Reliability of clinical competency evaluation list for novice physical and occupational therapists requiring assistance

YOSHIKIYO KANADA, RPT, PhD1, 2), HIROAKI SAKURAI, RPT, PhD1, 2)*, YOSHITO SUGIURA, RPT, MS3), YUDAI HIRANO, RPT, MS4), SOICHIRO KOYAMA, RPT, MS5), SHIGEO TANABE, RPT, PhD1, 2)

1) Fujita Health University School of Health Sciences: 1-98 Dengakugakubo, Kutsukake, Toyoake, Aichi 470-1192, Japan
2) Fujita Health University Graduate School of Health Sciences, Japan
3) Department of Rehabilitation, Health Care Service Facility for the Aged, Tobahouwaen, Japan
4) Tsujimura Surgical Hospital, Japan
5) Kawamura Hospital, Japan

Abstract. [Purpose] This study examined the reliability of a clinical ability evaluation table developed in a previous study with the aim of clarifying the abilities necessary for therapists to independently implement their duties. [Subjects and Methods] Forty-eight physical therapists with less than 2 years of clinical experience were targeted for evaluation, 48 main supervisors, and 48 sub-supervisors, 144 in total, were studied. [Results] The total score was lower when the evaluation was conducted by the target therapists themselves than when it was conducted by the main or sub-supervisors. Regarding the reliability of the total scores for the entire scale and each category, values representing the intra-rater reliability were higher when the evaluation was conducted by the target therapists or main supervisors, while there were marked differences between high and low values for each item. Regarding the inter-rater reliability, both the total scores for the entire scale and each category, as well as values for each item, were low. [Conclusion] Values representing the intra-rater reliability of the study table were low, indicating the necessity of further improvement.

Key words: Clinical competency, Evaluation, Reliability

INTRODUCTION

In Japan’s super-aging society, social demands for rehabilitation are increasing, and to meet these demands, the numbers of physical (PTs) and occupational (OTs) therapists have been rapidly increasing. The rapid increase in the number of therapists has consequently reduced years of experience of therapists in clinical environments, possibly leading to poorer-quality clinical services1). Under these circumstances, it is necessary to give more importance to post-graduate education in the workplace, in addition to improving school education systems. At present, post-graduate education is independently provided by associations, prefectural societies, or privately. Although some therapists mutually educate themselves in the workplace, it is generally difficult for them to establish organizations or systems to provide such education. Considering this a situation, school education should enable students to sufficiently learn about the fundamental items needed to achieve more specialized knowledge and skills after graduation, rather than the application of skills which should be focused on in post-graduation education. In short, systems are needed for teaching staff, clinical supervisors, and PTs and OTs in charge of post-graduation education to comprehensively provide standardized skills education before and after graduation2–6).

The majority of members of the Japanese Physical Therapy Association were young therapists in their twenties or thirties as of 20107), suggesting that insufficient provision of education due to a lack of supervising therapists may lead to decreases in the quality of therapists. In 2005, the minimum pre-graduation education goal specified in the Physical Therapy Education Guidelines was changed from becoming able to perform basic physical therapy to becoming able to perform basic physical therapy with some advice and supervision8). Furthermore, according to a survey which examined the status of clinical services provided by novice PTs immediately after graduation, therapists independently implement their duties only on limited occasions, and need supervisors’ advice7). Based on this, PTs immediately after certification possibly lack sufficient clinical abilities, and require post-graduation education, such as staff education in

*Corresponding author. Hiroaki Sakurai (E-mail: hsakurai@fujita-hu.ac.jp)

©2015 The Society of Physical Therapy Science. Published by IPEC Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (by-nc-nd) License <http://creativecommons.org/licenses/by-nc-nd/3.0/>.
the facilities they work in and opportunities to participate in training seminars.

In other professional areas, for example, in physician education, the provision of 2 year post-graduation clinical training has become compulsory, and the goals of such training have been determined. Post-graduation training goals and guidelines on supervision have also been established in nursing education. Based on the results of qualitative studies examining the behavioral goals for PTs graduating from training schools and those involving clinical supervisors, clinical evaluation scales have been developed in the United States. In some reports, it has been recommended that clinical evaluation scales have been developed in the United States. In particular, goal-setting is necessary for the continuous provision of specialized education. However, goals for therapists after certification have not yet been determined. For therapists after certification to achieve a certain level of competence, it is necessary for them to be able to independently perform their duties, and education guidelines and goals should be determined in consideration of this.

