An evaluation of the quality of post-stroke rehabilitation in Japan

Seungwon Jeong
Katsunori Kondo
Nariaki Shiraishi
Yusuke Inoue
Nihon Fukushi University, Nagoya, Japan

Context and objectives: This study was to assess the differences in rehabilitation outcomes between the different facilities in Japan, and to determine if there was any variation in patients’ functional recovery at hospital discharge across the different facilities.

Methods: This study focused on patients in recovery in the rehabilitation ward using the data of 680 patients from 12 hospitals after adjusting for triage at admission obtained from the Rehabilitation Patient Databank in Japan (issued in September, 2009) and compared the therapeutic results of each hospital. We estimate the expected value of levels of activities of daily living (ADL) at discharge for rehabilitation patients using regression analysis. Furthermore, we show the distribution of the expected improvement levels in ADL by hospitals.

Findings: At the time of admission, there were no differences among hospitals in their patients’ characters. However, outcomes differed widely among the hospitals. The differences in the participation of physicians registered as rehabilitation specialists, conference execution rate, amount of exercise per day, self-exercise without a therapist, and exercise in wards, were statistically significant differences between hospitals.

Conclusion: Due to the assessments for health care quality and the publication of results it is expected that health care providers will put in voluntary effort to improve their future health care services. Further studies should analyze the characteristics of high-performing hospitals.

Keywords: rehabilitation, quality of care, post-stroke

Introduction

Recently there have been growing concerns and requirements for health care quality in Japan. The introduction and implementation of pay for performance (P4P) programs for rehabilitation and recovery in 2008 was a significant advancement in the area of health care services. However, there still remains a gap involving a need for close examination of what factors may critically influence the rehabilitation medical service quality and the degree to which hospitals have made, or will make, efforts in achieving best practice in rehabilitation. To begin with, the assessment criteria for P4P set forth by the Japanese government are summarized as follows.

In Japan, P4P inpatient rehabilitation among stroke survivors began in 2008 with a primary objective to provide the services needed and improve post stroke patient functional recovery. Despite concerns over the effectiveness and quality of P4P by some stakeholders of the program in Japan, three standards were developed for the P4P inpatient rehabilitation program. They included: 1) >60% of participating stroke survivors should be discharged to the community; 2) >15% of severe stroke patients...
Material and population

In this retrospective study, patient level data were obtained from the Rehabilitation Patients Databank in Japan.

Hospitals contracted with the databank collected patient data twice per year, the first between January and February and the second between July and August, from April 2005 to March 2009. The patients who discharge from the hospital during the survey period are registered with the databank (consecutive data).

To ensure our outcomes of interest were comparable across the different facilities, we used the following inclusion and exclusion criteria to select our study cohort:

1. Drawing from the literature that acute patients and chronic patients (both of whom are included in the Rehabilitation Patients Databank) have different attributes, this study covered only one type of patients: chronic patients (21 hospitals, N = 1,519).

2. The effectiveness of chronic rehabilitation is affected by compound factors including the patient’s environment and treatment by medical practitioners, and thus this study attempted to use patient data with as similar conditions as possible in comparing different hospitals. To minimize the difference due to age in improvement rate in ADL, the study was conducted among patients aged 55 to 84 (N = 1,212).

3. By drawing on the literature that early onset was important for more effective rehabilitation for stroke patients, the interval between the stroke and the admission was confined to over 1 week and below 8 weeks (N = 1,201). The specific reason for choosing that interval was that patients treated within a week of the stroke were highly likely to be acute patients, who are more easily subject to compound factors that affect rehabilitation effects, and patients admitted after over 2 months were likely to have lost the opportunity of timely treatment, which could prevent accurate assessment of treatment effects albeit they were properly treated.

4. Similarly, the length of inpatient stay for this study were also confined to between over 1 week and below 8 weeks (N = 969) because patients who were discharged within a week were assumed to have stop being treated, and for patients who stayed in hospital over 2 months, there was a possibility that their length of stay increased for other reasons such as a complication.

5. Lastly, this study conducted a comparative analysis of patient-level treatment achievement in rehabilitation among hospitals. For fair comparison, hospitals whose submitted samples were too small were excluded and the cut-off was set as 15 cases available (12 hospitals, N = 680).
Variables of interest

The Functional Independence Measure (FIM™) is the most widely used functional assessment measure in the rehabilitation community, and it is regarded as most useful in assessing the patient progress during inpatient rehabilitation. The FIM is an 18-item, seven level ordinal scale. It is the product of an effort to resolve the long standing problem of lack of uniform measurement and data on disability and rehabilitation outcomes. The 18-item ordinal scale can be calculated into a summary score and two subscale scores of motor and cognitive function; the higher the score, the more independent a patient’s function. In this study, total FIM score was used to assess the patients’ functional recovery or as a proxy for patients’ daily living functions. Patient motor scale and cognitive scale were collected and analyzed at inpatient admission and discharge dates, respectively.

