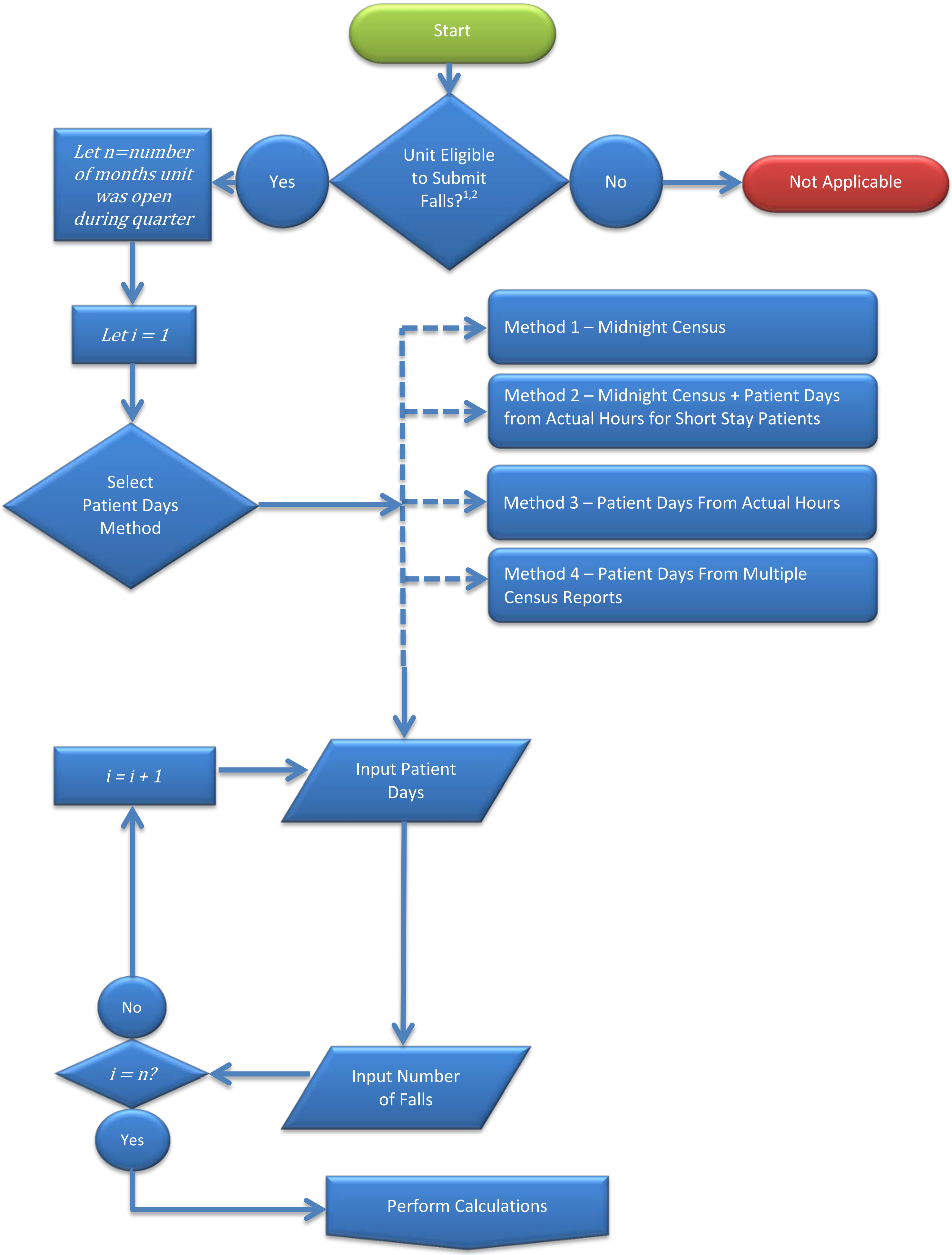
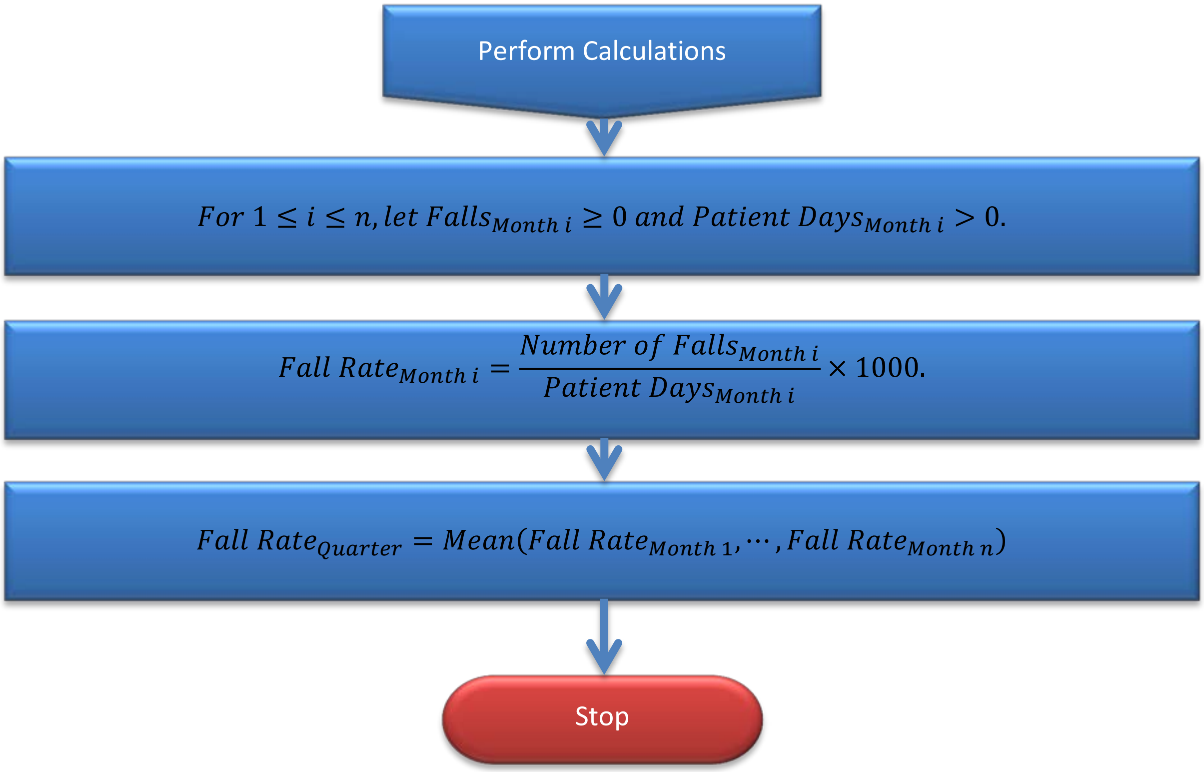
Fall Rate Algorithm/Measure Logic Flowchart



Fall Rate Algorithm/Measure Logic Flowchart



[1]Unit eligibility depends on NDNQI unit type designations. Eligible unit type designations are:

- Critical Care – Adult
- Step Down – Adult
- Medical – Adult
- Surgical – Adult
- Med-Surg Combined – Adult
- Critical Access
- Rehab – Adult

[2]Unit must have been open (patients present) at least 1 month during reporting period.

Measure #0141: Patient Falls

Scientific Supplement

4/10/2015

Table of Contents

Falls and Injury Falls Reliability Study

Final Report

By

Vincent Staggs, PhD

Emily Cramer, PhD

Executive Summary

We conducted two studies related to the scientific acceptability of the NDNQI total and injury fall rates. In the first study we assessed the reliability of these rates using a signal-to-noise analysis. Average unit-specific reliabilities, which are functions of a unit's fall rate and patient days, varied by unit type and ranged from 0.61 to 0.81. The proportion of variation in observed fall rates attributable to true differences among units ranged from 0.21 to 0.75. These measures of reliability cannot be used to determine the extent to which the two fall rates are measuring what they are intended to measure. Rather, they are measures of signal (true differences among unit fall rates) relative to expected randomness and other noise.

In the second studied under-reporting by surveying RNs about the proportions of falls reported on their unit. Based on median unit responses, the percentage of falls reported on NDNQI units exceeds 94% regardless of injury outcome. We found no evidence that non-injurious falls are reported less frequently than injurious falls.

Study 1

The purpose of this study was to assess the reliability of the NDNQI falls indicator by examining variation in annual rates of total and injurious falls per 1,000 patient days. We extracted 2013 falls data for six unit types: critical care, step-down, medical, surgical, medical-surgical, and rehabilitation. For each unit type we estimated the average reliability (across units) and intra-class correlation coefficient (ICC) for the annual total fall rate and annual injurious fall rate as described below.

The NDNQI calculates the annual total (injurious) fall rate for a unit by dividing the unit's count of total (injurious) falls for the year by its patient days for the year and multiplying the result by 1,000. More formally, let y_i be the observed annual count of total (injurious) falls for the *i*th unit, and let d_i be the unit's annual count of patient days divided by 1,000. The observed annual fall rate is given by $r_i = y_i/d_i$ (falls per 1,000 patient days).

The variation in observed annual fall rates can be divided into two components. A certain amount of variability is expected within each unit due to the randomness inherent in falls as a phenomenon. For example, a unit with a "true" fall rate of 3 falls per 1,000 patient days may not have 3 falls per 1,000 patient in a given 12-month period. This random variation does not provide any information about true fall rates differences among units and can be thought of as error variance or noise. The remaining variability, or signal, can be attributed to between-unit differences in fall rates. In equation form, we have:

Total variance = between-unit variance + error variance = signal + noise

Reliability can be defined as the proportion of total variation in the fall rate that is due to true differences among units, or equivalently, the proportion of total variation attributable to signal:

Reliability = between-unit variance / total variance = signal / (signal + noise)

Under the assumption that fall counts follow a Poisson distribution, we can use observed fall rates to estimate reliability as follows. Much of the statistical framework is from a method proposed by Hayes and Bennett[1] for estimating the coefficient of variation for cluster randomized trials in which the outcome is an incidence rate. Here we apply the method to fall rates, treating nursing units as clusters.

Let Y_i be a random variable denoting the count of falls for the *i*th unit during the time period of interest (one year in this study). Let λ_i be the *i*th unit's "true" fall rate, sampled from a distribution with mean λ (the "true" fall rate in the population of all units) and variance σ_λ^2 (the between-unit variance in "true" fall rates). We assume that Y_i, conditional on λ_i, has distribution Poisson($\lambda_i d_i$). It follows that the *i*th unit's fall rate in the time period of interest, Y_i/d_i, has conditional mean $E(Y_i/d_i \mid \lambda_i) = \lambda_i$ and conditional variance $Var(Y_i/d_i \mid \lambda_i) = \lambda_i/d_i$. This conditional variance is the random variation we would expect for a unit with "true" fall rate λ_i based on the randomness in Poisson counts. We estimate λ_i with the observed fall rate, r_i. The mean fall rate in the population, λ, is estimated by the observed overall fall rate, say r, computed as 1,000 times the sum of falls across units divided by the sum of patient days across units.

Unit fall rate reliability

The reliability of the fall rate for a given unit can be defined as $\sigma_\lambda^2/ (\sigma_\lambda^2 + \lambda_i/d_i)$. This measure takes a value on [0, 1] based on the size of the between-unit variance (signal) relative to the size of the error variance (noise). The reliability decreases as error variance increases relative to between-unit variance and approaches one as λ_i goes to zero or as d_i goes to infinity. Because the denominator in the formula is not the total variance in the fall rate, this measure should not be interpreted as the proportion of total fall rate variance accounted for by between-unit fall rate differences. It is simply a measure of signal relative to the specific unit's noise.

Estimating reliability requires estimates of λ_i and σ_λ^2. As noted above, we estimate λ_i with r_i. Following Hayes and Bennett,[1] we estimate σ_λ^2 as follows. We set s^2 equal to its expected value, yielding $s^2 = \sigma_\lambda^2 + [\lambda \times Ave(1/d_i)]$, where Ave() indicates the arithmetic mean. Then an estimate for σ_λ^2 is given by $s^2 - [r \times Ave(1/d_i)]$. Given this estimate we can estimate each unit's fall rate reliability using the formula given above. For a sense of the overall reliability of the falls measure we average reliability across units.

This method of estimating reliability is similar to the method outlined by Adams[2] for estimating the reliability of provider profile scores using the Gaussian linear mixed model. One difference is that under Adams' approach, σ_λ^2 is estimated using a mixed model with a random unit intercept and random error terms with known variances. In this study the normality assumption was inappropriate, and the assumption of independence between the random intercepts and random errors was questionable given that we estimated the error variances using the observed fall rates, so we opted for a simpler method requiring fewer assumptions.

Intra-class correlation coefficient

The intra-class correlation coefficient (ICC) can be defined as the proportion of total fall rate variance accounted for by difference among units in "true" fall rates. Taking the sample variance, s^2, as an estimate of total variance, we can estimate the ICC by dividing our estimate of σ_λ^2 by s^2. Unlike the unit-specific reliability measure defined above, the ICC is a single value characterizing the population.

Results

Table 1.

			Total fall rate			Injury fall rate		
Unit type	Units	Average patient days	Mean (SD)	Average reliability	ICC	Mean (SD)	Average reliability	ICC
Critical care	2,727	3,802	1.15 (0.98)	.64	.49	0.23 (0.35)	.65	.21
Step-down	1,856	6,643	2.90 (1.45)	.73	.63	0.64 (0.54)	.61	.39
Medical	2,371	7,791	3.45 (1.55)	.77	.72	0.77 (0.59)	.68	.57
Surgical	1,673	6,754	2.62 (1.43)	.76	.70	0.54 (0.53)	.70	.55

Medical-surgical	3,074	7,053	3.18 (1.99)	.76	.68	0.71 (0.55)	.62	.44
Rehabilitation	654	5,577	5.97 (2.86)	.81	.75	1.02 (0.84)	.68	.51

Discussion

It is important to note what the unit-specific reliability metric in this study is designed to measure, and what it is not designed to measure. The reliability estimates here reflect our ability to differentiate among units based on their true fall rates in the presence of noise, or as Adams[2] stated in the context of physician profiling, "Fundamentally, reliability is the measure of whether you can tell one physician, from another" (p. 10).

Much of the noise in fall rates is inherent in the randomness of the Poisson process and, as such, is unavoidable. It is clear from the reliability formula that a unit's reliability is a function of its fall rate and patient days. All else being equal, reliability decreases as the fall rate increases, and increases as the number of patient days increases.

The reliability metric here is *not* a measure of validity—i.e., it cannot be used to determine the extent to which the two fall rates are measuring what they are intended to measure, nor does it tell us how often falls are under-reported or whether nurses on different units define and report falls in a consistent way. Adams[2] provides a more detailed discussion of reliability in the context of quality measurement in healthcare.

The ICC estimate for each unit type reflects the size of between-unit variance as a proportion of total observed variance in fall rates. The estimated ICC can be expressed as $1 - [r \times \mathrm{Ave}(1/d_i)]/s^2$. Unit types with a higher average fall rate (r) or lower fall rate standard deviation

(s) tend to have lower ICC estimates, as do unit types (such as critical care) that have fewer patient days per unit (resulting in higher values of $Ave(1/d_i)$).

Study 2

The purpose of this study was to investigate reporting of falls in NDNQI hospitals by surveying RNs about the frequency with which falls of various types are reported on their unit. Following Blegen et al.,[3] we wrote five items for inclusion in the NDNQI RN Survey for administration in fall of 2013. The items were as follows:

1. Estimate the proportion of all patient falls actually occurring on the unit that were reported.

2. Of the falls occurring on your unit over the past 3 months that **did not result injury**, what proportion were reported?

3. Of the falls occurring on your unit over the past 3 months that resulted in **minor** injury (for example, pain, bruise, or abrasion), what proportion were reported?

4. Of the falls occurring on your unit over the past 3 months that resulted in **moderate** injury (for example, muscle/joint strain or an injury requiring suturing, skin glue, or splinting), what proportion were reported?

5. Of the falls occurring on your unit over the past 3 months that resulted in **major** injury (for example, internal injury, fracture, or an injury requiring surgery, casting, traction, or neurological consultation), what proportion were reported?

Respondents were presented the following response options for each item:

1. 0-10%
2. 11-20%
3. 21-30%
4. 31-40%
5. 41-50%
6. 51-60%
7. 61-70%
8. 71-80%
9. 81-90%
10. 91-100%

Data on these RN Survey items were collected from all units eligible to report falls to the NDNQI, including units serving ambulatory, psychiatric, pediatric, and neonatal populations. Thus the data collected reflect reporting of all falls meeting the NDNQI fall definition, including baby and child drops and developmental falls with injury. Because intentional fall events do not meet the NDNQI fall definition, they are not included in total or injury fall counts; hospitals report suspected intentional fall events separately, and the NDNQI tracks these events with a separate rate. However, the Survey item instructions did not specifically address intentional falls, so it is possible that responses reflect not only reporting of falls, but also, to a very limited degree, reporting of suspected intentional fall events.

