Western Journal of Nursing Research
2016, Vol. 38(1) 111–128

DOI: 10.1177/0193945914542851
wjn.sagepub.com
SAGE

Reliability and Validity of the NDNQI® Injury Falls Measure

Lili Garrard[1,2], Diane K. Boyle[3], Michael Simon[4,5], Nancy Dunton[1], and Byron Gajewski[1,2]

Abstract

Although remarkable efforts have been made to improve patient fall reporting through the utilization of standardized definitions, injury falls reporting has rarely been examined. This study used an overall intra-class correlation coefficient (ICC) estimate and factor analysis to assess the reliability and validity of the National Database of Nursing Quality Indicators® (NDNQI®) falls with injury measure. Data were collected from an online Fall Injury Level Survey that was administered to 1,159 NDNQI site coordinators (39.7% response rate; 91% registered nurses [RNs]). Estimated overall ICC was .85. Exploratory factor analysis (EFA) with a Promax rotation (root mean square error of approximation [RMSEA] = 0.053) identified three latent factors: No Injury, Minor Injury, and Moderate/Major Injuries. Final confirmatory factor analysis (CFA) assessment (comparative fit index [CFI] = 0.914, Tucker Lewis Index [TLI] = 0.910, RMSEA = 0.048) confirmed an acceptable model fit. Results provided strong evidence that the NDNQI falls with injury measure is reliable and valid in supporting hospitals' fall prevention efforts and future injurious falls research.

[1]University of Kansas School of Nursing, Kansas City, USA
[2]University of Kansas School of Medicine, Kansas City, USA
[3]University of Wyoming Fay W. Whitney School of Nursing, Laramie, USA
[4]University of Basel, Switzerland
[5]Inselspital Bern University Hospital, Switzerland

Corresponding Author:
Lili Garrard, University of Kansas School of Nursing, 3901 Rainbow Blvd., MS 3060, Kansas City, KS 66160, USA.
Email: lgarrard@kumc.edu

Keywords
injury falls, fall injury levels, reliability, NDNQI

Falls are common adverse events experienced by patients in hospitals and continue to pose challenges to health care quality. Fall reduction is identified as a patient safety priority in the United States (National Priorities Partnership, 2011). Approximately 30% of falls result in injury, particularly among older adults (Shorr et al., 2008). Injuries from falls burden hospitals and patients with increased costs due to longer lengths of stay and additional patient care costs (Currie, 2008). For older adults, the direct and indirect cost of injuries associated with falls is projected to reach US$54.9 billion (in year 2007 dollars) annually by 2020 (Centers for Disease Control and Prevention, 2013; Englander, Hodson, & Terregrossa, 1996). In an effort to promote patient safety, the National Quality Forum (NQF; 2011) named "patient death or serious injury associated with a fall while being cared for in a healthcare setting" (p. 9) as one of the health care Serious Reportable Events (SREs). Similarly, the Centers for Medicare & Medicaid Services (CMS) identified hospital falls and resulting trauma as one of the preventable Hospital-Acquired Conditions (HAC). Additional costs associated with HAC are no longer covered by Medicare for hospitals participating in the Inpatient Prospective Payment System (IPPS; CMS, 2012; Inouye, Brown, & Tinetti, 2009).

National Database of Nursing Quality Indicators® (NDNQI®) Fall and Falls With Injury Measures

NQF established a national framework to evaluate health care quality measurement and reporting. NQF's goals are to increase public awareness in quality performance, establish incentives for performance improvement, and provide national benchmarks (NQF, 2002; Simon, Klaus, Gajewski, & Dunton, 2013). Both patient fall and falls with injury have been endorsed by the NQF as national consensus measures since 2004 (NQF, 2004). The American Nurses Association (ANA), serving as the NQF steward for both measures, commissioned the NDNQI to conduct separate studies to assess the reliability of each fall measure in part to support their successful NQF re-endorsement in 2013 (NQF, 2013a, 2013b). NDNQI was established in 1998 by ANA to monitor nurse-sensitive quality indicators that are essential for patient safety and quality improvements in hospitals (Montalvo, 2007). NDNQI is a quality database that collects and evaluates unit-specific nurse-sensitive data from over 2,000 U.S. and international hospitals. Member hospitals of NDNQI benefit from regular reporting of nursing quality measures and various national comparison data that were shown to be helpful in quality improvement.

The NDNQI patient fall reliability study was conducted by Simon and colleagues (2013) to examine the agreement of fall classifications among staff in U.S. hospitals (sensitivity = 0.90, specificity = 0.88, mean probability for classifying a fall = 0.60). Based on the results of Simon's study, the NQF-endorsed NDNQI patient fall definition was revised to provide more standardized reporting of falls. Although remarkable efforts have been made to improve fall reporting, previous research has indicated a lack of standardized definition and methods of measuring and reporting falls-related injuries (Schwenk et al., 2012). As previously mentioned, injuries associated with falls increase the cost of health care substantially. Without standardized clinical guidelines for reporting injury falls, hospitals lack the ability to properly compare themselves with reliable national comparison data and to develop and implement cost-effective fall prevention plans. Given the financial impact on both hospitals and patients, correct classification of fall-related injuries is imperative, particularly being able to distinguish no injury and minor injuries from serious injuries. Correct classification will allow hospital fall prevention efforts to better target education, risk assessment, and prevention protocols. Thus, the need to evaluate standardized reporting of injury levels, the key to a reliable and valid injury falls measure, is apparent.