The authors previously conducted an interview and questionnaire-based study, involving those with experience of supervising other therapists to determine the abilities necessary to independently perform duties as a therapist identified as being necessary to perform professional duties and to develop a clinical ability evaluation table (evaluation table) covering those abilities which were extracted. The present study examined the reliability of the evaluation table, with the aim of effectively using it in the post-graduation education of therapists.

**SUBJECTS AND METHODS**

The evaluation table was developed, using the 55 items previously identified as the abilities necessary for therapists to independently perform their duties. To examine its reliability, therapists working in hospitals were asked to evaluate the clinical abilities of those under their supervision using it, adopting a 5-point rating method (total score: 0 to 220): 0=inappropriate assessment item; 1=professionally incompetent; 2=being able to understand and adopt appropriate actions to a certain extent under monitoring and supervision; 3=being able to accurately understand and adopt appropriate actions under monitoring and supervision; and 4=being able to accurately understand and adopt appropriate actions without supervision.

Therapists working in 3 medical facilities in the Tokai area, who consented to cooperate with the study, were studied. One and two of the study facilities specialized in acute and acute to post-acute care (including outpatient and visiting rehabilitation services), respectively. The 144 participants were categorized as follows: (1) 48 therapists targeted for evaluation who had less than 2 years of experience; (2) 48 main supervisors who were supervising the target therapists daily; and (3) 48 sub-supervisors who were assisting with their supervision. Both the main supervisors and sub-supervisors had 3 years or more experience. As ethical considerations, the participants were provided with oral and written explanations regarding the details of this study, and their participation in it was regarded as their consent.

The target therapists conducted self-evaluations, while the main supervisors and sub-supervisors objectively evaluated them, using the evaluation table. The target therapists and main supervisors were asked to repeat the evaluation twice at an interval of 1 week. Furthermore, the age, gender, and number of years after certification were examined as subjects’ characteristics.

For the analysis, the kappa coefficient was calculated to examine the intra-rater reliability between the target therapists and main supervisors, as well as the inter-rater reliability between the main and sub-supervisors for the 55 items. The intra-rater reliability between the target therapists and main supervisors based on the total scores of each category and the entire scale was examined using the intra-class correlation coefficient of: ICC (1, 1), while the inter-rater reliability between the main and sub-supervisors was calculated using ICC (2, 1). From the target therapists and main supervisors, data other than that related to intra-rater reliability were collected in the first evaluation. The significance level was chosen as 5%. SPSS Statistics Version 18 was used for the statistical analysis. This study was conducted with the approval of the Ethics Committee of Fujita Health University.

**RESULTS**

The number of years of experience was 1 or less for 28, and 2 or less for 20 of the target therapists, with a mean of 1.4 years. Years of experience among the main and sub-supervisors were 8.8 (range: 4 to 27), and 6.2 (3 to 13), respectively (Table 1). The intra-rater reliabilities for each of the 55 items based on the kappa coefficient are shown in Table 2. Ranges for the target therapists were 0.16 to 0.67 and for the main supervisors were 0.19 to 0.81. The inter-rater reliability between the main and sub-supervisors based on the kappa coefficient ranged from 0.00 to 0.32. The ICC (1, 1) representing the intra-rater reliability based on the total score was 0.81 and 0.85 for the target therapists and main supervisors, respectively. The ICC (2, 1) representing the inter-rater reliability between the main and sub-supervisors was 0.43. In both cases, clinical practice-related thoughts were less developed than basic attitudes and therapeutic skills.

**DISCUSSION**

To develop a clinical ability evaluation table, this study adopted a sequential exploratory strategy as part of a mixed method, in which qualitative and quantitative approaches are sequentially performed. The sequential exploratory strategy is mainly used to develop and examine measurement instruments. In our previous study, data were qualitatively collected, and a clinical ability evaluation table was developed based on the Delphi technique. The present study quantitatively analyzed the reliability of the evaluation table, and its findings are discussed below.

