改変版日本看護協会転倒転落リスクアセスメントツールの予測妥当性：後方視的コホート研究

The predictive validity of a modified Japanese Nursing Association fall risk assessment tool: A retrospective cohort study

Abstract

Background: Patient falls are the most common nursing care-related adverse event in hospitals. Extensive literature has been published on the predictive validity of fall risk assessment tools; however, there have been no studies examining the changes in predictive validity at different observation periods among hospital inpatients.

Objectives: To examine the predictive validity of a modified Japanese Nursing Association fall risk assessment tool and to compare its predictive validity at observation periods of 7, 14, 21, and 28 days.

Design: Retrospective cohort design.

Settings: Twelve wards of a 600-bed university hospital in Japan.

Participants: Patients 15 years and older admitted over a six-month period were enrolled. Patients were excluded if they were admitted to the intensive care unit or neuropsychiatry ward, had no fall risk assessment results within two days of admission, or had inconsistent assessment results.

Methods: Falls were observed for 28 days following admission. Predictive validity was evaluated using the area under the receiver operating curve, sensitivity, specificity, and positive and negative likelihood ratios at 7-, 14-, 21-, and 28-day observation points. Faller prevalence in each observation sample was adjusted for consistency using a bootstrap sampling method. All predictive validity indices were then recalculated and compared.

Results: A total of 4,144 patients were admitted and 67 patients fell (1.6% faller prevalence) within 28 days of admission. The modified Japanese Nursing Association fall risk assessment tool showed a sensitivity of 0.82, specificity of 0.71, positive likelihood ratio of 2.83, and negative likelihood ratio of 0.26 at a cut-point of ≥6, and the area under the receiver operating curve was 0.83. Predictive validity in the 7-day observation sample was significantly higher than the 14- and 28-day samples, but no significant difference was found relative to the 21-day observation sample.

Conclusions: The modified Japanese Nursing Association fall risk assessment tool demonstrated good predictive validity in a Japanese university hospital, but further evaluation is needed for other validity values and reliability. The findings from this study may indicate that predictive validity indices vary by the length of observation period and faller prevalence, but these findings need to be examined further.

Key Words: Accidental falls, Inpatients, Japan, Nursing assessment, Reproducibility of results, Risk assessment

1. INTRODUCTION

Falls account for approximately 40% of nursing care-related adverse events among hospital patients (D'Amour et al., 2014). While 16–34% of inpatient fallers suffer injuries, 1.5–3.9% of hospital falls result in fractures, intracranial hemorrhages, or even death (Schwendimann, Bühler et al., 2006; Waters et al., 2013).

A recently published guideline by the National Institute for Health and Care Excellence (NICE; 2013) included the following as risk factors for falls among hospitalized elderly patients: cognitive impairment, continence problems, fall history, unsuitable or missing footwear, medications, postural instability, mobility and/or balance problems, syncope syndrome, and visual impairment. Interventions targeting multiple risk factors identified by a fall risk assessment were shown to reduce the fall rate by 31% in hospital inpatients (Cameron et al., 2012). Chari et al. (2013) also reported that patients who were assessed for fall risks were 40% less likely to have fractures caused by falls compared to the patients whose fall risks were not assessed.

Fall risk assessment tools commonly used and evaluated in multiple hospitals are the St. Thomas Risk Assessment Tool in Falling Elderly Inpatients (STRATIFY; Oliver et al., 1997), the Morse Fall Scale (Morse et al., 1989), and the Hendrich II Fall Risk Model (Hendrich et al., 2003). Among these three tools, a meta-analysis by Aranda-Gallardo et al. (2013) found that the STRATIFY showed the greatest validity in acute hospital settings with sensitivity, specificity, positive likelihood ratio (LR), and negative LR of 0.80, 0.68, 2.47, and 0.33, respectively.

Unfortunately, none of the three tools is commonly used in Japanese hospitals. Furthermore, STRATIFY showed less than optimal validity (sensitivity of 0.65–0.68 and specificity of 0.75, and the area under the receiver operating curve [AUC] of 0.75–0.77) in a Japanese university hospital (Toyabe, 2010). In other Asian countries, the Morse Fall Scale (Chow et al, 2007; Kim et al., 2007; Kim et al., 2011), STRATIFY (Kim et al., 2007), and the Hendrich II Fall Risk Model (Kim et al., 2007) were examined either at acute care or rehabilitation hospitals, but no tool has exhibited both sensitivity and specificity of ≥0.70, which are the predictive validity criteria for fall risk assessment tools in clinical practice suggested by Oliver et al. (2004). To explain the low predictive validity in the study, Chow et al. (2007) discussed the possibility of differences in fall risks between Western and Asian populations. Since no published study has examined this difference, the low predictive validity of common fall risk assessment tools suggests the need to develop a tool for patients in Asian countries. Although the Japanese Nursing Association fall risk assessment tool was developed to address risk factors for Japanese patients and was modified for use at a Japanese university hospital, the tool contains multiple risk factors that are likely to apply to Korean inpatients (Kim et al., 2011). Modifications also included items related to treatment stage, patients' personality, and experiences in the hospital environment. These modified factors are not included in STRATIFY, the Morse Fall Scale, or the Hendrich II Fall Risk Model and could be common risk factors for patients in other Asian countries.

Previous systematic reviews and meta-analyses examining the validity of fall risk assessment tools have included studies with various lengths of observation periods for falls. The range was 10 days to two months in Oliver et al. (2008) and 6.8 to 14.6 days in Aranda-Gallardo et al. (2013). However, two recently published studies reported that predictive validity changed with longer observation periods (Bentzen et al., 2011; Duncan et al., 2012). A prospective cohort study by Bentzen et al. (2011) compared the predictive validity of three fall risk assessment methods (modified STRATIFY, staff judgment, and fall history) at 30, 90, and 180 days from assessment among residents in 18 nursing homes in Norway. In all three methods, sensitivity decreased while specificity increased during longer observation periods. Duncan et al. (2012) also prospectively examined the predictive validity of four balance tests at six and 12 months from an assessment of community dwellers with Parkinson's disease and found that predictive validity at six months was better than that at 12 months. Although hospital patients' fall risk factors would change more frequently than nursing home residents or community dwellers, no study has examined the predictive validity of fall risk assessment tools at different observation periods in a hospital setting.

The objectives of this study were to examine the predictive validity of the modified Japanese Nursing Association fall risk assessment tool currently used in a Japanese university hospital and to compare its predictive validity during 7-, 14-, 21-, and 28-day observation periods. The validation of the modified tool will enable other hospitals to adopt this improved assessment tool.

2. MATERIALS AND METHODS

2.1. Study design and setting

This retrospective cohort study was conducted in a 600-bed university hospital in Japan. The hospital is accredited by the Japanese Council for Quality Health Care. Twelve wards with 13 clinical specialties (internal medicine, surgery, neurosurgery, orthopedics, dermatology, urology, otorhinolaryngology, ophthalmology, radiology, anesthesiology, oral and maxillofacial surgery, obstetrics/gynecology, and pediatrics) were included. The intensive care unit and neuropsychiatry ward were excluded from the analyses because their patient-to-nurse ratio differed from the 7:1 ratio in the selected wards.

2.2. Patients

Patients 15 years and older admitted to the hospital over a six-month period (April 1, 2010–September 30, 2010) were included in the study. Patients' age criterion of ≥15 years was chosen based on the World Health Organization's (WHO) age classification of adults (2008). Specifically, this range covered the two adult age groups with high morbidity associated with falls, namely, ≥65 years and 15–29 years (WHO, 2012). Figure 1 depicts inclusion and exclusion criteria and the number of patients included in the analyses at 7, 14, 21, and 28 days of observation. Fall risk assessment results taken after the third day of admission were excluded to examine the predictive validity of the tool administered on admission. Patients were also excluded if they had inconsistent results in the fall risk assessment such as a difference in age on the administrative record and the item "age," chose "infant's developmental stage" although the patients were 15 years and older, or chose both "up ad lib" and "bedridden" and/or "requiring mobility assistance."

Most studies examining the validity of fall risk assessment tools include patients with different lengths of stay, and patients with no reported falls during hospitalization are considered non-fallers. This could result in potential bias because lengths of fall observation differ by patient. To reduce this risk of information bias, this study excluded patients who were discharged before the evaluation period. A total of 1,069 patients had <7 days of hospitalization and were excluded from the analysis at 7 days. Patients with <14, 21, and 28 days of hospitalization were also excluded from the analyses at 14, 21, and 28 days, respectively. Thus, 2,197 patients were included for analysis at 7 days, 1,249 patients at 14 days, 876 patients at 21 days, and 650 patients at 28 days.