Purpose

The purpose of the study was to investigate the reliability and validity of the NDNQI falls with injury measure by utilizing the NQF and NDNQI injury level definitions (NDNQI, 2010). The specific aims were to assess (a) the consistency of injury level assignment among raters of the fall injury scenarios and (b) the accuracy of correct injury level assignment. The information on the fall scenarios emulated those commonly found in adverse event or incident reports. Before the study began, approval was obtained from the University of Kansas Medical Center Human Subjects Committee.

Method

Design

Data collection for the injury falls reliability study followed a similar process to regular falls reporting to NDNQI by member hospitals. When a patient fall occurred in a hospital, a detailed incident report regarding the fall would be filed, including the hospital location of the fall; whether the fall was witnessed, self-reported, or assisted; medication administered to the patient; and any injuries observed at the time of the fall or during post-assessment. Based on the information collected on the incident reports, the fall prevention team would review the incident and determine whether it constituted a unit fall or not, and assign the proper injury level according to NDNQI definitions, which are described in a later section. A unit fall indicates that the event was a fall that occurred on a unit declared eligible by NDNQI for falls reporting. Once the incident had been thoroughly reviewed, it would be reported to NDNQI along with any other fall incidents on the same unit for the calculation of a unit fall rate.

Participants

Each NDNQI member hospital identifies a site coordinator whose primary responsibility is being a point of contact for all NDNQI-related activities. The NDNQI site coordinator serves a vital role in ensuring that all data collection and reporting adhere to NDNQI guidelines. Thus, the targeted survey population consisted of a convenience sample of site coordinators.

In total, 1,159 site coordinators were invited to participate and 461 responded, resulting in a 39.7% response rate. Among all respondents, 411 provided responses for all fall scenarios, which were considered as "complete" responses. Specific instructions for the site coordinators were provided in an email invitation. Because fall prevention programs in hospitals are often viewed as an inter-professional team effort, other hospital staff who serve as final decision makers about injury levels were also asked to be consulted while completing the survey. The most important aspect of the survey was that respondents must assign each scenario to a fall injury level using the NDNQI definitions. A typical respondent was a registered nurse (RN; 91%), held a masters or higher degree (60%), and worked in nursing management (40%) or quality improvement (31%).

Survey Development

A Fall Injury Level Survey was generated using a convenience sample of de-identified incident reports from NDNQI hospitals and NDNQI guidelines on injury levels. Each scenario went through rigorous revisions after being reviewed by hospital and NDNQI staff members who were involved in patient fall–related activities. This process was critical to ensure the content validity of the fall scenarios on the survey. Twenty fall scenarios were selected as candidates for the final survey.

Two senior NDNQI staff members served as fall experts for determining the correct classification of injury levels in the 20 fall scenarios. Both experts were masters prepared RNs with over 30 years of clinical experience and who provided daily guidance for NDNQI hospitals on classifying actual falls. The experts scored the fall scenarios independently and reached 100% agreement on classification after discussions. Five scenarios were excluded from this study as they were identified by the experts as not a fall or not a unit fall according to the NDNQI fall definition. Thus, the NDNQI experts' judgment was considered the correct injury level classification and deemed to be the "gold" standard. The final Injury Fall Level Survey consisted of 15 fall scenarios, and the distributions of the scenarios were as follows: 6 non-injurious falls, 3 minor injury falls, 3 moderate injury falls, 3 major injury falls, and 0 death resulting from a fall (Table 1). Having the experts' gold standard was a crucial first step for subsequent statistical analysis. Table 1 shows an abbreviated description and the expert classification for each of the scenarios.

To address the first aim, survey participants were asked to classify the injury level of each scenario according to the NDNQI definitions. Also, questions were included in the survey about the respondents' characteristics such as professional background, highest education level, and current work department within the hospital. The Fall Injury Level Survey was conducted online using the survey tool Zoomerang (http://www.zoom-erang.com).

NDNQI Fall and Injury Level Definitions

The NQF-endorsed NDNQI fall and injury level definitions were given in the survey to assist respondents with injury level classifications for the fall scenarios (NDNQI, 2010). A fall was defined as

> an unplanned descent to the floor (or extension of the floor, e.g., trash can or other equipment) with or without injury to the patient, and occurs on an eligible reporting nursing unit. All types of falls are to be included whether they result from physiological reasons (fainting) or environmental reasons (slippery floor). Include assisted falls—when a staff member attempts to minimize the impact of the fall. Exclude falls by visitors, students, and staff members; falls on other units not eligible for reporting; falls of patients from eligible reporting units, however patient was not on unit at time of the fall (e.g., patient falls in radiology department). (p. 13)

Injury levels are reported to NDNQI (2010) based on the following guidelines:

Table 1. Expert Injury Level Classification and Mean Scale Score of Fall Scenarios.