Prediction of Functional Independence Measure score at the time of discharge

There are two methods for calculating the improvement rate in FIM. The first is a method where the rate is calculated by taking the difference between the admission and discharge score and dividing that by the length of stay. This only provides a general sketch of how much improvement took place for any patients without considerations of individual patient characteristics. The second method involves taking the difference between the estimated scores based on individual patient characteristics obtained at admission, and the actual scores shown at discharge. This enables difference in treatment outcomes among hospitals to be compared on the same case mix basis, that is, between patients with the same conditions at the start, and this method was adopted for this study. The following is the applied calculation method for FIM scores at discharge.

Motor FIM scores and cognitive FIM scores at admission were respectively computed, controlling for variables affecting the outcomes regardless of rehabilitation care quality experienced by stroke patients, such as age, number of days from the stroke until admission, and modified Rankin Scale score before the stroke. Estimated scores were obtained from regression analysis with FIM score at discharge as a dependent variable and motor FIM score, cognitive FIM score at admission, age, number of days from the stroke until admission, and modified Rankin Scale score before stroke as independent variables (Table 1).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>The prediction equation for Functional Independence Measurement™ score at the time of discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable: FIM™ score at the time of discharge</td>
<td>B</td>
</tr>
<tr>
<td>Motor FIM at admission</td>
<td>0.432</td>
</tr>
<tr>
<td>Cognitive FIM at admission</td>
<td>0.439</td>
</tr>
<tr>
<td>Age</td>
<td>-0.093</td>
</tr>
<tr>
<td>Number of days from the onset of the stroke until admission</td>
<td>-0.068</td>
</tr>
<tr>
<td>Modified Rankin Scale score before stroke</td>
<td>-0.049</td>
</tr>
</tbody>
</table>

Notes: Estimated scores were obtained from regression analysis with FIM score at discharge as a dependent variable and motor FIM score, cognitive FIM score at admission, age, number of days from the stroke until admission, and modified Rankin Scale score before stroke, and diagnosis as independent variables. Diagnosis was excluded by regression model.

Abbreviations: P, probability; R², coefficient of determination; FIM, Functional Independence Measurement.

It was found that those variables used in the estimating equation had an accountability of about 60% with the coefficient of determination (R²) being 0.649. This suggested that the rest (35%) was accounted for by certain processes and practices involved in rehabilitation care service, which was the concerned area of this study in comparing differences in performance of different hospitals.

Therapeutic achievement

The therapeutic achievement in this study was defined as the difference between the predicted-discharge FIM score and the actual FIM score measured. We analyzed this difference with a 95% confidence interval (CI).

Statistical analysis

Evaluation of therapeutic achievement for each patient using estimated values from regression analysis

To evaluate the therapeutic achievement for each patient, the calculation of FIM gain was made by taking the difference between the actual FIM score and the estimated FIM score obtained from the equation in Table 1. The whole distribution of scores was divided into thirds to classify subject patients into three groups.

The range of therapeutic achievement for each patient was divided into three groups: good (a group whose prediction value was higher than its measurement value: over 6); fair (a group whose prediction value was supposed to be equal to its measurement value: from -4 up to 5); and poor (a group whose measurement value is below its prediction value: over 5). Each group’s distribution was examined.
Comparison between hospitals using cluster analysis
After confirming if there is any difference in the shape of distribution for FIM Gain Group 3 (good, fair, poor) between hospitals, high performing hospitals and low performing hospitals were respectively identified and accordingly assigned to each cluster in order to test what factors influence the difference. The group rated as “poor” in terms of FIM gain was used as a criterion for cluster analysis.

By carrying out cluster analysis with the characters of patient distribution in each hospital, this research classified hospitals according to performance into a high performance group or low performance group.

Results
Distribution of therapeutic achievements in each hospital
While focusing on the databank data (issued in September, 2009) in patients in recovery in the rehabilitation ward, this study adjusted for triage at admission and compared the therapeutic results of each hospital by using the data of 680 patients from 12 hospitals as described above.

