ORIGINAL ARTICLE

Model-based Prediction of Length of Stay for Rehabilitating Stroke Patients

*Chien-Lin Lin,[1] Pao-Hsuan Lin,[2,3] Li-Wei Chou,[1] Shou-Jen Lan,[4] Nai-Hsin Meng,[1] Sui-Foon Lo,[1] Hong-Dar Isaac Wu[5]**

Background/Purpose: Accurate length-of-stay (LOS) estimates have an impact on medical costs for stroke patients. Most studies have reported only descriptive sample means or have provided linear-model-based estimates for LOS. This study calculated point and interval estimates by treating hospital discharge as an event, and utilizing the proportional hazards (PH) model to provide the estimation of hospital stay for first-ever stroke patients in a rehabilitation department of a clinical center.

Methods: Pairwise analysis for correlations between age, sex, comorbidity status, modified Barthel index (MBI) and functional independence measure (FIM) was performed. These explanatory variables are used in the K-sample comparisons, the χ^2 test for association, the PH regression analysis, and log-transformed linear (LTL) regression.

Results: The PH model gave a prediction on estimated mean LOS, with an absolute bias of 0.85 days, by combining MBI and FIM into a single variable, or a bias of 1.15 days and 1.16 days with MBI and FIM variables, respectively. The LTL-based estimation generated a bias of 5.91 days. The PH model has relatively shorter confidence intervals than those obtained by sample-mean and LTL methods.

Conclusion: We recommend using the PH model for predicting mean LOS when the PH assumption for patients with different clinical characteristics is satisfied. However, the proposed method only applies to rehabilitating stroke patients. [*J Formos Med Assoc* 2009;108(8):653–662]

Key Words: length of stay, linear models, log-transform, proportional hazards models, rehabilitation

Since Taiwan instituted national health insurance 14 years ago, close to 98% of the population is covered by the program.[1] The program uses a universal budget payment method to control medical care costs. Payment for excessive length of hospital stay (LOS) in relation to each disease is closely monitored, as LOS has a significant impact on the healthcare budget. Cerebrovascular disease was the second leading cause of death in Taiwan in 1984–2003 (crude mortalities, 53.5–78.4/10^5 person-years),[1] and rehabilitating stroke patients often have longer mean LOS than that for all diseases combined (9.4 days in Taiwan, 2003).[1] Therefore, it is essential to investigate the principal factors that affect LOS in order to manage healthcare costs. Although LOS is a factor in determining inpatient short-term prognosis, it may also be a direct or indirect indicator of long-term survival.[2–4] Accurate LOS estimates for stroke patients and their families are important. These LOS estimates allow nursing home networks to prepare for delivering appropriate after-discharge home care.

[1]Department of Rehabilitation, China Medical University Hospital, [2]Biostatistics Center, and [3]Department of Public Health, China Medical University, [4]Graduate Institute of Health Administration, Asia University, and [5]Department of Applied Mathematics and Institute of Statistics, National Chung-Hsing University, Taiwan.

Received: March 19, 2008
Revised: August 20, 2008
Accepted: March 8, 2009

***Correspondence to:** Dr Hong-Dar Isaac Wu, Department of Applied Mathematics and Institute of Statistics, National Chung-Hsing University, 250 Kuo-Kuang Road, Taichung 402, Taiwan.
E-mail: honda@amath.nchu.edu.tw

For hospitals, precise prediction of LOS facilitates flexibility in managing bed occupancy. The effectiveness of various treatments and clinical management can be investigated by analysis of mean LOS. For the national insurance authority, surveying LOS between hospitals at the same or different levels, comparing LOS between areas, assessing the cost-effectiveness of current treatment strategies, and preparing randomized trials for outcome-oriented evaluations also depends on accurate LOS predictions.[2,5–8]

Estimation of mean LOS for stroke patients can be based on sample means[9] (which is "model-free"), according to specific categories defined by age, sex, comorbidity, and patient-modified Barthel index (MBI), functional independence measure (FIM), and other measures of function.[10] Unless the available sample size is extremely large, calculating the sample mean with its large-sample property for prediction of LOS is not efficient. On the other hand, obtaining a model-based prediction is appealing, in that it facilitates unified comparisons between different hospitals, and renders meaningful monitoring of medical resources in a national healthcare insurance system.