2.3. Instrument

The fall risk assessment tool used in the study (Appendix 1) was a modified version of the

Japanese Nursing Association fall risk assessment tool first introduced in 1999 that continues to be recommended by the Japanese Nursing Association (1999, 2007; Appendix 2). The original version consisted of 32 items, including age, sex, history of falls and/or syncope, sensory impairment, functional disability, mobility, cognition, medications, and elimination. The Japanese Nursing Association fall risk assessment tool is most commonly used in Japanese hospitals, but its predictive validity has not been well evaluated. The modified tool consisted of 35 items including age, fall history, sensory impairment, motor functions, mobility, cognition, medications, elimination, treatment stage, personality, and environment. The items "9 years old or younger" and "infant's developmental stage (rolling over, crawling, etc.)" were not used for the analyses. The fall risk scores ranged from 0 to 21. Higher scores indicated a higher risk of falling. The hospital's risk management committee modified the tool based on submitted incident reports and clinical expertise. The modified tool had been used in the hospital since 2000. All of the hospital nurses were oriented to the use of this tool upon employment. This is the first time the modified Japanese Nursing Association tool has been statistically analyzed to examine its validity.

2.4. Fall definition and identification

A fall was defined as "inadvertently coming to rest on the ground, floor or other lower level, excluding intentional change in position to rest in furniture, wall or other objects" (WHO, 2007, p. 1). The researcher reviewed descriptions of the fall events in the incident reports to ensure that reported events matched the definition. A faller was defined as a hospital inpatient with a reported fall at least once during the 28 days following admission.

2.5. Data collection

Inpatient administrative records, fall risk assessment results, and fall incident reports were combined to create a database for analyses. Information on age, sex, admission and discharge dates, and clinical specialty were obtained from the inpatient administrative records. The patients' fall risks were evaluated by registered nurses on the ward within two days of admission and recorded in an electronic charting system. Hospital staff electronically submitted incident reports as part of their typical duties, but they were not aware of the fall definition used in this study.

2.6. Statistical analysis

Nominal variables were described using frequencies and proportions. Distributions of numeric variables were examined using histograms and the Kolmogorov-Sminov test. Since age, length of hospital stay, and fall risk assessment score showed skewed distributions, these were described using median, range, or first and third inter-quartiles.

Patient characteristics for observation periods were examined using Pearson's chi-square test for nominal variables and the Kruskal-Wallis test for numeric variables with skewed distributions. The total assessment scores of the modified Japanese Nursing Association tool were compared between fallers and non-fallers at 28 days of admission using the Mann-Whitney test. Additionally, each item of the modified Japanese Nursing Association tool was examined using Pearson's chi-square test with Yate's correction applied when expected frequencies were <5 in more than 20% of the cells (Peacock & Peacock, 2011).

The receiver operating characteristic curve (ROC), AUC, sensitivity, specificity, and positive and negative LRs were used to examine the predictive validity of the fall risk assessment tool. The AUC equals 1.0 for perfect accuracy and 0.5 for a tool that is equivalent to randomly selecting high or low risk of falling (Hanley & McNeil, 1982). Sensitivity refers to the number of fallers who were correctly identified by the tool divided by the total number of fallers.

Specificity refers to the number of non-fallers who were correctly identified by the tool divided by the total number of non-fallers (Peacock & Peacock, 2011). Values of >0.7 for sensitivity and specificity were considered to have enough accuracy for a fall assessment tool (Oliver et al., 2004). Positive LR was defined as sensitivity / (1 - specificity), whereas negative LR was defined as (1 - sensitivity) / specificity (Simel et al., 1991). Higher positive LRs and lower negative LRs were interpreted as having better accuracy. For these analyses, patients were observed for falls during the period of 7, 14, 21, and 28 days of hospitalization, respectively. This 7-day increment was chosen as the time span for a periodic reassessment, scheduled every 7 days during hospitalization. An optimal cut-off point to calculate sensitivity, specificity, and LRs was determined to meet both sensitivity and specificity greater than 0.7 (Oliver et al., 2004), with more importance placed on sensitivity than specificity for the entire sample of 3,266 patients.

Several researchers have claimed that the prevalence of target condition (e.g., presence of disease) influences sensitivity and specificity (Brenner & Gefeller, 1997; Leeflang et al, 2009, 2013; Li & Fine, 2011). Because LRs, ROC, and AUC are calculated using sensitivity and specificity, these values can also be influenced by prevalence (Brenner & Gefeller, 1997; Li & Fine, 2011). Thus, this study attempted to adjust faller prevalence among samples with different observation periods using a bootstrap sampling method. Bootstrap sampling entails random sampling with replacement (StataCorp, 2013), stratified by fall occurrence. Faller prevalence in the study population was 1.6% (67 / 4,144), so bootstrap sampling was completed for each observation period that had a faller prevalence of 1.6% (16 fallers and 984 non-fallers). Because the numbers of non-fallers was less than 984 in the 21-day and 28-day observation samples, the non-fallers of these two periods were first duplicated and then bootstrap samples were obtained. Predictive validity values were also calculated using these bootstrap samples. Finally, differences in AUCs with the bootstrap samples of 1,000 patients between observation periods were compared by Pearson's chi-square test.

All statistical analyses were conducted using SPSS version 19 (IBM Corp, Armonk, NY), MedCalc Statistical Software version 13.0 (MedCalc Software, Ostend, Belgium), and Stata/MP12.1 (Stata Corporation, College Station, TX). *P*-values of less than 0.05 were considered statistically significant. Associated 95% confidence intervals (CIs) were also calculated wherever applicable.

2.7. Ethical considerations

The Ethics Committee for Epidemiological Research of the researcher's university approved the study and it was conducted consistent with the Declaration of Helsinki (World Medical Association, 2008) and the Ethical Guidelines for Epidemiologic Studies by Japan's Ministry of Health, Labour and Welfare and Ministry of Education, Culture, Sports, Science and Technology (2008). Because all study data were collected from existing hospital records, the need for informed consent was waived. Consistent with the Japanese guidelines, a notification of the research, including the study objectives and contents, was posted to the bulletin boards in inpatient wards and outpatient areas to disclose research information.

3. RESULTS

3.1. Patient characteristics

Patients' characteristics are shown in Table 1. A total of 4,144 patients (median age 60 years, 50.7% female) were admitted during the six-month period of April to September 2010. A total of 67 patients (1.6%) fell within 28 days of admission. Fallers were significantly older (median 72 years, range 29–91 vs. median of 60 years, range 15–97, $p < 0.001$) and had longer hospital stays (median 37 days, range 4–411 days vs. median of 9 days, range 1–355 days, $p < 0.001$) than non-fallers. Fall rates (fallers/inpatients) were highest in orthopedics (3.4%), followed by dermatology (2.8%), and internal medicine (2.6%).

Patients without fall risk assessment data collected within two days of admission (n = 701) and patients with inaccurate assessment data (n = 177) were excluded from further analyses (Figure 1). Excluded patients (n = 878) included 47.7% females, had a median age of 64 years (interquartile range 52–74 years), and a median length of hospitalization of 9 days (interquartile range 5–21 days).

In the comparison of observation periods, there was a significant difference in age (p < 0.001) and length of stay (p < 0.001). While internal medicine, surgeries, and obstetrics/gynecology were the three largest clinical specialties included in the 7-day observation, orthopedics replaced obstetrics/gynecology in the other three observation periods. There were 18, 27, 30, and 32 fallers in the 7-, 14-, 21-, and 28-day observation samples, respectively.

3.2. Assessment results of the modified Japanese Nursing Association tool

Assessment results of the modified Japanese Nursing Association tool among 3,266 patients and comparisons between fallers and non-fallers are shown in Table 2. From the modified Japanese Nursing Association tool, 21 items showed significantly larger proportions among fallers compared to non-fallers. Use of hypnotics and/or tranquilizers and diuretics was higher among fallers than non-fallers, but were not statistically significant. “Up ad lib” was the only item with a significantly larger proportion among non-fallers than fallers. Comparing the assessment results by observation periods, the longer observation periods exhibited significantly higher scores and larger proportions in 19 items and showed significantly smaller proportions in “up ad lib” (Supplementary Table 1).

3.3. Predictive validity of the modified Japanese Nursing Association tool and

determination of cut-off point

Using 3,266 patients (49 fallers) across observation periods, two cut-off points showed both sensitivity and specificity greater than 0.7 (sensitivity 0.82, specificity 0.71 at ≥6; sensitivity 0.71, specificity 0.79 at ≥7) (Supplementary Table 2; Supplementary Figure 1). Since greater importance was placed on sensitivity than specificity, a cut-off point of ≥6 was chosen to calculate predictive values in each observation period. At this cut-off point, positive LR was 2.83 and negative LR was 0.26. The AUC was 0.83 (95% CI: 0.79–0.87).

3.4. Predictive validity in the 7-, 14-, 21-, and 28-day observation samples

Predictive validity indices in the original samples and bootstrap samples with consistent faller prevalence by observation periods are shown in Table 3. In original samples, sensitivity, specificity, and LRs exhibited higher accuracy in the 7-day observation sample, and AUCs were lower when the sample observation periods became longer. Using bootstrap samples adjusted for faller prevalence, ROC curves and AUCs showed greater accuracy in the 7-day bootstrap sample than other observation periods (Figure 2). The AUC of the 7-day bootstrap sample was significantly greater than AUCs in 14-day ($p = 0.04$) and 28-day ($p = 0.002$) bootstrap samples, but did not differ significantly in the 21-day bootstrap sample ($p = 0.08$).