Sample

The sample was limited to respondents from U.S. hospital units for which responses to the Survey falls items were available from at least 5 RNs and at least 25% of the eligible RNs on the unit, and for which falls data were reported for quarters 2-4 of 2013. Responses to the falls items were available from 58,514 RNs on 2,596 units meeting these inclusion criteria. We removed data for two groups of respondents. First, we excluded 448 respondents (0.8%) who indicated that the percentage of total falls reported on their unit was 91-100%, but the percentage reported in each of the four subcategories (non-injury, minor injury, moderate injury, and major injury) was in the 0-10% range. Then we excluded 74 respondents (0.1%) who chose 0-10% for total falls reported on their unit but 91-100% for each of the four subcategories.

As a result of these exclusions, 4 of the original 2,596 units dropped below the minimum of 5 responses and were removed from the sample. The final sample comprised 57,870 RNs (98.9% of the original 58,514) from 2,592 nursing units in 365 hospitals. The average unit response rate (percent of unit RNs responding) was 74.5%.

Analysis

During the study respondents reported some confusion about how to respond for a period in which there were no falls, or no falls of a particular type—cases in which the proportion of falls reported is meaningless because its denominator is zero. Apparently many respondents chose the 0-10% response in these cases. Many units had a bimodal distribution of 0-10% responses and 91-100% responses to the total falls item, with relatively few responses in between. On average, 88.2% of a unit's respondents chose 91-100%, 2.4% chose 81-90%, and 6.4% chose 0-10%, leaving only 3.0% of responses for the other 7 response options combined. For 1,065 units (41.1%), at least 50% of respondents chose 91-100% but at least 5% chose 0-10%.

As further evidence that some RNs chose the 0-10% option to indicate no falls, there was a significant negative correlation between a unit's 2013 fall rate and the proportion of its respondents choosing 0-10% for the total falls item (Spearman $r_s = -0.33$, p-value < .001). Consistent with this finding, respondents from the four neonatal unit types, which had the lowest 2013 total fall rates, were much more likely to choose the 0-10% option for total falls than respondents from unit types with higher fall rates. The percentage of neonatal unit respondents choosing 0-10% ranged from 21.2% to 33.8%, whereas the average percentages on adult medical, surgical, and rehabilitation units, respectively, were 4.5%, 6.0%, and 2.8%. This is the pattern one would expect if respondents were choosing 0-10% to indicate that no falls were observed during the time period in question.

Not surprisingly, the estimated intra-class correlation coefficient for responses to the total falls item was only 0.09, indicating that 91% of variability was due to within- rather than between-unit differences. Given the bimodal distribution of the data and the high within-unit variability, we used unit medians to assess fall reporting. The rationale was that nurses on each unit were all estimating the same parameters (percentages of falls reported on their unit), and the median provides an aggregate measure of their ratings that is not influenced by outliers.

For the purpose of estimating unit fall reporting rates we set each response equal to the midpoint of the corresponding interval (5 for 0-10%, 15.5 for 11-20%, etc.) and then found each unit's median response for each item. We computed means and standard deviations of the unit median responses for the entire sample and for each of the six highest-volume unit types: critical care, step-down, medical, surgical, medical-surgical, and rehabilitation. Together the six high-volume unit types accounted for 84.6% of patient days and 87.6% of total falls reported by NDNQI hospitals in 2013.

Results

Results are shown in Table 2. Because we converted responses to midpoints, the maximum possible unit median was 95.5%.

Table 2. Means and standard deviations for unit medians of percentages of falls reported.

		Fall item				
Unit type	Units	Total	No injury	Minor injury	Moderate injury	Major injury
All	2,592	94.8 (6.8)	94.7 (7.7)	94.5 (8.3)	94.3 (9.4)	94.2 (9.8)
Critical care	473	94.8 (6.8)	94.5 (8.2)	94.3 (9.4)	94.1 (9.8)	93.9 (11.3)
Step-down	336	95.0 (5.8)	94.7 (7.8)	94.7 (7.1)	94.7 (8.5)	94.6 (8.9)
Medical	403	95.3 (4.5)	95.0 (6.0)	95.1 (5.4)	94.8 (7.1)	94.7 (7.6)
Surgical	332	95.0 (6.6)	95.0 (6.6)	94.7 (8.0)	94.7 (8.3)	94.6 (8.7)
Medical-surgical	533	95.4 (2.0)	95.2 (4.0)	95.0 (5.7)	94.7 (7.5)	94.7 (7.5)

Rehabilitation	91	95.5 (0.0)	95.5 (0.0)	95.5 (0.0)	94.5 (9.5)	94.2 (10.0)

We also computed the percentage of respondents who chose the 91-100% option for each item, both for the total sample and for the six highest-volume unit types. As shown in Table 3, percentages tended to be slightly higher for the six specific unit types than for the sample as a whole. This is likely due to the higher frequency of 0-10% responses for low-fall unit types in the overall sample. Given respondents' confusion about how to respond when there were no falls on their unit during the period in question, the percentages shown are almost certainly underestimates.

Table 3. Percentage of unit respondents choosing 91-100% option.

Unit type	Units	Fall item				
		Total	No injury	Minor injury	Moderate injury	Major injury
All	2,592	87.9	84.8	82.9	82.5	82.8
Critical care	473	88.1	85.5	83.2	82.9	83.5
Step-down	336	88.8	85.6	83.7	83.5	83.7
Medical	403	89.3	85.8	84.2	83.3	83.5
Surgical	332	88.5	85.7	83.5	82.8	82.9
Medical-surgical	533	89.6	86.3	83.8	83.5	83.8
Rehabilitation	91	94.0	90.7	85.2	81.8	81.7

Discussion

Based on responses to survey items with a similar 10-interval response scale, Blegen et al.[3] estimated that 77% of total falls, 72% of non-injury falls, 77% of minor falls, and 80% of major falls are reported. Our results suggest that levels of reporting for NDNQI hospitals are substantially higher, exceeding 94% regardless of injury outcome. Given that the maximum response was 95.5% even for units reporting 100% of falls, true reporting rates may be even higher.

For a closer comparison with the results of the Blegen et al. study,[3] we could ignore the problem of respondents incorrectly choosing the 0-10% response when there were zero falls observed (the reason for our median-based approach) and simply estimate fall reporting rates by setting each response equal to the midpoint of the corresponding interval and taking the average. This yields reporting rates for the entire sample of 87.8%, 86.8%, 84.3%, 83.2%, and 83.0% for total, non-injury, minor injury, moderate injury, and major injury falls, respectively. Again, these numbers suggest higher reporting rates in NDNQI hospitals.

Unlike Blegen et al.,[3] we did not find any evidence that injury falls are more likely to be reported than non-injury falls. On the contrary, reporting rates appeared to be lower for injury than for non-injury falls. This may be another artifact of respondents choosing the 0-10% option when no falls of the given type were observed. Injury falls occur less frequently than non-injury falls, so one would expect a greater proportion of erroneous 0-10% responses, and thus lower reporting rates, for injury falls. Consistent with this hypothesis that 0-10% responses were more common for falls occurring with less frequency, reporting rates seemed to decrease with the severity of fall injury.

References

1. Hayes, R. J., & Bennett, S. (1999). Simple sample size calculation for cluster-randomized trials. International journal of epidemiology, 28(2), 319-326.

2. Adams J. L. (2009). The reliability of provider profiling: a tutorial. RAND Corporation, Santa Monica.

3. Blegen MA, Vaughn T, Pepper G, et al. Patient and staff safety: voluntary reporting. Am J Med Qual 2004;19(2):67-74.

Brief Literature Review: Prevalence and Significance of Inpatient Falls and Injury Falls

A patient fall is defined as an unintentional and sudden descent to the floor with or without injury to the patient (NDNQI, 2012). It is one of the most reported adverse events, and a leading cause of injury among adults in inpatient settings (Degelau et al., 2012). Patient falls are a serious healthcare problem that has a severe impact upon individuals, families, and healthcare systems (Dykes, Carroll, Hurley, Benoit, & Middleton, 2009). It leads to increased lengths of stay, higher number of malpractice lawsuits, and more than $4,000 in additional charges per hospitalization (Inouye, Brown, & Tinetti, 2009). On the patient level, a fall may result in injury and a fear of falling, which can lead to reduced mobility and eventually loss of function and further falls (Greenberg, 2011). Families and caregivers are also affected by patient falls as they hold the burden of caring for the patient (Kuzuya et al., 2006).

Hospital falls are prevalent, and affect a wide range of patients in different units. Observational studies revealed that the rate of hospital falls range between 3.3 to 11.5 falls/1000 patient/days [pd] (Currie, 2008; Fischer et al., 2005; Hitcho et al., 2004; Krauss et al., 2007; Quigley et al., 2009). These rates vary by the type of hospital unit (Mion, 2012). Higher rates of inpatient falls occur in geriatric, neurology and rehabilitation units (Degelau et al., 2012; Moin, 2012). A possible interpretation of this phenomenon is that patients of these units are more susceptible to falling due to having older age, longer hospital stay, and having a neurological condition. These factors contribute to added risk of falling among patients in hospitals (Nakai, Akeda, & Kawabata, 2006).

Furthermore, a recent study -used data collected from the NDNQI- showed that falls rate among 1263 hospitals was (3.56 /1000 pd) (Bouldin et al., 2013). About one quarter of the reported falls resulted in an injury (26.1% with a rate of 0.93/1000 pd). In another study, approximately 30% of hospital falls resulted in injury (Stevens, 2004). The highest number of incident falls occurred in the medical unit (4.03/1000 pd) and the lowest in the surgery unit (2.76/1000 pd; Bouldin et al., 2013).

Injury falls may result in fatal and non-fatal injuries ranging from minor lacerations to severe head injuries (WHO, 2012). The majority of fall-related injuries are non-fatal. The injury levels in falls are classified into 5 degrees: none, minor, moderate, major, and death (NQF, 2011). Injury falls lead to increased health care utilization and expenditure and thus overload the health care systems and providers. In 2008, the Centers for Medicare and Medicaid (CMS) declared that it would no longer reimburse hospitals for costs of hospitalized patient falls, which increases the financial burden on hospitals (Inouye, Brown, & Tinetti, 2009).

The financial burden of hospital falls is significant. It includes costs of care, length of stay, and legal liability. The costs of medical care for patient falls are high. A single fall without serious injury costs the hospital an average of additional $3,500 (Boushon et al, 2012). Whereas, falls with severe injury are the most costly for hospitals. It can cost $13,316 (Wong et al, 2011) and up to $ 27,000 (Boushon et al, 2012) of additional expenses. Although staggering, these numbers underestimate the actual magnitude of inpatient falls. It continues to cost millions of dollars each year not to mention the intangible costs due to reduced quality of life for both patients and their caregivers.

References

Bouldin, E. L., Andresen, E. M., Dunton, N. E., Simon, M., Waters, T. M., Liu, M., . . . Shorr, R. I. (2013). Falls among adult patients hospitalized in the United States: prevalence and trends. *J Patient Saf*, *9*(1), 13-17.

Boushon, B., Nielsen, G., Quigley, P., Rutherford, P., Taylor, J., Shannon, D., & Rita, S. (2012). How-to Guide: Reducing Patient Injuries from Falls. Cambridge, MA: Institute for Healthcare Improvement. Retrieved from http://www.ihconline.org/UserDocs/Pages/HowtoGuideReducingPatientInjuriesfromFalls_copy.pdf

Centers for Disease Control and Prevention, (2013): National Center for Injury Prevention and Control. Web–based Injury Statistics Query and Reporting System (WISQARS) [online]. Retreived from http://www.cdc.gov/injury/wisqars/

Currie, L. (2008). Fall and Injury Prevention. In R. G. Hughes (Ed.), Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville (MD).

Degelau, J., Belz, M., Bungum, L., Flavin, PL., Harper, C., Leys, K., Lundquist, L., Webb, B., Institute for Clinical Systems Improvement (ICSI) (2012). Prevention of falls (acute care). Health care protocol. Bloomington (MN): Institute for Clinical Systems Improvement (ICSI). Retrieved from http://www.qualitymeasures.ahrq.gov/content.aspx?id=36945

Dykes, P. C., Carroll, D. L., Hurley, A. C., Benoit, A., & Middleton, B. (2009). Why do patients in acute care hospitals fall? Can falls be prevented? *J Nurs Adm*, *39*(6), 299-304

Fischer, I. D., Krauss, M. J., Dunagan, W. C., Birge, S., Hitcho, E., Johnson, S., . . . Fraser, V. J. (2005). Patterns and predictors of inpatient falls and fall-related injuries in a large academic hospital. Infect Control Hosp Epidemiol, 26(10), 822-827. doi: 10.1086/502500

Greenberg, S., (2011). Assessment of Fear of Falling in Older Adults: The Falls Efficacy Scale-International (FES-I). Try This: Best Practices in Nursing Care to Older Adult, 29. Retrieved from http://consultgerirn.org/uploads/File/trythis/try_this_29.pdf

Hitcho, E. B., Krauss, M. J., Birge, S., Claiborne Dunagan, W., Fischer, I., Johnson, S., . . . Fraser, V. J. (2004). Characteristics and circumstances of falls in a hospital setting: a prospective analysis. Journal of General Internal Medicine, 19(7), 732-739. doi: 10.1111/j.1525-1497.2004.30387.x

Inouye, S. K., Brown, C. J., & Tinetti, M. E. (2009). Medicare nonpayment, hospital falls, and unintended consequences. *N Engl J Med*, *360*(23), 2390-2393.

Krauss, M. J., Nguyen, S. L., Dunagan, W. C., Birge, S., Costantinou, E., Johnson, S., . . . Fraser, V. J. (2007). Circumstances of patient falls and injuries in 9 hospitals in a midwestern healthcare system. Infect Control Hosp Epidemiol, 28(5), 544-550. doi: 10.1086/513725

Kuzuya, M., Masuda, Y., Hirakawa, Y., Iwata, M., Enoki, H., Hasegawa, J., . . . Iguchi, A. (2006). Falls of the elderly are associated with burden of caregivers in the community. *Int J Geriatr Psychiatry*, *21*(8), 740-745. doi: 10.1002/gps.1554

Mion, L., (2012). Patient Falls in the Hospital Setting: A Persistant Problem. Sizewise. Retreived from http://www.sizewise.net/getattachment/b4bf1db2-9799-44ee-bde606d7c491dcfc/Patient_Falls.aspx

Nakai, A., Akeda, M., & Kawabata, I. (2006). Incidence and risk factors for inpatient falls in an academic acute-care hospital. *J Nippon Med Sch*, *73*(5), 265-270

National Database for Nursing Quality Indicators, (2012). Guidelines for Data Collection on the American Nurses Association's National Quality Forum Endorsed Measures. P 13, Retrieved from http://www.nursingquality.org/Content/Documents/NQF-Data-Collection-Guidelines.pdf

National Quality Forum, (2011). Falls with injury: measure submission and evaluation worksheet. Retrieved from file:///C:/Users/ajaber/Downloads/0202.pdf

Quigley, P. A., Hahm, B., Collazo, S., Gibson, W., Janzen, S., Powell-Cope, G., . . . White, S. V. (2009). Reducing serious injury from falls in two veterans' hospital medical-surgical units. J Nurs Care Qual, 24(1), 33-41.

Stevens, J. A., (2004). Falls among older adults-risk factors and prevention strategies. In Falls Free: Promoting a National Falls Prevention Action Plan. Washington, DC: National Council on Aging.

WHO, (2012). Falls. Retreived from http://www.who.int/mediacentre/factsheets/fs344/en/index.html

Wong, C. A., Recktenwald, A. J., Jones, M. L., Waterman, B. M., Bollini, M. L., & Dunagan, W. C. (2011). The cost of serious fall-related injuries at three Midwestern hospitals. *Jt Comm J Qual Patient Saf*, *37*(2), 81-87.