Fall Scenario		Expert Classification	Mean Scale Score (95% CI)
S1[a]	Pt. found sitting on bathroom floor. Steri-strips applied to lacerations on elbow.	Moderate	2.72 [2.26, 3.17]
S2[a]	Pt. lost balance and fell backward. Complained of low back pain. MD ordered Dilaudid and heat packs applied. X-rays negative for fracture or displacement.	Minor	2.02 [1.50, 2.55]
S3[a]	Pt. was found on floor lying next to bed after a loud sound heard from room. No signs/symptoms of injury at that time and at 24 hr post event.	None	1.21 [0.63, 1.79]
S4	Pt. reported to nurse that she "hurt her arm" during fall when walking to BR. No signs of injury and had full ROM. Tylenol administered.	None	1.59 [1.09, 2.09]
S5[a]	Pt. stated he tripped on IV pump power cord and fell. No pain or other injury at the time of the fall or 24 hr post fall.	None	1.05 [0.82, 1.27]
S6[a]	Pt. reported she fell out of a chair to floor while reaching for a book on bedside table. Her NG tube was pulled out, but no other pain or signs of injury 24 hr post fall. MD said to leave NG tube out.	None	1.10 [0.77, 1.42]
S7[a]	Pt. states she fell on knees while reaching for shoes. No injury noted at the time. The next day (15 hr later) pt. complained of R knee pain. X-ray negative, ice, and ACE bandage applied.	Minor	2.04 [1.68, 2.39]
S8[a]	Pt. found on floor. Complained of pain on R side of head, R elbow, and knees. Pt. states he is dizzy, neuro checks found reduced R hand grasp. Small subdural hematoma found on CT scan and pt. transferred to ICU.	Major	3.91 [3.58, 4.25]
S9[a]	Pt. reported he tripped with walker on door jam and fell. Pt. denies pain or other symptoms. Chest X-rays prior to fall indicated a recent rib fracture. Pain meds given 4 hr prior to deep breathing exercises.	None	1.54 [0.54, 2.54]
S10[a]	Pt. found on BR floor and states she hit head. Small laceration on forehead and bandaid applied. Also complained of low back pain, CT of head and lumbar back negative for fracture or hematomas. Pt. given acetaminophen.	Minor	2.09 (1.73, 2.46)
S11[b]	Pt. found unconscious on BR floor after a loud sound heard from room. Large amount of blood on BR floor, sink, and R side of head. Does not respond to painful stimuli, pupils dilated, no B/P, weak and thready pulse. Code blue activated and CPR performed for 15 min without success.	Moderate	4.88 [4.32, 5.43]
S12[a]	While pt. was assisted to BR with gait belt he became dizzy. While trying to lower pt. to the toilet, he became limp and was lowered to the floor. He arm struck the handrail and started swelling. X-ray revealed closed fracture of ulna and a cast was applied.	Major	3.75 [3.27, 4.23]

(continued)

Table 1. (continued)

Fall Scenario		Expert Classification	Mean Scale Score (95% CI)
S13[a]	Pt. walked unassisted to BR after returned to room from EGD. Pt. states he fell to floor after trying to get back in bed. He complained of pain in R ankle. X-ray revealed distal fracture and a cast was applied. After 3 days, pt. complained of numbness and tingling in foot and toes appear blue/purple with swelling. Cast removed 17 hr later by MD and no pedal pulses. Pt. taken to OR for immediate amputation.	Major	3.97 [3.78, 4.16]
S14[b]	Pt. lost balance and fell to floor during transfer from commode to bed. Six staff helped lift pt. with bath blankets to bed and blankets ripped and pt. fell against side rails. Pt. treated for 5 inch abrasion to lumbar area. X-ray of lumbar revealed small compression fracture and treated with back brace.	None	3.60 [2.99, 4.21]
S15[a]	Pt. became dizzy while walking to BR with assistance. Nurse assisted patient to the floor. Pt. sustained 4 inch skin tear on R forearm during the decent. Steri-strips and Kerlix bandage applied.	Moderate	2.66 [2.18, 3.15]

Note. Injury level scale: 1 = *none*, 2 = *minor*, 3 = *moderate*, 4 = *major*, 5 = *death*. CI = confidence interval; Pt. = patient; MD = medical doctor; BR = bathroom; ROM = range of motion; IV = intravenous therapy; NG = nasogastric; R= right; ACE = all cotton elastic (a bandage brand name); CT = computerized tomography; ICU = intensive care unit; B/P = blood pressure; CPR = cardiopulmonary resuscitation; EGD = esophagogastroduodenoscopy; OR = operating room; CFA = confirmatory factor analysis.
a. Final scenario selected by CFA.
b. Complex scenario.

None—patient had no injuries (no signs or symptoms) resulting from the fall, if an x-ray, CT scan or other post fall evaluation results in a finding of no injury

Minor—resulted in application of a dressing, ice, cleaning of a wound, limb elevation, topical medication, pain, bruise or abrasion

Moderate—resulted in suturing, application of steri-strips/skin glue, splinting or muscle/joint strain

Major—resulted in surgery, casting, traction, required consultation for neurological (basilar skull fracture, small subdural hematoma) or internal injury (rib fracture, small liver laceration) or patients with coagulopathy who receive blood products as a result of a fall

Death—the patient died as a result of injuries sustained from the fall (not from physiologic events causing the fall). (pp. 14-15)

Analysis

Coding of responses. Each respondent selected one out of five injury levels according to NDNQI definitions for each of the 15 fall scenarios described in the survey. The response options were coded as 1 “none,” 2 “minor,” 3 “moderate,” 4 “major,” and 5 “death.” The correct injury level for each scenario was the gold standard set by the experts’ classification as described above. Based on the gold standard, all participant responses were further classified as 1 “correct” and 0 “incorrect,” for all 15 fall scenarios. The data file containing the recoded dichotomous data for the 15 fall scenarios served as the main analysis file for all statistical analyses used in this study.