Natural choices for model-based LOS analysis include: log-transformed linear (LTL) regression and median regression, both of which account for distributional skewedness.[11,12] A major limitation of these approaches is that patients can be discharged for numerous reasons (e.g. cure, transfer, or death[11,13]). Thus, the observed LOS may be a right-censored datum, which indicates that a direct result for the mean LOS estimate may be an underestimate. However, if discharge from hospital is treated as an event-time variable and proportional hazards (PH) regression is applied,[14] mean LOS can be obtained based on the theory of event-history data analysis. Notably, the PH model has been used widely when analyzing outpatient mortality or survival.[3,4,9,15]

The present study compared two analyses of mean LOS prediction: sample means, and the PH model. However, LTL assessment was also compared in order to illustrate its deviation to underestimate the mean LOS. Patients enrolled in the analysis had experienced a cerebrovascular attack and were recruited from the Rehabilitation Department at China Medical University Hospital (CMUH) in central Taiwan.

Materials and Methods

Patients and data collection

We enrolled 586 patients who had experienced their first stroke, with cerebral hemorrhage or cerebral infarction, from a 1400-bed medical center at CMUH between January 1, 1997 and February 28, 2005. The patients were recruited from neurology, internal medicine, and emergency resuscitation departments and transferred (or re-hospitalized) to the Department of Rehabilitation at CMUH. Transfer date was set as the zero time point for event history analysis. The baseline data collected from hospital records were as follows: age, sex, coexistence of diabetes mellitus and/or hypertension, history of stroke and severe injury, and family disease history. These data were recorded typically within the first 6 hours of hospitalization for an acute-stage event. For patients who were admitted for rehabilitation, which comprised physical, occupational or speech therapy, MBI and FIM questionnaires were administered within 24 hours of transfer. The MBI and FIM scores represent the generic severity of disability of inpatients, and have been applied widely in stroke research and various medical fields.[10,17–20] Furthermore, the change in score following treatment is indicative of patient improvement.[18] These scores are recorded routinely for patients in various departments, particularly at the Rehabilitation Department of CMUH. The data were ascertained from a computerized databank. Patients who had a previous event other than stroke, such as trauma or head injury, were excluded, as were patients who had received rehabilitation. Figure 1 presents a brief description of the process used to collect data.

When an event-history analysis is used, patients are discharged as a result of curative treatment, and not for death or other causes. For outpatients

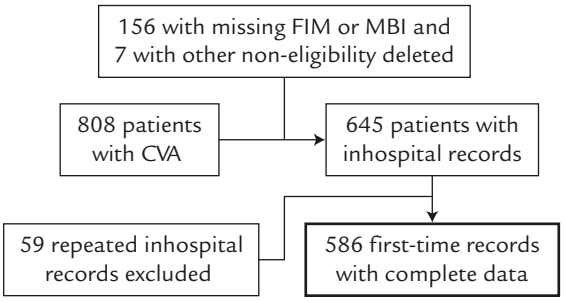

Figure 1. Data processing with retrospective exclusion of non-eligible inpatient records. FIM = functional independence measure; MBI = modified Barthel index; CVA = cerebrovascular accident.

who were transferred to other hospitals and then returned, the first LOS at CMUH was treated as a right-censored observation. As a reference for comparison, sample means and 95% confidence intervals (CIs) unique to each category were calculated. The following two model-based estimates were considered.

Cox PH model

$$h(t;Z) = h_0(t)\exp(\beta_1 \times age + \beta_2 \times sex + \beta_3 \times MBI + \beta_4 \times FIM\ldots),$$

where h(t;Z) is the incidence (or hazard) function of event time with covariate Z (which is a vector of age, sex, MBI and FIM), and $h_0(t)$ is an unknown baseline incidence function. A simple estimate of mean survival, for an individual with a covariate-vector Z, is $\int \hat{S}_Z(t)dt$, where

$$\hat{S}_Z(t) = \exp\{-\int \hat{h}_0(t)\ \exp(\hat{\beta}^T Z)dt\}.$$

In the above expression, $\beta^T Z = \beta_1 \times age + \beta_2 \times sex + \beta_3 \times MBI + \beta_4 \times FIM\ldots$. Estimations for parameters β and $h_0(t)$ follow standard statistical principles. The corresponding CIs are generated as $(\int S_{L,PH}(t)dt,\ \int S_{U,PH}(t)dt)$, where $S_{L,PH}$ and $S_{U,PH}$ are upper and lower confidence limits, respectively, derived from the formulas offered in Klein and Moeschberger.[16]

LTL regression

$$\log(LOS + 1) = \beta_0 + \beta_1 \times age + \beta_2 \times sex + \beta_3 \times MBI + \beta_4 \times FIM\ldots.$$

To eliminate any heterogeneity effect, the variables age, FIM and MBI were categorized into several groups (Table 1). Finally, the estimate of mean log(LOS + 1) was calculated easily using conventional linear model theory with the CI, denoted as (C_l, C_u); thus, the transformed CI was $(\exp(C_l) - 1, \exp(C_u) - 1)$. Note that the underestimate based on the LTL model was attributed to the log- and exponential-transforms procedure.