*Evidence of the Impact of Hospital **Falls/Injury Falls** on Patients, Caregivers, and Health Care Systems*

Factor	Author (year)	Sample	Result
Vulnerable Populations	Fhon et al (2013)	240 elderly people (≥ 60 years)	Higher levels of falls were found among women, younger subjects (60 to 79 years). Risk factors also include previous falls in which 26.8% were victims of 1 to 2 falls.
	Lavedan (2014)	640 elderly people (≥ 75 years)	Basic disability (OR=2.17), depressive symptoms (OR=1.67) and fear of falling (OR=2.53) were associated with falls. Having these factors increases the risk of falling
	Oliver et al (2010)	Review	Common risk factors among inpatient population include a history of falling, muscle weakness, agitation and confusion, urinary incontinence or frequency, sedative medication, and postural hypotension
	Ranaweera et al (2013)	1200 people (≥ 65 years)	Risk factors for falls include falls in the previous year (OR 4.67), high disability level (OR 2.04) and high house risk level (OR 1.68).
	Lee & Stokic (2008)	1472 patients admitted to a rehabilitation center	Diagnosis of stroke and amputation, age between 41 and 50 yrs, lower cognitive FIM scores, and a large number of medical comorbidities (> or =9) were associated with a high risk for fall.
	Centers for Disease Control (2006)	Patients aged 65 and older treated in EDs between 1993-2003 and 2001-2005.	Sex disparities on fatal falls, overall fall rate, and hip fractures as the result of falls. Fall rates were higher for whites than for other races.
	Stevens & Sogolow (2005)	22,560 adults (≥ 65 years)	Fall rates (OR 2.3) as well as the rate of injury resulting and hospitalization (OR 2.3) from a fall were higher among women.
Financial Costs	Jorgensen (2011)	Report in a nursing journal	By 2020, the estimated direct and indirect costs of injuries related to falls will reach $54.9 billion.
	Boushon et al (2012)	IHI guide – reducing falls	A fall without serious injury costs the hospital an average of additional $3,500, whereas a fall with severe injury costs $27,000.

	Wong et al (2011)	57 hospital inpatients at 3 hospitals	Falls with severe injury costs the hospital $13,316 and up of additional expenses
	Heinrich et al (2010)	Systematic review	Mean costs per fall victim, per fall and per fall-related hospitalization ranged from 1,059 to 42,840 USD and depended on fall severity.
Health Outcomes	Stevens (2004)	Report/review	Approximately 30% of hospital falls result in injury
	Oliver et al (2010)	Review	About 1% - 3% of falls in hospitals result in fracture, but even minor injuries can cause distress and delay rehabilitation
	Bouldin et al (2013)	NDNQI data from 1263 hospitals across the United States between 2006-2008	26.1% of reported hospital falls resulted in an injury
	Fhon et al (2013)	240 elderly people (≥ 60 years)	54.2% of subjects who fell suffered scratches and 78% were afraid of suffering a new fall.
Caregivers	Faes et al (2010)	10 fall patients and 10 caregivers.	Caregivers reported a fear of their care recipient falling. Some caregivers rated the consequences of their CRs' cognitive problems as more burdensome than their falls and believed that a prevention programme would not be useful because of the CRs' cognitive impairment, physical problems, age and personalities.
Length of stay	Fitzpatrick (2011)	Report in a nursing journal	Injury falls lead to as much as a 61% increase in patient-care costs and lengthen a patient's hospital stay
	Dunne et al (2014)	292 patients with 330 controls in an acute care hospital	Patients who had an in-hospital fall had an average length of stay of 37.2 days compared to 25.7 days in the controls who did not fell.
Hospital Unit	Patman et al., (2011)	190 patient who survived their acute care stay	Thirty-two patients (17%) fell at least once on the in-patient wards following their ICU stay
	Lee & Stokic (2008)	1472 patients admitted to a rehabilitation center	One hundred forty patients (9.5%) fell at least once during their inpatient rehabilitation stay in a tertiary care rehabilitation center

Epidemiology (Incidence and prevalence rates)	Krauss et al (2007)	7,082 inpatients who fell in 9 hospitals between 2001-2003	Average falls rate (with and without injury) in hospitals is about 4.75 per 1000 patient days.
	Oliver et al (2010)	Review	The hospital fall rates are between 3 and 5 falls per 1000 bed-days.
	Bouldin et al (2013)	NDNQI data from 1263 hospitals across the United States between 2006-2008	Falls rate among 1263 hospitals was 3.56 per 1000 patient days. The highest number of incident falls occurred in the medical unit (4.03/1000 pd) and the lowest in the surgery unit (2.76/1000 pd)
Prevention and Intervention	Oliver et al (2010)	Review	The literature suggests that the most appropriate approach to fall prevention in the hospital environment includes multifactorial interventions with multiprofessional input. Single interventions that may reduce falls include delirium avoidance programs, reducing sedative and hypnotic medication, in-depth patient education, and sustained exercise programs.

References

Bouldin, E. L., Andresen, E. M., Dunton, N. E., Simon, M., Waters, T. M., Liu, M., . . . Shorr, R. I. (2013). Falls among adult patients hospitalized in the United States: prevalence and trends. *J Patient Saf, 9*(1), 13-17.

Boushon, B., Nielsen, G., Quigley, P., Rutherford, P., Taylor, J., Shannon, D., & Rita, S. (2012). How-to Guide: Reducing Patient Injuries from Falls. Cambridge, MA: Institute for Healthcare Improvement. Retrieved from http://www.ihconline.org/UserDocs/Pages/HowtoGuideReducingPatientInjuriesfromFalls_copy.pdf

Centers for Disease Control (CDC; 2006). Fatalities and injuries from fall among older adults – United States, 1993-2003 and 2001-2005. *Morbidity and Mortality Weekly Report*, 55(45), 1221-1224.

Degelau, J., Belz, M., Bungum, L., Flavin, PL., Harper, C., Leys, K., Lundquist, L., Webb, B., Institute for Clinical Systems Improvement (ICSI) (2012). Prevention of falls (acute care). Health care protocol. Bloomington (MN): Institute for Clinical Systems Improvement (ICSI). Retrieved from http://www.qualitymeasures.ahrq.gov/content.aspx?id=36945

Dunne, T. J., Gaboury, I., & Ashe, M. C. (2014). Falls in hospital increase length of stay regardless of degree of harm. *J Eval Clin Pract, 20*(4), 396-400. doi: 10.1111/jep.1214

Faes, M. C., Reelick, M. F., Joosten-Weyn Banningh, L. W., Gier, M., Esselink, R. A., & Olde Rikkert, M. G. (2010). Qualitative study on the impact of falling in frail older persons and family caregivers: foundations for an intervention to prevent falls. *Aging Ment Health, 14*(7), 834-842. doi: 10.1080/1360786100378182

Fhon, J. R., Rosset, I., Freitas, C. P., Silva, A. O., Santos, J. L., & Rodrigues, R. A. (2013). Prevalence of falls among frail elderly adults. *Rev Saude Publica, 47*(2), 266-273. doi: 10.1590/s0034-8910.2013047003468

Fitzpatrich, M. A., (2011). Meeting the challenge of falls reduction. American Nurse Today (Supplement), March 2011, P 1.

Heinrich, S., Rapp, K., Rissmann, U., Becker, C., & Konig, H. H. (2010). Cost of falls in old age: a systematic review. *Osteoporos Int, 21*(6), 891-902. doi: 10.1007/s00198-009-1100-1

Jorgenson, J., (2011). Meeting the challenge of falls reduction. American Nurse Today (Supplement), March 2011, P 2-3.

Krauss, M.J., Nguyen, S.L., Dunagan, W.C., Birge, S., Costantinou, E., Johnson, S., et al. (2007). Circumstances of patient falls and injuries in 9 hospitals in a midwestern healthcare system. Infection Control and Hospital Epidemiology, 28(5), 544-550.

Lavedan Santamaria, A., Jurschik Gimenez, P., Botigue Satorra, T., Nuin Orrio, C., & Viladrosa Montoy, M. (2014). [Prevalence and associated factors of falls in community-dwelling elderly.]. *Aten Primaria*. doi: 10.1016/j.aprim.2014.07.012

Lee, J. E., & Stokic, D. S. (2008). Risk factors for falls during inpatient rehabilitation. *Am J Phys Med Rehabil, 87*(5), 341-350; quiz 351, 422. doi: 10.1097/PHM.0b013e31816ddc01

Oliver, D., Healey, F., & Haines, T. P. (2010). Preventing falls and fall-related injuries in hospitals. *Clin Geriatr Med, 26*(4), 645-692. doi: 10.1016/j.cger.2010.06.00

Patman, S. M., Dennis, D., & Hill, K. (2011). The incidence of falls in intensive care survivors. *Aust Crit Care, 24*(3), 167-174. doi: 10.1016/j.aucc.2011.06.001

Ranaweera, A. D., Fonseka, P., PattiyaArachchi, A., & Siribaddana, S. H. (2013). Incidence and risk factors of falls among the elderly in the District of Colombo. *Ceylon Med J, 58*(3), 100-106. doi: 10.4038/cmj.v58i3.5080

Stevens, J. A. (2004). Falls among older adults-risk factors and prevention strategies. *In falls free: Promoting a national falls prevention action plan* (pp.318). Washington, DC: National Council on Aging.

Stevens, J.A. & Sogolow. (2005). Gender differences for non-fatal unintentional fall related injures among older adults. *Inju ry Prevention*, 11, 115-110. Doi: 10.1136/ip.2004.005835.

Wong, C. A., Recktenwald, A. J., Jones, M. L., Waterman, B. M., Bollini, M. L., & Dunagan, W. C. (2011). The cost of serious fall-related injuries at three Midwestern hospitals. *Jt Comm J Qual Patient Saf, 37*(2), 81-87.

National Database of Nursing Quality Indicators®

Methods Development Project

Final Report to

The American Nurses Association

August 3, 2011

Prepared by the National Database of Nursing Quality Indicators

School of Nursing

University of Kansas Medical Center

Executive Summary

NDNQI staff have developed methodologies for bringing mixed acuity units into the database for reporting on clinical and staffing indicators and for creating hospital-level indicators. The process used in the development of these methodologies involved literature review, advice from methodological experts, new data collection, and statistical simulation. Key findings included:

- There were many barriers to creating risk or acuity-adjusted unit-level measures.
    - Acuity data are not available for units.
    - Risk data for specific outcomes would require a large, and perhaps burdensome, increase in data collection and reporting for participating hospitals.
- A method for creating categories of mixed acuity units was developed based on Medicare billing days. The method has face and criterion validity.
- Six types of mixed acuity units were identified for adult or pediatric populations:
    - Mixed Acuity III: Units with at least 50% critical care patient days each month.
    - Mixed Acuity II: Units with at least 25% critical care days each month or at least 50% step-down days each month; includes only units not meeting the criterion for Mixed Acuity III.
    - Mixed Acuity I: Units not meeting the criteria for Mixed Acuity III or II.
    - Burn units
    - Bone Marrow Transplant units
    - Critical Access units
- Rolling up unit-level indicators should take into account the unit composition of a hospital, as well as the size of each unit. Results should be meaningful to users.
- Hospital level indicators can be developed from a weighted average or unit-level z-score. The weights are based on the staffed bed size of each unit. For hospital reporting, the z-scores are translated back into the original metric of the indicator.
- Data collection and reporting for mixed acuity units and reporting for hospital-level indicators may be implemented at the direction of the American Nurses Association.

Background and Objectives

The National Database of Nursing Quality Indicators® (NDNQI®) is a proprietary database of the American Nurses Association (ANA). NDNQI collects and evaluates nursing-sensitive data from 1,800 hospitals in the United States

Hospitals that participate in NDNQI receive quarterly unit-level comparative data reports. NDNQI comparative data are currently stratified by unit type and hospital characteristics. Thus, each unit can compare its nurse staffing and patient outcomes to national percentiles for similar units. Hospitals value the NDNQI reports and want to have all of their units included. Mixed acuity units are currently excluded from NDNQI as they are too diverse to be included in comparative data. NDNQI has identified a methodologically and conceptually sound method of acuity adjustment or acuity stratification to allow inclusion of mixed acuity units in NDNQI reports. To date, most risk or acuity adjustment work has been done at the hospital level, using patient characteristics, diagnostic related groups, or the hospital case mix index.

NDNQI also was asked to create a methodology to produce hospital-level measures. Hospital executives want summary measures of performance for their entire facility. Consumers and oversight organizations want information to make conclusions about a facility's nursing quality. Typically, hospital-level reporting is either based on patient-level data or weighted averages of unit-level data. The patient-level approach is divergent from NDNQI's unit-based data collection. Incorporating the unit orientation into the measurement model is consistent with NDNQI's focus on nursing unit performance. A methodology was needed that was based on unit level data and takes into account the fact that hospitals may not submit data on all eligible units and hospitals vary in unit composition.

The methods development project had two main goals:

1. Develop a method for unit-level acuity adjustment of nursing indicators
2. Develop a method to calculate hospital-level indicators

Development Process

The project originally had two goals:

1. Develop adjustment methods for unit data on nurse staffing and patient outcomes.
2. Develop a method for rolling-up unit-level data to hospital totals.

The requirements for adjustments and roll-up methodologies that are consistent with the NDNQI paradigm include:

1. The nursing care unit should continue to be the unit of observation.
2. Post-adjustment results should be interpretable by clinicians
3. Data used in adjustment must be comparable across sites.
4. If adjustment is to occur with every reporting period, data must be available for update on a quarterly basis and within 45 days of the end of a calendar quarter.
5. Within a group of effective methods, the method requiring the least additional respondent burden should be selected.
6. Adjustment and summary methods should have attributional and face validity.

The development process consisted of four phases:

1. Literature review
2. Advisory panel input
3. Data collection
4. Statistical methods development

The report concludes with a description of a work order for implementation of the adjustment and roll-up methodologies.

Literature Review

The project began with a literature review of risk and acuity adjustment as they might apply to the inclusion of mixed acuity units into NDNQI reporting. It also covered the development of hospital-level measures that maintain NDNQI's unit-level focus. The review identified the following information:

Adjusting Staffing Measures

- Risk adjustment is a statistical method that accounts for patient characteristics known to the correlated with a particular outcome measure. In general, outcome measures are adjusted for risk.
- Acuity adjustment refers to the use of a composite variable reflecting the level of patients' need for care. In general, staffing measures are adjusted for acuity.
- Adjustment variables that are themselves measures of quality should be excluded in order to avoid over-adjustment.
- There was disagreement in the literature on which specific variables to use for adjustment.
    - A hospital's case mix index is the average of its patients' diagnosis-related group (DRG) weights. DRG weights reflect the relative resource consumption. Each patient is assigned one DRG for their entire inpatient stay. The level of measurement for DRGs is hospital, not nursing care unit. Further, the extent to which DRG weights capture *nursing* resource consumption is questionable. DRG weights do not explicitly account for variation in need for daily nursing.
    - Nursing intensity weights (NIW) were developed as a nursing-specific refinement of the DRG payment system. Nursing intensity weights reflect the relative level of nursing care needed by a typical patient in each diagnosis-related group. NIW, being based on DRGs, are at the hospital-level, rather than the nursing care unit-level.
    - Numerous proprietary software programs classify patients' severity of illness or project a unit's nursing workload. There is no predominant patient classification system in use across the U.S., so comparable data are not available.

Adjusting Patient Outcomes

- Outcome-specific adjustments are needed for pressure ulcers, falls, and other patient outcomes. Published examples of adjusted outcomes typically rely on patient-level measures of risk specific to each outcome, as opposed to more general illness severity or acuity measures.
- Adjusting outcomes based on patient characteristics would place a large data collection burden on hospitals that do not yet have electronic health records. Further, the data are protected health information which hospitals may be reluctant to release. Finally, the evidence for the use of PHI data producing effective adjustments was limited.

Hospital-Level Reporting

- Hospital level reporting typically is done by counting all of the various conditions or events that occur in the hospital and then dividing by the number of patients, days, device days, etc. across the hospital.
- This approach is patient-focused and abandons the unit-performance perspective that is the primary unit of analysis for NDNQI.
- Minnick (2000) noted that nursing-sensitive indicators typically vary more across units in the same hospital than across hospitals. Hospital summary measures of pressure ulcers, patient falls, and other indicators that mask variation between units within the organization are not as meaningful as unit-type comparisons across hospitals.
- In 2009, the National Quality Forum's report on Voluntary Consensus Standards for Nursing-Sensitive Care: An Initial Performance Measure Set

Expert Panel

ANA established an expert advisory panel to provide guidance on the data and methodological issues of adjustment and roll-up. Panelists brought forth a variety of issues, but no consensus was established on methodological approach. The following issues were identified.

- NDNQI's unit-based approach may not provide stable measures, much less enough data for unit-level risk adjustment.
- Risk adjusters must have high levels of sensitivity and specificity. More research would need to be done to identify which, if any, adjusters have these measurement properties.

