PRESSURE ULCER RISK AND PREVENTION: EXAMINING THE INTER-RATER RELIABILITY OF THE NATIONAL DATABASE OF NURSING QUALITY INDICATORS® (NDNQI)

By

Shirley Moore Waugh

Submitted to the graduate degree program in Nursing and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy

_____

Chairperson: Sandra Bergquist-Beringer

_____

Karen Wambach

_____

Shin Hye Park

_____

Jianghua He

_____

Lynne Connelly

Date Defended: April 6, 2015

The Dissertation Committee for Shirley Moore Waugh

certifies that this is the approved version of the following dissertation:

PRESSURE ULCER RISK AND PREVENTION: EXAMINING THE INTER-RATER RELIABILITY OF THE NATIONAL DATABASE OF NURSING QUALITY INDICATORS® (NDNQI)

_____

Chairperson Sandra Bergquist-Beringer

Date approved: April 25, 2015

ABSTRACT

Measuring and reporting performance have become the norm. The purpose of this descriptive multi-site (*N* = 36 NDNQI-participating hospitals) study was to examine the reliability of the National Database of Nursing Quality Indicators® (NDNQI®) pressure ulcer (PrU) risk and prevention measures. This is the first known study to examine the inter-rater reliability of these measures.

Data for Part 1 of this two-part study were extracted from 1,637 patient records by 120 raters. One rater at each hospital was considered the "expert". Agreement between the expert and non-expert raters was calculated for the *risk* measures. Among the patients, 530 were "at risk" for PrU, and included in calculations of agreement for the *prevention* measures. In Part 2, raters completed an online survey about the methods they use to collect these data.

Cohen's kappa values varied widely within and across hospitals. Because most patients were assessed for PrU risk, and those at risk received prevention, the prevalence of a "Yes" response was high suggesting prevalence-adjusted kappa (PAK) may be a better estimate of inter-rater reliability than Cohen's kappa. PAK values for: *Skin assessment*, PAK = .977, 95% CI [.966 – .989]; *Risk assessment*, PAK = .978, 95% CI [.964 –.993]; *Time since last risk assessment*, PAK = .790, 95% CI [.729 – .852]; *Risk assessment scale*, PAK = .997, 95% CI [.991 – 1.0]; *Risk status*, PAK = .877, 95% CI [.838 – .917]; *Any prevention*, PAK = .856, 95% [.769 – .943]; *Skin assessment documented*, PAK = .956, 95% CI [.904 – 1.0]; and *Pressure-redistribution surface use*, PAK = .839, 95% CI [.763 – .916] indicated substantial to near perfect agreement. PAK values for: *Routine repositioning*, PAK = .577, 95% CI [.494 – .661];

Nutritional support, PAK = .500, 95% CI [.418 – .581]; and *Moisture management*, PAK = .556, 95% CI [.469 – .643] indicated moderate agreement.

Results provide support for the reliability of all (5) PrU *risk* measures, and three of six *prevention* measures. Areas of disagreement between the expert and non-expert raters should direct education to improve reliability. Results of the online survey suggest raters need further training on the NDNQI guidelines for PrU data collection.

ACKNOWLEDGEMENTS

Thank you to my dissertation committee, Dr. Sandra Bergquist-Beringer, Dr. Karen Wambach, Dr. Shin Hye Park, Dr. Jianghua He (Wendy), and Dr. Lynne Connelly, for their time and guidance. Special appreciation is given to my program advisor and dissertation committee chair, Dr. Bergquist-Beringer, for her expertise, tireless efforts, high standards, prompt feedback, and attention to detail. Without her, this dissertation could not have taken place. Special appreciation is also given to Dr. He for her patience, kindness, and guidance while overseeing my statistical analyses. I would be remiss if I did not acknowledge other faculty members who taught and mentored me here at the University of Kansas, as well as at Washburn University. Thank you to the many graduate biostatistics students who helped me along the way. They developed the Excel table I used to calculate the Cohen's kappa and prevalence-adjusted kappa values. My appreciation is extended to the NDNQI staff who oversaw this research. They were always available to me, and their expertise in NDNQI research is truly impressive. My gratitude is given to the hospitals who participated in this study.

Others who made it possible for me to complete this journey include my classmates. Without your friendship and support I could not have earned this degree. Finally, thank you to my family and friends for "picking up the slack", understanding my absence, and expecting me to succeed. I would like to dedicate this research to the nurses who provide these pressure ulcer risk and prevention interventions. They are the heart and soul of nursing.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I: INTRODUCTION

". . . if he has a bed-sore, it is generally the fault not of the disease, but of the nursing."