By considering each patient’s conditions at admission, their FIM score would be likely to judge the improvement of ADL predicted at the time of discharge then compared against the measurement value at discharge. As a result, the FIM score showed a distribution range of 24.4 at a 95% CI of $-21.2$, with $-0.28$ on average (standard deviation [SD]: 12.88). On the basis of the results, the patients were divided into three different groups: “good”, “fair”, and “poor”. Each hospital’s distribution was examined. As a result, it was found that there were statistically significant differences in the ratio of distribution of “poor” among the hospitals (Table 2). Notably, while the group for the poor levels of the FIM score accounted for only 7.7% of patients in Hospital A, the ratio of the group for the poor levels in Hospital L was up to 60.5% of patients.

Analysis of factors affecting the difference in therapeutic achievements at each hospital
Although the FIM score was predicted at the time of discharge in the same conditions by adjusting the triage at admission, each hospital showed a variety of differences in distribution from the measurement value. Therefore, by using cluster analysis to analyze factors leading to the different distribution of therapeutic achievements in each hospital, the study divided all the hospitals into two different groups according to therapeutic achievement while retaining the actual condition of health care services provided by each hospital group (Figure 1).

Group one was a hospital group with superior quality of health care (high performance group) and the other group was a hospital group with lower quality of health care (low performance group). The following elements were all analyzed: differences in the FIM improvement score, FIM improvement rate per day, the ratio of discharged patients to home, length of stay, participation of physician registered as a rehabilitation specialist to JARM (Japanese Association of Rehabilitation Medicine), conference execution rate, amount of exercise (physical therapist [PT], occupational therapist [OT], and speech

<table>
<thead>
<tr>
<th>Hospital</th>
<th>N</th>
<th>Good (%)</th>
<th>Fair (%)</th>
<th>Poor (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>13</td>
<td>6 (46.2)</td>
<td>6 (46.2)</td>
<td>1 (7.7)</td>
</tr>
<tr>
<td>B</td>
<td>32</td>
<td>10 (31.3)</td>
<td>16 (50.0)</td>
<td>6 (18.8)</td>
</tr>
<tr>
<td>C</td>
<td>21</td>
<td>6 (28.6)</td>
<td>11 (52.4)</td>
<td>4 (19.0)</td>
</tr>
<tr>
<td>D</td>
<td>30</td>
<td>12 (40.0)</td>
<td>11 (36.7)</td>
<td>7 (23.3)</td>
</tr>
<tr>
<td>E</td>
<td>8</td>
<td>2 (25.0)</td>
<td>4 (50.0)</td>
<td>2 (25.0)</td>
</tr>
<tr>
<td>F</td>
<td>68</td>
<td>22 (32.4)</td>
<td>28 (41.2)</td>
<td>18 (26.5)</td>
</tr>
<tr>
<td>G</td>
<td>52</td>
<td>24 (46.2)</td>
<td>13 (25.0)</td>
<td>15 (28.8)</td>
</tr>
<tr>
<td>H</td>
<td>256</td>
<td>105 (41.0)</td>
<td>71 (27.7)</td>
<td>80 (31.3)</td>
</tr>
<tr>
<td>I</td>
<td>13</td>
<td>2 (15.4)</td>
<td>6 (46.2)</td>
<td>5 (38.5)</td>
</tr>
<tr>
<td>J</td>
<td>13</td>
<td>2 (15.4)</td>
<td>5 (38.5)</td>
<td>6 (46.2)</td>
</tr>
<tr>
<td>K</td>
<td>22</td>
<td>3 (13.6)</td>
<td>7 (31.8)</td>
<td>12 (54.5)</td>
</tr>
<tr>
<td>L</td>
<td>76</td>
<td>7 (9.2)</td>
<td>23 (30.3)</td>
<td>46 (60.5)</td>
</tr>
</tbody>
</table>

Notes: “Good” (a group whose prediction value was higher than its measurement value: over 6); “Fair” (a group whose prediction value was supposed to be equal to its measurement value: from 4 up to 5); “Poor” (a group whose measurement value is below its prediction value: over 5).
therapist [ST], in total) per day, self-exercise without a therapist, and exercise in wards. In all the items, except the ratio of discharged patients to home, there were statistically significant differences between those two groups (Table 3).

### Improvement in Functional Independence Measure score
In the FIM score, widely used as criteria for ADL, the high performance hospital group improved by 30.3% on average at the time of discharge than at admission, while the low performance hospital group improved by 19.0% on average. In the FIM improvement rate per day, the high performance hospitals group improved by 0.36%, while the low performance hospitals group improved by 0.17%.