4. DISCUSSION

4.1. Predictive validity of the modified Japanese Nursing Association fall risk assessment tool

This study was the first to examine the predictive validity of the modified Japanese Nursing Association fall risk assessment tool; this tool was found to be adequate for use in clinical practice according to Oliver et al.'s (2004) criteria. Although the predictive validity of commonly used fall risk assessment tools were not examined in this study, the sensitivity of

0.82 and specificity of 0.71 found for the modified Japanese Nursing Association tool are higher than for the STRATIFY (sensitivity 0.80, specificity 0.68), the Morse Fall Scale (sensitivity 0.76, specificity 0.68), and the Hendrich II Fall Risk Model (sensitivity 0.63, specificity 0.64) among acute hospital patients in Asia, North America, Europe, and Australia as reported in the meta-analysis of studies conducted by Aranda-Gallardo et al. (2013). Faller prevalence of the studies included in the meta-analysis were 0.9–16.3% for STRATIFY (Barker et al., 2011; Milisen et al., 2007; Papaioannou et al., 2004; Vassallo et al., 2005; Walsh et al., 2011), 2.1–19.9% for the Morse Fall Scale (Chapman et al., 2011; Kim et al., 2011; Schwendimann, De Geest et al., 2006, 2007), and 1.1–7.8% for the Hendrich II Fall Risk Model (Chapman et al., 2011; Ivziku et al., 2011; Kim et al., 2007; Lovallo et al., 2010). The predictive validity is also higher than the predictive validity of the STRATIFY, as evaluated in another university hospital in Japan with a faller prevalence of 3.4% (Toyabe, 2010). All studies, except those by Milisen et al. (2007; 0.9% in surgical wards) and Kim et al. (2007; 1.1%), reported higher faller prevalence than the current study. Assuming that higher prevalence increases sensitivity and decreases specificity as described by Brenner and Gefeller (1997), sensitivity of the current study remains higher, but specificity could be lower than the above-mentioned studies. However, how the prevalence influences sensitivity and specificity was reported to be different between studies (Leeflang et al., 2009). One of the explanations for why the modified Japanese Nursing Association tool resulted in higher validity was that it consisted of more comprehensive fall risk factors than the other three tools (STRATIFY: five items, Morse: six items, Hendrich II: eight items). Another explanation may be that fall risk assessment tools have been reported to show the highest validity in the setting where the tool was developed (Oliver et al., 2004). The studies included in the meta-analysis

(Aranda-Gallardo et al., 2013) and Toyabe's study (2010) were conducted in external facilities.

From the modified Japanese Nursing Association tool, 21 items showed significantly larger proportions among fallers compared to non-fallers. Most of the items were common to the risk factors identified by NICE (2013) guidelines such as fall history, visual impairment, postural instability, mobility/balance problems, cognitive problems, medications (analgesics, laxatives, antihypertensives), and continence problems. Antihypertensive medications, commode chair use, decreased strength, and personality traits were not included in the original Japanese Nursing Association tool, but showed significant association with fallers. Patients' personality traits were also not included in the STRATIFY (Oliver et al., 1997), the Morse Fall Scale (Morse et al., 1989), the Hendrich II Fall Risk Model (Hendrich et al., 2003), or the NICE guidelines. Although the subjectivity of personality assessment on the modified Japanese Nursing Association tool could be a concern, assessing patient personality could be a new approach to fall prevention. A multi-site study by Meade et al. (2006) reported an approximate 50% reduction in patient falls with hourly nursing rounds. While the authors did not extend their discussion to the reasons for this reduction, it is possible that this approach was effective with patients who were hesitant to use a call light for assistance. Since "up ad lib" was the only item of which non-fallers exhibited a significantly larger proportion, its removal from the tool should be considered.

Use of hypnotics and/or tranquilizers is a well-known risk factor for falls (Bloch et al., 2011; Woolcott et al., 2009), but no significant association was found in this study. A possible explanation for this difference could be that this study did not include hypnotics or sedatives administered after the initial fall risk assessment. Prevalence of hypnotics or sedative use in this study (15.6%) was low compared to a reported prevalence of 27.6% in a multicenter study at 44 general hospitals in Japan (Enomoto et al., 2010).

In clinical practice, the modified Japanese Nursing Association tool may be useful in similar practice settings and among those with similar patient characteristics, but it should not solely be used to predict falls during hospitalization. Instead, it should be used to assess patients' fall risks and trigger nurses' actions for fall prevention as recommended by NICE (2013). Specifically, the guideline states that all patients 65 years and above and patients aged 50–64 years who, based on clinical judgment, are thought to have a higher risk for falls, should be assessed for modifiable risk factors. Regarding clinical judgment, a systematic review and meta-analysis by Haines et al. (2007) and two studies published after the review article (Milisen et al., 2012; Webster et al., 2010) reported that nurses' clinical judgment was equivalent to commonly used tools such as the STRATIFY or the Morse Fall Scale. However, one area of concern is that nurses' background information such as years of nursing experience was not described in these studies. How nurses' education or experience influences their clinical judgment on patients' fall risk has not been studied. Additional studies would be required to conclude that nurses would be able to perform fall risk assessments equivalent to a tool consisting of fall risk factors.

According to the root cause analyses of voluntary, reported, and sentinel events to the Joint Commission (2014), assessment was the most frequently reported root cause of fall-related events resulting in patient death or permanent loss of function, followed by leadership and communication. To guide nurses' clinical judgments and actions, the use of a tool to assess patients' fall risk would benefit patients, especially in hospitals with nurses with a different skill mix. Neither the modified Japanese Nursing Association tool nor other commonly used fall risk assessment tools provide interventions linked to the fall risks. A prospective multicenter cohort study conducted in the Netherlands by Van Gaal et al. (2014) reported that, although 19% of the patients were identified as "high risk" using the STRATIFY, no high risk patients received adequate fall preventive care. Adding recommended interventions for each risk factor assessed for patients may trigger nurses' actions regarding identified fall risk factors.

4.2. Influence of the length of observation and faller prevalence on predictive validity

In the original and bootstrap samples, the 7-day observation samples showed higher accuracy than the 14-, 21-, and 28-day observation samples in terms of sensitivity, specificity, positive LR, negative LR, and AUC. Statistically, the AUC of the 7-day bootstrap sample was significantly higher than the 14-day and 28-day bootstrap samples, although the difference was insignificant in the 21-day bootstrap sample. Currently, the researcher is unaware of published studies examining the predictive validity of fall risk assessment tools in different observation periods in a hospital setting; however, there are two studies conducted with nursing home residents (Bentzen et al., 2011) and community-dwellers with Parkinson's disease (Duncan et al., 2012). Consistent with the current study, both studies reported that overall predictive validity was lower in samples with longer observation periods, although faller prevalence was not adjusted in these studies. Differences between the present study and these others were that specificity and the positive LR were lower among samples with longer observation periods in this study and the study by Duncan et al. (2012), but they were higher in the study by Bentzen et al (2011). This may be because of differences in patient characteristics that most of the patients in the current study were acutely ill on admission. Consequently, their fall risk factors may have been resolved in tandem with their recovery. Furthermore, Duncan et al. (2012) compared longer observation periods (six-month and 12-month) than Bentzen et al. (2011; 30

days, 90 days, 180 days), so changes in individuals' fall risk (balance) may have influenced accuracy. Conversely, the patients in Bentzen's (2011) study were nursing home residents and their risk factors may not have changed greatly over the 180-day observation period.

In this study, prevalence of fallers in each observation sample was adjusted to have a consistent proportion using a bootstrap sampling method, and predictive validity indices were recalculated. In the 7-day, 14-day, and 21-day observation samples, sensitivities and negative LRs changed as described by Brenner and Gefeller (1997), but specificity and positive LRs changed in the opposite direction or did not change. In the 28-day observation sample, only positive LR changed as depicted by Brenner and Gefeller (1997). This difference might be attributed to the differences in patient characteristics between samples. The influence of prevalence on predictive validity indices has been reported to vary between studies (Leeflang et al., 2009).

As a whole, these results indicate that the predictive validity indices of fall risk assessment tools may differ not only by the study settings but also by faller prevalence and observation periods. Further studies are needed to confirm that the findings are observed using other fall risk assessment tools and other hospital settings, but researchers and clinicians may need to consider faller prevalence and the length of observation as the factors influencing predictive validity values. In addition, researchers may want to determine observation periods for falls according to their patients' characteristics for more accurate evaluation of the predictive validity of fall risk assessment tools.