Florence Nightingale (1860)

The problem of pressure ulcers (PrUs) is well-documented. While not all PrUs are avoidable, most are (Black et al., 2011), thus they are now considered to be the result of poor patient quality of care. A number of non-profit and governmental agencies have worked to decrease PrU occurrence through programs that promote use of interventions to prevent them. Measurement of these interventions is necessary to evaluating quality of care initiatives toward decreasing PrU occurrence. No studies were found that examined the reliability of PrU risk and prevention measures.

Pressure Ulcers

Pressure ulcers are extremely painful (Rastinehad, 2006) and decrease quality of life (Gorecki et al., 2009). They are associated with increased mortality (Lyder et al., 2012; Russo, Steiner, & Spector, 2008), increased likelihood of readmission within 30 days of discharge, and increased hospital length of stay (11.2 days compared to 4.8 days) (Lyder et al., 2012). In 2006, 503,300 hospital stays included a Stage III or Stage IV PrU diagnosis (Russo et al., 2008). These patients were more likely to be discharged to a long-term care facility compared to those without a PrU. Almost 75% of those with a hospital-acquired PrU (HAPU) are 65 years of age and older.

While treatment costs associated with PrUs varies according to severity of the wound, a Stage III or Stage IV PrU can take several months to more than two years to heal, at an estimated cost of $105,846 (Chan et al., 2012) to more than $240,000 (Schessel, Ger, & Oddsen, 2012) per PrU. Overall costs for all PrU stages are estimated to exceed $4200/month (Chan et al., 2012).

These costs do not include lost wages. Treatment costs far exceed costs associated with PrU prevention (Padula, Mishra, Makic & Sullivan, 2011; Pham et al., 2011; Spetz, Brown, Aydin, & Donaldson, 2013).

The Institute of Healthcare Improvements' 5 Million Lives Campaign encouraged hospitals to adopt interventions aimed at preventing patient harm, including pressure ulcers (Ayello & Lyder, 2008; Duncan, 2007). The goal of this 2007 – 2008 campaign was zero PrU occurrences. Healthy People 2020 set the goal of eliminating PrU hospitalizations among older adults by 10% (Department of Health and Human Services, 2007). One of the most influential PrU prevention initiatives was the Center for Medicare and Medicaid Services' (CMS) decision to deny payment for the cost of treating Stage III and Stage IV hospital-acquired PrUs effective October 1, 2008; private payers followed suit (Blue Cross & Blue Shield, 2012). More recently, the Partnership for Patients Initiative, a public-private partnership of physicians, nurses, hospitals, employers, patients, and governmental agencies; identified PrUs as one of ten core patient safety focus areas. This initiative's goal was to reduce the 2010 rate of hospital-acquired Stage III or Stage IV pressure ulcers by 40% (CMS, n. d.). Performance of "multicomponent interventions to reduce pressure ulcers" was among the top ten strategies "strongly recommended" (p. 366) for immediate implementation by an international panel of 21 stakeholders and evaluation experts (Shekelle et al., 2013). This list was based on evaluations of the effectiveness of these strategies.

Definition and Classification

A pressure ulcer is defined as a "localized injury to the skin and/or underlying tissue usually over a bony prominence, as a result of pressure, or pressure in combination with shear. A number of contributing or confounding factors are also associated with pressure ulcers; the

significance of these factors is yet to be elucidated" (NPUAP/EPUAP, 2009, p. 19). The National Pressure Ulcer Advisory Panel (NPUAP), (2012) defines and classifies pressure ulcers according to the amount of tissue damage (para 5-10).

Category/Stage I: Non-blanchable erythema

Intact skin with non-blanchable redness of a localized area usually over a bony prominence. Darkly pigmented skin may not have visible blanching; its color may differ from the surrounding area. The area may be painful, firm, soft, warmer or cooler as compared to adjacent tissue. Category I may be difficult to detect in individuals with dark skin tones. May indicate "at risk" persons.

Category/Stage II: Partial thickness

Partial thickness loss of dermis presenting as a shallow open ulcer with a red pink wound bed, without slough. May also present as an intact or open/ruptured serum-filled or sero-sanginous filled blister. Presents as a shiny or dry shallow ulcer without slough or bruising [bruising may indicate deep tissue injury]. This category should not be used to describe skin tears, tape burns, incontinence associated dermatitis, maceration or excoriation.