The relationship between the range of kappa coefficients and reliability has been reported as follows: 0.81 to 1.00, almost perfect; 0.61 to 0.80, substantial; and 0.41 to 0.60,
### Table 1. Subjects’ characteristics

| Items                        | Target therapists | Main supervisors | Sub-supervisors   |
|------------------------------|-------------------|------------------|-------------------|
| Gender                       | Male 18           | Male 8.8±4.7     | Male 6.2±2.7      |
|                             | Female 30         | (4–27)           | (3–13)            |
| Physical therapist           | 27                |                  |                   |
| Occupational therapist       | 21                |                  |                   |
| Age                          | 23.2±1.5          |                  |                   |
| The number of years of experience | 1.4±0.5     |                  |                   |
| First year                   | 28                |                  |                   |
| Second year                  | 20                |                  |                   |

### Table 2. Intra- and inter-rater reliability of a clinical ability evaluation table for therapists

| Evaluation items                                                                 | Among target therapists | Among main supervisors | Between main and sub-supervisors |
|----------------------------------------------------------------------------------|-------------------------|------------------------|---------------------------------|
| <Basic attitudes>                                                               |                         |                        |                                 |
| We are dressed appropriately as a member of society                             | 0.25                    | 0.54                   | 0.00                            |
| Using appropriate language as a member of society                               | 0.55                    | 0.70                   | 0.00                            |
| Adhering to appointed times and deadlines                                       | 0.58                    | 0.36                   | 0.00                            |
| Complying with rules in the workplace                                          | 0.22                    | 0.57                   | 0.28                            |
| Understanding the role and duties of the therapist as a team member             | 0.29                    | 0.52                   | 0.20                            |
| Adopting appropriate actions in consideration of the role of the therapist as a team member | 0.35                    | 0.81                   | 0.09                            |
| Contributing to the improvement of coordination as a team member                | 0.47                    | 0.62                   | 0.19                            |
| Efficiently performing duties to complete them within working hours             | 0.50                    | 0.49                   | 0.01                            |
| Appropriately understanding and considering confidentiality and personal information management | 0.43                    | 0.45                   | 0.00                            |
| Performing appropriate infection control measures (including washing hands)     | 0.24                    | 0.28                   | 0.01                            |
| Performing appropriate equipment management (before and after use)              | 0.31                    | 0.46                   | 0.00                            |
| Performing treatment, with a sense of responsibility                            | 0.49                    | 0.52                   | 0.23                            |
| Appropriately managing the therapist’s own physical condition and schedule and avoiding interference with his/her duties | 0.53                    | 0.19                   | 0.09                            |
| Appropriately implementing reporting, communication, and consultation procedures (developing and expressing the therapist’s own thoughts) in all times | 0.39                    | 0.44                   | 0.23                            |
| Identifying problems which are difficult to independently address              | 0.46                    | 0.62                   | 0.22                            |
| Consulting about problems which are difficult to independently address with appropriate persons in appropriate situations | 0.31                    | 0.34                   | 0.21                            |
| Seriously accepting and addressing issues noted by the supervisor or the therapist’s own failures | 0.48                    | 0.57                   | 0.21                            |
| Developing positive attitudes and making efforts to achieve knowledge and skills | 0.41                    | 0.42                   | 0.25                            |
| Performing treatment and duties based on learned outcomes and experience        | 0.38                    | 0.45                   | 0.24                            |
| <Therapeutic skills>                                                           |                         |                        |                                 |
| Adopting appropriate measures, such as life-saving techniques, to manage sudden changes in patients’ conditions | 0.19                    | 0.63                   | 0.05                            |
| Appropriately dealing with individual patients in consideration of their symptoms | 0.29                    | 0.48                   | 0.00                            |
| Using appropriate verbal or non-verbal communication methods for individual patients | 0.50                    | 0.56                   | 0.28                            |
| Showing empathy when communicating with patients in consideration of their psychological conditions | 0.19                    | 0.55                   | 0.26                            |
| Appropriately listening to patients and their families to clarify their needs   | 0.41                    | 0.47                   | 0.00                            |
| Having the medical knowledge necessary for a therapist                         | 0.53                    | 0.42                   | 0.22                            |
| Selecting appropriate evaluation items for individual patients                  | 0.38                    | 0.44                   | 0.10                            |
| Performing vital (blood pressure and heart rate) measurements, according to each situation | 0.35                    | 0.56                   | 0.05                            |
moderate\(^{19}\). ICC values of 0.8 or greater, 0.7 or greater, and 0.6 or greater are defined as “excellent”, “good” and “acceptable” respectively\(^{20}\). In the present study, the total scores for each of the 3 categories and the entire scale, representing the intra-rater reliability, were higher when the 55 items were self-evaluated by the target therapists or objectively evaluated by their main supervisors. However, the values markedly varied among the items, presumably due to some unclear expressions, resulting in differences in the interpretation of some items among raters.