- Nursing intensity weights are based on AP-DRGs, which reflect the care provided for an entire hospital stay, not the care provided on a specific unit. Current NIW have 2-3 levels within each DRG to distinguish critical vs. acute care. DRGs reflect patients at their sickest during the hospital stay.
- Use of patient classification systems may eventually be mandated nationally, but they vary across sites and are generally not well validated.
- Admissions-Transfer and Discharge data is another important predictor of workload that varies widely across units.
- Data for risk adjustment should be based on admission risk assessments. DRGs are not admission-based.
- Mixed acuity units use billing levels that correspond to existing NDNQI unit type strata. Validity of unit type as a proxy for acuity is uncertain.
- Rolling up different sized units in a hospital creates a weighted average, which complicates statistical issues. Some units will have a higher put-through than others so their average occupancy rate will be lower but the total number of people seen might be high.
- The statistical issues for dealing with weighted averages are pretty straight forward but the average needs to be interpretable.
- Any adjustment method will make mistakes. How high you set the standard for accuracy is subjective. Public reporting requires a higher standard than internal use for quality improvement.

Revised Study Purpose

After the completion of the discovery portion of the project, NDNQI researchers concluded that the original scope of work should be modified to reflect the information from the literature review and expert panel.

Due to questions about feasibility and advisability of various approaches to adjustment, we limited our objective to the primary goal of our adjustment work, namely to have a way to make meaningful comparisons of mixed acuity units so that these units can begin submitting data on NDNDQI clinical and staffing indicators.

The method we propose for hospital roll-up is a straightforward weighting technique. The original scope of work stated that we would collect data from a variety of hospitals to test alternative approaches. Because there are a wide variety of unit profiles within hospitals and unpredictable patterns of non-reporting, the sample size required to test all combinations of circumstances would be very large. Yet, even a large sample might not encounter all possible situations. A more satisfactory, and efficient, approach was to conduct the development tests with simulated data. It is quite feasible to vary systematically the composition of unit profiles and patterns of missing data.

Results

Mixed Acuity Units

The purpose of this study was to develop acuity adjustment methods to allow meaningful comparisons of mixed acuity units' patient outcomes, and to compare the feasibility and effectiveness of these methods using data collected from a sample of NDNQI hospitals. Due to questions about the feasibility and advisability of various approaches to adjustment (discussed below), we limited our objectives to the primary goal of finding a way to make meaningful comparisons of mixed acuity units so that these units can begin submitting data on NDNDQI quarterly indicators. During the discovery activities for unit-based acuity adjustments a variety of issues were uncovered. We concluded the following:

- Adjustment of staffing measures is not warranted. Staffing measures themselves are a measure of acuity, so adjusting them for acuity is analogous to adjusting SAT scores for test-taker IQ.
- Device-related infection measures are based on at-risk patients and are thus already adjusted for acuity to a large degree.
- Gathering patient clinical admission data for use in risk or acuity adjustment would represent a sizeable data collection burden for hospitals, as well as a substantial expansion in the scope of NDNQI's data collection. Even if certain clinical variables proved useful for adjusting NDNQI outcome measures, any adjustment method based on these variables could be used only for hospitals willing and able to provide these data to NDNQI; other hospitals' mixed acuity units would continue to be excluded from participation in quarterly indicators. Moreover, a patient's risk and acuity often change over the course of the hospital stay due to surgery, healthcare-acquired infection, effectiveness of treatment, etc.
- It was argued in the first Advisory Council meeting that one-day prevalence measures, which are unstable because they are based on data for only one day out of every 90, should not undergo unit-level risk adjustment. We disagree that the instability of these measures is a reason not to adjust them for risk or acuity, but we advise against adjustment using patient risk or acuity variables for the reasons stated above.

The mixed acuity unit study is described below. We propose creating new acuity-based unit types using the proportions of patient billing days at the critical, step-down, and standard levels of care, as well as comparison groups based on unit sub-specialty for unit types for which this is practicable.

NDNQI member hospitals with at least one mixed acuity unit enrolled were invited to participate in a survey on a volunteer basis. No specific sample size was targeted, as the study does not involve statistical inference. Hospitals submitted data for March, 2011, on nursing care hours, patient days, and (in the case of adult units) patient falls. Data were also submitted for a single pressure ulcer prevalence survey conducted in 2011 (prior to the data submission deadline of May 6).

In addition, hospitals reported for each mixed acuity unit the number of patient days billed by level of care/acuity for March, 2011. NDNQI sent a customized Excel spreadsheet to each participating site coordinator for recording these data. Patient days were classified using the following billing levels:

1. Critical/intensive care (highest level of care).
2. Step-down/intermediate/progressive care. This may also be called transitional care.
3. Standard/routine care (e.g. medical or surgical care).
4. Rehabilitation care.
5. Skilled nursing/sub-acute care.
6. Inpatient hospice care. This level applies only to patients who have been discharged from acute care.
7. Short-stay observational care. This includes 23-hour observational care.

Adult Mixed Acuity Units. Patient days by billing level were reported by 52 adult mixed acuity units. Critical, step-down, and standard care days made up the bulk of the days for these units. Five units reported at least one hospice or rehabilitation day, but these days did not make up more than 1.1% of the total patient days for any unit. One unit reported skilled nursing days, which accounted for 27.8% of its total days, the remainder being standard care days. Most units billed one or more short stay days; the average percentage of short stay days was 2.3%.

All 52 units billed at least 45% of their patient days at a single level, and 48 billed over 50% of their days at one level. Three units met NDNQI's 90% criterion for an existing unit type for the month of the study. For all 52 units, one or two billing levels accounted for at least 75% of patient days. Some units combined primarily critical and step-down care days, some combined step-down and standard care days, and some combined critical and standard care days (in some cases with no step-down days; some hospitals do not have a billing level between critical and standard care).

Pediatric Mixed Acuity Units. Of the ten pediatric mixed acuity units that submitted data on patient billing days, nine billed at least 60% of their patient days at either the critical, step-down, or standard care level. No rehabilitation, skilled nursing, or hospice days were billed. Short stay days accounted for less than 8% of patient days for all units save one, for which they accounted for 36%.

Other Unit Types. Four critical access units submitted data for the study. Their proportions of days at the critical, step-down, standard, skilled nursing, and short stay levels varied widely. Data were also submitted for one bone marrow transplant unit, one neonatal mixed acuity unit, and one obstetrics unit.

Approaches to Acuity Adjustment

Adjusting Outcomes Using Patient Days by Billing Level. If NDNQI collected data on patient days at various billing levels from all participating units, these data could be used to adjust each unit's scores on the NDNQI measures using an appropriate regression model. Each measure could be regressed on the proportions of patient days billed at the critical, step-down, and standard levels to obtain a predicted value on the measure for each unit, and this could be compared to the unit's observed value on the measure to compute an acuity-adjusted score. This would require regular collection of data on billing days so that adjusted scores could be reported to hospitals each month/quarter.

There are two problems with this approach. First, not all hospitals use the same billing levels, or even the same number of billing levels, and reconciling the various sets of levels would be difficult. And second, it would be burdensome for hospitals would submit these data on a regular basis.

Using RNHPPD as a Proxy for Acuity. Another option is to use RN hours per patient day (RNHPPD)—a variable on which NQNDI already collects data—as a proxy for acuity to classify units and/or compute acuity-adjusted scores on NDNQI measures. In an internal study conducted in 2010 using four quarters of data from critical, step-down, medical, surgical, and medical/surgical units, RNHPPD was shown to contain much of the same information as unit type. As shown in Table 1, critical, step-down, and standard care units differ markedly in RNHPPD. Using a general linear model, unit type was found to account for 82% of the variation in RNHPPD. Moreover, in linear models in which unit type was included as a predictor of fall rate or rate of hospital-acquired pressure ulcers, the addition of RNHPPD as a predictor resulted in virtually no change in the proportion of variance accounted for by the models.

Table 1

RN Hours per Patient Day by Unit Type (Data from 2009-2010)

Unit Type	Mean	SD
Critical Care	15.3	*2.6*
Step Down	7.7	*1.9*

Med	5.8	*1.4*
Surg	6.0	*1.4*
Med-Surg	5.9	*1.5*

There is further evidence from the present study that RNHPPD is linked to patient acuity. As part of the study, units submitted data on three staffing variables—RNHPPD, total nursing care hours per patient day (TNHPPD), and skill mix (percentage of nursing care hours provided by RNs). The correlations between these variables and the proportions of patient days billed at the critical, step-down, standard, skilled nursing, hospice, and short stay levels are shown in Table 2. Both RNHPPD and TNHPPD were positively correlated with the proportions of days at the critical and hospice care levels and negatively correlated with the proportions at the standard, skilled nursing, and short stay levels. Correlations with skill mix were generally much weaker.

Table 2

Correlations of Staffing Variables with Proportions of Patient Days at Billing Levels (N = 59)

Billing Level	Critical Care	Step-down	Standard	Skilled Nursing	Hospice	Short Stay
Correlation with RNHPPD	0.73	0.11	–0.59	–0.42	0.38	–0.37
Correlation with TNHPPD	0.65	0.17	–0.56	–0.44	0.25	–0.40
Correlation with Skill Mix	0.44	–0.07	–0.25	–0.27	0.30	–0.09

Given the strong associations between RNHPPD and both unit type and patient acuity, RNHPPD could be used to adjust scores on NDNQI outcome measures in the same way as proportions of patient days at various billing levels (as described above). However, RNHPPD is associated not only with acuity but also with quality of care, and adjusting quality measures for RNHPPD (or other staffing variables) would involve more than adjustment for acuity. The unintended consequence of such an adjustment would be that units with higher RNHPPD would have their scores favorably adjusted regardless of the acuity of their patients or the quality of care provided, and whereas current NDNQI outcome measures reflect both quality of care and patient acuity, these adjusted scores would reflect some combination of quality of care, patient acuity, and staffing, making them difficult to interpret as indicators of quality.

Alternatively, RNHPPD could be used to define new unit types for mixed acuity units. For example, a mixed acuity unit with RNHPPD between the critical care and step-down RNHPPD means could be assigned to a critical/step-down mixed unit type. While this would be less problematic than using RNHPPD to adjust outcome measures, defining unit types based on staffing would tend to favor units with high RNHPPD relative to patient acuity by placing them in comparison groups with higher-acuity units that have RNHPPD in the same range.

Proposed Method for Mixed Acuity Units. We propose using the billing-days data collected in this study to create new acuity-based unit types for mixed acuity units, and to define these unit types using simple rules similar to the 90% classification rule we currently use.

A number of factors were considered in creating these new unit types. First, it is important to have enough types to ensure within-type homogeneity, but not so many types that some have too few units to serve as a meaningful comparison group. Second, the rules for classifying units must be easy to understand and based on numbers that site coordinators can accurately estimate if they do not have exact data. Third, the new unit types should reflect the proportion of billing days at the critical, step-down, and standard care levels, as well as the RNHPPD, reported by the units in this study. And fourth, the mean RNHPPD and ulcer rates for the units assigned to the new unit types should fit reasonably well in an acuity-ordered list of unit type means, including the means for the existing critical care, step-down, and medical/surgical types.

With these considerations in mind, three new unit types were created for adult units that do not meet the 90% criterion for an existing unit type. They are defined as follows:

1. Mixed Acuity III: Units with at least 50% critical care patient days each month.
2. Mixed Acuity II: Units with at least 25% critical care days each month or at least 50% step-down days each month; includes only units not meeting the criterion for Mixed Acuity III.
3. Mixed Acuity I: Units not meeting the criteria for Mixed Acuity III or II.

We propose to use the same classification scheme to create three new unit types for pediatric mixed acuity units.

The definitions were used to assign each of the adult units in this study to a new unit type. Descriptive statistics for RNHPPD and unit-acquired pressure ulcer (UAPU) rate were calculated by unit type both for the new unit types and, using data from the first quarter of 2011, for the existing critical care, step-down, and med/surg unit types. As shown in Table 3, RNHPPD and UAPU rates are generally ordered as one would expect, with higher numbers for higher-acuity unit types.

Table 3

Descriptive Statistics for RNHPPD and UAPU by Unit Type

Unit Type	***RNHPPD***			***UAPU Rate (%)***		
	N	**Mean**	***SD***	**N**	**Mean**	***SD***
Critical Care	2267	15.1	*2.8*	2121	6.4	*9.5*
Mixed Acuity III	11	14.3	*2.7*	10	4.2	*5.7*
Mixed Acuity II	19	11.1	*2.2*	15	3.7	*8.8*
Step-down	1547	7.7	*2.0*	1417	2.6	*5.3*
Mixed Acuity I	13	7.3	*1.6*	12	0.7	*1.6*
Med/Surg	2429	5.8	*1.5*	2194	2.0	*4.0*

In addition to the six new unit types (three adult, three pediatric) defined above, we propose to introduce the following: Burn unit (adult), burn unit (pediatric), bone marrow transplant unit (adult), and bone marrow transplant (pediatric). These units, along with critical access units, should be allowed to submit data on appropriate indicators and to receive quarterly reports with the understanding that there may be a high degree of within-group heterogeneity for their unit type.

Hospital-Level Indicators

The purpose of this portion of the study was to develop and compare several methods for measuring hospital-level performance on NDNQI indicators. The primary challenge in measuring NDNQI hospital performance is that hospitals differ in the number and type of nursing units they comprise. This makes meaningful comparison of hospitals difficult, even among hospitals of the same size and teaching status. Several hospital-level measures are presented below. We propose a method in which indicator scores are adjusted for unit type and unit size before being aggregated to the hospital level.

The rate of hospital-acquired pressure ulcers (HAPUs) is used as an example in this study without loss of generality; the methods described here can be applied to any indicator computed as a rate or proportion. For method demonstration and comparison, ulcer data were simulated for three fictitious hospitals, each with 1,000 patients assessed for pressure ulcers.

The simulated data are shown in Table 4. Note that the hospitals differ in both the types and sizes of their component units. Hospital 1 has one ICU, one step-down unit, and no rehabilitation unit. Hospital 2 has more ICU and step-down patients than Hospital 1 and has rehabilitation patients. Hospital 3 has the most ICU, step-down, and rehabilitation patients of the three hospitals. In addition, there are differences among the hospitals in the number and size of their medical, surgical, and medical/surgical units.

Table 4

Hospital-Acquired Pressure Ulcers by Unit

	Hospital 1			Hospital 2			Hospital 3		
Unit	**HAPUs**	**Patients**	**Rate**	**HAPUs**	**Patients**	**Rate**	**HAPUs**	**Patients**	**Rate**
ICU 1	9	50	0.180	15	65	0.231	7	40	0.175
ICU 2	-	-	-	-	-	-	6	40	0.150
Step-down 1	12	120	0.100	7	74	0.095	8	78	0.103
Step-down 2	-	-	-	10	74	0.135	7	78	0.090
Medical 1	4	105	0.038	3	86	0.035	6	78	0.077
Medical 2	9	105	0.086	4	86	0.047	5	78	0.064
Medical 3	-	-	-	-	-	-	3	78	0.038
Surgical 1	1	95	0.011	1	84	0.012	0	69	0.000
Surgical 2	-	-	-	0	84	0.000	0	69	0.000
Med/Surg 1	7	105	0.067	5	93	0.054	5	78	0.064
Med/Surg 2	5	105	0.048	7	93	0.075	6	78	0.077
Med/Surg 3	8	105	0.076	8	95	0.084	5	78	0.064
Med/Surg 4	8	105	0.076	2	95	0.021	2	78	0.026
Med/Surg 5	7	105	0.067	-	-	-	-	-	-
Rehab	-	-	-	8	71	0.113	10	80	0.125
Totals	70	1000	-	70	1000	-	70	1000	-

Several methods for calculating a hospital-level ulcer rate measure are described below. In Table 5, which is intended to serve as a visual aid, unit-level data and the values used to compute the hospital measures for Hospital 2 are shown, along with ulcer rate means and standard deviations for the six unit types (based on third quarter 2010 NDNQI data). A comparison of the sample hospitals on the various measures is provided in Table 3.

Method 1. An overall hospital ulcer rate can be computed by dividing the total number of patients in the hospital who have a hospital-acquired pressure ulcer by the total number of patients assessed:

$$(HAPUs_1 + HAPUs_2 + . . . + HAPUs_M)/(n_1 + n_2 + . . . + n_M),$$

where the units in the hospital are numbered $j = 1, 2, . . . , M$; $HAPUs_j$ is the count of HAPUs on the *j*th unit; and n_j is the count of patients assessed on the *j*th unit.

As shown in Table 6, the overall ulcer rate for the three hospitals in the study is 7.0%. Under this method differences among hospitals in the types of units they comprise are ignored, and performance of the three hospitals in preventing pressure ulcers appears to be equal.

Method 2. A raw average of unit ulcer rates can be computed by summing the unit ulcer rates within a hospital and dividing by the number of units:

$$(HAPUs_1/n_1 + HAPUs_2/n_2 + . . . + HAPUs_M/n_M)/M.$$

As with Method 1, the three hospitals appear to be performing equally, each having an average unit ulcer rate of 7.5%. This number is higher than the overall hospital rate because the measure does not account for differences among units in the number of patients assessed, allowing small units with high rates to exert disproportionate influence on the average unit rate. For example, in Hospital 2 the ICU ulcer rate of 23.1%, which is based on 65 patients, is given the same weight as the low surgical unit rates, which are based on 84 patients each (see Table 2).