The inter-relationship between LOS and potential predictors of LOS were investigated preliminarily by comparing means and medians using descriptive statistics and Wilcoxon's rank-sum test (or their multiple-sample counterparts). As an intermediate step, the confounding structure of predictors was identified by the joint distributions between the most significant variables using correlation coefficients and the χ^2 test for association. Finally, the PH and LTL models were implemented to estimate regression coefficients, rate ratios, and associated mean LOS. A value of $p < 0.05$ was considered statistically significant.

Results

Bivariate analysis of the confounding variables

Sex was not significant in predicting LOS (Table 1). Male and female patients had similar values for mean and median LOS, and other percentiles (rank-sum tests, $p = 0.262$). Age, however, was significant; multiple comparison tests revealed that patients aged < 50 and ≥ 80 years were statistically equivalent in mean/median LOS. Comorbidity status and physical therapy, occupational therapy and/or speech therapy had no predictive power. The most significant variables were MBI and FIM scores (both $p < 0.001$).

To determine the validity of model-based prediction, the correlation structure was examined. This assessment helped establish a group of possible explanatory variables. As age ($p = 0.047$), MBI and FIM were the most significant variables (Table 1), their pairwise associations were examined (Tables 2 and 3). The joint distribution of FIM and

Table 1. K-sample (K ≥ 2) comparison for LOS of 586 stroke patients with different sex, age, comorbidity, physical therapy, occupational therapy, speech therapy, MBI score and FIM score

Variables		Statistics						Test K–W*
		n	Mean	SD	Q1	Q2	Q3	
Sex	Female	236	31.6	19.7	17.0	28.0	43.0	0.262
	Male	350	29.9	19.4	14.0	28.0	41.0	
Age (yr)	< 50	105	28.3	19.3	13.0	23.0	41.0	0.047
	50–64	189	33.1	19.9	19.0	30.0	45.0	
	65–79	250	30.4	19.6	16.0	28.0	42.0	
	≥ 80	42	26.1	17.2	12.0	25.0	34.0	
Comorbidity	None	226	31.0	20.3	15.0	28.5	43.0	0.850
	DM	28	31.6	20.0	17.5	30.5	40.0	
	HYP	257	30.5	18.7	16.0	27.0	43.0	
	DM + HYP	75	29.3	20.2	14.0	26.0	40.0	
PT	Yes	581	30.4	19.5	15.0	28.0	42.0	0.064
	No	5	47.8	23.7	34.0	37.0	50.0	
OT	Yes	555	30.3	19.3	15.0	28.0	42.0	0.335
	No	31	34.8	23.1	17.0	31.0	47.0	
ST	Yes	330	31.3	19.8	17.0	28.0	41.0	0.329
	No	256	29.6	19.1	14.0	28.0	43.0	
MBI	0	154	34.5	19.0	20.0	31.5	46.0	< 0.001
	5–30	352	30.8	20.2	16.0	28.0	42.0	
	≥ 35	80	21.9	14.8	10.5	18.5	30.5	
FIM	< 29	146	36.7	21.7	21.0	33.0	48.0	< 0.001
	29–63	286	31.4	19.2	16.0	29.0	43.0	
	≥ 64	154	23.3	15.4	11.0	21.0	33.0	

**Kruskal–Wallis test, reduces to Wilcoxon's rank sum test when K = 2. n = sample size; SD = standard deviation; Q1, Q2, and Q3 = 25%, 50% (median), and 75% points; K–W = Kruskal–Wallis; DM = diabetes mellitus; HYP = hypertension; PT = physiotherapy; OT = occupational therapy; ST = speech therapy; MBI = modified Barthel index; FIM = functional independence measure.*

Table 2. Joint distribution (or cross classification) between FIM and MBI scores at patient admission*