4.3 Study limitations

This study has a number of limitations. First, it was conducted only in a single hospital in Japan; therefore, the findings cannot be generalized to other patient groups or settings. Future studies should include multiple hospitals to validate the assessment tool. Second, selection bias may be present, as excluded patients were significantly older than the patients included for the analyses. Underrepresentation of older patients may have influenced the association between some items of the tool and fallers. Third, the possibility of under-reporting falls should be considered for the following reasons: the falls were identified only by incident reports and the hospital staff was unaware of the definition of falls used in this study. When using incident reports alone to identify hospital falls, 28–40% of falls are missed (Hill et al., 2010; Shorr et al., 2008). Under-reporting falls could result in underestimation of sensitivity and overestimation of specificity, thus influencing other predicative validity values accordingly. It is recommended that future research should use multiple sources to identify fall occurrences, such as combining chart review and incident reports (Hill et al., 2010; Shorr et al., 2008). Moreover, prospective studies will allow researchers to inform hospital staff of their definition of falls prior to the study. Furthermore, inter-observer reliability on fall identification ought to be tested after informing staff of the fall definition. Fourth, this study only examined the predictive validity of the tool and did not evaluate other aspects such as content, concurrent, and convergent validities. Additionally, inter-rater and intra-rater reliability of the tool were not tested. Because multiple raters were used in this study to assess patient fall risks and some items pertaining to elimination, personality, and environment require subjective evaluation, the possibility of measurement errors must be considered. Consequently, further research is needed to evaluate the validity and reliability of the tool. Finally, this study examined neither the Japanese Nursing Association tool nor the modified Japanese Nursing Association tool to the fullest extent. Specifically, two items of the modified Japanese Nursing Association tool intended to assess children's fall risks were excluded from analyses. As a result, further research is needed that includes all items of the tool and patients aged <15 years.

5. CONCLUSIONS

The modified Japanese Nursing Association fall risk assessment tool showed good predictive validity in the Japanese university hospital. Predictive validity in the 7-day observation sample was significantly higher than 14- and 28-day samples, but no significant difference was found for the 21-day observation sample. The influence of faller prevalence and observation period on the predictive validity of fall risk assessment tools requires further investigation.

ACKNOWLEDGEMENT

The author would like to thank Ms. Takako Shimoji for coordinating data collection and providing information on the development of the modified fall risk assessment tool and it use in the hospital.

References

Aranda-Gallardo, M., Morales-Asencio, J.M., Canca-Sanchez, J.C., Barrero-Sojo, S., Perez-Jimenez, C., Morales-Fernandez, A., de Luna-Rodriguez, M.E., Moya-Suarez A.B., Mora-Banderas, A.M., 2013. Instruments for assessing the risk of falls in acute hospitalized patients: a systematic review and meta-analysis. BMC Health Serv. Res., 13(1), 122–137. doi: 10.1186/1472-6963-13-122

Barker, A., Kamar, J., Graco, M., Lawlor, V., Hill, K., 2011. Adding value to the STRATIFY falls risk assessment in acute hospitals. J. Adv. Nurs., 67(2), 450–457. doi: 10.1111/j.1365-2648.2010.05503.x

Bentzen, H., Bergland, A., Forsén, L., 2011. Diagnostic accuracy of three types of fall risk methods for predicting falls in nursing homes. Aging Clin. Exp. Res., 23(3), 187–195.

Bloch, F., Thibaud, M., Dugué, B., Brèque, C., Rigaud, A., Kemoun, G., 2011. Psychotropic drugs and falls in the elderly people: updated literature review and meta-analysis. J. Aging Health, 23(2), 329–346. doi: 10.1177/0898264310381277

Brenner, H., Gefeller, O., 1997. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. Stat. Med., 16(9), 981–991.

Cameron, I.D., Gillespie, L.D., Robertson, M.C., Murray, G.R., Hill, K.D., Cumming, R.G., Kerse, N., 2012. Interventions for preventing falls in older people in care facilities and hospitals. Cochrane Database Syst. Rev., 12: CD005465. doi: 10.1002/14651858.CD005465.pub3.

Chapman, J., Bachand, D., Hyrkäs, K., 2011. Testing the sensitivity, specificity and feasibility of four falls risk assessment tools in a clinical setting. J. Nurs. Manag., 19(1), 133–142. doi: 10.1111/j.1365-2834.2010.01218.x

Chari, S., McRae, P., Varghese, P., Ferrar, K., Haines, T.P., 2013. Predictors of fracture from falls reported in hospital and residential care facilities: a cross-sectional study. BMJ Open, 3(8), e002948. doi: 10.1136/bmjopen-2013-002948

Chow, S.K.Y., Lai, C.K.Y., Wong, T.K.S., Suen, L.K.P., Kong, S.K.F., Chan, C.K., Wong I.Y., 2007. Evaluation of the Morse Fall Scale: applicability in Chinese hospital populations. Int. J. Nurs. Stud., 44(4), 556–565.

D'Amour, D., Dubois, C.A., Tchouaket, E., Clarke, S., Blais, R., 2014. The occurrence of adverse events potentially attributable to nursing care in medical units: cross sectional record review. Int. J. Nurs. Stud., 51(6), 882–891. doi: 10.1016/j.ijnurstu.2013.10.017

Duncan, R.P., Leddy, A.L., Cavanaugh, J.T., Dibble, L.E., Ellis, T.D., Ford, M.P., Foreman, K.B., Earhart, G.M., 2012. Accuracy of fall prediction in Parkinson's disease: six-month and 12-month prospective analyses. Parkinson's Disease, 2012, 1–7. doi: 10.1155/2012/237673

Enomoto, M., Tsutsui, T., Higashino, S., Otaga, M., Higuchi, S., Aritake, S., Hida, A., Tamura, M., Matsuura, M., Kaneita, Y., Takahashi, K., Mishima, K., 2010. Sleep-related problems and use of hypnotics in inpatients of acute hospital wards. Gen. Hosp. Psychiatry, 32(3), 276–283. doi: 10.1016/j.genhosppsych.2010.01.015

Haines, T.P., Hill, K., Walsh, W., Osborne, R., 2007. Design-related bias in hospital fall risk screening tool predictive accuracy evaluations: systematic review and meta-analysis. J. Gerontol. A Biol. Sci. Med. Sci., 62(6), 664–672.

Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology, 143(1), 29–36.

Hendrich, A.L., Bender, P.S., Nyhuis, A., 2003. Validation of the Hendrich II Fall Risk Model: a large concurrent case/control study of hospitalized patients. Appl. Nurs. Res., 16(1), 9–21.

Hill, A., Hoffmann, T., Hill, K.D., Oliver, D., Beer, C., Mcphail, S., Brauer, S., Haines, T.P., 2010. Measuring falls events in acute hospitals—A comparison of three reporting methods to identify missing data in the hospital reporting system. J. Amer. Geriatric Soc., 58(7), 1347–1352. doi: 10.1111/j.1532-5415.2010.02856.x

Ivziku, D., Matarese, M., Pedone, C., 2011. Predictive validity of the Hendrich fall risk model II in an acute geriatric unit. Int. J. Nurs. Stud., 48(4), 468–474. doi: 10.1016/j.ijnurstu.2010.09.002

Japanese Nursing Association, 1999. Systematic approach to prevent medical errors: risk management guideline for nurse managers. https://direct.nurse.or.jp/jna_system/guidline/riskmanagement/index.asp [In Japanese, Access limited to Japanese Nursing Association members]. [Accessed 5/11/2011].

Japanese Nursing Association, 2007. Standard of nursing practice. Japanese Nursing Association Press, Tokyo. [In Japanese].

The Joint Commission, 2014. Sentinel event data — Root causes by event type (2004-2013). http://www.jointcommission.org/sentinel_event.aspx [Accessed 5/8/2014].

Kim, E.A., Mordiffi, S.Z., Bee, W.H., Devi, K., Evans, D., 2007. Evaluation of three fall-risk assessment tools in an acute care setting, J. Adv. Nurs., 60(4), 427–435.

Kim, K.S., Kim, J.A., Choi, Y.K., Kim, Y.J., Park, M.H., Kim, H.Y., Song, M.S., 2011. A comparative study on the validity of fall risk assessment scales in Korean hospitals. Asian Nurs. Res., 5(1), 28–37. doi: 10.1016/S1976-1317(11)60011-X

Leeflang, M.M., Bossuyt, P.M., Irwig, L. 2009. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. J. Clin. Epidemiol., 62(1), 5–12. doi: 10.1016/j.clinepi.2008.04.007

Leeflang, M.M., Rutjes, A.W., Reitsma, J.B., Hooft, L., Bossuyt, P.M., 2013. Variation of a test's sensitivity and specificity with disease prevalence. Canad. Med. Assoc. J., 185(11), E537–544. doi: 10.1503/cmaj.121286

Li, J., Fine, J.P., 2011. Assessing the dependence of sensitivity and specificity on prevalence in meta-analysis. Biostat., 12(4), 710–722. doi: 10.1093/biostatistics/kxr008

Lovallo, C., Rolandi, S., Rossetti, A.M., Lusignani, M., 2010. Accidental falls in hospital inpatients: evaluation of sensitivity and specificity of two risk assessment tools. J. Adv. Nurs., 66(3), 690–696. doi: 10.1111/j.1365-2648.2009.05231.x

Meade, C.M., Bursell, A.L., Ketelsen, L., 2006. Effects of nursing rounds on patients' call light use, satisfaction, and safety. Am. J. Nurs., 106(9), 58–70.