### Ratio of discharged patients
The high performance hospital group was 80.5% while the low performance hospital group was 76.6%, which shows there was no statistically significant difference between those two groups.

### Average length of stay
The average length of stay for the high performance hospital group was 113.5 days, which was surprisingly longer compared with 86.4 days in the low performance hospital group.

### Participation of physicians registered as rehabilitation specialists to the Japanese Association of Rehabilitation Medicine
Sixty percent of rehabilitation specialists in the high performance hospital group participated compared with 31.9% in the low performance hospital group. The conference execution rate (more than once a week and by more than three types of specialists) showed 69.2% in the high performance hospital group carried out such an execution, while 29.9% in the low performance hospital group carried out such an execution.

### Total daily exercise
Measuring the total amount of exercise with the PT, OP, and ST per day showed that the high performance hospital group spent 82.9 minutes on average per day exercising.

### Self-exercise without therapists
In investigating the frequency of self-exercise without therapists, we found that 78.5% of the high performance hospital group carried out such exercises compared with 40.4% of the low performance hospital group.

### Exercise in wards
Measurement of exercise in the wards (compared to exercise in the rehabilitation care room), showed that 92.8% of the high performance hospital group carried out such exercises compared with 40.4% of the low performance hospital group.

---

**Table 3** Results of analyzing the differences between the high performance hospital group and the low performance hospital group

<table>
<thead>
<tr>
<th></th>
<th>High performance hospital group</th>
<th>Low performance hospital group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean SD [95% CI]</td>
<td>Mean SD [95% CI]</td>
<td></td>
</tr>
<tr>
<td><strong>FIM™ improvement</strong></td>
<td>30.3 points 15.0 [14.32–8.35]</td>
<td>19.0 points 18.2</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td><strong>FIM improved per day</strong></td>
<td>0.36 points 0.14 [-0.22–0.02]</td>
<td>0.17 points 0.23</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td><strong>The ratio of discharged patients to home</strong></td>
<td>80.5% 60% [18.2–32.3]</td>
<td>76.6% 31.9%</td>
<td>0.154</td>
</tr>
<tr>
<td><strong>Length of stay</strong></td>
<td>86.4 days 36.9 [18.2–32.3]</td>
<td>113.49 days 37.4</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td><strong>Participation of rehabilitation specialist</strong></td>
<td>60% 69.2%</td>
<td>31.9% 29.9%</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td><strong>Conference execution rate</strong></td>
<td>82.9 min 70.6 min</td>
<td>82.9 min 70.6 min</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td><strong>Self-exercise without therapists</strong></td>
<td>78.5% 82.9%</td>
<td>38.3% 40.4%</td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td><strong>Exercise in wards</strong></td>
<td>92.8% 82.9% [18.2–32.3]</td>
<td>40.4% 70.6 min [18.2–32.3]</td>
<td>P &lt; 0.001</td>
</tr>
</tbody>
</table>

Notes: The conference execution rate is based on the fact that people with more than three types of occupation took part in the conference more than once a week. However, the definition of the conference execution rate was not clear in the data provided before 2008.

Abbreviations: SD, standard deviation; CI, confidence interval; P, probability; FIM, Functional Independence Measure; OT, occupational therapist; PT, physical therapist; ST, speech therapist.
Discussion
In previous research on the effects of rehabilitation interventions, increased functional skills on admission to rehabilitation, early initiation of rehabilitation services, and rehabilitation in an interdisciplinary versus multidisciplinary setting have been shown to improve ADL.11 In this study we examined whether such factors differed between hospitals. In particular, we looked for factors that are important in Japan. We used a two-step method to find out whether there were differences between hospitals in the distributions of improvements in ADL. First, we used patients’ data to predict what FIM score would be at the time of discharge. Then, as the index of ADL improvement, we used the difference between the predicted and the actual FIM at the time of discharge. Next, we looked for differences between those hospitals with relatively greater ADL improvement and those with relatively less ADL improvement.

Although FIM score was predicted at the time of discharge in the same case mix conditions by adjusting for triage at admission, each hospital showed a variety of differences from the measurement value. As a result of analyzing which factors produced the differences among the hospitals, it was found that there was a difference in the pattern that each hospital provides health care services.

We examined each hospital’s structure and process of providing health care services represented by the participation of physicians registered as a rehabilitation specialists to JARM and the execution of conferences, and found that the hospital group with high performance showed participation and execution rates over two times more than in the hospitals group with low performance. This result confirms prior research, and emphasizes the importance of providing rehabilitative health care services by PT, OT, ST, nurses, and ward staff in addition to rehabilitation specialists (Stroke Unit).12–19

In the amount of exercise per day, self-exercise rates, and ward-exercise execution rates, it was found that the hospital group with high performance had more patients exercising than the other hospital group with low performance.