Patient no.		MBI					Total
		0	5	10–20	25–30	≥ 35	
FIM	< 20	43 (84.3%)	2 (3.9%)	5 (9.8%)	1 (2.0%)	0 (0.0%)	51
	20–28	46 (48.4%)	7 (7.4%)	39 (41.1%)	3 (3.2%)	0 (0.0%)	95
	29–44	53 (37.1%)	19 (13.3%)	54 (37.8%)	13 (9.1%)	4 (2.8%)	143
	45–63	11 (7.7%)	15 (10.5%)	83 (58.0%)	24 (16.8%)	10 (7.0%)	143
	64–80	1 (1.1%)	3 (3.4%)	25 (28.4%)	35 (39.8%)	24 (27.3%)	88
	≥ 81	0 (0.0%)	1 (1.5%)	9 (13.6%)	14 (21.2%)	42 (63.6%)	66
Total		154 (26.3%)	47 (8.0%)	215 (36.7%)	90 (15.4%)	80 (13.7%)	586

**Pearson's and Spearman's correlation coefficients between FIM and MBI (at entry) were 0.727 ($p < 0.001$) and 0.725 ($p < 0.001$), respectively. FIM = functional independence measure; MBI = modified Barthel index.*

MBI scores was expressed by box-plots (Figure 2, FIMs with respect to different MBI groups) and a cross-classified table (Table 2). For different age groups, individual MBI and FIM scores were compared (Table 3). Analytical results indicated that although MBI and FIM were measures for different aspects of a stroke patient's condition, they were highly correlated. When these scores were

Table 3. Joint distributions between age and MBI score (upper panel), and age and FIM score (lower panel)

Patient no.		Age (yr)				*p**
		<50	50–64	65–79	≥80	
MBI	0	22 (14.3%)	42 (27.3%)	70 (45.5%)	20 (13.0%)	<0.001
	5–30	59 (16.8%)	115 (32.7%)	157 (44.6%)	21 (6.0%)	
	≥35	24 (30.0%)	32 (40.0%)	23 (28.8%)	1 (1.3%)	
FIM	<29	18 (12.3%)	45 (30.8%)	70 (48.0%)	13 (8.9%)	0.005
	29–63	45 (15.7%)	95 (33.2%)	121 (42.3%)	25 (8.7%)	
	≥64	42 (27.3%)	49 (31.8%)	59 (38.3%)	4 (2.6%)	

**By conventional χ^2 test for association (Fisher's exact test gave similar results but is not reported). MBI = modified Barthel index; FIM = functional independence measure.*

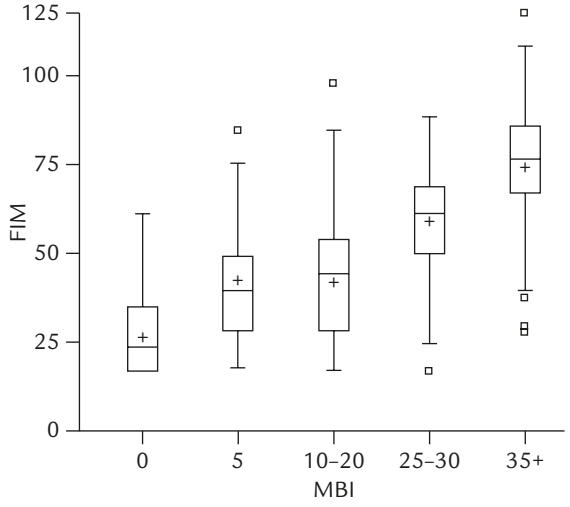

Figure 2. Box plot of FIM scores according to different MBI groups, showing the relationship between individual patient MBI and FIM scores at admission. FIM = functional independence measure; MBI = modified Barthel index.

treated as continuous variables or categorized into several groups, Pearson's and Spearman's correlation coefficients were both approximately 0.73 (both $p < 0.001$; χ^2 test, $p < 0.001$). Older patients had lower MBI and FIM scores. In particular, 28.4% and 30.8% of patients aged ≥ 65 years had MBI = 0 and FIM < 29, respectively (Table 3); however, 24.9% and 26.3% of the entire sample (586 patients) had MBI = 0 and FIM < 29, respectively. To investigate the relationship between LOS and MBI/FIM, a ceiling effect of the functional scores may result in a heterogeneous effect in subsequent regression analysis. According to the data, this problem was negligible for FIM, and because those with MBI > 35 comprised a small group, the heterogeneity effect within this group was averaged out (because of grouping). In summary, the MBI and FIM scores grouped in this analysis reduced bias during regression analysis.