Moreover, like the overall hospital rate (Method 1), this measure ignores differences among hospitals in the types of units they comprise. For example, whereas the averages for Hospitals 2 and 3 include a rehabilitation unit rate, the average for Hospital 1 does not.

Method 3. Method 2 can be adjusted to control for differences in unit size by weighting each unit's ulcer rate by its number of patients assessed, summing these weighted rates, and dividing by the total number of patients assessed for the hospital. This is equivalent to weighting each unit ulcer rate by that unit's proportion of patients assessed and summing these weighted rates**.** This measure is identical to the overall hospital rate described under Method 1 and does not account for differences among hospitals in the types of units they comprise.

Method 4. The ulcer rate for each unit can be adjusted for unit type by subtracting the average ulcer rate for units of that type and then dividing by the standard deviation of the ulcer rates for units of that type. The resulting z-score is the difference, in standard deviations, of the unit's ulcer rate from the average ulcer rate for units of that type. For example (see Table 4), the ICU in Hospital 2 has an ulcer rate of 23.1%, which is about 1.7 standard deviations above the average ulcer rate for ICUs (6.7%).

These z-scores, which are all on the same metric, can be averaged for each hospital to yield an average unit z-score: $(z_1 + z_2 + . . . + z_M)/M$, where z_j is the z-score for the *j*th unit. As shown in Table 5, the average unit in Hospital 2 has an ulcer rate slightly over one-half a standard deviation above the mean rate for its unit type, while unit rates in Hospitals 1 and 3 average 0.69 and 0.54 standard deviations, respectively, above their unit type means (see Table

5). Like the raw average of unit ulcer rates (Method 2), the average unit z-score does not account for differences in unit size.

Method 5. A weighted average of unit z-scores can be computed by weighting each unit's z-score (defined under Method 4) by its number of patients assessed, summing these weighted scores, and dividing by the total number of patients for the hospital:

$$(z_1n_1 + z_2n_2 + . . . + z_Mn_M)/(n_1 + n_2 + . . . + n_M).$$

Under this method, which takes into account both the types and sizes of each hospital's units, Hospital 1 loses its advantage of having the fewest ICU patients and rehabilitation patients, and Hospital 3 is no longer penalized for having the greatest number of ICU and rehabilitation patients. As shown in Table 3, the score on this measure for Hospital 1 was 0.67, while the score for the other two hospitals was 0.52.

The weighted average of z-scores can be converted to the ulcer rate metric by multiplying by the ulcer rate standard deviation for units of all types (equal to 0.070 for quarter three of 2010) and adding the overall unit ulcer rate mean (0.038). Hospitals 2 and 3 have similar adjusted rates (7.4% and 7.5%, respectively), while the rate for Hospital 1 is a full percentage point higher (8.5%). These rates are higher than the overall hospital ulcer rates, reflecting adjustments for unit type and unit size.

Table 5

Examples of Unit and Hospital Measures for Hospital 2

Unit	Pts	Rate	Z-score	Weighted Z-score	Avg Rate for Unit Type	SD for Unit Type
ICU 1	65	0.231	1.67	0.109	0.067	0.098
Step-down 1	74	0.095	0.87	0.064	0.038	0.065
Step-down 2	74	0.135	1.49	0.111		
Medical 1	86	0.035	0.10	0.009	0.030	0.049
Medical 2	86	0.047	0.34	0.029		
Surgical 1	84	0.012	–0.20	–0.017	0.020	0.040
Surgical 2	84	0	–0.50	–0.042		
Med/Surg 1	93	0.054	0.49	0.045	0.029	0.051
Med/Surg 2	93	0.075	0.91	0.084		
Med/Surg 3	95	0.084	1.08	0.103		
Med/Surg 4	95	0.021	–0.16	–0.015		
Rehab	71	0.113	0.57	0.041	0.048	0.113
Average Unit Rate		0.075				
Average Unit Z-score			0.555			
Weighted Z-score Average				0.520		
Weighted Z-score Average on Ulcer Rate Metric				0.074		

Table 6

Comparison of Hospitals by Measure

	Measure	Hospital 1	Hospital 2	Hospital 3
Method 1	Hospital Rate	0.070	0.070	0.070
Method 2	Average Unit Rate	0.075	0.075	0.075
Method 4	Average Unit Z-score	0.69	0.56	0.54
Method 5	Weighted Z-score Average	0.67	0.52	0.52

	Weighted Z-score Average (Ulcer Rate Metric)	0.085	0.074	0.075

Proposed Method. We propose reporting to hospitals the weighted z-score average (Method 5). While this adjusted rate may not be transparent to all hospital users, since it will not be the same as the raw rate they have internally, it can be used to track a hospital's performance across time and, unlike the other measures considered, allows for meaningful comparison (e.g. percentile ranking) of hospitals within a given comparison group. It should be noted that this adjusted rate is a relative measure, affected not only by the performance of the other NDNQI hospitals but also by changes in the set of hospitals and units reporting data to NDNQI; however, given the number of units participating in NDNQI, we expect unit type averages and standard deviations to be quite stable across time for most indicators.

Discussion

This study resulted in a method for incorporating mixed acuity units into NDNQI data collection and reporting and producing hospital-level indicators, capabilities desired by participating hospitals. These enhancements can be implemented with minimal increase to respondent burden.

The information gathered for the methods development study underscored the difficulty of creating risk adjusted patient outcome measure for acute care units, while maintaining a low respondent burden and transparency for users. The analysis of billing days data from the special study illustrated that mixed acuity unit types can be created that are consistent with NDNQI's existing unit classification scheme. The consistency indicates criterion validity. The stratification of mixed acuity units into types I, Ii and III produced indicator results that were intermediate between existing unit types. Implementation of the mixed acuity unit types required hospitals to go through unit enrollment with the NDNQI liaisons. Hospitals should monitor changes to patient populations on mixed acuity units as signified by changes in the proportion of billing days by payment level. Updates to mixed acuity unit classifications would become part of the site coordinators' responsibility for unit classification maintenance.

The method developed for producing hospital-level indicators accounts for both the unit composition of hospitals, as well as variation in the size of the unit. Implementation of the method requires only that hospitals report on the number of staffed beds per unit. The resulting indicators are in the native metric of the original indicator and thus have a high level of transparency for report users.

NDNQI
Quality Improvement Solutions from ANA

Report of Patient Days Reliability Study

Prepared for the American Nurses Association

By

Chenjuan (Tina) Ma, PhD

Byron Gajewski, PhD

Emily Cramer, PhD

Nancy Dunton, PhD

January 2014

Table of Content

Abstract

Patient days measure the length of patients' exposure to the health care settings and reflect patient care workload for nurses and other healthcare professionals. A review of patient days literature highlighted the importance of collecting accurate patient days data for the sake of both clinical practice and health service research. The National Database of Nursing Quality Indicators (NDNQI) has been collecting patient days data from member hospitals since its establishment in 1998. Researchers from the NDNQI have made tremendous efforts in order to ensure high quality of its patient days data, including reliability studies.

The study described in this report is the second patient days reliability study assessing the quality of the NDNQI patient days. Our patient days reliability study consist of two phases: (1) a multi-site patient census study and (2) a site coordinator survey. The multi-site patient census survey was conducted to address specific aim one, to identify the agreement between the NDNQI quarterly submitted patient days data and the patient days data from the multi-site patient census study (gold standard). The site coordinator survey was designed to investigate the current practices and other issues related to patient days data collection and data accuracy in NDNQI member hospitals. Comparisons of the two survey samples to the NDNQI population suggested both the survey samples well represented the NDNQI population.

Participating units in the multi-site patient census study were asked to count the number of patients on their units every 2 hours over a 24-hour period on each of the 7 randomly assigned data collection days in June 2013. 224 units from 58 hospitals that submitted patient days data at least for 4 days were included in our final analyses.

Among these 224 units, 41% used midnight census (Method 1) to collection patient days data for NDNQI quarterly patient days data submission, 40% used midnight censes for inpatients with actual hours for short stay patients (Method 2), 5% used actual hours for both inpatients and short stay patients (M4), and 13% used multiple census reports.

Our reliability analyses using one-way random effect models were conducted overall and by data collection methods. The results indicated a high reliability of the NDNQI patient days data. The overall intra-class correlation coefficient (ICC) between the multi-site patient census study and the matched NDNQI data was 0.95 with a 95% C.I. of 0.94-0.96. The method-specific ICCs ranged from 0.94 to 0.96, and none of their lower boundary of 95% C.I. was below 0.88. Additional analysis also suggested that different data collection method can affect the accuracy of patient days data.

The site coordinator survey found that data accuracy (42%) and availability of the data sources (37%) were the two main factors in the selection of patient days data collection in NDNQI member hospitals. Surveyed site coordinators expressed that the data collection method in use was appropriateness in both units without short stay patients (86%) and units with short stay patients (76%). They also reported that the accurate level of their patient days data for NDNQI as excellent or good for units without short stay patients (87%) and for units with short stay patients (76%).

This reliability study concludes that the current data collection methods used by the NDNQI member hospitals are appropriate. The patient days data collected using these methods is with a good reliability. Meanwhile it suggests that given the increase of observational/short stay patients and the nationwide implementation of electronic healthcare records, NDNQI should modify their patient days data collection methods timely to help member hospitals collect accurate patient days data, such as including patient counts at a different time point in addition to the midnight census.

Part I: Introduction

Patient days as an indicator in health services research and clinical practice directly measures the length of a patient's stay in a health care setting. It also reflects the patient care workload for nurses and other healthcare professionals and is related to health care cost. Furthermore, patient day indicator is an important element in generating nursing and patient outcome indicators measuring nurse staffing levels and quality of health care. Therefore there is a necessity in collecting reliable patient days data for the sake of both clinical practice and health care research.

The National Database of Nursing Quality Indicators (NDNQI) was founded in 1998 by the American Nursing Association with the mission of aiding nurses in efforts of improving care quality and patient safety (Montalvo, 1997). The NDNQI has collected patient days data since its establishment. Patient days indicator is one of the most important indicators collected by the NDNQI. It is used to calculate the nurse staffing indicator of nursing hours per patient day as well as several nurse sensitive patient outcome indicators, including patient falls and hospital infections. Nurse staffing and patient outcome data are included in the NDNQI member hospital reports to assist hospital executives in their policy-making process, and are frequently used in health care research.

To ensure high quality of the NDNQI patient days data and the indicators generated from it, a research team from the NDNQI conducted the first study to assess the reliability of the patient days data in 2008-2009. Our study described in this report is the second patient days reliability study assessing the quality of NDNQI patient days data. The specific aims of our study are:

Aim 1: To determine the reliability of the NDNQI patient days data by identifying the Intraclass Correlation Coefficient between the NDNQI patient days data and a unique dataset collected from a specially designed multi-site patient census study.

Aim 2: To investigate site coordinators' perspectives on the NDNQI patient days data collection practices and data accuracy via a site coordinator survey.

This study has been approved by the Institutional Review Board of the University of Kansas Medical Center.

Part II Background

This session begins with a synthesis of the literature on patient days with a focus of publications since NDNQI's previous patient days reliability study in 2008-2009. Both methodological studies and investigations including patient days as a study variable were reviewed. The literature review is followed by a detailed description of the NDNQI patient days indicator.

Patient days in research literature

A most basic and important question regarding an indicator is its reliability. Only two published papers were found in PubMed and CINAHL that examined methodological issues of patient days data collection methods and data accuracy at unit-level in the past 5 years (Simon, Yankovskyy, Klaus, Gajewski, & Dunton, 2011; Simon, Yankovskyy, & Dunton, 2010). Findings from these two studies indicate that different patient days collection methods, unit types, and the proportion of short stay patients (SSPs) influence the accuracy of patient days data. Simon, et al. also suggests inaccuracy in patient days data jeopardizes the reliability of reported association between nursing and patient outcomes, because patient days measure is a block stone in generating various nursing and patient outcome indicators.

A review of literature revealed that despite limited research on the quality of patient days data, patient days have been very frequently used in health services research. Researchers have used patient days as the denominator to generate nursing care hours per patient day which reflects the nurse staffing level (Kalisch, Friese, Choi, & Rochman, 2011; Lerner, 2013; Li et al., 2011; Staggs, 2013; Twigg, Duffield, Bremner, Rapley, & Finn, 2011) and to create patient outcome measures (e.g. patient fall) (Flynn, Liang, Dickson, Xie, & Suh, 2012; He, Dunton, & Staggs, 2012; Nedved, Chaudhry, Pilipczuk, & Shah, 2012). Researchers also used patient days in studies examining the association between nursing and patient outcomes (Kalisch, Tschannen, & Lee, 2012; Lake, Shang, Klaus, & Dunton, 2010; Wilson, Bremner, Hauck, & Finn, 2011). Although not conclusive, some researchers have suggested that nursing hours per patient day at a unit-level is a better indicator of nurse staffing levels with higher predictive power in patient outcomes research, compared to alternative nurse staffing measures from administrative data (Van den Heede et al., 2009).Therefore, the quality of patient days data plays an important role in nursing and patient outcomes research. To ensure the reliability of the findings from this research, accurate patient days data are required.

In summary, the findings from our literature review regarding the scarcity of research on measuring patient days and the wide use of patient days data in clinical practice and health services research highlight the importance and necessity of collecting reliable patient days data.

NDNQI patient days indicator

Patient days indicator is one of first collected indicators by the NDNQI. Patient days have been measured at different levels, such as unit-level or hospital-level (Carayon et al., 2013; Chen, Sexton, Kaye, & Anderson, 2009; Staggs & He, 2013). The NDNQI patient days data were collected at unit-level. Registered units from NDNQI member hospitals are required to report monthly patient days data once in each quarter using one of the methods defined by the NDNQI. The NDNQI patient days data have also been used to calculate other NDNQI indicators including but not limited to the nursing care hours per patient day and patient falls.

NDNQI researchers have been making continuous efforts to improve the quality of NDNQI patient days data. One example is creating and timely updating of the NDNQI Patient Days Guidelines for Data Collection. In the Guidelines, each available method for collecting patient days data is detailed. These methods are based on either patient census counts, actual patient hours, or a combination of the two. Patient Days Guidelines instruct NDNQI member hospitals to select the method that will provide the most accurate patient days data from their units, depending on the frequency of patient turnover on the unit, the presence of short stay patients, and availability of data sources.

In addition, the NDNQI researchers conducted a reliability study in 2008-2009 to assess the accuracy of patient days data. Findings from this study indicated that the NDNQI quarterly patient days data demonstrated high reliability overall as well as method-specifically for four of the five available data collection methods then. The exception was for method 3, Midnight census and patient days from average hours for short stay patients (Simon et al., 2011). Following suggestions from this study, method 3 was removed from the Patient Days Guidelines. To date, there are four patient days data collection methods used by NDNQI member units. These methods and their definitions are described in **Table1**.

Table 1 NDNQI patient days data collection methods and definitions

Meth	**Definition**
M1	Midnight census
M2	Midnight census + actual hours for short stay patients
M4	Patient days from actual hours for all patients

M5	Patient days from multiple census reports

M3 was removed based on the findings from a reliability study in 2008-2009.

Meanwhile, we also realize that some potential differences may exist between the NDNQI patient days data collected using the four data collection methods and the patient days data in the real world. **Figure 1** illustrates a unit with five example patients counted in two hours census counts. **Table 2** shows the unit's patient days data derived by each collection methods. According to the definitions, Method 4 reports patient days based on the actual hours and gives the most accurate estimation of patient days. All the other methods underestimate or overestimate patient days. As shown in **Figure 1**, the estimated patient days with different collection methods are affected by the occurrence of short stay or observational patients (<23 hours on the unit) and their exact admission/discharge times. In addition, the average length of patients' stay which is related to patient turnover may also play a role in accurately measuring patient days.

Figure 1. Example patient days

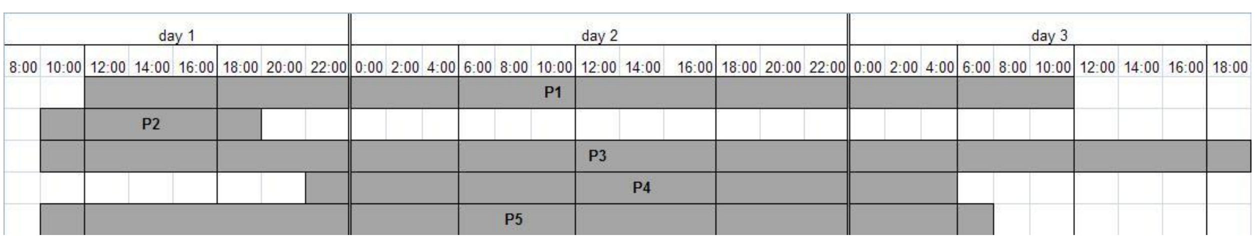


Table 2 Patient days by data collection methods of the illustrating example

Method	Patien	Patie
M1: Midnight census	192	8.0
M2: Midnight census for inpatients + actually hours for short stay patients	202	8.4
M4: Patient days from actual hours for all	194	8.1

M5: Patient days from multiple census counts	200	8.3

In addition, there have been many dynamic changes in the hospital care settings, which may have affected the ways of collecting patient days data and consequently influence the accuracy of patient days data. Among the most important changes are, for example, the increase in the number of short stay patients in some units and increasing implementation of electronic health records (EHRs). The EHRs dramatically increase the availability of more accurate patient days data within hospitals.