Milisen, K., Coussement, J., Flamaing, J., Vlaeyen, E., Schwendimann, R., Dejaeger, E., Surmont, K., Boonen, S., 2012. Fall prediction according to nurses' clinical judgment: differences between medical, surgical, and geriatric wards. J. Am. Geriatr. Soc., 60(6), 1115–1121. doi: 10.1111/j.1532-5415.2012.03957.x

Milisen, K., Staelens, N., Schwendimann, R., De Paepe, L., Verhaeghe, J., Braes, T., Boonen, S., Pelemans, W., Kressig, R. W., Dejaeger, E., 2007. Fall prediction in inpatients by bedside nurses using the St. Thomas's Risk Assessment Tool in Falling Elderly Inpatients (STRATIFY) instrument: a multicenter study. J. Am. Geriatr. Soc., 55(5), 725–733.

Ministry of Health, Labour and Welfare, Ministry of Education, Culture, Sports, Science and Technology, 2008. Ethical guidelines for epidemiologic studies. http://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/ekigaku/0504sisin.html [In Japanese]. [Accessed 5/9/2011].

Morse, J.M., Morse, R.M., Tylko, S.J., 1989. Development of a scale to identify the fall-prone patient. Canad. J. Aging, 8(4), 366–377. doi: 10.1017/S0714980800008576

National Institute for Health and Care Excellence, 2013. Falls: assessment and prevention of falls in older people. NICE clinical guideline, 161. http://www.nice.org.uk/guidance/cg161 [Accessed 6/27/2013].

Oliver, D., Britton, M., Seed, P., Martin, F.C., Hopper, A.H., 1997. Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: case-control and cohort studies. BMJ, 315(7115), 1049–1053.

Oliver, D., Daly, F., Martin, F.C., McMurdo, M.E., 2004. Risk factors and risk assessment tools for falls in hospital in-patients: a systematic review. Age Aging, 33(2), 122–130.

Oliver, D., Papaioannou, A., Giangregorio, L., Thabane, L., Reizgys, K., Foster, G., 2008. A systematic review and meta-analysis of studies using the STRATIFY tool for prediction of falls in hospital patients: how well does it work? Age Aging, 37(6), 621–627. doi: 10.1093/ageing/afn203

Papaioannou, A., Parkinson, W., Cook, R., Ferko, N., Coker, E., Adachi, J.D., 2004. Prediction of falls using a risk assessment tool in the acute care setting. BMC Med., 2,1. doi: 10.1186/1741-7015-2-1

Peacock, J.L., Peacock, P.J., 2011. Oxford handbook of medical statistics. Oxford University Press, New York.

Schwendimann, R., Bühler, H., De Geest, S., Milisen, K., 2006. Falls and consequent injuries in hospitalized patients: effects of an interdisciplinary falls prevention program. BMC Health Serv. Res., 6(1), 69–76. doi: 10.1186/1472-6963-6-69

Schwendimann, R., De Geest, S., Milisen, K., 2006. Evaluation of the Morse Fall Scale in hospitalized patients. Age Ageing, 35(3), 311–313. doi: 10.1093/ageing/afh066

Schwendimann, R., De Geest, S., Milisen, K., 2007. Screening older patients at risk for falling during hospitalization. Int. J. Inj. Contr. Saf. Promot., 14(1), 64–65. doi: 10.1080/17457300600775536

Shorr, R.I., Mion, L.C., Chandler, A.M., Rosenblatt, L.C., Lynch, D., Kessler, L.A., 2008. Improving the capture of fall events in hospitals: combining a service for evaluating inpatient falls with an incident report system. J. Am. Geriatr. Soc., 56(4), 701–704.

Simel, D.L., Samsa, G.P., Matchar, D.B., 1991. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. J. Clin. Epidemiol., 44(8), 763–770.

StataCorp., 2013. Stata 13 Base Reference Manual. Stata Press, College Station, Texas.

Toyabe, S., 2010. World Health Organization fracture risk assessment tool in the assessment of fractures after falls in hospital. BMC Health Serv. Res., 10(1), 106–118. doi: 10.1186/1472-6963-10-106

Van Gaal, B.G., Schoonhoven, L., Mintjes-de Groot, J.A., Defloor, T., Habets, H., Voss, A., van Achterberg, T., Koopmans, R.T., 2014. Concurrent incidence of adverse events in hospitals and nursing homes. J. Nurs. Scholarsh., 46(3), 187–198. doi: 10.1111/jnu.12063

Vassallo, M., Stockdale, R., Sharma, J.C., Briggs, R., Allen, S., 2005. A comparative study of the use of four fall risk assessment tools on acute medical wards. J. Am. Geriatr. Soc., 53(6), 1034–1038. doi: 10.1111/j.1532-5415.2005.53316.x

Walsh, W., Hill, K.D., Bennell, K., Vu, M., Haines, T.P., 2011. Local adaptation and evaluation of a falls risk prevention approach in acute hospitals. Int. J. Qual. Health Care, 23(2), 134–141. doi: 10.1093/intqhc/mzq075

Waters, T.M., Chandler, A.M., Mion, L.C., Daniels, M.J., Kessler, L.A., Miller, S.T., Shorr, R.I., 2013. Use of International Classification of Disease, Ninth Revision, Clinical

Modification, codes to identify inpatient fall-related injuries. J. Am. Geriatr. Soc., 61(12), 2186–2191. doi: 10.1111/jgs.12539

Webster, J., Courtney, M., Marsh, N., Gale, C., Abbott, B., Mackenzie-Ross, A., McRae, P., 2010. The STRATIFY tool and clinical judgment were poor predictors of falling in an acute hospital setting. J. Clin. Epidemiol., 63(1), 109–113. doi: 10.1016/j.clinepi.2009.02.003

Woolcott, J.C., Richardson, K.J., Wiens, M.O., Patel, B., Marin, J., Khan, K.M., Marra, C.A., 2009. Meta-analysis of the impact of 9 medication classes on falls in elderly persons. Arch. Intern. Med., 169(21), 1952–1960. doi: 10.1001.archinternmed.2009.357

World Health Organization, 2007. WHO global report on falls prevention in older age. http://www.who.int/ageing/publications/Falls_prevention7March.pdf [Accessed 5/8/2011].

World Health Organization, 2008. The global burden of disease: 2004 update. http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf [Accessed 5/8/2011].

World Health Organization, 2012. Fact sheet No. 344: falls. http://www.who.int/mediacentre/factsheets/fs344/en/ [Accessed 2/12/2014].

World Medical Association, 2008. Declaration of Helsinki: Ethical principles for medical research involving human subjects. http://www.wma.net/en/30publications/10policies/b3/17c.pdf [Accessed 5/9/2011]