Regression models

Table 4 presents PH univariate analysis results with the regressors sex, age, MBI, MBI-diff, FIM, and FIM-diff. MBI-diff and FIM-diff represented the differences in FIM and MBI scores for dates of discharge and hospitalization. These two difference scores were time-dependent covariates that had a dynamic meaning, in that the patients and/or attending physicians assessed improvement, to determine whether a patient should have been discharged. Therefore, score changes were not suitable for use as predictors for further estimation of mean LOS and other outcome variables, which were unknown at admission. A rate ratio (RR) > 1 indicated likely discharge compared with that for the reference group. The functioning scores and their differences were much more important than age and sex (Table 4). Large MBI and FIM scores at admission shortened LOS. The trends in RR for MBI (1.40 and 2.19) and FIM (1.23 and 2.20) were both significant. Conversely, patients who had a large difference in FIM or MBI scores usually had long LOS (RR < 1). This indicated that clinical improvement in patient functioning was a result of effective rehabilitation, thereby encouraging a prolonged LOS. However, those who did not obtain functional improvement tended to be discharged early.

Table 4. Rate-ratio estimate for various explanatory variables using the univariate Cox proportional hazards model

Variable		RR	95% CI	*p*
Sex	Male	1.08	0.91–1.28	0.369
	Female	1	–	–
Age (yr)	< 50	0.78	0.53–1.13	0.183
	50–64	0.66	0.47–0.93	0.017
	65–79	0.79	0.57–1.11	0.170
	≥ 80	1	–	–
MBI	0	1	–	–
	5–30	1.40	1.14–1.73	0.002
	≥ 35	2.19	1.73–2.78	< 0.001
MBI-diff	≤ 0	1	–	–
	1–14	0.75	0.57–0.99	0.039
	15–29	0.61	0.47–0.81	0.001
	≥ 30	0.58	0.41–0.82	0.002
FIM	< 29	1	–	–
	29–63	1.23	1.01–1.49	0.040
	≥ 64	2.20	1.63–2.95	< 0.001
FIM-diff	≤ 0	1	–	–
	1–2	0.84	0.65–1.09	0.187
	3–11	0.74	0.58–0.94	0.016
	≥ 12	0.75	0.55–1.03	0.074

RR = rate ratio; CI = confidence interval; MBI = modified Barthel index; MBI-diff = difference in MBI score for dates of discharge and hospitalization; FIM = functional independence measure; FIM-diff = difference in FIM score for dates of discharge and hospitalization.

Table 5 presents the mean LOS estimates and associated CIs for the three methods (sample-mean, PH, and LTL). Subgroups were characterized by age versus a variable with MBI/FIM combined. MBI and FIM were combined for the following reasons. First, MBI and FIM were highly correlated, and second, many patients had an MBI of 0 at admission. Patients with MBI = 0 were divided into two groups: those with FIM < 29 and those with FIM ≥ 29. The group with MBI 5–30 consisted of three subgroups (Table 2), who did not differ significantly for incidence of discharge. Briefly, the categorization (on age, MBI and FIM) in Table 5 produced the smallest absolute bias. The MBI/FIM combined analysis could be ignored, and only a single variable was utilized, FIM or MBI, to simplify analysis. Finally, patients aged ≥ 80 years were excluded from analysis of PH-based prediction, as most elderly patients in Taiwan had other concerns about their discharge, such as whether they would be cared for by their families, or be transferred to a nursing home. Exclusion of elderly patients resulted in a non-proportional-hazards phenomenon when compared with other groups. The sample mean estimates for the first MBI/FIM group (MBI = 0 and FIM < 29) were 30.7, 39.1 and 32.5 days for the three age groups (Table 5). The corresponding means estimated by the PH model were 32.4, 35.8 and 32.0 days, respectively. All absolute biases (1.7, 3.3, and 0.5 days) were tolerable. For the other entries, the meaning was similar. Generally, for the group 5 ≤ MBI ≤ 30, the PH model had a very precise prediction, with absolute biases of 0.1, 0.4, and 0.6. Conversely, the LTL model estimated mean LOS of 26.5, 32.8, and 27.5 days for the first group (MBI = 0 and FIM < 29), which resulted in a large absolute bias and wide CIs. Overall, by taking the weighted average according to the sample size of each entry, the PH model obtained a mean absolute bias of

Table 5. Prediction of mean LOS and the corresponding 95% CIs from the Cox PH model compared with the method of naïve sample means and that based on an LTL regression model