In Japan, all residents are provided with social insurance for medical expenses. For treatments including rehabilitation, PT, OT, and ST, the amount of time given to each treatment must be reported, and reimbursement is limited to the cost of 180 minutes per day. However, that limit was not set on the basis of any evidence. This study provides evidence of the effects of PT, OT, ST, exercise in wards, and self-exercise without therapists.

Patients exercised for about 12 to 13 minutes more per day in the hospital group with high performance than in the hospital group with low performance. This is a relatively large difference even when allowing that the criterion for health care treatment puts a limit on exercise per day at 180 minutes in Japan. While this data is not explored in this study, if the results of cluster analysis as shown in Figure 1 were divided into three groups, Hospital A would be in a class of its own. Hospital A was a medical organization that showed far superior therapeutic achievements to all the other hospitals. We asked Hospital A for an interview and found that since the hospital had not been open to the public for a long time, there were more health care staff than patients. Our interview found that regardless of the actual time period prescribed by health care criterion, the health care staff actively encouraged patients to take more exercise.

Considering that the limitation of exercise to 180 minutes per day in Japan is far from the standards introduced by the US stroke rehabilitation guidelines,19 we anticipate that improvements in results will be seen if hospitals increase the amount and frequency of exercise beyond the maximum limit required in Japan.

In the same context, the increases in self-exercise execution rates and ward-exercise execution rates are due to escalating therapeutic achievements that are possible. In addition, therapeutic achievements show the importance of daily exercise for patients rather than by separating the wards from physical therapy rooms and exercising only for rehabilitation. In order to promote and increase the total amount of patient exercise and so on in wards, the health care patterns in hospital teams needs to be systematized more concretely and realistically.

The present finding that more intense training was associated with greater improvement in ADL is consistent with the results previously reported.20–23 Meanwhile, there was no significant difference in the ratio of discharged patients to home between the hospital group with superior therapeutic achievements and the hospital group with inferior therapeutic achievements. A re-examination of the appropriateness of the ratio of discharged patients to home, one of the P4P assessment criteria applied to recovery rehabilitation wards presently in Japan, may show whether it is an appropriate criterion.

There is a high possibility that there may exist a number of confounding factors in addition to the endeavor of medical teams and the hospital system likely to affect patient decision-making in discharging from hospital to home. For
instance, when the conditions of patients were the same at the time of discharge, there was a 20.2% difference in the ratio of discharge to patients’ homes, depending on their different environments, such as patients’ domestic circumstance, and whether or not they could be nursed.

Limitations of the findings
The Rehabilitation Patient Databank used in this study went through seven updates over 3 years, with measures taken to ensure its reliability. Although there was a manual of data-entry procedures, we can not be absolutely sure that all of the physicians and therapists who entered data into the system used the same criteria in their evaluations. In addition, we studied variables related to rehabilitation therapy and ADL, but we were not able to control for medical findings other than those related to post-stroke rehabilitation. However, all of the hospitals that participated in this study actively cooperated with efforts to improve the quality of rehabilitation treatment in Japan. We divided the hospitals into high-performing and low-performing groups, and there may be some limit to the generalizability of these findings to all rehabilitation hospitals in Japan.

Conclusion
This study focused on patients in recovery in the rehabilitation ward using the data of 680 patients from 12 hospitals after adjusting for triage at admission obtained from the databank (issued in September, 2009) in Japan and compared the therapeutic results of each hospital. The comparison found that there were statistically significant differences in the results of therapeutic achievement in each hospital. In the hospital group with superior therapeutic achievements, there were greater amounts of exercise and a higher participation rate of rehabilitation specialists compared with the other hospital groups with inferior therapeutic achievements, which suggests the importance of frequent health care treatments and exercise.

Due to the assessments for health care quality and the publication of results conducted in a number of first world countries such as the US and the UK, it is expected that health care providers will put in voluntary effort to improve their future health care services. Nevertheless, prior to such effort of medical providers, the development of criteria to assess health care quality and to verify their validity is required.

Acknowledgments
This study was funded by Pfizer Health Research Foundation Japan (2008). The Rehabilitation Patients Data Bank of Japan was developed with financial support from the Health Labour Sciences Research Grant of Japan (2007–2009).

Disclosure
The authors report no conflicts of interest in this work.

References