Reliability and validity analysis. The reliability of a measure is the “ability to produce similar results when repeated measurements are made under identical conditions” (Bordens & Abbott, 2011, p. 130). One common practice to assess the reliability of a target, under the influence of judgments made by a group of respondents is to calculate the intra-class correlation coefficient (ICC). ICC is calculated as the proportion of the total variance that is due to the true variance from raters (Skrondal & Rabe-Hesketh, 2004). For this study, the fall scenarios were treated as targets and the survey participants as raters. An overall ICC could be used to describe the between-scenario variation of injury level assignment. A high ICC would indicate that the majority of the variance was due to differences among the scenarios, which implied that the difference within each scenario, influenced by raters, was small. Thus, the raters had a high consistency of injury fall classification for each scenario. In this study, the overall ICC estimate was interpreted as excellent (around .90), very good (around .80), and adequate (around .70), following general guidelines provided by Kline (2011). The overall reliability estimate computation was performed using SPSS software version 20.

In addition to reliability, the validity of the fall scenarios also was assessed. The validity of a measure is defined as “the extent to which it measures what you intend it to measure” (Bordens & Abbott, 2011, p. 133). For the 15 fall scenarios, it was important to assess the construct validity of the scenarios. In other words, the goal was to determine if the fall scenarios could appropriately predict the severity of injury falls by assessing the accuracy of correct injury level assignment. A decision needed to be made after examining the proportion of respondents selecting the exactly correct injury level and selecting the correct injury level within one response option, both with a 95% confidence interval. Two scenarios (S11 and S14) were very complex, which might have caused a large proportion of the respondents to choose the wrong

Table 2. 95% Confidence Interval for the Proportion of Exactly Correct and Correct Within One Injury Level.

Fall Scenario[a]	Exactly Correct (%)	Correct Within One Injury Level (%)
S1	[67.17, 75.44]	[100.00, 100.00]
S2	[69.97, 78.04]	[98.59, 100.08]
S3	[82.47, 88.95]	[98.94, 100.17]
S4	[36.52, 45.63]	[98.94, 100.17]
S5	[93.62, 97.45]	[99.34, 100.21]
S6	[88.66, 93.89]	[98.23, 99.98]
S7	[84.09, 90.30]	[100.00, 100.00]
S8	[90.11, 95.00]	[97.89, 99.86]
S9	[69.56, 77.81]	[78.56, 85.74]
S10	[84.27, 90.50]	[98.91, 100.18]
S11[b]	[0.00, 0.68]	[3.39, 7.70]
S12	[73.30, 81.18]	[96.90, 99.42]
S13	[94.53, 98.08]	[100.00, 100.00]
S14[b]	[0.43, 2.82]	[1.75, 5.21]
S15	[61.83, 70.73]	[98.90, 100.18]

a. Abbreviated descriptions of the scenarios are summarized in Table 1.
b. Complex scenario.

injury level (Table 2). Given the psychometric difficulties, a decision was made to eliminate these two scenarios from the construct validity analysis.

Thirteen fall scenarios remained for assessment of construct validity, which was approached with a two-stage factor analysis using only complete responses. An exploratory factor analysis (EFA) was the logical first step to explore the possible latent factor structure of the injury levels among the fall scenarios. Once the latent factor structure was identified from EFA, it was necessary to verify the factor structure by using a confirmatory factor analysis (CFA) with structural equation modeling. Factor analysis is a correlation-oriented approach that aims to reproduce the inter-correlation among the variables. Several types of correlations exist; however, due to the nature of dichotomous data in this study, tetrachoric correlation was the most appropriate correlational method to serve as the basis of the factor analysis. Unlike Pearson's correlation for continuous data, using tetrachoric correlation allowed us to estimate correlations among dichotomously measured variables as if the variables were made on a continuous scale.

The construct validity computations were all performed using Mplus software version 5.21 (Muthén & Muthén, 1998-2009). Mplus is an advanced statistical software recognized for its powerful ability to fit various latent variable models. Following recommendations by MacCallum, Roznowski, and Necowitz (1992), the main analysis file with 411 complete responses were randomly split into comparable training (196 responses, 47.7%) and validation (215 responses, 52.3%) data sets to avoid capitalization on chance concerns. An EFA with categorical factor indicators was conducted using the training data set in Mplus, which conveniently incorporated tetrachoric correlation into the analysis. Traditional factor extraction, such as Kaiser's criterion, has been widely accepted for suggesting factors with an eigenvalue greater than 1 as common factors. Eigenvalues are often interpreted as the variances extracted by the common factors. However, eigenvalues based Kaiser's criterion should not be used solely to determine the number of factors due to over-extraction concerns. Another requirement for including items in a specific factor was that the individual items must meet a criterion of at least 0.30 in absolute value for factor loading to be retained. Additional model fit can be evaluated by using the root mean square error of approximation (RMSEA), and a RMSEA value around 0.05 or less usually indicates an acceptable model fit. As latent factors were identified, a CFA with categorical factor indicators using structural equation modeling was performed on the validation data set to confirm the factor structure demonstrated in the EFA step. Several statistical indices such as the comparative fit index (CFI; around 0.9 or higher), Tucker Lewis Index (TLI; around 0.9 or higher), and RMSEA (around 0.05 or less) were used to assess the final model fit.