To summarize, both the literature review and the description of the NDNQI patient days data suggest a second reliability study assessing the quality of the NDNQI patient days data is needed.

Part III: Study Methods and Results

This reliability study was conducted in two phases, the Multi-Site Patient Census Study (Phase I) and the Site Coordinator Survey (Phase II). Phase I was designed to address the specific aim one and the phase II for specific aim two. In this session, we described the research methods and results for each phase respectively.

Phase I: Multi-Site Patient Census Study

Methods

Multi-site patient census study

To address our primary aim of assessing the reliability of the NDNQI quarterly collected patient days data, we designed a multi-site patient census survey to collect patient days data that reflects the patient days in reality. In this multi-site survey, we asked site coordinators of participating units to collect data on seven randomly assigned days during June of 2013. Each of these seven days also represented one day of a week. Therefore each hospital had a random pattern of data collection days. On data collection days, patients on each unit were counted every two hours for a period of 24 hours on each of the designated data collection day (**Appendix A**). All the collected data from each participating unit were then entered into an online database, from which researchers from the NDNQI can assess to and download the data.

Ideally the best way to assess the quality of NDNQI patient days data is to compare it to a patient days dataset based on actual check-in and check-out time of each patient cared for on a unit. However, this method will require the collection of individual patient data and substantially increase respondents' burden, which makes it not feasible. Our survey did not obtain patients' actual admission and discharge time on a unit; however, the aforementioned **Figure 1 and Table 2** display that our survey design (counting patient every two hours on seven randomly selected days) could generate very similar, if not exactly the same, data as in the real world with a huge decrease of workload to nurses. The data collected via this survey was considered as the gold standard in the reliability analyses.

Sample calculation

A sample size calculation was completed before the survey was sent out. Given the findings from last NDNQI patient days reliability study (Simon et al., 2011), an intraclass correlation coefficient (ICC) of 0.90 with an estimated error margin of 0.05 were chose to calculated the minimum desired sample size, which is 220 units (Giraudeau & Mary, 2001).

Data analysis

Patient days data from the multi-site patient census survey were calculated and aggregated for each unit to reflect the actual patient days of participating units in June of 2013. The average number of patients on each data collection day on each unit was first calculated. Then this average number was multiplied by the number of occurrences of this weekday (e.g. 5 Saturdays) in June. Finally all the products were added up to indicate the patient days of each unit for the month of June 2013. The NDNQI patient days data that are submitted quarterly was obtained from the NDNQI server for the same month of the same units and was used for comparison to the survey data in the reliability analyses.

We first described the response rates of the multi-site patient census survey. We then conducted a descriptive comparison of the survey data and the NDNQI patient days data by data collection methods, unit characteristics and hospital characteristics to identify whether the survey sample well represented the NDNQI patient days data. We identified the level of agreement presented by ICC (1,1) between multi-site patient census data and the NDNQI patient days data using one-way random effect models. Finally, we conducted an exploratory analysis to identify factors that might influence the patient days data collection and accuracy. All the analyses were completed using STATA 12.0 (StataCorp, College Station, TX).

Results

Sample

Initial call for participation in this multi-site patient census study was sent out by email to 700 randomly selected hospitals from a total of 1793 NDNQI member hospitals which are currently active and have units submitting patient days data in the past three years. Sixty six hospitals agreed to participate in this study. From these 66 hospitals, 282 units were randomly selected. During the data collection period, 234 units in 63 hospitals attempted to submit patient days data. Two hundred and thirty units from 60 hospitals submitted data (including both complete and incomplete submissions). For analysis purpose, only units with data for at least 4 days out of the 7 selected data collection days were included. Our final multi-site patient census survey data consisted of 224 units from 58 hospitals. Two units were further excluded from the comparative analysis between the survey data and NDNQI data due to the lack of NDNQI patient days data in the survey month. **Table 3** summarizes the response rates for hospitals, units, and data collection days.

Table 3. Response rates of hospitals, units, and data collection days

	Enrolled	Any data submission	4 or more days of data submission	Response rate (%)
Hospitals	66	60	58	90
Units	282	230	224	82
Data collection days	1,974	1,573	1,552	80

Table 4 compares the units in the survey sample with all units in NDNQI member hospitals for which patient days data were submitted by patient days data collection methods, patient population, hospital bed size, teaching status, and magnet status. There were slightly more small (defined by bed size) non-teaching non-magnet hospitals in our survey sample than the NDNQI population, yet the largest difference was no more than 11%. The distributions of patient days data collection methods and patient population, the two most important factors related to the quality of patient days data, were very similar across the survey sample and the NDNQI population. Thus, despite some differences, we are confident our survey sample well represented the NDNQI population.

Table 4. Comparison of survey sample and NDNQI population by collection methods, patient population, and hospital characteristics

	Survey sample (n=224)		NDNQI population (n=15,368)	
	N	%	N	%
Collection method				
M1 - MC*	91	40.63	6,661	43.34
M2 - MC + PDs** from actual hours for SSPs***	90	40.18	5,103	33.21
M4 - PDs from act. hrs for inpatients and SSPs	12	5.36	1,391	9.05
M5 - PDs from multiple census reports	29	12.95	2,213	14.4
Patient population				
Adult inpatient	168	75.00	11,670	75.94
Neonatal inpatient	14	6.25	768	5.00
Pediatric inpatient	18	8.04	1,155	7.52
Psychiatric	9	4.02	1,139	7.41
Rehab inpatient	15	6.70	636	4.14
Hospital bed size				
< 100	37	16.52	1,332	8.67
100-199	74	33.04	3,289	21.40
200-299	30	13.39	3,359	21.86
300-399	28	12.50	2,451	15.95
400-499	20	8.93	1,692	11.01
>=500	35	15.63	3,245	21.12
Teaching status				
Academic Medical Center	30	13.39	3,286	21.38
Teaching	7	3	6	4

	2	2.14	,298	0.98
Non-teaching	1 22	5 4.46	5 ,784	3 7.64
Magnet status				
Non-Magnet	1 27	5 6.7	7 ,210	4 6.92
Applicant	4 0	1 7.86	3 ,005	1 9.55
Magnet	5 7	2 5.45	5 ,153	3 3.53

*Midnight census; **Patient days; ***Short stay patients

Reliability analysis

Table 4 displays the distribution of overall patient days and patient days by data collection methods in the survey sample and the matched NDNQI sample. Overall, the patient days data from our multi-site patient census survey and the matched NDNQI patient days data had almost the same distribution. When stratified by data collection methods, the largest difference (difference between means) was observed in Method 5, collecting patient day data from multiple census reports. The matched NDNQI data had an average of approximately 23 more days per month in June 2013 than the survey sample. The second largest difference existed when using method 4, collecting patient day data from actual hours for both inpatients and short stay patients, for which the matched NDNQI data had 21 days less per month on average. Using method 1, mid-night census only, and method 2, mid-night census and actual hours for short stay patients, both resulted in a 7-day difference but with reverse directions.

Table 4. Comparison of distributions of patient days between the survey sample and the matched NDNQI sample (n=213)*

	SURVEY		NDNQI	
	Mean (SD)	**Range**	**Mean (SD)**	**Range**
Overall	525.15 (289.23)	22.33-1672.42	526.74 (292.98)	33.00-1717.00
M1 - MC	551.65 (312.02)	22.33-1569.08	544.90 (313.31)	33.00-1591.00
M2 - MC + PDs from actual hours for SSPs	496.57 (253.57)	97.17-1106.75	503.69 (257.67)	92.00-1147.00
M4 - PDs from act. hrs for inpatients and SSPs	524.92 (415.39)	143.50-1672.42	504.26 (436.14)	142.00-1717.00
M5 - PDs from multiple census reports	534.12 (265.54)	52.67-1144.08	556.85 (265.24)	56.00-1148.00

To achieve a more accurate comparison, only units with 7-day data from the multi-site patient census survey were included.

The one-way random-effects ICC analyses identifying the overall and method-specific agreement of the patient days data between the survey sample and the matched NDNQI sample were conducted on units with at least 4 days of data submission (n=224) as well as on units submitted data on all 7 data collection days (n=213). No significant differences were observed in the results and here we only reported the results from the analysis including all units submitting data for 4 days or more. The results are displayed in **Table 6.** The reported ICC is the absolute-agreement ICC, which reflects the proportion of between units variance in total variance. In other words, the larger the ICC, the higher the agreement between the survey sample and the matched NDNQI sample.

Table 5. ICC of patient days data from the survey sample and the matched NDNQI sample

	N	**ICC**	**95% C.I.**
Overall	224	0.95	0.94-0.96
M1 - MC	91	0.96	0.94-0.97
M2 - MC + PDs from actual hours for SSPs	90	0.94	0.91-0.96
M4 - PDs from act. hrs for inpatients and SSPs	12	0.96	0.88-0.99
M5 - PDs from multiple census reports	29	0.96	0.92-0.98

Our analyses indicated excellent agreement of the patient days data between the survey sample and the matched NDNQI sample. The overall ICC (1, 1) between the survey sample and the matched NDNQI sample was 0.95 (95%C.I.: 0.94-0.96), and the method-specific ICCs (1, 1) ranged from 0.94-0.96. All of them are above 0.8, which is usually considered as a cut-off point for excellent agreement (Forbes & Taunton, 1994;

Hughes & Anderson, 1994). Furthermore, given the representativeness of the survey sample compared with the NDNQI population (**Table 4**) and the ICCs (overall and method-specific), our study indicated a high reliability of the NDNQI patient days indicator.

In addition to the ICC analyses, an examination of factors that may be related to the difference between the survey data and the NDNQI data was conducted. The study factor included data collection method, unit type, and degree of short stay patients (DSSP). In this analysis, the different between the two data was defined as the ratio of the NDNQI data to the survey data for each unit. The calculation of DSSP used information from a single item question in the survey asking whether there have been short stay patient(s) during that data collection day. The DSSP for each unit was the proportion of days with existence of short stay patient(s) out of the total data collection days (at least 4 days or more).

Figure 2 displays the distribution of the ratio of the NDNQI patient days to patient days in the survey sample. This overall difference and the method-specific differences are summarized in **Table 7**. Results from one-way variance analysis with Bonferroni correction for multiple comparisons confirmed that reporting methods had an influence on the accuracy of patient day data. No significant association was found between the NDNQI data to survey data ratio and the DSSP and between the the NDNQI data to survey data ratio and unit type.

Figure 2. Distribution of the NDNQI data to survey data ratios on patient days

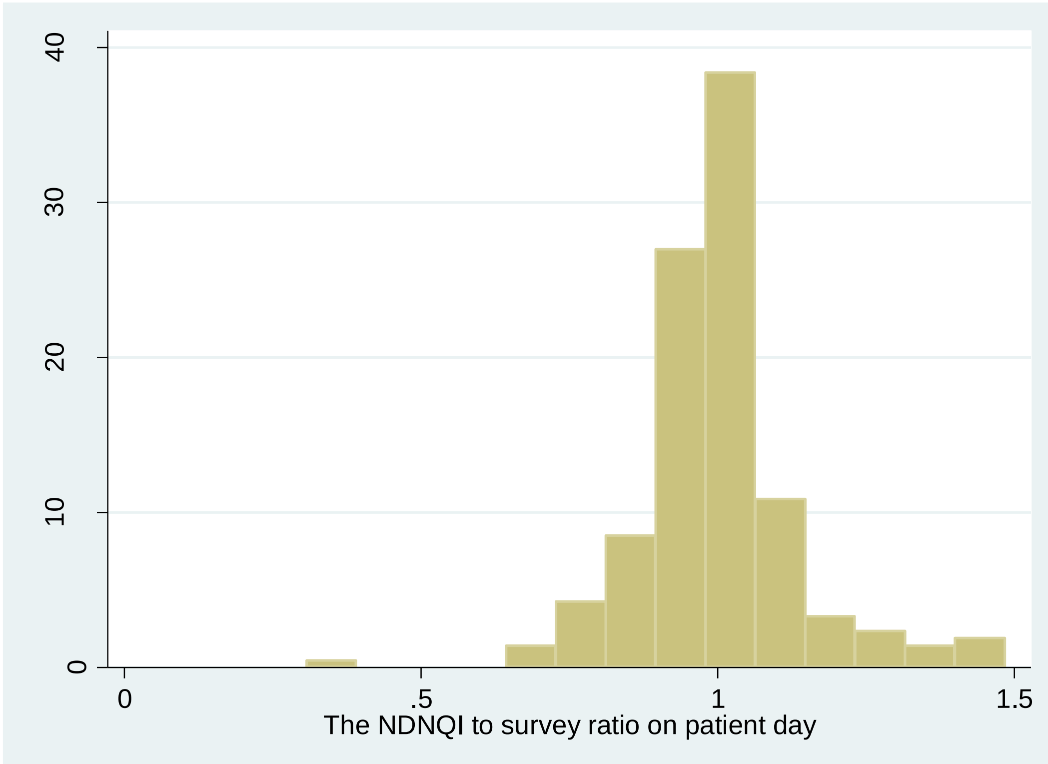


Table 7. Distribution of the NDNQI data to survey data ratios on patient days

		Mean (SD)	Range
Overall	19	1.01 (0.16)	0.31-1.74
M1 - MC	0	1.00 (0.13)	0.67-1.48
M2 - MC + PDs from actual hours for SSPs	9	1.00 (0.13)	0.74-1.48
M4 - PDs from actual hours for inpatients and SSPs	2	0.97 (0.26)	0.31-1.32
M5 - PDs from multiple census reports	8	1.11 (0.23)	0.88-1.74

Notes: MC: Midnight census; PDs: patient days; SSPs: short stay patients

Phase II: Site Coordinator Survey

Method

Site Coordinator Survey

In addition to the multi-site patient census survey, an online survey was designed to collection information from site coordinators (SCs) about their experience and perception of current patient days data collecting practice and other related issues in their hospitals (**Appendix B**). This information will assist NDNQI in verifying the reliability and validity of its patient days data received from member hospitals and provides informative evidence for future efforts in improving data collection. The survey was revised from the questionnaire used in the 2008-2009 NDNQI Patient Days Reliability Study Site Coordinator Survey. The face validity of this survey was assessed by NDNQI staff, including members from research team, analyst team, and liaison team.

SCs were chosen as the targeted population in this survey because of their vital role in ensuring high quality data collection and submission to NDNQI. SCs' responsibilities can be summarized in three areas. First, SCs are the primary point of contact between the NDNQI and member hospitals regarding any quarterly and annually NDNQI data collection and submission. Second, SCs are the coordinator in the process of unit enrollment/determination in NDNQI data submission and perform on-going unit monitoring and supervising of data collection and submission activities. Third, SCs are responsible for maintaining the security of his/her hospitals website by authorizing who is allowed access within his/her facility. The majority of the SCs are responsible for one hospital, however, some of them may coordinate two or more hospitals. For those who are in charge of more than one hospital, they were asked to choose the first hospital by alphabetical order of their hospitals' name to complete the survey.

In late June and early July 2013 the online survey was constructed and tested using REDCap before public access by SCs. RedCap is developed by Vanderbilty University and is a secure, web-based application designed exclusively to support data collection for research purpose (Harris et al., 2009). Five hundred and sixty three SCs were randomly selected from a total of 1,492 SCs representing 1793 active NDNQI member hospitals that have unit(s) submitting patient days data in the past three years. An invitation letter along with detailed explanation of the research purpose and other related information was sent to those randomly selected SCs on July 10, 2013. SCs were asked to complete the online survey within a two-week survey period. A follow up email was sent out to remind SCs about survey deadline.

Data analysis

Descriptive statistics of the enrollment and response rates were first analyzed. It was followed by a comparison of hospital characteristics between the survey hospitals and the NDNQI member hospitals to determine the representativeness of this SC survey. Finally a detailed descriptive analysis of SCs' perception of different aspects of the patient days data collection practice in their hospitals.

