the differences between the prediction formulas were clearly demonstrated by statistical methods.

For multiple regression analysis, the multiple correlation coefficient is often used as a method to compare accuracy. However, we reason that rather than showing prediction accuracy only in the group used for constructing the multiple regression formula, it is desirable to include a validation group and to demonstrate the prediction accuracy in that group. This requires using a comparison method other than the multiple correlation coefficient, because it cannot be used in the validation group. In this study, we performed comparisons using the intraclass correlation coefficient, which showed that the three prediction formulas had almost the same accuracy. Therefore, we further compared the formulas using the absolute value of residual obtained by subtracting the predicted value from the measured value. Previous studies performed comparisons using residual, square of residual, and the absolute value of residual [4, 8, 17−19]. The purpose of this study was to compare prediction formulas, and, since determining which formula predicts values closer to the actual scores is more important than finding which formula has the smaller systematic error, we used the absolute value of residual for comparison.

Characteristics of the outcome prediction methods in this study

A previous study showed that when predicting ADL score at discharge from ADL score on admission to a Kaifukuki rehabilitation ward, the relationship is not linear [4]. The reason is that when using mFIM as an outcome measure, due to the influence of the floor effect (mFIM on admission lower than 13 points cannot be obtained) and the ceiling effect (mFIM at discharge higher than 91 points cannot be achieved), the degree of ADL improvement of patients needing a moderate level of assistance is greater than that of other patients. Various methods have been proposed to solve this non-linearity problem, such as modifying the multiple regression analysis and using a non-linear regression equation [4, 20, 21]. In this study, we made various modifications to the multiple regression analysis with an easy-to-understand logical structure of prediction, and compared the formulas constructed.

According to Sonoda et al. [4] and Inouye [20], the accuracy of a multiple regression equation is improved when the model is brought closer to linearity. In the present study, prediction formula S, which is a simple multiple regression equation obtained from the scatter plot of measured values versus predicted values, yielded a sigmoid curve. On the other hand, prediction formula R and prediction formula E were models closer to linearity, and their prediction accuracy was higher than that of prediction formula S.

Both prediction formula R and prediction formula E conceivably correct the ceiling effect of mFIM on admission. First, for prediction formula R, the reciprocal of mFIM for patients with high score on admission becomes close to 0, and hence has less influence on the predicted value. In the case of prediction formula E, we first predicted mFIM effectiveness by multiple regression analysis, and then used this value to calculate mFIM at discharge. Since the values of mFIM effectiveness range from 0 to 1, in the latter stage of calculation [predicted mFIM effectiveness × (91 − mFIM on admission)] the values obtained for almost all the patients with high mFIM on admission become small, and consequently the gap between the measured value and the predicted value is also diminished.

For prediction formula R, because the minimum value of mFIM is 13 and not 0, using the reciprocal of mFIM probably has little influence on the floor effect. In addition, a study has shown that mFIM effectiveness cannot correct the low level of outcome when mFIM on admission is in the low range (13−48 points) [22]. In the present study also, the predicted value of mFIM at discharge tended to be lower than the measured value when mFIM on admission was in the low range (Figure 3), showing that the floor effect was not corrected.

As discussed above, in the present comparison of the three prediction formulas, mitigation of the ceiling effect in patients with high mFIM on admission is the main reason for the better results obtained from prediction formula R and prediction formula E compared with prediction formula S. As for the possible contributing factors for no difference between prediction formula R and prediction formula E, there is a limit to the applicability of data transformation to bring the prediction model toward linearity when using multiple regression analysis, and a certain number of patients who are affected by inhibition factors probably reduce the prediction accuracy.

Future directions

In this study, the intraclass correlation coefficient remained at around 0.9 despite the modifications used to generate prediction formula R and prediction formula E. To explore further modifications to the multiple regression analysis, piecewise multiple regression analysis that applies different multiple regression equations to the low mFIM group and high mFIM group [23] may improve the prediction accuracy in the low mFIM group.