For oblique rotations (correlated factors, for example, Promax), the concept of the proportion of variance explained by a factor is complex and less intuitive. Factor solutions provided by a Varimax rotation (uncorrelated factors) are often very similar to the Promax solutions. Thus, the Varimax factor solutions can be used as a proxy to compute the variability explained by a given factor under the Promax setting. The proportion of variance explained by a factor can be calculated as the sum of squared factor loadings on the assigned factor divided by the number of fall scenarios assigned to that particular factor. In addition, Mplus also provides estimates for the proportion of variance in each fall scenario, explained by their assigned factor.

Results

Reliability

The variance within each scenario was 0.252 and the variance between the 15 fall scenarios was 1.479, resulting in an overall ICC (1, 1) of .85, which was between "very good" and "excellent" according to the general guidelines provided by Kline (2011). The ICC (1, 1) indicated a substantial reliability of the fall scenarios and a high consistency of injury level assignment among the respondents for each scenario. The mean scale scores with 95% confidence intervals for all 15 fall scenarios are summarized in Table 1. The variance between the scenarios was much larger than the variance within each scenario, which echoed the results of the overall ICC estimate and indicated a high reliability.

As mentioned above, two scenarios (S11 and S14) were very complex and were excluded from further analysis. After exclusion, the overall ICC (1, 1) for the remaining 13 scenarios was re-calculated to be .82, which still maintained a very good reliability and was suitable for the validity analysis.

Validity

During the initial EFA conducted on the training data set, six factors with eigenvalues greater than 1 were suggested based on Kaiser's criterion (eigenvalues: 3.556, 2.807, 1.582, 1.195, 1.171, 1.079, 0.692, 0.654, 0.353, 0.284, 0.094, −0.163, −0.305), but only three factors could be extracted successfully, indicating an over-extraction based on Kaiser's criterion. Factor loadings of the three-factor model were further clarified after applying a Promax rotation for correlated factors, resulting in a RMSEA of 0.053, which indicated an acceptable model fit. All scenarios loaded over 0.30 on the assigned factors. The aim of the EFA was to identify underlying factor structure that could be used to predict the severity of injury falls. The results indicated three latent factors: ability associated with classifying non-injurious falls (No Injury), ability associated with classifying minor injury falls (Minor Injury), and ability associated with classifying moderate or major injury falls (Moderate/Major Injuries; Table 3).

With the validation data set, the CFA model was specified using the three factors measured by the 13 scenarios, with each scenario assigned to the relevant factor. The goal was to identify and retain scenarios that contributed most to respondents' ability associated with injury fall classifications. Estimates of the pattern coefficients representing the direct effects of the factors on the scenarios ranged from −0.021 to 0.950 (Figure 1a). Several statistical indices were used to determine the adequacy of model fit to the data. Results from the initial CFA assessment did not indicate a good model fit (CFI = 0.868, TLI = 0.863, RMSEA = 0.055). Pattern coefficient estimates for all scenarios were statistically significant (p value < .05) with the exception of Scenario 4 (−0.021, p value = .851) and Scenario 13 (0.395, p value = .051). The pattern coefficient estimate for Scenario 13 can be considered as marginally significant and we decided to keep this scenario in the model. The

Table 3. Factor Loadings After Promax Rotation for Three-Factor Structure With Injury Levels.

Fall Scenario[a]	No Injury	Minor Injury	Moderate/Major Injuries	Injury Level
S1	−0.078	0.239	**0.801**	Moderate
S8	−0.173	−0.014	**0.535**	Major
S12	0.104	−0.005	**0.643**	Major
S13	0.005	−0.144	**0.883**	Major
S15	0.094	0.028	**0.715**	Moderate
S2	−0.019	**0.778**	0.033	Minor
S7	0.059	**0.504**	−0.005	Minor
S10	0.197	**0.312**	−0.274	Minor
S3	**0.448**	0.444	0.233	None
S4	**0.758**	−0.244	0.008	None
S5	**0.873**	0.64	−0.023	None
S6	**0.684**	0.393	−0.064	None
S9	**0.311**	0.03	0.082	None

a. Abbreviated descriptions of the scenarios are summarized in Table 1.
Note. Highest factor loading for each fall scenario is in bold.

CFA model was re-fitted after removing Scenario 4 and the confirmed structure remained the same (Figure 1b). The final CFA assessment confirmed an acceptable model fit and supported the hypothesis that a relationship exists between the 12 final fall scenarios (Table 1) and the three underlying latent factors (CFI = 0.914, TLI = 0.910, RMSEA = 0.048).