Results

Sample

Two hundred and ninety three SCs who received an invitation of participation letter responded to the online survey, which resulted in a response rate of 52%. Among the 293 respondents, 279 SCs representing 279 hospitals completed the online survey, with 44% of them have been in the SC position for 3 years or more. In our analysis, only information from the 279 completed surveys was used.

One concern of our study sample was whether it was representative of the NDNQI member hospitals. To address this concern, a comparison of the survey respondents and the NDNQI population regarding their hospital characteristics, such as hospital type and bed size, was conducted. The results are presented in **Table 8**. The comparison analysis revealed minor differences between the NDNQI population and survey sample in term of hospital characteristics. For example, there were slightly fewer large hospitals with a bed size of 300 or more and more non-teaching hospitals in the survey population than the NDNQI population. However these differences are not expected to weaken the representativeness of this SC survey.

Table 8. Comparison of NDNQI population and SC survey by hospital characteristics

	NDNQI* (n=15,368)		**Survey (n=279)**	
	N	**%**	**N**	**%**
Hospital Type				
General	14,325	93.21	238	85.3
Pediatric	497	3.23	10	3.58
Critical Access	69	0.45	9	3.23
Rehabilitation	110	0.72	9	3.23
Phsychiatric	88	0.57	2	0.72
Other	279	1.82	20	7.12
Hospital Size				
0-24	139	0.90	5	1.79
25-49	290	1.89	20	7.17
50-74	34	2.2	19	6.8

	3	3		1
75-99	560	3.64	24	8.60
100-199	3,289	21.40	77	27.60
200-299	3,359	21.86	63	22.58
300-399	2,451	15.95	27	9.68
400-499	1,692	11.01	18	6.45
500 or more	3,245	21.12	26	9.32
Teaching Status				
Academic Medical Center	3,286	21.38	34	12.19
Teaching	6,298	40.98	114	40.86
Non-teaching	5,784	37.64	131	46.95
Magnet Status				
Yes	5,153	33.53	68	24.37
No	10,215	66.47	211	75.63

*NDNQI hospitals with unit(s) submitting patient day data in June 2013

Data collection method selection and the determining factor

In addition to unit managers and/or SCs, administrative and/or financial department played an important role in determining the method of collecting and reporting patient days (**Figure 3**). Approximately 34% of the survey respondents reported that the reporting method were decided by administrative department and/or financial department without their involvement; and another 33% of them indicated that the reporting method was a collaborative production of unit manager, SC, and administrative/financial department. The decision of which data collection method to use was usually based on data accuracy (43%) and the availability of patient days data sources (37%) (**Figure 4**).

Figure 3. Data collection method selection (%)

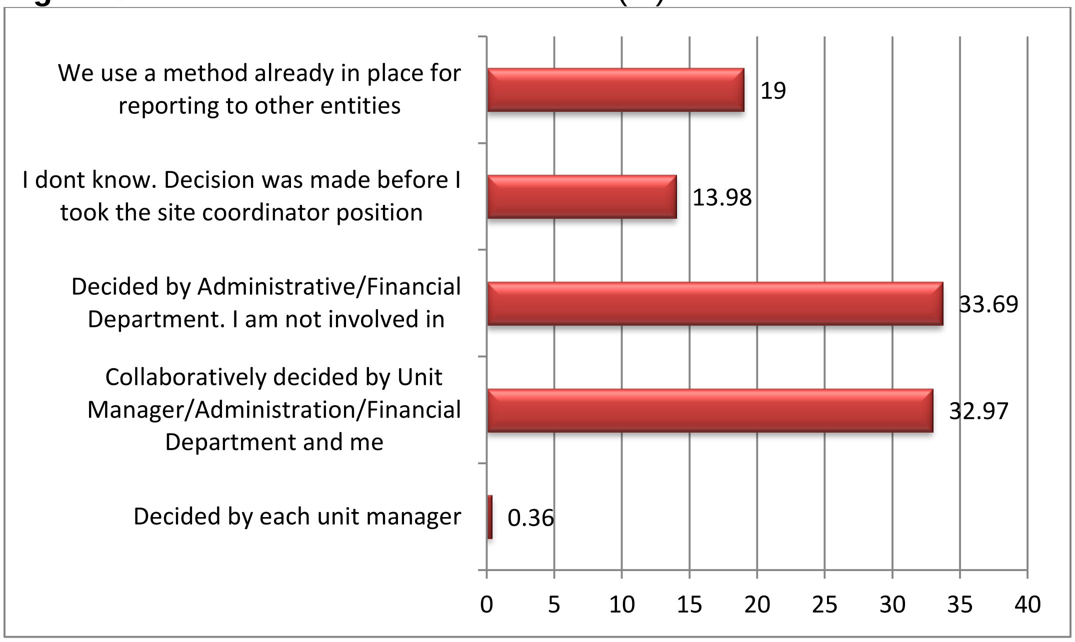
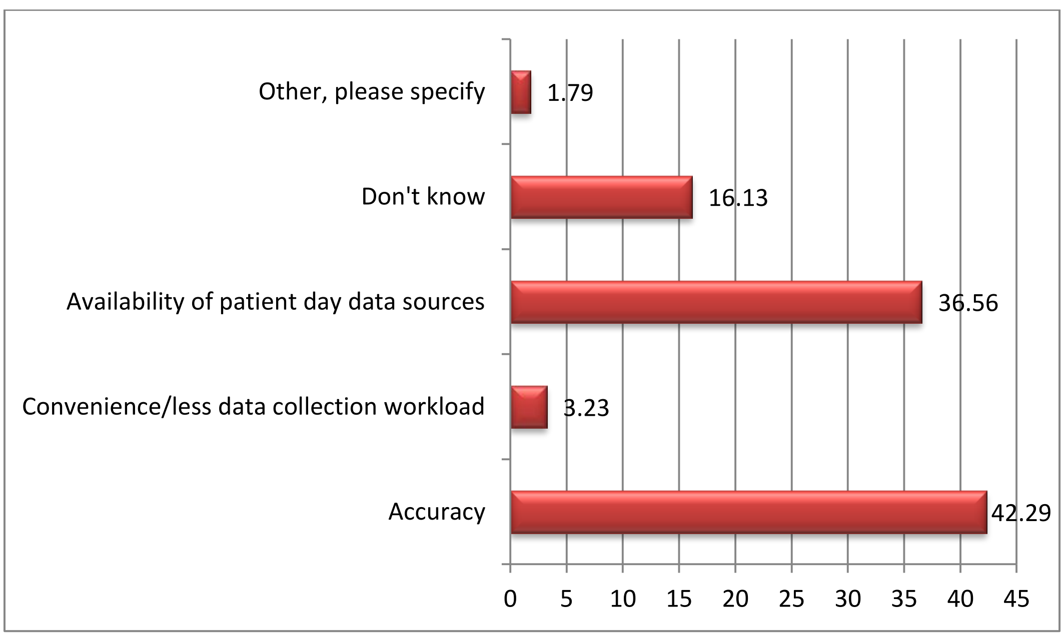


Figure 4. Most important reason in data collection method selection (%)



Appropriateness of data collection method and accuracy of patient days data

SCs were asked in the survey about their perceptions of the appropriateness of their data collection method and the accuracy of their submitted patient days data by units with and without short stay patients respectively. Because previous NDNQI patient days reliability study suggested that the presence and proportion of short stay patients could affect data accuracy. SCs indicated that the patient day collection method used in units without short stay patients were very appropriate (42%) or appropriate (45%), and in units with short stay patients were very appropriate (34%) or appropriate (42%) (**Figure 5**). In term of the accuracy of the submitted patient days data, 50% of the SCs reported it as excellent and 37% reported it as good in units without short stay patients; and 38% indicated it as excellent and 38% indicated it as good in units with short stay patients (**Figure 6**). Regardless of the presence of short stay patients or not, over 10% of the SCs were not sure about the appropriateness of method and the accuracy of submitted data.

Utilization of NDNQI Guidelines and data verification approaches

We asked SCs questions about the use of NDNQI Guidelines for Data Collection manual or online tutorial and approaches for data verification as indirect indicators of the overall quality of patient days data. As shown in **Figure 7**, roughly 37% of the SCs reported that they referred to the NDNQI Guideline or online tutorial for data collection at least once a quarter.

Figure 5. Appropriateness of data collection method by units with/without short stay patients (%)

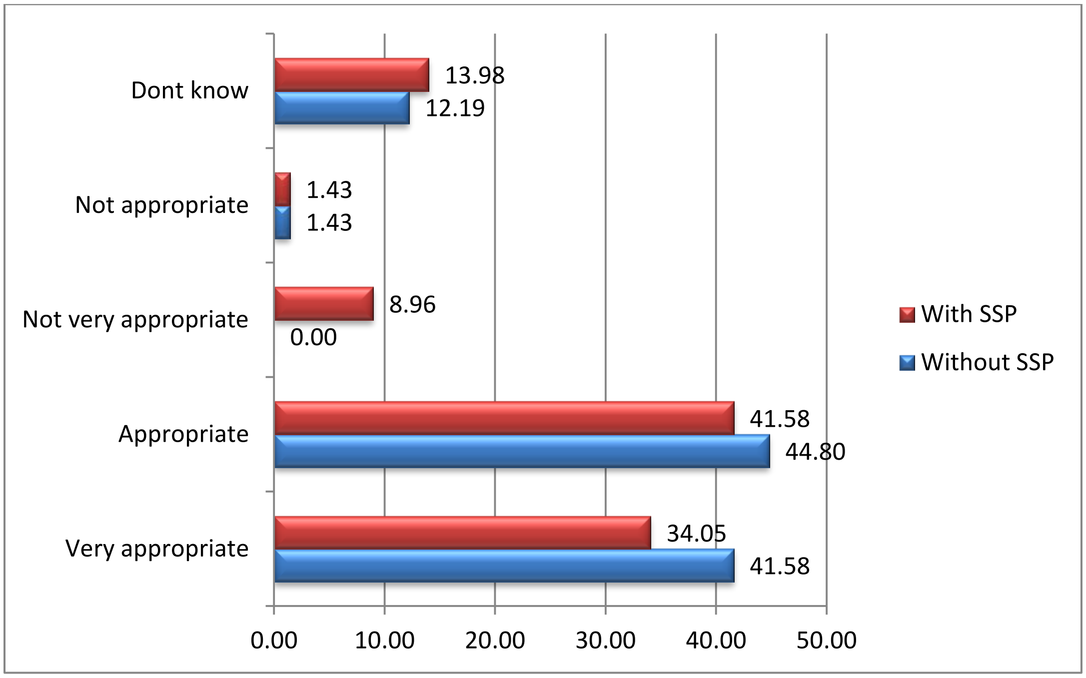
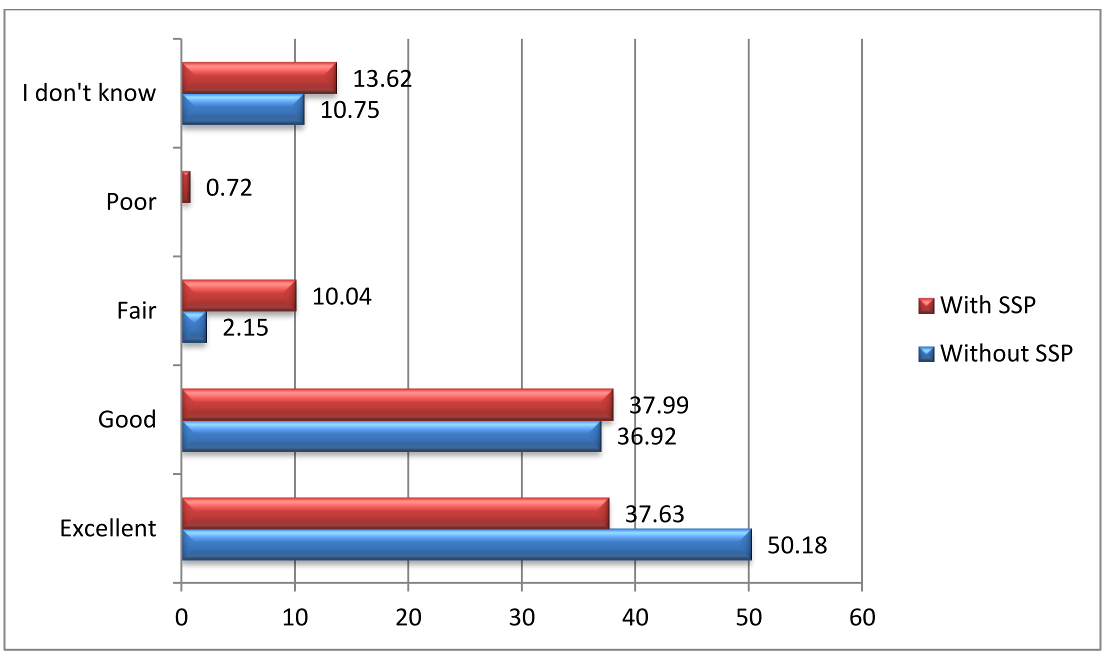


Figure 6. Accuracy of patient days data by units with/without short stay patients (%)



In addition to the NDNQI Guidelines and online tutorials, we also asked SCs about other data verification approaches. **Figure 8** shows the number of verification approaches used and **Figure 9** presents how frequently these approaches were used. Approximately three in four of the SCs reported that their hospitals used at least one method to verify patient days data before submission, and about 38% of the SCs indicated their hospitals used 2-4 methods for data verification. Among the listed four data verification methods, “comparing each unit’s value to earlier quarters’ data” was most frequently employed (56%), followed by “comparing to values used in other reports” (43%, **Figure 9**). There were 23% of the SCs reported that they did not use other verification and submitted data as received.

Figure 7. Utilization of NDNQI Guidelines or online tutorial for data collection (%)

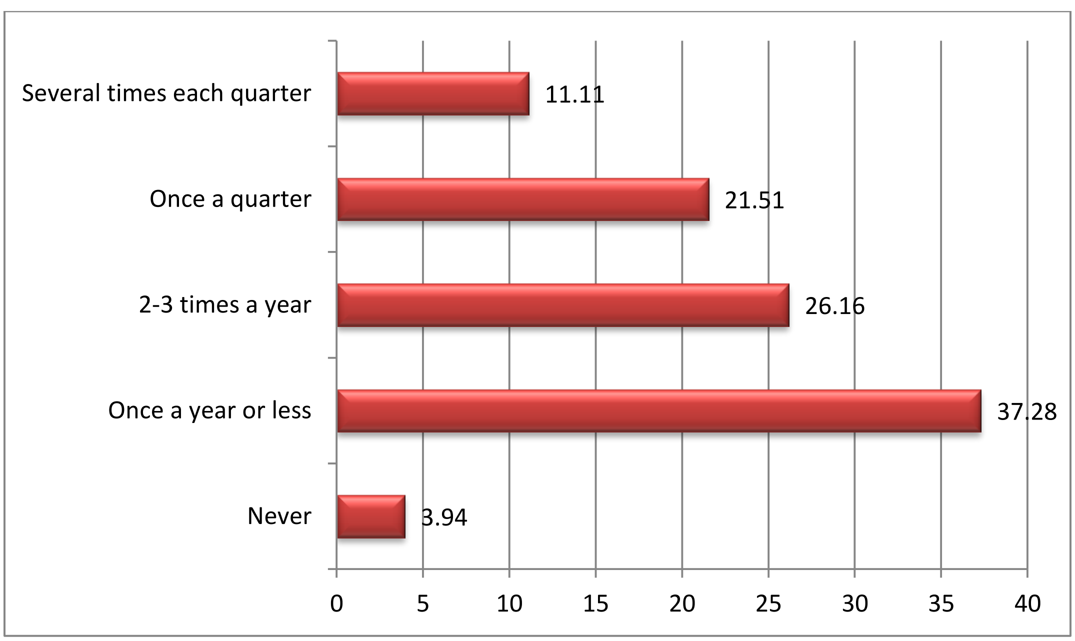


Figure 8. Number of data verification methods used (%)