		Age (yr)								
		< 50			≥ 50, < 65			≥ 65, < 80		
		mean	lcl	ucl	mean	lcl	ucl	mean	lcl	ucl
MBI = 0 and	*n*	10			27			42		
FIM < 29	Mean	30.7	21.5	39.9	39.1	30.6	47.7	32.5	26.6	38.4
	PH	32.4	29.0	35.8	35.8	32.0	39.6	32.0	29.3	34.7
	LTL	26.5	21.3	32.7	32.8	27.3	39.4	27.5	23.1	32.7
MBI = 0 and	*n*	12			15			28		
FIM ≥ 29	Mean	33.3	21.2	45.5	38.4	27.2	49.6	37.7	30.1	45.3
	PH	35.3	30.7	39.8	39.0	33.4	44.5	34.8	30.9	38.6
	LTL	28.4	22.5	35.7	35.2	28.4	43.6	29.5	24.1	36.2
5 ≤ MBI ≤ 30	*n*	59			115			157		
	Mean	29.8	24.5	35.1	33.4	29.7	37.2	30.1	27.0	33.3
	PH	29.9	27.8	31.9	33.0	30.8	35.2	29.5	28.0	30.9
	LTL	22.6	19.4	26.4	28.1	25.0	31.5	23.5	21.2	26.1
MBI ≥ 35	*n*	24			32			23		
	Mean	20.9	13.4	28.3	24.3	19.5	29.2	19.3	13.3	25.2
	PH	20.8	19.8	21.7	22.9	21.7	24.1	20.5	19.6	21.4
	LTL	15.2	12.4	18.6	19.0	15.8	22.8	15.9	13.1	19.2

LOS = length of stay; CI = confidence interval; PH = proportional hazards; LTL = log-transformed linear; lcl = lower 95% confidence limit; ucl = upper 95% confidence limit; MBI = modified Barthel index; FIM = functional independence measure.

0.85 days, and that obtained with the LTL model was 5.91 days. The LTL-based analysis had a systemic bias that could only be avoided using *ad hoc* and posterior adjustments. Additionally, PH-based CIs of each category were markedly shorter than those of LTL analysis and sample mean estimates.

The effectiveness in using Cox's PH technique as a building block for predicting LOS depends on the proportional hazards assumption. This assumption can be examined readily using standard statistical packages (e.g. SAS 8.2 and S-Plus 4.5). For an illustration, only the Kaplan–Meier (KM) estimates for the survivor function for different MBI and FIM groups are shown (Figure 3). In both of the curves, proportionality was acceptable. The same KM plot for different ages (excluding ≥ 80 years) had a similar type and was omitted. In conclusion, the estimates based on Cox's PH model were satisfactory for obtaining mean LOS prediction.

Discussion

This study addressed the need to better predict the LOS of patients during inpatient stroke rehabilitation, which is an important medical and economic issue. PH regression was utilized for the following reasons: (1) PH regression provides convenient explanations regarding the intensity of event of discharge for different patients, and can be implemented easily using various statistical packages. Moreover, PH regression is efficient; i.e. it has short CIs. (2) PH regression can achieve a unified assessment of LOS for intra- and inter-hospital, and multilevel comparisons. (3) The hazard-regression model can be extended to a multivariate setting such that short-term events (e.g. LOS) and long-term events (e.g. mortality) can be modeled together in a general framework.

However, the PH model assumes proportionality, which is a strong condition that the among-group incidences may not satisfy. Consequently,

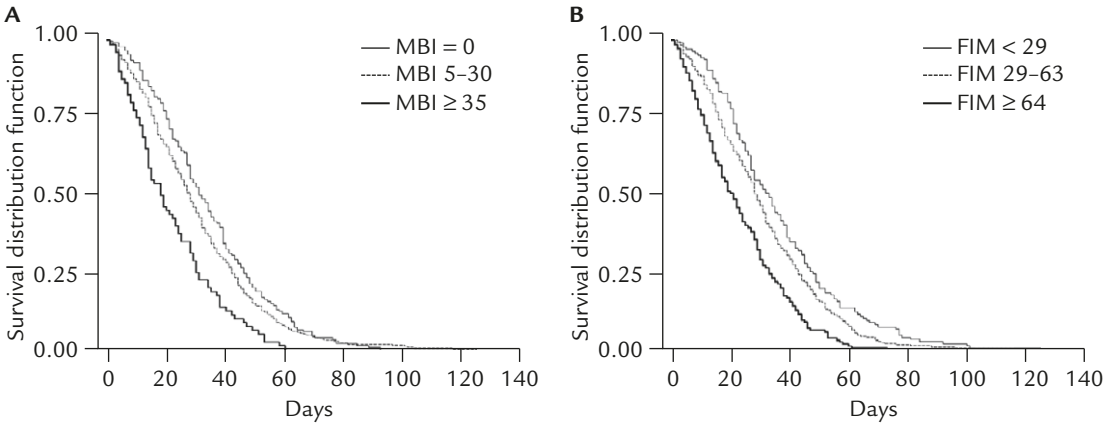

Figure 3. Kaplan–Meier survival estimates for different (A) MBI and (B) FIM groups. MBI = modified Barthel index; FIM = functional independence measure.