As mentioned above, Varimax factor solutions were used as a proxy to calculate the variability explained by the three correlated latent factors. Results from the Varimax rotation are not reported here due to the high degree of similarity with the Promax rotation solutions. The proportion of variance explained by the No Injury, Minor Injury, and Moderate/Major Injuries factors were 52.3%, 31.9%, and 46.7%, respectively. In addition, the ability associated with classifying non-injurious falls accounted for 53.6%, 34.8%, 31.8%, 15.6%, and 49.3% of the proportion of variance in Scenarios 1, 8, 12, 13, and 15, respectively. The variability in Scenarios 2, 7, and 10, explained by the ability associated with classifying minor injury falls, were 33.7%, 36.0%, and 56.1%, respectively. Finally, the ability associated with classifying moderate or major injury falls accounted for 57.8%, 90.5%, 66.9%, and 11.7% of the variability in Scenarios 3, 5, 6, and 9, respectively.

The construct validity analysis findings indicated that the final 12 fall scenarios from the survey resulted in appropriate latent structures for predicting

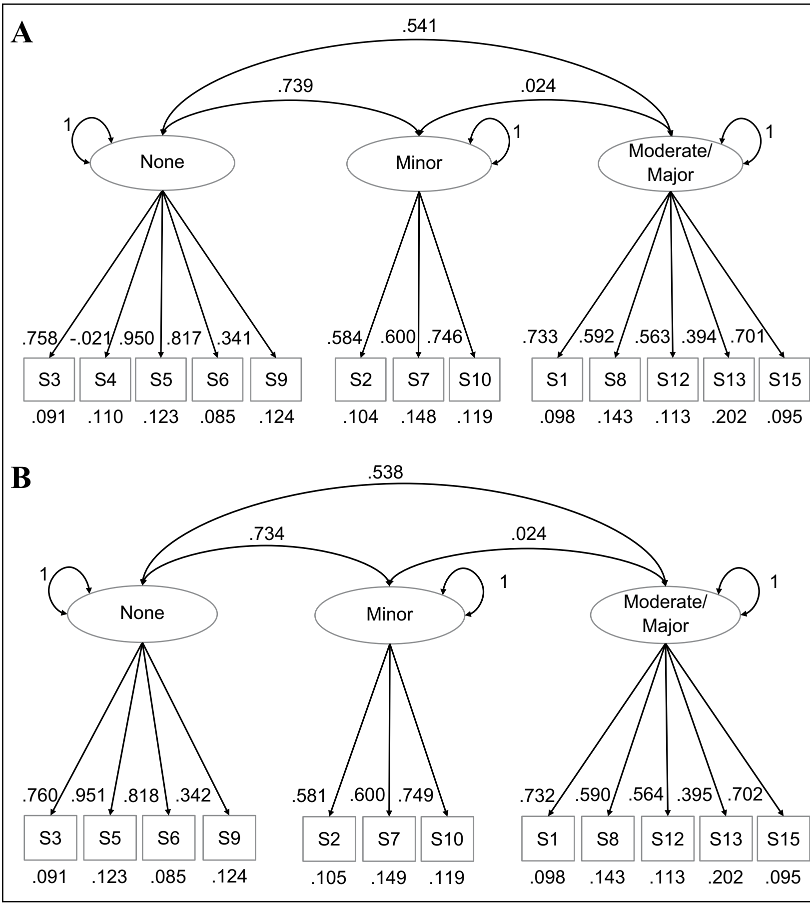

Figure 1. (A) Initial CFA model and (B) final CFA model.
Note. CFA = confirmatory factor analysis.

the severity of the injury falls, and thus supporting the validity or accuracy of injury level classifications made by survey respondents for all 12 final fall scenarios.

Discussion

The overall ICC estimate for the 15 fall scenarios fell between very good and excellent, indicating high consistency of injury level classifications among respondents for each fall scenario. Results provided strong evidence for the reliability of the NDNQI falls with injury measure. Construct validity was also confirmed, resulting in 12 final fall scenarios with 4 non-injurious falls, 3 minor injury falls, 2 moderate injury falls, and 3 major injury falls. The 12 final fall scenarios represented a reliable and valid approach to evaluate respondent fall injury level classification ability.

From the results of the construct validity analysis, it was apparent that the scenarios clustered very well into the three distinct categories. However, the correlations among the three latent factors exhibited a very interesting pattern, which could be presented as poor (Minor vs. Moderate/Major = .024, p value = .810), average (None vs. Moderate/Major = .538, p value $< .05$), and good (None vs. Minor = .734, p value $< .05$). The pattern in the factor correlation estimates merited further investigation. The poor correlation (.024) between Minor Injury and Moderate/Major Injuries could be interpreted, such that the respondents' ability to correctly classify minor injuries did not imply that they would also have the same ability to correctly classify moderate or major injuries, and vice versa. This finding is rather concerning and can indicate several potential issues, such as confusion over the definitions, ambiguity of the incident reports, or bias introduced from both the patient and fall evaluator's perspectives. On the contrary, it is certainly encouraging to see that the respondents had average ability to correctly distinguish no injury from moderate or major injuries, and vice versa. Moreover, the respondents had a good ability to correctly classify no injury from minor injuries, and vice versa. The overall results can be viewed as an indication that more education or training is needed for correctly identifying all injury levels, particularly the moderate or major injury falls, as these types of fall scenarios are rare. The clarity of the injury level definitions also needs to be further reviewed to minimize potential classification challenges. In addition, although the construct validity assessed respondents' ability to distinguish among No Injury, Minor Injury, and Moderate/Major Injuries, the ability to distinguish injury levels within the global category of Moderate/Major Injuries remains unknown and requires further investigation.