Fig. 1

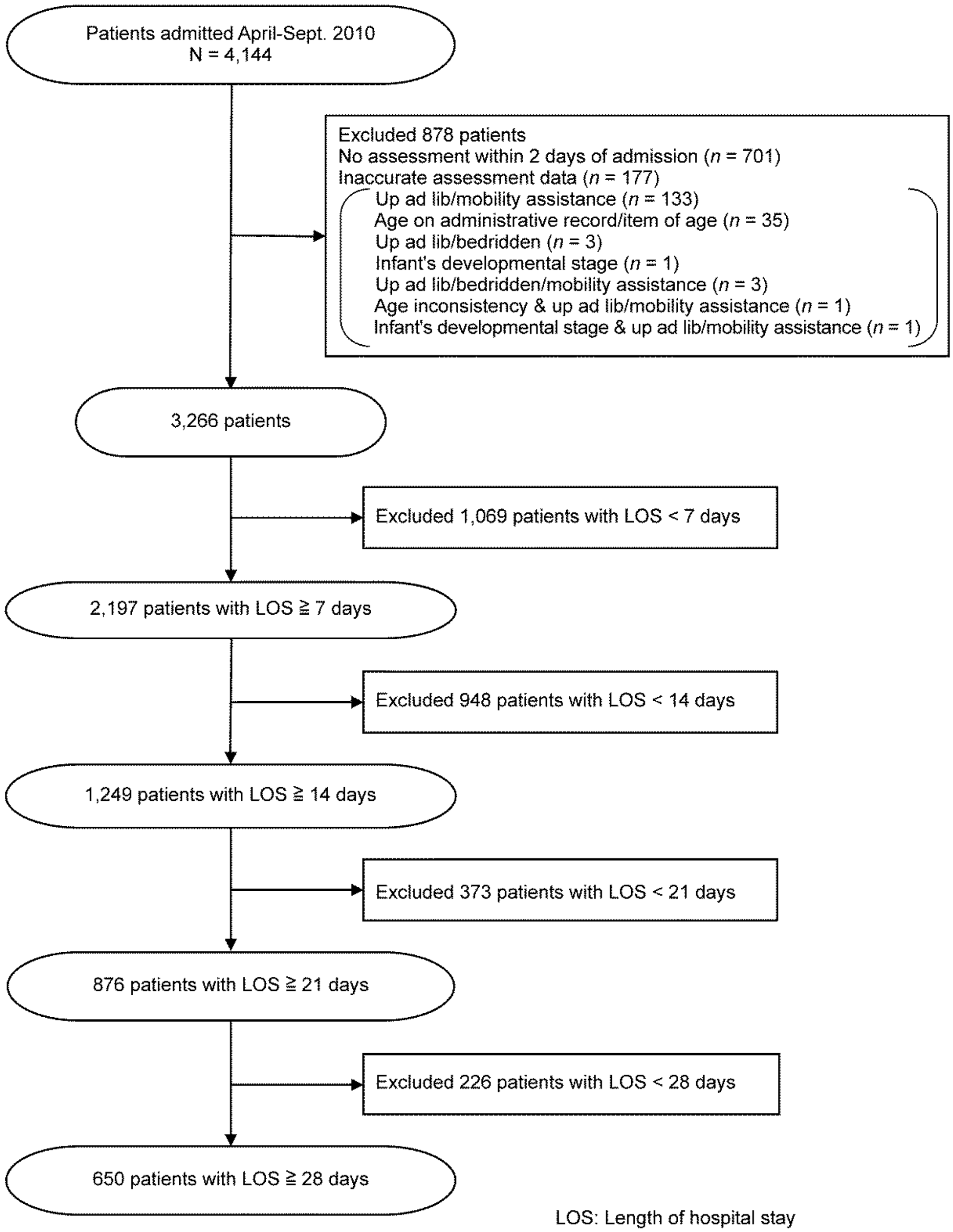


Figure 1. Flowchart of patient selection

Fig. 2

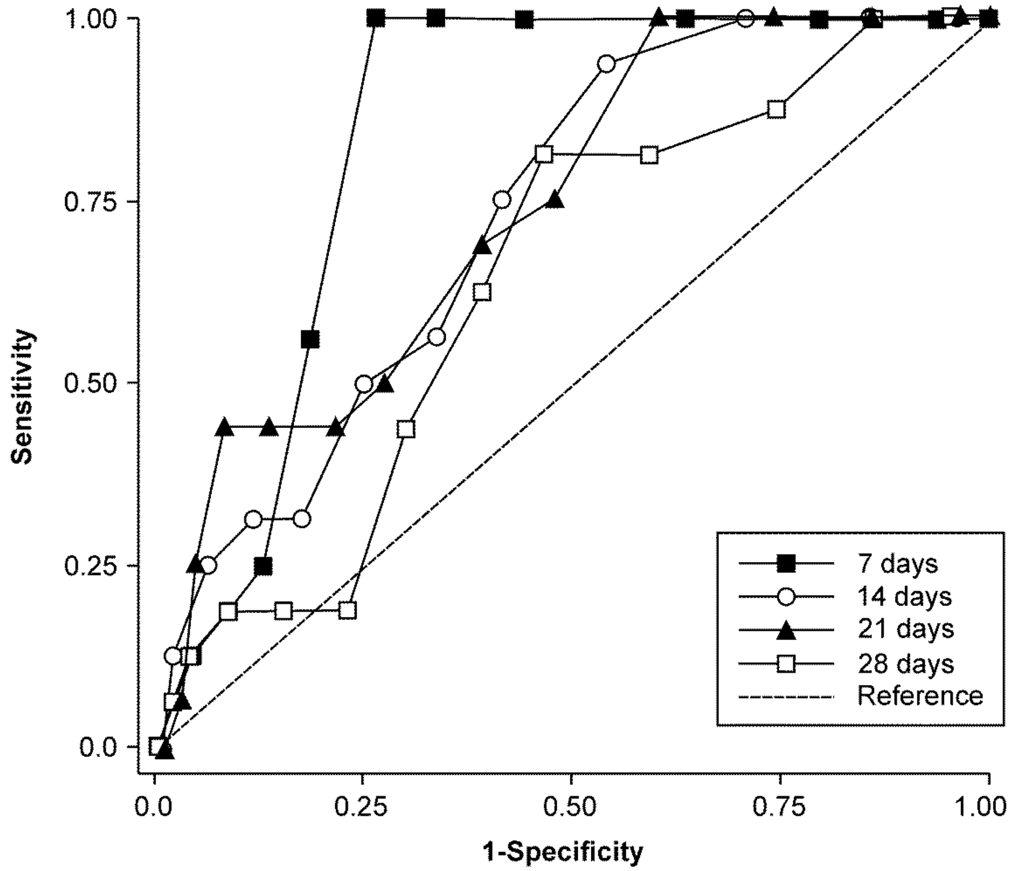


Figure 2. Receiver operating characteristics (ROC) curves of bootstrap samples of 1,000 patients by 7-, 14-, 21-, and 28-day observation periods. Area under the ROC curves (AUCs) and 95% confidence intervals were 0.84 (0.80–0.88), 0.73 (0.63–0.83), 0.74 (0.63–0.84), and 0.64 (0.52–0.76) in 7-, 14-, 21-, and 28-day bootstrap samples, respectively.

Note: Differences in AUCs with the bootstrap samples between observation periods by Pearson's chi-square test: 7 and 14 days ($p = 0.04$), 7 and 21 days ($p = 0.08$), 7 and 28 days ($p = 0.002$), 14 and 21 days ($p = 0.90$), and 21 and 28 days ($p = 0.22$).

Table 1. Patient characteristics

Characteristics	Total (*n* = 4,144)	Patients after exclusion (*n* = 3,266)	Observation periods				*p*-value
			7-day (*n* = 2197)	14-day (*n* = 1249)	21-day (*n* = 876)	28-day (*n* = 650)	
Age (years), median (Q1, Q3)	60 (43, 72)	59 (40, 71)	61 (43, 72)	62 (49, 73)	63 (50, 74)	64 (50, 74)	<0.001[a]
Age 15–29 years, *n* (%)	406 (9.8)	357 (10.9)	218 (9.9)	95 (7.6)	61 (7.0)	49 (7.5)	<0.001[b]
Age 30–49 years, *n* (%)	924 (22.3)	781 (23.9)	476 (21.7)	231 (18.5)	151 (17.2)	106 (16.3)	
Age 50–64 years, *n* (%)	1,119 (27.0)	862 (26.4)	589 (26.8)	347 (27.8)	239 (27.3)	172 (26.5)	
Age 65+ years, *n* (%)	1,695 (40.9)	1,266 (38.8)	914 (41.6)	576 (46.1)	425 (48.5)	323 (49.7)	
Female, *n* (%)	2,099 (50.7)	1,680 (51.4)	1,087 (49.5)	582 (46.6)	400 (45.7)	293 (45.1)	0.09[b]
Median length of stay (days), (Q1, Q3)	10 (5, 22)	10 (5, 22)	15 (9, 32)	28 (19, 49)	39 (27, 61)	47 (36, 71)	<0.001[a]
Clinical specialty, *n* (%)							
Internal Medicine	784 (18.9)	609 (18.6)	459 (20.9)	329 (26.3)	250(28.5)	176 (27.1)	
Surgeries	732 (17.7)	494 (15.1)	366 (16.7)	233 (18.7)	164 (18.7)	118 (18.2)	
Ophthalmology	596 (14.4)	373 (11.4)	209 (9.5)	63 (5.0)	19 (2.2)	10 (1.5)	
Obstetrics/Gynecology	585 (14.1)	551 (16.9)	289 (13.2)	108 (8.6)	77 (8.8)	60 (9.2)	
Urology	308 (7.4)	270 (8.3)	166 (7.6)	70 (5.6)	38 (4.3)	24 (3.7)	
Otorhinolaryngology	298 (7.2)	251 (7.7)	198 (9.0)	88 (7.0)	63 (7.2)	49 (7.5)	
Orthopedics	234 (5.6)	199 (6.1)	170 (7.7)	137 (11.0)	110 (12.6)	82 (12.6)	
Dermatology	180 (4.3)	157 (4.8)	104 (4.7)	68 (5.4)	43 (4.9)	32 (4.9)	
Maxillofacial Surgery	155 (3.7)	131 (4.0)	95 (4.3)	45 (3.6)	30 (3.4)	25 (3.8)	