Potential directions for improving prediction accuracy include conversion of FIM score into an interval scale by Rasch analysis [24] and prediction by adding temporal data. As examples of methods to supplement temporal data, using FIM improvement at one month after admission as the independent variable in multiple regression analysis [12], and prediction of mFIM at discharge from FIM on admission and at 2−6 weeks after admission from a logarithmic curve [21] have been reported. However, there is a concern regarding mixing methods, with the result that one may get closer to the goal of rehabilitation by waiting to collect temporal data; it is important to consider the trade-off between improvement of prediction accuracy and delay of prediction time. Furthermore, when adding an inhibition factor to the variables, it is important to consider which patient group is adversely affected by the inhibition factor [25]. In the future, neural networks [26] and AI may be applied to outcome prediction.

When applying such innovative methods, the approach that we used in this study for comparing accuracy of prediction formulas, that is, including the same variables in the models, dividing subjects into a prediction formula construction group and a validation group, and comparing the residual and intraclass correlation coefficient, may be useful.

Conclusion

With post-stroke patients hospitalized in a Kaifukuki rehabilitation ward as subjects, we compared three prediction formulas obtained from multiple regression analyses using the same variables. When prediction accuracy was assessed using the absolute value of residual obtained by subtracting the predicted value from the measured value, the prediction accuracy for prediction formula R that used the reciprocal of mFIM and for prediction formula E calculated via predicted FIM effectiveness was higher than for prediction formula S obtained from conventional multiple regression analysis. These results prove that the prediction accuracy can be improved by data transformation that brings the prediction model closer to linearity. The findings of the present study may be used to develop a more accurate outcome prediction formula by improving the prediction method and the independent variables used in the regression model.

References

1. Data management service of the Uniform Data System for Medical Rehabilitation and the Center for Functional Assessment Research: Guide for use of the Uniform Data Set for Medical Rehabilitation. version 3.1, State University of New York at Buffalo, Buffalo, 1990.
2. Meyer MJ, Pereira S, McClure A, Teasell R, Thind A, Koval J, et al. A systematic review of studies reporting multivariable models to predict functional outcomes after post-stroke inpatient rehabilitation. Disabil Rehabil 2015;