The majority of fall scenarios had about 70% to 90% of respondents selecting the exactly correct injury level with the exception of three scenarios (S4, S11, and S14). Specifically S11 and S14 had close to 0% of the respondents being exactly correct (Table 2). When the requirement was relaxed to allow within one injury level, S11 and S14 still remained very low with less than 10% of the respondents being correct (Table 2). The sequence of events in Scenario 11 made it unclear whether the fall caused the patient death or the death caused the fall. In Scenario 14, the patient fell and then was dropped by the staff as they attempted to assist the patient back to bed, leading to confusion about the injury level assignment. These two scenarios were considered to be very complex, which resulted in a wide variance of injury level assignment among the respondents. Thus, both fall scenarios were excluded from the construct validity analysis for psychometric difficulties. The complex fall scenarios (e.g., S11 and S14) need to be examined carefully and debriefed by the fall prevention team, and when necessary, expert consultations should be considered to help prevent bias by the fall evaluator. In addition, concerns can arise with patient self-reported falls (e.g., S4) because this type of fall is often not observed and hard to validate without evidence; thus, potential bias could be introduced from both the patient and fall evaluator's perspectives.

One limitation of this study comes from the usage of incident reports to help design the online survey. Previous research by Shorr and colleagues (2008) pointed out that using incident reports alone contributes to the underreporting of both injurious and non-injurious falls in hospitals. Potential bias could be introduced by using a convenience sample of de-identified incident reports, which are not representative for all fall scenarios that patients experience daily in hospitals. Although all fall scenarios went through rigorous revisions to ensure their clinical reality, it remains unclear how frequent these scenarios occur. Perhaps more scenarios need to be developed to cover the full spectrum of NDNQI injury classifications.

Another limitation of this study comes from the sample selection bias. The primary audience for the survey was a convenience sample of NDNQI site coordinators. Comparing with the general population of U.S. hospitals, NDNQI consists of more Magnet® designated, not-for-profit, larger, and higher case-mix index (CMI) hospitals (Lake, Shang, Klaus, & Dunton, 2010). The general profile of NDNQI hospitals may include more hospital resources, which play an important role in establishing training for staff and fall prevention programs. Being the primary respondent of the survey (68%), NDNQI site coordinators are constantly informed on new updates to NDNQI guidelines and definitions. They are most familiar with NDNQI frameworks and thus may represent a more "trained" group of hospital staff in regard to standardized data collection and reporting. The ability of correct injury level classification across other hospital staff involved in fall-related activities still remains unclear and needs to be further evaluated.

In this study, the reliability and validity of the NDNQI falls with injury measure was evaluated and findings supported the successful re-endorsement by NQF. The NDNQI site coordinators demonstrated high consistency in classifying injury levels for specific fall scenarios, according to NDNQI definitions. The Falls Injury Level Survey with the final 12 fall scenarios was shown to be valid in assessing respondents' abilities to predict the severity of the injury falls, particularly among non-injurious falls, minor injury falls, and moderate or major injury falls. Hospital site coordinators are encouraged to continue contacting NDNQI for assistance with the classification of complex fall scenarios and patient self-reported fall scenarios. Findings of this study also supported rationales for revising the standardized NDNQI falls and injury level definitions to include additional types of falls and provide more clarification on injuries.

An implication from this study is that the Falls Injury Level Survey can be utilized in the future as a training tool for hospital staff that serve as final decision makers on injury levels. Researchers at NDNQI launched a well-known and comprehensive Pressure Ulcer Identification and Staging Training Program in 2009, which can be used to guide the development of a falls with injury training tool (Bergquist-Beringer et al., 2009; Bergquist-Beringer, Gajewski, Dunton, & Klaus, 2011; Gajewski, Hart, Bergquist-Beringer, & Dunton, 2007; Hart, Bergquist, Gajewski, & Dunton, 2006). In addition, because the NDNQI injury falls measure is NQF-endorsed, standardized injury level definitions are available to the public domain. A recent article published by Mion and colleagues (2012) utilized NDNQI injury level definitions as part of their retrospective study for determining potential predictors and outcomes of injurious falls among a cohort of hospital patients. The NDNQI injury falls measure provides a reliable and valid tool for non-NDNQI hospitals and external researchers to support future quality improvement efforts and injurious falls research.

Acknowledgments

Chenjuan Ma and Yu Jiang provided helpful comments on an earlier version of this article. Jan Davidson provided clinical consultation with summarizing the fall scenarios.

Declaration of Conflicts of Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was conducted under a contract with American Nurses Associations.

References

Bergquist-Beringer, S., Davidson, J., Agosto, C., Linde, N. K., Abel, M., Spurling, K., & Christopher, A. (2009). Evaluation of the National Database of Nursing Quality

Indicators (NDNQI) training program on pressure ulcers. *Journal of Continuing Education in Nursing*, *40*, 252-258; quiz 259-260, 279.