Neurosurgery	146 (3.5)	126 (3.9)	85 (3.9)	63 (5.0)	41 (4.7)	34 (5.2)
Radiology	70 (1.4)	58 (1.8)	40 (1.8)	36 (2.9)	34 (3.9)	34 (5.2)
Pediatrics	40 (1.0)	37 (1.1)	14 (0.6)	9 (0.7)	7 (0.8)	6 (0.9)
Anesthesiology	16 (0.4)	10 (0.3)	2 (0.1)	0 (0.0)	0 (0.0)	0 (0.0)

Q1: first quartile, Q3: third quartile
[a]Kruskal-Wallis test, [b]Pearson's χ^2 test

Table 2. Modified Japanese Nursing Association fall risk assessment tool and falls within 28 days of admission

Variables	Total (*n* = 3,266)	28 days of admission: Fallers (*n* = 49)	28 days of admission: Non-fallers (*n* = 3217)	*p*-value
Fall risk assessment score, median (Q1, Q3)	4 (2, 6)	8 (6, 11)	4 (2, 6)	< 0.001[a]
Fall risk assessment tool items, *n* (%)				
Age ≧ 65, *n* (%)	1,266 (38.8)	28 (57.1)	1,238 (38.5)	0.008
Fall history within 1 year	314 (9.6)	18 (36.7)	296 (9.2)	< 0.001
Visual impairment	319 (9.8)	12 (24.5)	307 (9.5)	0.001
Hearing impairment	220 (6.7)	6 (12.2)	214 (6.7)	0.21
Motor functions				
Impaired extremities	518 (15.9)	18 (36.7)	500 (15.5)	< 0.001
Bone/joint problems	202 (6.2)	9 (18.4)	193 (6.0)	0.001
Muscle weakness	733 (22.4)	27 (55.1)	706 (21.9)	< 0.001
Mobility				
Up ad lib	2,592 (79.4)	21 (42.9)	2,571 (79.9)	< 0.001
Unstable when standing/walking	486 (14.9)	21 (42.9)	465 (14.5)	< 0.001
Use of mobility assistive devices	477 (14.6)	23 (46.9)	454 (14.1)	< 0.001
Requiring mobility assistance	409 (12.5)	20 (40.8)	389 (12.1)	< 0.001
Bedridden but able to move extremities	93 (2.8)	4 (8.2)	89 (2.8)	0.07
Cast, IV lines, or tubes	577 (17.7)	7 (14.3)	570 (17.7)	0.53
Cognition				
Feeling restless	424 (13.0)	9 (18.4)	415 (12.9)	0.26
Forgetful	273 (8.4)	13 (26.5)	260 (8.1)	< 0.001
Impaired judgment/understanding	163 (5.0)	10 (20.4)	153 (4.8)	< 0.001
Unable to use a call light	80 (2.4)	3 (6.1)	77 (2.4)	0.23
Medication				
Analgesics	417 (12.8)	19 (38.8)	398 (12.4)	< 0.001
Laxatives	219 (6.7)	9 (18.4)	210 (6.5)	0.003
Diuretics	247 (7.6)	7 (14.3)	240 (7.5)	0.13
Chemotherapy	122 (3.7)	2 (4.1)	120 (3.7)	1.00
Antiparkinsonian	20 (0.6)	1 (2.0)	19 (0.6)	0.71
Hypnotics/tranquilizers	511 (15.6)	12 (24.5)	499 (15.5)	0.09
Antihypertensive	994 (30.4)	22 (44.9)	972 (30.2)	0.03
Elimination				
Toileting ≥2x per night	1,176 (36.0)	32 (65.3)	1,144 (35.6)	< 0.001
Urinary/bowel incontinence	148 (4.5)	7 (14.3)	141 (4.4)	0.003
Requiring toileting assistance	316 (9.7)	11 (22.4)	305 (9.5)	0.005
Commode chair use	74 (2.3)	6 (12.2)	68 (2.1)	<0.001

Treatment stage				
Rehabilitation	48 (1.5)	2 (4.1)	46 (1.4)	0.35
Anemia/orthostatic hypotension	236 (7.2)	6 (12.2)	230 (7.1)	0.28
Decreased strength	217 (6.6)	8 (16.3)	209 (6.5)	0.01
Surgery within 3 days	102 (3.1)	2 (4.1)	100 (3.1)	1.00
Personality				
Hesitant to use a call light	266 (8.1)	13 (26.5)	253 (7.9)	< 0.001
Does not like to depend on others	189 (5.8)	8 (16.3)	181 (5.6)	0.004
New to the hospital environment	1,435 (43.9)	25 (51.0)	1,410 (43.8)	0.31

Q1: first quartile, Q3: third quartile
Pearson's χ^2 test, except [a]Mann-Whitney's test

Table 3. Predictive validity indices of the modified Japanese Nursing Association fall risk assessment tool in 7-, 14-, 21-, and 28-day observation samples

	Original samples						Bootstrap samples of 1,000 patients (faller prevalence 1.6%)				
Obs. periods	Prev., %	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	-LR (95% CI)	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	-LR (95% CI)	AUC (95% CI)
7-day	0.8	0.83 (0.59–0.96)	0.66 (0.64–0.68)	2.45 (2.0–3.0)	0.25 (0.1–0.7)	0.81 (0.74–0.88)	1.00 (0.79–1.00)	0.66 (0.63–0.69)	2.97 (2.7–3.2)	0.00	0.84 (0.80–0.88)
14-day	2.2	0.78 (0.58–0.91)	0.59 (0.56–0.62)	1.9 (1.5–2.3)	0.38 (0.2–0.8)	0.77 (0.69–0.84)	0.75 (0.48–0.93)	0.58 (0.55–0.62)	1.80 (1.3–2.4)	0.43 (0.2–1.0)	0.73 (0.63–0.83)
21-day	3.4	0.77 (0.58–0.90)	0.55 (0.52–0.58)	1.7 (1.4–2.1)	0.42 (0.2–0.8)	0.72 (0.64–0.80)	0.75 (0.48–0.93)	0.52 (0.49–0.55)	1.57 (1.2–2.1)	0.48 (0.2–1.1)	0.74 (0.63–0.84)
28-day	4.9	0.78 (0.60–0.91)	0.53 (0.49–0.57)	1.65 (1.4–2.1)	0.41 (0.2–0.8)	0.71 (0.63–0.79)	0.81 (0.54–0.96)	0.53 (0.50–0.57)	1.74 (1.4–2.2)	0.35 (0.1–1.0)	0.64 (0.52–0.76)

Obs. periods: observation periods, Prev: Prevalence of fallers
+LR: Positive likelihood ratio, -LR: Negative likelihood ratio, AUC: area under the receiver operating characteristic curve

Appendix 1. Modified Japanese Nursing Association fall risk assessment tool

Items		Score
Age	65 years and older[a] or 9 years and younger [b]	1
History	Fall history within 1 year	1
Sensory Functions	Visual impairment that affects daily life	1
	Hearing impairment that affects daily life	
Motor Functions	Problems in the extremities (paralysis, numbness)	1
	Bone and/or joint problem	
	Muscle weakness	1
Mobility	Up ad lib	1
	Unstable when standing and/or walking	1
	Use wheelchair, cane, or walker	1
	Requires mobility assistance	1
	Bedridden but able to move extremities	
	Cast, IV lines, or other tubes	1
	Infant's developmental stage (roll over, crawl, etc.)[b]	1
Cognition	Feeling restless because of anxiety or worried about something	1
	Recently feeling forgetful	
	Impaired judgment and/or understanding	
	Unable to use a call light	1
Medications	Analgesics	1
	Laxatives	
	Diuretics	

	Chemotherapy	
	Antiparkinsonians	
	Hypnotics/tranquilizers	1
	Antihypertensives	
Elimination	Waking up more than 2 times at night for toileting	1
	Urinary and/or bowel incontinence	1
	Requiring toileting assistance	
	Using a commode chair[a]	
Treatment stage	In rehabilitation stage[a]	1
	Anemia and/or orthostatic hypotension[a]	1
	Decreased strength due to fever, diarrhea, vomiting, etc.[a]	
	Within 3 days of surgery[a]	1
Personality	Hesitant to use a call light to ask for nurse's help[a]	1
	Does not like to depend on others[a]	
Environment	New to the ward or hospital environment[a]	1