imposing proportionality unavoidably introduces a bias to mean LOS estimates. To remedy this fault, as many confounders as possible must be collected at an early stage of a patient's hospitalization. Moreover, 17 (2.9%) right-censored observations existed. If they were further considered, sample means would be slightly larger than those calculated in the present study, which would generate a larger absolute bias than 0.85 days for PH estimates.

To validate the model employed with continuously distributed variables, the dataset can be split randomly into a training set and a test set for analysis.[21] On the other hand, the variables in this study are all categorized, so that an easy-to-use table can be prepared for clinicians. With this purpose in mind, a model with a parsimonious number of variables was constructed. Table 5 potentially offers such parsimony. Nevertheless, the results in Table 5 are not dogmatic. The MBI and FIM are essentially measuring similar activities. FIM has cognitive tasks in addition to the motor activities seen on MBI. FIM is used traditionally in an inpatient setting and scored by therapists, and MBI is used typically for monitoring outpatients. In the present study, FIM was probably better to be used for prediction. Thus, if the combined MBI/FIM variable was to be replaced by a single FIM variable, the absolute bias of mean LOS prediction was 1.16 days (data not shown).

Other variables were analyzed in the present study. For example, the differences between MBI and FIM scores are important to clinical practitioners. These differences are correlated strongly with LOS. Notably, LOS for stroke patients in rehabilitation was correlated positively with MBI and FIM differences (Table 4). However, this correlation was not predictive at hospitalization. That is, physicians usually determine a patient's prognosis and daily condition when deciding whether to discharge a patient; patients themselves and their families sometimes request discharge as a result of self-assessed improvement. Consequently, prior knowledge of a patient's progress is supposed to be unknown at hospitalization, in order to predict the possible LOS by the score differences.

In many studies, stroke type (such as cerebral hemorrhage, cerebral infarction, and transient ischemic attack) is a very significant predictor of LOS, if LOS is defined as length of the entire hospital stay. The term "entire" implies that it contains the acute stage—therefore, the time of admission to the neurology, internal medicine and emergency resuscitation departments is defined as "time zero". In the present study, however, the defined LOS contained only the period from admission to the rehabilitation department to discharge, so that it could be viewed as length of stay in the rehabilitation department. In that manner, that which was closely relevant to the characteristics at the acute stage will possibly decay during the rehabilitation stage. Certainly, it is still an interesting issue to be clarified, and is more likely

to be complicated by the causes of death that occur at this stage.

Comorbidity status may have an impact on LOS and subsequent survival.[22,23] In a previous study of patients who experienced their first stroke, comorbidity was a confounder and an effect-modifier. To deal with this phenomenon by modeling LOS through PH-based regression, three approaches can be considered: a PH model with interaction terms; a PH model combined with a stratified analysis; and a stratified PH model that uses comorbidity as an index that stratifies the baseline cumulative incidence. These approaches are more complex statistical approaches. To simplify the present study, they were not utilized.

Finally, the MBI and FIM scores were limited by their reproducibility, although these scores have been well-tested in previous studies.[10,18] In future studies, MBI and FIM scores and other variables obtained via questionnaires should be examined for interrater reliability, so that uncertainty caused by sampling properties can be reduced, and the impact of measurement errors can also be assessed. Moreover, because this was a retrospective study, some important variables could not be tracked, including the National Institutes of Health Stroke Scale, the history of diseases related to cerebrovascular diseases and rehabilitation therapy, and risk factors such as smoking, alcohol consumption, other comorbidities, and various serum biochemical indicators. These data should also be collected uniformly to increase prediction accuracy. Therefore, the results obtained in the present study should be confined to a population such as that defined in this study.

Acknowledgments

The authors would like to thank the National Science Council (NSC 92-2118-M-039-001), China Medical University (CMU 92-PH-04), and China Medical University Hospital (DMR-95-081), Taiwan, for financially supporting this research; and Dr Mu-Jung Kao, Dr Pey-Yu Yang, and Dr Ting-I Hang for offering part of the data.