37: 1316–23.
3. Heinemann AW, Linacre JM, Wright BD, Hamilton BB, Granger C. Prediction of rehabilitation outcomes with disability measures. Arch Phys Med Rehabil 1994; 75: 133–43.
4. Sonoda S, Saitoh E, Nagai S, Okuyama Y, Suzuki T, Suzuki M. Stroke outcome prediction using reciprocal number of initial activities of daily living status. J Stroke Cerebrovasc Dis 2005; 14: 8–11.
5. Koh GC, Chen CH, Petrella R, Thind A. Rehabilitation impact indices and their independent predictors: a systematic review. BMJ Open 2013; 3: e003483.
6. Tokunaga M, Watanabe S, Sonoda S. A method of calculating functional independence measure at discharge from functional independence measure effectiveness predicted by multiple regression analysis has a high degree of predictive accuracy. J Stroke Cerebrovasc Dis 2017; 26: 1923–8.
7. Tokunaga M, Tori K, Eguchi H, Kado Y, Ikejima Y, Ushijima M, et al. The stratification of motor FIM and cognitive FIM and the creation of four prediction formulas to enable higher prediction accuracy of multiple linear regression analysis with motor FIM gain as the objective variable—an analysis of the Japan Rehabilitation Database. Jpn J Compr Rehabil Sci 2017; 8: 21–9.
8. Tokunaga M, Taniguchi M, Nakakado K, Mihono T, Okido A, Ushijima T, et al. Assessment of the effects of factors in stroke rehabilitation using eight multiple regression analyses—an analysis of the Japan Rehabilitation Database—. Jpn J Compr Rehabil Sci 2015; 6: 78–85.
9. Imada Y, Tokunaga M, Fukunaga K, Sannomiya K, Inoue R, Hamasaki H, et al. Relationship between cognitive FIM score and motor FIM gain in patients with stroke in a Kaifukuki rehabilitation ward. Jpn J Compr Rehabil Sci 2014; 5: 12–8.
10. Liu M, Domen K, Chino N. Comorbidity measures for stroke outcome research: a preliminary study. Arch Phys Med Rehabil 1997; 78: 166–72.
11. Sonoda S, Saitoh E, Nagai S, Kawakita M, Kanada Y. Full-time integrated treatment program, a new system for stroke rehabilitation in Japan: comparison with conventional rehabilitation. Am J Phys Med Rehabil 2004; 83: 88–93.
12. Tokunaga M, Katsura K, Tokisato K, Honda S, Nakanishi T, Takai S, et al. Increasing the prediction accuracy of FIM gain by adding FIM improvement for one month from admission to the explanatory variables in multiple regression analyses. Jpn J Compr Rehabil Sci 2017; 8: 16–20.
13. Jeong S, Inoue Y, Kondo K, Matsumoto D, Shiraishi N. Formula for predicting FIM for stroke patients at discharge from an acute ward or convalescent rehabilitation ward. Jpn J Compr Rehabil Sci 2014; 5: 19–25.
14. Iwai N, Aoyagi Y. Discharge index and prediction for stroke patients in the post-acute stage. Jpn J Compr Rehabil Sci 2012; 3: 37–41.
15. Mutai H, Furukawa T, Araki K, Misawa K, Hanihara T. Factors associated with functional recovery and home discharge in stroke patients admitted to a convalescent rehabilitation ward. Geristr Gerontol Int 2012; 12: 215–22.
16. Tokunaga M, Nishikawa M, Matsumoto A, Nanbu S, Nakagawa A, Maeda Y, et al. The Relationship between measured values and values predicted using multiple regression analysis for mean motor FIM at discharge —a study at 13 Kaifukuki rehabilitation hospitals for stroke patients in the Japan Rehabilitation Database—. Jpn J Compr Rehabil Sci 2015; 6: 86–90.
17. Tokunaga M, Sannomiya K, Nakanishi R, Yonemitsu H. The external validity of multiple regression analyses predicting discharge FIM score in patients with stroke hospitalized in Kaifukuki rehabilitation wards—analysis of the Japan Rehabilitation Database—. Jpn J Compr Rehabil Sci 2015; 6:14–20.
18. Jeong S, Kondo K, Shiraishi N, Inoue Y. An evaluation of post-stroke rehabilitation in Japan. Clinical Audit 2010; 2: 59–66.
19. Tokunaga M, Ikeda Y, Inoue M, Kodama J, Sakamoto M, Nagatomo M, et al. Multiple regression analysis stratified by age and FIM at admission. J Clin Rehabil 2015; 24: 828–34. Japanese.
20. Inouye M. Predicting outcomes of patients in Japan after acute stroke using a simple model. Am J Phys Med Rehabil 2001; 80: 645–9.
21. Koyam T, Matsumoto K, Okuno T, Domen K. A new method for predicting functional recovery of stroke patients with hemiplegia: logarithmic modelling. Clin Rehabil 2005; 19: 779–89.
22. Tokunaga M, Nakanishi R, Watanabe S, Maeshiro I, Hyakudome A, Sakamoto K, et al. Corrected FIM effectiveness as an index independent of FIM score on admission. Jpn J Compr Rehabil Sci 2014; 5: 7–11.
23. Sonoda S, Saitoh E, Domen K. Chino N. Prognostication of stroke patients using the Stroke Impairment Assessment Set and the Functional Independence Measure. In: Chino N, Melvin JL, eds. Functional Evaluation of Stroke Patients. 1st ed. Tokyo: Springer-Verlag Tokyo; 1996. p.103–114.
24. Linacre JM. The structure and stability of the Functional Independence Measure. Arch Phys Med Rehabil 1994; 75: 127–32.
25. Okamoto S, Sonoda S, Watanabe M, Okazaki H, Yagihashi K, Okuyama Y. Relationship between Functional Independence Measure (FIM) score on admission and influence of inhibitive factors in a comprehensive inpatient stroke rehabilitation ward. Jpn J Compr Rehabil Sci 2018; 9: 59–65.
26. Sonoda S, Saitoh E, Tsujiuchi K, Suzuki M, Domen K, Chino N. Stroke outcome prediction using neural networks. Sougou Rehabil 1995; 23: 499–504.