Bergquist-Beringer, S., Gajewski, B., Dunton, N., & Klaus, S. (2011). The reliability of the National Database of Nursing Quality Indicators pressure ulcer indicator: A triangulation approach. *Journal of Nursing Care Quality*, *26*, 292-301. doi:10.1097/NCQ.0b013e3182169452

Bordens, K., & Abbott, B. B. (2011). *Research design and methods: A process approach* (8th ed.). New York, NY: McGraw-Hill Higher Education.

Centers for Disease Control and Prevention. (2013). *Costs of falls among older adults*. Retrieved from http://www.cdc.gov/homeandrecreationalsafety/falls/fallcost.html

Centers for Medicare & Medicaid Services. (2012). *Hospital-acquired conditions (HAC) in acute inpatient prospective payment system (IPPS) hospitals*. Retrieved from http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HospitalAcq-Cond/Downloads/HACFactsheet.pdf

Currie, L. (2008). Fall and injury prevention. In R. G. Hughes (Ed.), *Patient safety and quality: An evidence-based handbook for nurses*. Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www.ahrq.gov/professionals/clinicians-providers/resources/nursing/resources/nurseshdbk/nurseshdbk.pdf

Englander, F., Hodson, T. J., & Terregrossa, R. A. (1996). Economic dimensions of slip and fall injuries. *Journal of Forensic Sciences*, *41*, 733-746.

Gajewski, B. J., Hart, S., Bergquist-Beringer, S., & Dunton, N. (2007). Inter-rater reliability of pressure ulcer staging: Ordinal probit Bayesian hierarchical model that allows for uncertain rater response. *Statistics in Medicine*, *26*, 4602-4618. doi:10.1002/Sim.2877

Hart, S., Bergquist, S., Gajewski, B., & Dunton, N. (2006). Reliability testing of the National Database of Nursing Quality Indicators pressure ulcer indicator. *Journal of Nursing Care Quality*, *21*, 256-265.

Inouye, S. K., Brown, C. J., & Tinetti, M. E. (2009). Medicare nonpayment, hospital falls, and unintended consequences. *New England Journal of Medicine*, *360*, 2390-2393. doi:10.1056/NEJMp0900963

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Lake, E. T., Shang, J., Klaus, S., & Dunton, N. E. (2010). Patient falls: Association with hospital Magnet status and nursing unit staffing. *Research in Nursing & Health*, *33*, 413-425. doi:10.1002/nur.20399

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*, 490-504.

Mion, L. C., Chandler, A. M., Waters, T. M., Dietrich, M. S., Kessler, L. A., Miller, S. T., & Shorr, R. I. (2012). Is it possible to identify risks for injurious falls in hospitalized patients? *Joint Commission Journal on Quality and Patient Safety*, *38*, 408-413.

Montalvo, I. (2007). The National Database of Nursing Quality Indicators (NDNQI®). *The Online Journal of Issues in Nursing*, *12*(3). Retrieved from http://www.nursingworld.org/MainMenuCategories/ANAMarketplace/ANAPeriodicals/OJIN/TableofContents/Volume122007/No3Sept07/NursingQualityIndicators.aspx?%3E

Muthén, L. K., & Muthén, B. O. (1998-2009). *Mplus user's guide* (5th ed.). Los Angeles, CA: Author.

National Database of Nursing Quality Indicators. (2010). *Guidelines for data collection on the American Nurses Association's National Quality Forum endorsed measures*. Kansas City: University of Kansas Medical Center.

National Priorities Partnership. (2011). *Input to the Secretary of Health and Human Services on priorities for the national quality strategy*. Washington, DC: National Quality Forum. Retrieved from http://www.qualityforum.org/WorkArea/linkit.aspx?ItemID=68238

National Quality Forum. (2002). *A national framework for healthcare quality measurement and reporting: A consensus report*. Retrieved from http://www.qualityforum.org/Publications/2002/07/A_National_Framework_for_Healthcare_Quality_Measurement_and_Reporting.aspx

National Quality Forum. (2004). *National voluntary consensus standards for nursing-sensitive care: An initial performance measure set*. Retrieved from http://www.qualityforum.org/Publications/2004/10/National_Voluntary_Consensus_Standards_for_Nursing-Sensitive_Care__An_Initial_Performance_Measure_Set.aspx

National Quality Forum. (2011). *Serious reportable events in healthcare—2011 update: A consensus report*. Retrieved from http://www.qualityforum.org/Publications/2011/12/Serious_Reportable_Events_in_Healthcare_2011.aspx

National Quality Forum. (2013a). *Falls with injury*. Retrieved from http://www.qualityforum.org/QPS/0202

National Quality Forum. (2013b). *Patient fall rate*. Retrieved from http://www.qualityforum.org/QPS/0141

Schwenk, M., Lauenroth, A., Stock, C., Moreno, R. R., Oster, P., McHugh, G., . . . Hauer, K. (2012). Definitions and methods of measuring and reporting on injurious falls in randomised controlled fall prevention trials: A systematic review. *BMC Medical Research Methodology*, *12*, Article 50. doi:10.1186/1471-2288-12-50

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. *Journal of the American Geriatrics Society*, *56*, 701-704.

Simon, M., Klaus, S., Gajewski, B. J., & Dunton, N. (2013). Agreement of fall classifications among staff in US hospitals. *Nursing Research*, *62*, 74-81.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Boca Raton, FL: CRC Press.