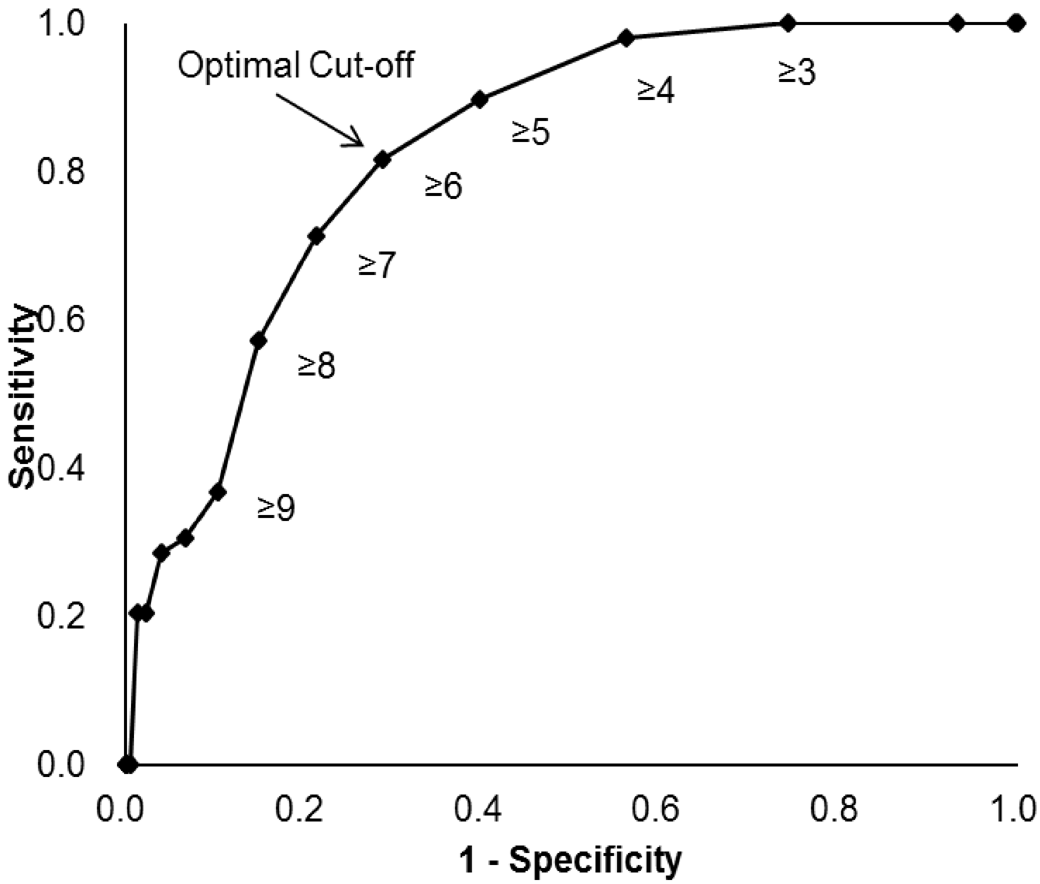
[a]Not included in the Japanese Nursing Association's tool. [b]Not included in the analyses.

Appendix 2. Japanese Nursing Association fall risk assessment tool

Classification	**Characteristics**	**Score**
Age	70 years and older or 9 years and younger	2
Sex	Male	1
History	History of falls History of syncope	2
Sensory Functions	Visual impairment Hearing impairment	1
Functional disability	Paralysis, numbness Bone and/or joint problem (contracture, deformity)	3
Mobility	Muscle weakness Uses wheelchair, cane, or walker Requires mobility assistance Unstable Bedridden	3
Cognition	Disorientation, altered level of consciousness, confusion Dementia Impaired judgment and/or understanding Restlessness Difficulty with learning due to memory impairment	4
Medications	Analgesics	1
	Antiparkinsonians	1
	Narcotics	1
	Diuretics	1
	Hypnotics/tranquilizers	1
	Laxatives	1
	Chemotherapy	1
Elimination	Urinary and/or bowel incontinence	2
	Requires toileting assistance	2
	Frequent urination	2
	Urinary catheter	2
	Night time toileting	2
	Patient room located far from the toilet	2

Patients are assessed on admission, 1 week from admission, and when condition changes

Fig. S1



Supplementary Figure 1. Receiver operating characteristics curve and the optimal cut-off point (n = 3,266)

Supplementary Table 1.

Assessment results of the modified Japanese Nursing Association fall risk assessment tool by observation period

Variables	Observation periods				*p*-value
	7-day (*n* = 2197)	**14-day (*n* = 1249)**	**21-day (*n* = 876)**	**28-day (*n* = 650)**	
Fall risk assessment score, median (Q1, Q3)	4 (3, 7)	5 (3, 7)	5 (3, 8)	5 (4, 8)	<0.001[a]
Fall risk assessment tool items, *n* (%)					
Age ≧65	914 (41.6)	576 (46.1)	425 (48.5)	323 (49.7)	<0.001
Fall history within 1 year	243 (11.1)	170 (13.6)	129 (14.7)	104 (16.0)	0.002
Visual impairment	233 (10.6)	134 (10.7)	101 (11.5)	79 (12.2)	0.66
Hearing impairment	179 (8.1)	120 (9.6)	89 (10.2)	68 (10.5)	0.15
Motor functions					
Impaired extremities	390 (17.8)	275 (22.0)	216 (24.7)	174 (26.8)	<0.001
Bone/joint problems	157 (7.1)	116 (9.3)	93 (10.6)	71 (10.9)	0.002
Muscle weakness	573 (26.1)	400 (32.0)	304 (34.7)	239 (36.8)	<0.001
Mobility					
Up ad lib	1670 (76.0)	859 (68.8)	569 (65.0)	402 (61.8)	<0.001
Unstable when standing/walking	365 (16.6)	240 (19.2)	184 (21.0)	145 (22.3)	0.002
Use of mobility assistive devices	387 (17.6)	287 (23.0)	224 (25.6)	181 (27.8)	<0.001
Requiring mobility assistance	337 (15.3)	246 (19.7)	185 (21.1)	153 (23.5)	<0.001
Bedridden but able to move extremities	82 (3.7)	65 (5.2)	55 (6.3)	45 (6.9)	0.001
Cast, IV lines, or tubes	375 (17.1)	224 (17.9)	168 (19.2)	135 (20.8)	0.15
Cognition					
Feeling restless	312 (14.2)	204 (16.3)	154 (17.6)	120 (18.5)	0.02
Forgetful	196 (8.9)	139 (11.1)	103 (11.8)	87 (13.4)	0.004
Impaired judgment/understanding	124 (5.6)	98 (7.8)	79 (9.0)	61 (9.4)	0.001

Unable to use a call light	65 (3.0)	50 (4.0)	42 (4.8)	33 (5.1)	0.02
Medication					
Analgesics	329 (15.0)	233 (18.7)	185 (21.1)	142 (21.8)	<0.001
Laxatives	154 (7.0)	90 (7.2)	65 (7.4)	51 (7.8)	0.90
Diuretics	211 (9.6)	149 (11.9)	123 (14.0)	98 (15.1)	<0.001
Chemotherapy	69 (3.1)	29 (2.3)	22 (2.5)	19 (2.9)	0.51
Antiparkinsonian	17 (0.8)	10 (0.8)	6 (0.7)	4 (0.6)	0.97
Hypnotics/tranquilizers	386 (17.6)	252 (20.2)	196 (22.4)	145 (22.3)	0.004
Antihypertensive	723 (32.9)	446 (35.7)	315 (36.0)	235 (36.2)	0.18
Elimination					
Toileting ≥2x per night	836 (38.1)	504 (40.4)	379 (43.3)	284 (43.7)	0.012
Urinary/bowel incontinence	123 (5.6)	86 (6.9)	66 (7.5)	57 (8.8)	0.02
Requiring toileting assistance	263 (12.0)	189 (15.1)	146 (16.7)	116 (17.8)	<0.001
Commode chair use	61 (2.8)	42 (3.4)	33 (3.8)	26 (4.0)	0.32
Treatment stage					
Rehabilitation	37 (1.7)	27 (2.2)	17 (1.9)	14 (2.2)	0.75
Anemia/orthostatic hypotension	179 (8.1)	109 (8.7)	80 (9.1)	63 (9.7)	0.60
Decreased strength	167 (7.6)	114 (9.1)	87 (9.9)	67 (10.3)	0.06
Surgery within 3 days	73 (3.3)	45 (3.6)	32 (3.7)	28 (4.3)	0.70
Personality					
Hesitate to use a call light	195 (8.9)	128 (10.2)	83 (9.5)	66 (10.2)	0.55
Does not like to depend on others	142 (6.5)	96 (7.7)	74 (8.4)	50 (7.7)	0.22
New to the hospital environment	983 (44.7)	553 (44.3)	399 (45.5)	307 (47.2)	0.63

Q1: first quartile, Q3: third quartile
P-values obtained by Pearson's χ^2 test, except [a] Kruskal-Wallis test

Supplementary Table 2. Sensitivity and specificity by cut-off points (n = 3,266)

Cut-off Point	Sensitivity	Specificity
≥ 1	1.00	0.003
≥ 2	1.00	0.068
≥ 3	1.00	0.258
≥ 4	0.980	0.438
≥ 5	0.898	0.602
≥ **6**	**0.816**	**0.712**
≥ 7	0.714	0.786
≥ 8	0.571	0.851
≥ 9	0.367	0.896
≥ 10	0.306	0.933
≥ 11	0.286	0.960
≥ 12	0.204	0.977
≥ 13	0.020	0.987
≥ 14	0.00	0.995
≥ 15	0.00	0.998
≥ 16	0.00	0.999
≥ 17	0.00	0.999
≥ 18	0.00	100