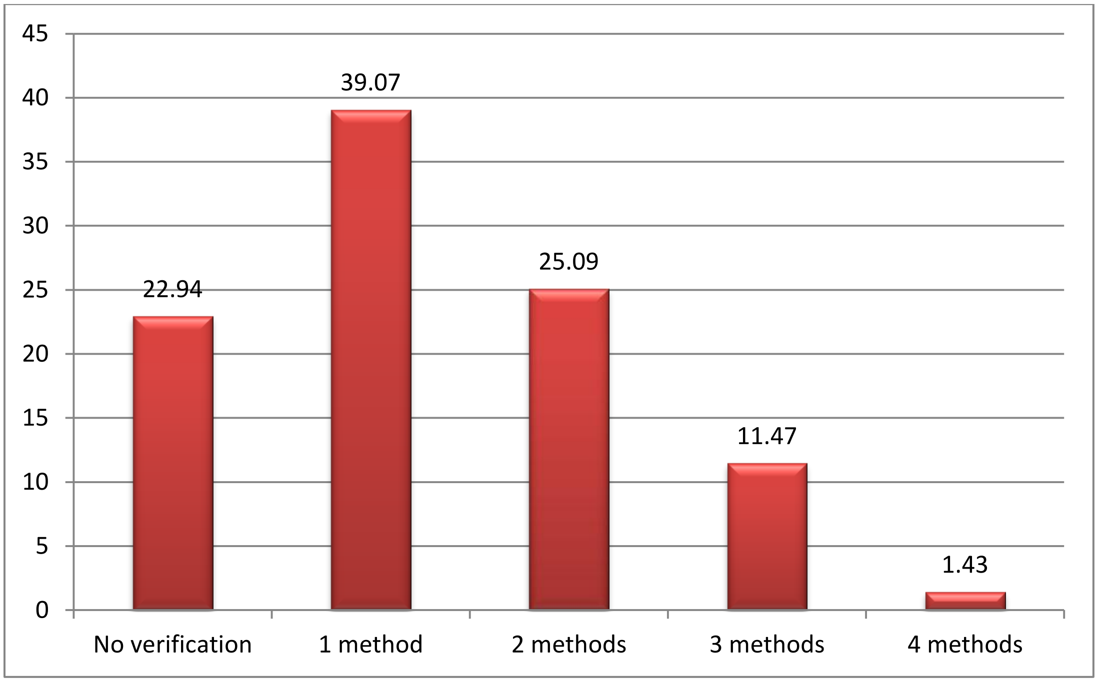
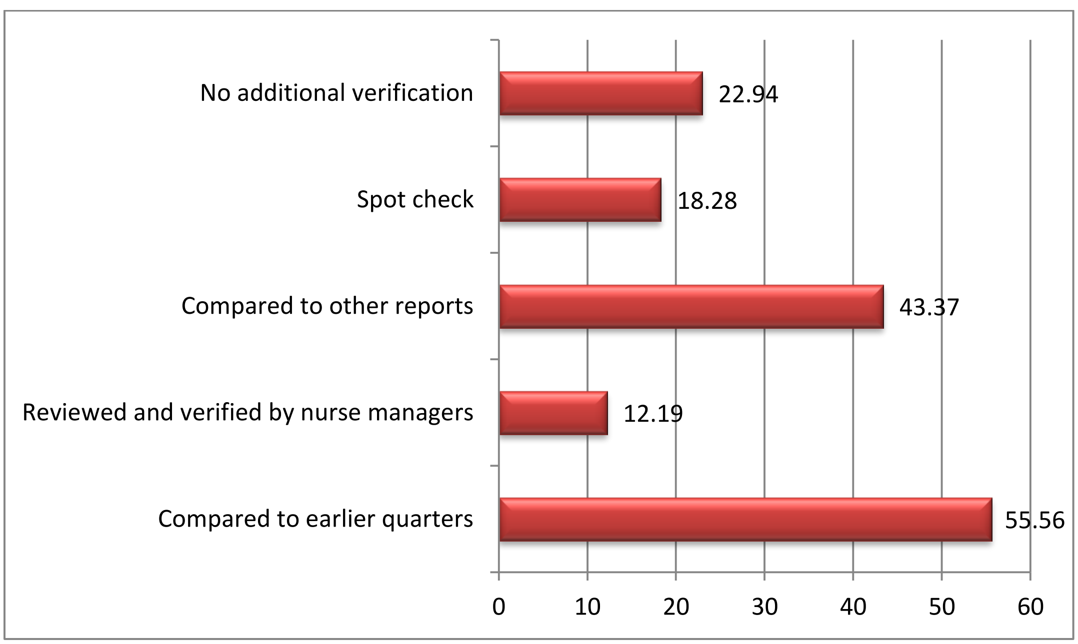


Figure 9. Utilization of different data verification methods (%)

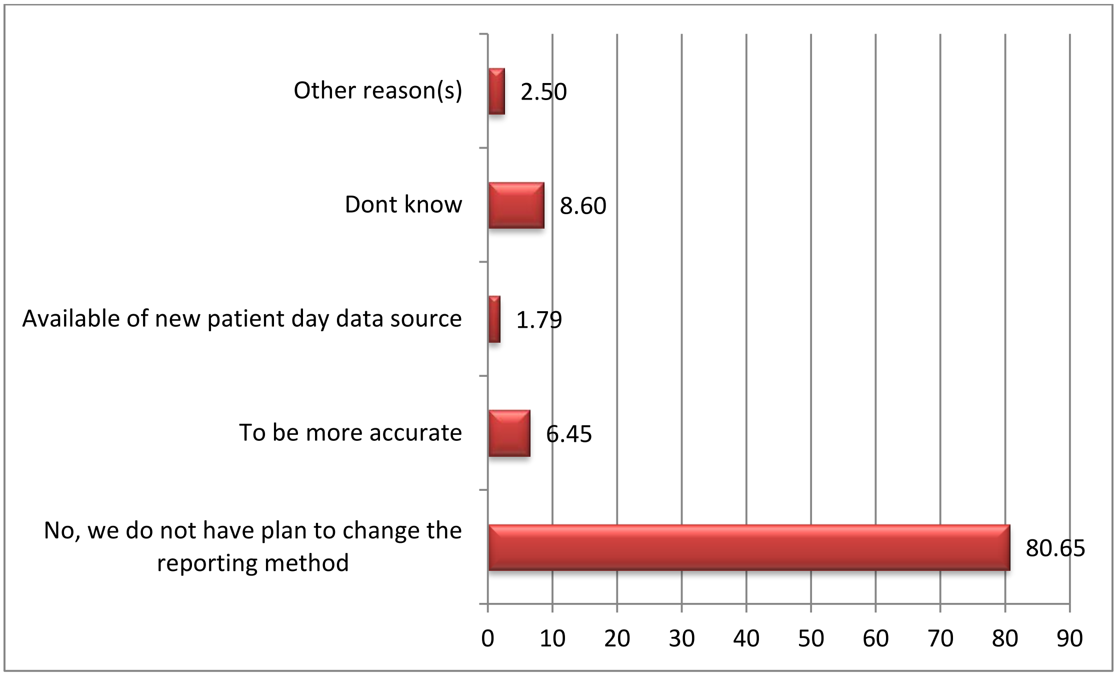


Other issues related to patient days data collection

When asked about future plans to change patient days data collection method within the next six months, about 81% of the SCs expressed that they did not plan to change the method (**Figure 10**). Among those who indicated a plan of changing the methods (11% or 30 SCs), the reasons for change varied. These reasons included “to be more accurate”, “available of new patient day data sources,” “changing units with sd and cc beds”, and “increase of short stay or observational patients on units.”

SCs were also asked to report how their hospitals track actual patient hours from short stay patients, “medical records (electronic, paper, or both),” “administrative/financial department data,” “others, please specify” or “do not know.” Over half of the SCs (57%) indicated that the administrative/financial department provided hours from short stay patients for NDNQI data submission. Approximately 29% reported they used medical records. Among those who reported “other,” the answer they gave varied and included “other patient hour tracking programs/systems, such as patient classification system,” “tracked by quality data personnel,” “specific ADT information,” “bed control program that counts bed at noon and midnight,” “use 4 data points in 24 hours to capture patients in a bed,” “time tracked by nursing supervisors,” and “do not submit short stay hours or not applicable,”

Figure 10. Plan to change data collectin method within the 6 months (%)



Part IV: Conclusion

The purpose of this study is to examine the reliability of the NDNQI quarterly patient days data collected via four different methods by comparing it to a dataset collected through a uniquely designed multi-site patient census study. Our analysis indicates excellent reliability in general as well as method-specific. All the identified ICCs (1, 1) were above 0.9 (**Table 6**) and none of their low boundary of 95% CI is below 0.88 (Lee, Koh, & Ong, 1989), which indicates no revision is required at the current moment NDNQI patient day data collection methods.

This study also confirmed that using different data collection methods can result in different levels of data accuracy. This finding is as expected and consistent with other literature (Beswick, Hill, & Anderson, 2010; Simon et al., 2011). However, it should be noted that this difference only exists in Method 5 (patient days from multiple census reports). Furthermore, results from the linear regression model indicated that reported patient days using Method 5 were likely to be overestimated.

We did not find a relationship between the DSSP and the difference between survey data and the NDNQI data (measured as the ratio of the NDNQI data to the multi-site patient census survey data). This should not be simply considered as DSSP does not influence patient days data collection method and accuracy. Our analyses showed that units with higher proportions of short stay patients were more likely to select the reporting methods tracking actual hours of the short stay patients. It should be noted that although the calculated DSSP partially reflected the situation of short stay patients on study units, it did not fully depict the picture of the proportion of patient days attributable to short stay patients. SCs also expressed in the survey that they felt more confident in both the appropriateness of the data collection method in use and the accuracy of the collected data on units without short stay patients compared to units with short stay patients (**Figure 5 and Figure 6**). These reasons may also explain that why no relationship was detected between unit types and the data accuracy.

Data accuracy is the main driving factor in selecting patient days data collection method; however method selection and data accuracy is also limited due to the availability of patient days data sources. Theoretically, units using Method 1 (midnight census) should be only those without short stay patients. Our analyses showed that 44% of the units with Method 1 (midnight census) have some short stay patients. Findings from the SC survey supported this conclusion as well (**Figure 4**).

Our study suggested that the NDNQI should pay close attention to the content in the NDNQI data collection guideline and/or online tutorials and maybe provide several optional approaches for units to validating data before submission, in order to help units collect accurate patient day data (**Figure 7, Figure 8 and Figure 9**). SCs expressed the frequent use of the NDNQI guidelines and online tutorial during data collection. Other information such as how short stay patients can influence patient days measurement (and associated factors like total nursing care hours per patient day or falls) and how to conduct pilot studies to collect patient days for short stay patients could also be added into the NDNQI data collection guidelines.

Midnight census is the most frequently used method in collecting patient days data. Yet, it is the method least likely to result in accurate patient day data at the presence of short stay patients (Beswick et al., 2010). One of the approaches that can better capture the existence of short stay patients without a dramatic increase in workload is the midnight/noon census method. There are units that have used this method for patient days data collection according to the SC survey. Moreover, evidence has suggested that collecting the counts of patient admissions and discharges along with the collection of patient days data can better reflect the care workload for nurses (Beswick et al., 2010).

Given the increasing use of the Electronic Health Records (EHRs) in health care settings, units should be encouraged to use EHRs which can provides more accurate information regarding patients' stay on a unit. The SC survey showed that

EHRs is not used (29%) or only partially used (32%) in reporting patient days data, although 93% of the SCs claimed that their hospitals had EHRs. This also requires a more active role of the SCs coordinating data collection activities between the units and department in charge of the EHRs.

This study has a couple of limitations. First, we were not able to obtain the exact admission and discharge time of each patient on participating units. However, as illustrated in **Figure 7** and described in on page 6, collecting patient counts every two hours does not substantially increase workload for nurses while generate patient days data that is very similar as using admission and discharge time. Another limitation is that the multi-site patient census survey did not collect information on the proportion of patient days from short stay patients. However, the proportion of short stay patients is not the primary aim of this study; and we were able to create a variable that partially reflects the existence of short stay patients by asking whether there were short stay patient(s) on each data collection day.

References

Beswick, Susan, Hill, Pamela D, & Anderson, Mary Ann. (2010). Comparison of nurse workload approaches. *Journal of Nursing Management, 18*(5), 592-598.

Carayon, Pascale, Wetterneck, Tosha B, Cartmill, Randi, Blosky, Mary Ann, Brown, Roger, Kim, Robert, . . . Wood, Kenneth E. (2013). Characterising the complexity of medication safety using a human factors approach: an observational study in two intensive care units. *BMJ quality & safety*, bmjqs-2013-001828.

Chen, L. F., Sexton, D. J., Kaye, K. S., & Anderson, D. J. (2009). Patient-days: a better measure of incidence of occupational bloodborne exposures. *Am J Infect Control, 37*(7), 534-540. doi: 10.1016/j.ajic.2009.02.004

Flynn, Linda, Liang, Yulan, Dickson, Geri L, Xie, Minge, & Suh, Dong-Churl. (2012). Nurses' practice environments, error interception practices, and inpatient medication errors. *Journal of Nursing Scholarship, 44*(2), 180-186.

Forbes, S, & Taunton, RL. (1994). Reliability of aggregated organizational data: An evaluation of five empirical indices. *Journal of Nursing Measurement, 2*(1), 37-48.

Giraudeau, B, & Mary, JY. (2001). Planning a reproducibility study: how many subjects and how many replicates per subject for an expected width of the 95 per cent confidence interval of the intraclass correlation coefficient. *Statistics in medicine, 20*(21), 3205-3214.

Harris, Paul A, Taylor, Robert, Thielke, Robert, Payne, Jonathon, Gonzalez, Nathaniel, & Conde, Jose G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics, 42*(2), 377-381.

He, Jianghua, Dunton, Nancy, & Staggs, Vincent. (2012). Unit-level time trends in inpatient fall rates of US hospitals. *Medical Care, 50*(9), 801-807.

Hughes, L.C., & Anderson, R.A. (1994). Issues regarding aggregation of data in nursing systems research. *Journal of Nursing Measurement, 2*(1), 79-101.

Kalisch, Beatrice, Friese, Christopher R, Choi, Seung Hee, & Rochman, Monica. (2011). Hospital nurse staffing: choice of measure matters. *Medical Care, 49*(8), 775.

Kalisch, Beatrice J, Tschannen, Dana, & Lee, Kyung Hee. (2012). Missed nursing care, staffing, and patient falls. *Journal of Nursing Care Quality, 27*(1), 6-12.

Lake, Eileen T, Shang, Jingjing, Klaus, Susan, & Dunton, Nancy E. (2010). Patient falls: association with hospital Magnet status and nursing unit staffing. *Research in nursing & health, 33*(5), 413-425.

Lee, James, Koh, David, & Ong, CN. (1989). Statistical evaluation of agreement between two methods for measuring a quantitative variable. *Computers in biology and medicine, 19*(1), 61-70.

Lerner, Nancy B. (2013). The Relationship Between Nursing Staff Levels, Skill Mix, and Deficiencies in Maryland Nursing Homes. *The health care manager, 32*(2), 123-128.

Li, Yu-Fang, Wong, Edwin S, Sales, Anne E, Sharp, Nancy D, Needleman, Jack, Maciejewski, Matthew L, . . . Liu, Chuan-Fen. (2011). Nurse staffing and patient care costs in acute inpatient nursing units. *Medical Care, 49*(8), 708-715.

Montalvo, I. (2007). The National Database of Nursing Quality Indicators (NDNQI). The Online Journal of Issues in Nursing, 12(3),

Nedved, Patricia, Chaudhry, Rabhea, Pilipczuk, Dina, & Shah, Shital. (2012). Impact of the Unit-Based Patient Safety Officer. *Journal of Nursing Administration, 42*(9), 431-434.

Simon, M., Yankovskyy, E., Klaus, S., Gajewski, B., & Dunton, N. (2011). Midnight census revisited: Reliability of patient day measurements in US hospital units. *Int J Nurs Stud, 48*(1), 56-61. doi: 10.1016/j.ijnurstu.2010.07.002

Simon, M., Yankovskyy, Y., & Dunton, N. (2010). Solving the mystery of patient days and the midnight census. *Nurs Manage, 41*(2), 12-14. doi: 10.1097/01.numa.0000368560.92652.66

Staggs, Vincent S. (2013). Nurse staffing, RN mix, and assault rates on psychiatric units. *Research in nursing & health, 36*(1), 26-37.

Staggs, Vincent S, & He, Jianghua. (2013). Recent Trends in Hospital Nurse Staffing in the United States. *The Journal of nursing administration*.

Twigg, D., Duffield, C., Bremner, A., Rapley, P., & Finn, J. (2011). The impact of the nursing hours per patient day (NHPPD) staffing method on patient outcomes: a retrospective analysis of patient and staffing data. *Int J Nurs Stud, 48*(5), 540-548. doi: 10.1016/j.ijnurstu.2010.07.013

Van den Heede, Koen, Sermeus, Walter, Diya, Luwis, Clarke, Sean P, Lesaffre, Emmanuel, Vleugels, Arthur, & Aiken, Linda H. (2009). Nurse staffing and patient outcomes in Belgian acute hospitals: cross-sectional analysis of administrative data. *International Journal of Nursing Studies, 46*(7), 928-939.

Wilson, Sally, Bremner, Alexandra, Hauck, Yvonne, & Finn, Judith. (2011). The effect of nurse staffing on clinical outcomes of children in hospital: a systematic review. *International Journal of Evidence‐Based Healthcare, 9*(2), 97-121.