References

1. Department of Health, Executive Yuan, Taiwan. *Statistics and Publications*. Available at: http://www.doh.gov.tw/EN2006/ [Date accessed: July 11, 2009]
2. Rønning OM, Guldvog B. Stroke units versus general medical wards, I: twelve- and eighteen-month survival: a randomized, controlled trial. *Stroke* 1998;29:58–62.
3. Thourani VH, Weintraub WS, Guyton RA, et al. Outcomes and long-term survival for patients undergoing mitral valve repair versus replacement: effect of age and concomitant coronary artery bypass grafting. *Circulation* 2003;108:298–304.
4. Toumpoulis IK, Anagnostopoulos CE, DeRose JJ Jr, et al. The impact of deep sternal wound infection on long-term survival after artery bypass grafting. *Chest* 2005;127:464–71.
5. Lanska DJ. Length of hospital stay for cerebrovascular disease in the United States: Professional Activity Study, 1963–1991. *J Neurol Sci* 1994;127:214–20.
6. McKendry M, McGloin H, Saberi D, et al. Randomised controlled trial assessing the impact of a nurse delivered, flow monitored protocol for optimization of circulatory status after cardiac surgery. *Br Med J* 2004;329:258–61.
7. Caro JJ, Huybrechts KF, Kelley HE, for the Stroke Economic Analysis Group. Predicting treatment costs after acute ischemic stroke on the basis of patient characteristics at presentation and early dysfunction. *Stroke* 2001;32:100–6.
8. Van Exel NJ, Koopmanschap MA, Reimer WS, et al. Cost-effectiveness of integrated stroke services. *Q J Med* 2005;98:415–25.
9. Grieve R, Hutton J, Bhalla A, et al. A comparison of the costs and survival of hospital-admitted stroke patients across Europe. *Stroke* 2001;32:1684–91.
10. Kwon S, Hartzema AG, Duncan PW, et al. Disability measures in stroke: relationship among the Barthel Index, the Functional Independence Measure, and the Modified Rankin Scale. *Stroke* 2004;35:918–23.
11. Chang KC, Tseng MC, Weng HH, et al. Prediction of length of stay of first-ever ischemic stroke. *Stroke* 2002;33:2670–4.
12. Lee AH, Fung WK, Fu B. Analyzing hospital length of stay: mean or median regression? *Med Care* 2003;41:681–6.
13. Chang KC, Tseng MC. Costs of acute care of first-ever ischemic stroke in Taiwan. *Stroke* 2003;34:e219–21.
14. Cox DR. Regression models and life tables (with discussion). *J R Stat Soc B* 1972;34:187–220.
15. Wang SL, Pan WH, Lee MC, et al. Predictors of survival among elders suffering strokes in Taiwan. *Stroke* 2000;31:2354–60.
16. Klein JP, Moeschberger ML. *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer, 1997:259–60 (Chapter 8).
17. Feigenson JS, McDowell FH, Meese P, et al. Factors influencing outcome and length of stay in a stroke rehabilitation

unit. Part 1: analysis of 248 unscreened patients—medical and functional prognostic indicators. *Stroke* 1997;8:651–6.
18. Van Der Putten JJ, Hobart JC, Freeman JA, et al. Measuring change in disability after inpatient rehabilitation: comparison of the responsiveness of the Barthel Index and the Functional Independence Measure. *J Neurol Neurosurg Psychiatry* 1999;66:480–4.
19. Weimar C, Kurth T, Kraywinkel K, et al, for the Germany Stroke Data Bank Collaborators. Assessment of functioning and disability after ischemic stroke. *Stroke* 2002;33: 2053–9.
20. Bode RK, Heinemann AW, Semik P, et al. Relative importance of rehabilitation therapy characteristics on functional outcomes for persons with stroke. *Stroke* 2004;35: 2537–42.
21. Jiménez RE, Lam RM, Marot M, et al. Observed–predicted length of stay for an acute psychiatric department, as an indicator of inpatient care inefficiencies. Retrospective case-series study. *BMC Health Serv Res* 2004;4:4–13.
22. Rochon PA, Katz JN, Morrow LA, et al. Comorbid illness is associated with survival and length of hospital stay in patients with chronic disability: a prospective comparison of three comorbidity indices. *Med Care* 1996;34:1093–101.
23. Liu M, Domen K, Chino N. Comorbidity measure for stroke outcome research: a preliminary study. *Arch Phys Med Rehabil* 1997;78:166–72.