The scores for each item and category, as well as the total score, representing the inter-rater reliability between the main and sub-supervisors, were low in all cases. This may be explained by the current facility system, in which a one-to-one relationship is established between target therapists and main supervisors, while sub-supervisors assisting with supervision observe the former only within a limited time frame, and are therefore less able to accurately evaluate their abilities. Similarly, values representing the inter-rater reliability were low, possibly due to the selection of main and sub-supervisors, focusing on their supervising positions and years of experience, resulting in variations in the quality of supervision and rating among them. Both the ICC (1, 1) representing the intra-rater reliability between self- and

### Table 2. Continued

| Evaluation items                                                                 | Among target therapists | Among main supervisors | Between main and sub-supervisors |
|---------------------------------------------------------------------------------|-------------------------|------------------------|---------------------------------|
| Appropriately (and also accurately, efficiently) conducting medical interviews with patients | 0.51                    | 0.48                   | 0.06                            |
| Appropriately (and also accurately, efficiently) examining reflexes             | 0.57                    | 0.73                   | 0.12                            |
| Appropriately (and also accurately, efficiently) conducting orthopedic examination | 0.51                    | 0.73                   | 0.11                            |
| Appropriately (and also accurately, efficiently) evaluating pain                | 0.39                    | 0.75                   | 0.14                            |
| Appropriately (and also accurately, efficiently) evaluating coordination        | 0.55                    | 0.62                   | 0.25                            |
| Appropriately (and also accurately, efficiently) evaluating muscle tone         | 0.60                    | 0.51                   | 0.28                            |
| Appropriately (and also accurately, efficiently) measuring the range of motion   | 0.65                    | 0.60                   | 0.00                            |
| Appropriately (and also accurately, efficiently) evaluating the muscle strength  | 0.46                    | 0.59                   | 0.09                            |
| Appropriately (and also accurately, efficiently) conducting sensory examination  | 0.29                    | 0.66                   | 0.00                            |
| Appropriately (and also accurately, efficiently) evaluating the motor function of patients with paralysis (using the SIAS and Brunnstrom Stage Test) | 0.67                    | 0.65                   | 0.32                            |
| Appropriately (and also accurately and efficiently) evaluating ADL (using instruments, such as the FIM and Barthel index) | 0.29                    | 0.70                   | 0.23                            |

<Clinical practice-related thoughts>

Clarifying individual patients’ general characteristics
Identifying individual patients’ possible risks based on the results of examination
Logically examining the causes of problems in movements or activities of daily living
Developing treatment programs to achieve goals (also referring to literature)
 Safely implementing treatment programs
 Safely handling treatment devices
 Appropriately managing risks related to medical accidents, such as tube removal and bleeding
 Appropriately managing risks related to falls
 Providing appropriate range-of-motion training
 Providing appropriate muscle-strengthening training
 Providing appropriate assistance and guidance for the maintenance of sitting positions
 Providing appropriate assistance and guidance for the maintenance of standing positions
 Providing appropriate assistance and guidance for standing from a seat
 Providing appropriate assistance and guidance for transfer
 Providing appropriate assistance and guidance for gait training
 Continuously evaluating (and observing) patients in the progress of treatment

Criteria:
Ratings
4=being able to accurately understand and adopt appropriate actions without supervision
3=being able to accurately understand and adopt appropriate actions under monitoring and supervision
2=being able to understand and adopt appropriate actions to a certain extent under monitoring and supervision
1=professionally incompetent
0=inappropriate assessment item
objective evaluations by the target therapists and main supervisors, respectively, and ICC (2, 1) representing the inter-rater reliability between objective evaluations by the main and sub-supervisors were lower for clinical practice-related thoughts than those for basic attitudes and therapeutic skills, indicating the necessity of observing target therapists in more detail to address this category, as it regards the process of thought performed in actual clinical situations, covering diverse conditions.

When focusing on the selection of evaluation items, the developed clinical ability evaluation table is likely to have a certain level of content validity. Therefore, the study items may be appropriate goals for novice therapists and effectively usable as part of post-graduation education. However, considering the low inter-rater reliability values, it may be desirable to conduct evaluations using among multiple raters. If time-dependent changes in novice therapists could be clarified, without differences in ratings among raters, the usability of the evaluation table would be further enhanced. Therefore, the improvement of the evaluation table’s inter-rater reliability is an issue which needs to be addressed. In line with this, it may be necessary to revise expressions in the table, make organizational and systematic arrangements, and establish appropriate methods for enabling sub-supervisors to perform supervision similar to main supervisors, rather than only assisting them.

As a follow up to this study, methods to effectively use the evaluation table for post-graduation education will be examined to determine education guidelines and goals for therapists to become able to independently perform their duties, supervise novice therapists, and achieve knowledge and skills covering specific domains, such as specialized or certified therapies.

REFERENCES

1) Yoshino J, Futawatari T, Otani K, et al.: Qualitative study of the primary goals of continuing education in physical therapy. J Jpn Phys Ther Assoc, 2010, 37: 410–416.
2) Kanada Y, Sakurai H, Sugiura Y, et al.: Standardizing the assessment of the clinical abilities of physical therapists and occupational therapists using OSCE. J Phys Ther Sci, 2012, 24: 985–989. [CrossRef]
3) Sakurai H, Kanada Y, Sugiura Y, et al.: Standardization of clinical competence evaluation in the education of physical therapists and occupational therapists—establishment of an OSCE compliant education system. J Phys Ther Sci, 2013, 25: 101–107. [CrossRef]
4) Sakurai H, Kanada Y, Sugiura Y, et al.: Standardization of clinical skill evaluation in physical/occupational therapist education: effects of introduction of an education system using OSCE. J Phys Ther Sci, 2013, 25: 1071–1077. [Medline] [CrossRef]
5) Sakurai H, Kanada Y, Sugiura Y, et al.: Reliability of the OSCE for physical and occupational therapists. J Phys Ther Sci, 2014, 26: 1147–1152. [Medline] [CrossRef]
6) Sakurai H, Kanada Y, Sugiura Y, et al.: Standardization of clinical skill evaluation in physical/occupational therapist education: OSCE-based clinical skill evaluation for physical and occupational therapists. J Phys Ther Sci, 2014, 26: 1387–1397. [Medline] [CrossRef]
7) Japanese Physical Therapy Association: Physical Therapy White Paper. Tokyo: Shinohara publication, 2010.
8) Japanese Physical Therapy Association: A Guide to Clinical Training Education, 5th ed. Tokyo: Japanese Physical Therapy Association, 2005.
9) Ministry of Health, Labour, and Welfare: Core Curriculum for Postgraduate Clinical Training. http://www.mhlw.go.jp/shingi/2002/06/0627-3f.html (Accessed Jan. 8, 2014)
10) Ministry of Health, Labour, and Welfare: Novice Nurse Training Guidelines. http://www.mhlw.go.jp/bunya/iryou/ushirase/dl/100210-3.pdf (Accessed Jan. 8, 2014)
11) Jette DU, Bertoni A, Coots R, et al.: Clinical instructors’ perceptions of behaviors that comprise entry-level clinical performance in physical therapist students: a qualitative study. Phys Ther, 2007, 87: 833–843. [Medline] [CrossRef]
12) Fitzgerald LM, Delitto A, Irrgang JJ: Validation of the clinical internship evaluation tool. Phys Ther, 2007, 87: 844–860. [Medline] [CrossRef]
13) Japan Society for Medical Education: Medical Education White Paper. Tokyo: Shinohara publication, 2006.
14) Rauscher L, Greenfield BH: Advancements in contemporary physical therapy research: use of mixed methods designs. Phys Ther, 2009, 89: 91–100. [Medline] [CrossRef]
15) Creswell JW: Study Designs—Qualitative, Quantitative, and Mixed Methods. Tokyo: Japan Nursing Association publication society, 2007, pp 235–255.
16) Kerlinger FN: Foundations of Behavioral Research, 3rd ed. New York: Rinehart and Winston, 1986, pp 16–17.
17) Sumson T: The delphi technique: an adaptive research tool. Br J Occup Ther, 1998, 61: 153–156. [CrossRef]
18) Hammell KW: Qualitative Research in Evidence-Based Rehabilitation. New York: Churchill Livingstone, 2004, pp 136–137.
19) Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics, 1977, 33: 159–174. [Medline] [CrossRef]
20) Kuwabara Y, Saito T, Inagaki Y: [Evaluation of intra- and inter-observer reliability]. Kokyu To Junkan, 1993, 41: 945–952. [Medline]