Category/Stage III: Full thickness skin loss

Full thickness tissue loss. Subcutaneous fat may be visible but bone, tendon or muscle are *not* exposed. Slough may be present but does not obscure the depth of tissue loss. *May* include undermining and tunneling. The depth of a Category/Stage III pressure ulcer varies by anatomical location. The bridge of the nose, ear, occiput and malleolus do not have (adipose) subcutaneous tissue and Category/Stage III ulcers can be shallow. In contrast, areas of significant adiposity can develop extremely deep Category/Stage III pressure ulcers. Bone/tendon is not visible or directly palpable.

Category/Stage IV: Full thickness tissue loss

Full thickness tissue loss with exposed bone, tendon or muscle. Slough or eschar may be present. Often includes undermining and tunneling. The depth of a Category/Stage IV pressure ulcer varies by anatomical location. The bridge of the nose, ear, occiput and malleolus do not have (adipose) subcutaneous tissue and these ulcers can be shallow. Category/Stage IV ulcers can extend into muscle and/or supporting structures (e.g., fascia, tendon or joint capsule) making osteomyelitis or osteitis likely to occur. Exposed bone/muscle is visible or directly palpable.

Unstageable/Unclassified: Full thickness skin or tissue loss – depth unknown

Full thickness tissue loss in which actual depth of the ulcer is completely obscured by slough (yellow, tan, gray, green or brown) and/or eschar (tan, brown or black) in the wound bed. Until enough slough and/or eschar are removed to expose the base of the wound, the true depth cannot be determined; but it will be either a Category/Stage III or IV. Stable (dry, adherent, intact without erythema or fluctuance) eschar on the heels serves as "the body's natural (biological) cover" and should not be removed.

Suspected Deep Tissue Injury – depth unknown

Purple or maroon localized area of discolored intact skin or blood-filled blister due to damage of underlying soft tissue from pressure and/or ***shear.*** The area may be preceded by tissue that is painful, firm, mushy, boggy, warmer or cooler as compared to adjacent tissue. Deep tissue injury may be difficult to detect in individuals with dark skin tones. Evolution may include a thin blister over a dark wound bed. The wound may further evolve and become covered by thin eschar. Evolution may be rapid exposing additional layers of tissue even with optimal treatment.

Prevalence and Incidence

Fortunately, once highly prevalent, hospital-acquired pressure ulcer (HAPU) rates have decreased over time. Claims data from the nearly 40 million Medicare patients who were hospitalized from 2009 through 2011 revealed 8,812 Stage III, Stage IV, or unstageable HAPUs (0.65 cases/1,000 patients) (HealthGrades, 2013). This HAPU rate is notably lower than the 2007 through 2009 rate of 368,261 Stage III, IV, or unstageable HAPUs among 14 million Medicare hospitalizations (26.64 cases/1,000 patients)—which cost Medicare $1.99 billion (Reed & May, 2011). Others have reported similar trends. The International Pressure Ulcer Prevalence Survey™ reported U.S. hospital-acquired Stage II or higher PrU (HAPU2+) prevalence rates were 6.4% in 2006 compared to 5.0% in 2009 (VanGilder, Amlung, Harrison, & Meyer, 2009), and 3.6% in 2012 (VanGilder, Lachenbruch, Harrison, & Meyer, 2013). The percentages that these 2012 HAPUs were classified as were Stage I (33.8%), Stage II (37.3%), Stage III (3.3%), Stage IV (1.2%), unstageable (9.2%), suspected deep tissue injury (13.6%), indeterminable (1.1%), and stage not collected (0.5%). These data highlight the degree to which HAPU rates are underestimated by indicators that do not count all PrU categories/stages.

Data on *all stages* of HAPUs showed similar trends. These overall rates decreased each year from a high of 10.4% in 2003 to a low of 1.8% in 2010 (Stotts, Brown, Donaldson, Aydin, & Fridman, 2013). The National Database of Nursing Quality Indicators® (NDNQI®) PrU data collected in 2010 revealed a HAPU rate of 3.6% (Bergquist-Beringer, Dong, He, & Dunton, 2013); compared to 6.5% from 2006 through 2007 (Bergquist-Beringer, Gajewski, & Davidson, 2012). He, Staggs, Bergquist-Beringer, and Dunton (2013) reported a statistically significant decrease in HAPU rates from 2004 to 2007, and even more so from 2008 to 2011. Reduction in HAPU occurrence has been attributed to national initiatives (O'Reilly, 2008) and prevention

programs (Chou et al., 2013; Niederhauser et al., 2012, Soban, Hempel, Munjas, Miles, & Rubenstein, 2011; Sullivan & Schoelles, 2013) that promote use of the interventions to prevent them, such as the PrU risk and prevention measures examined in this study.

Clinical Guidelines for Pressure Ulcer Prevention

Clinical guidelines for PrU prevention provide practitioners with concise instructions for reducing PrUs that are based on current scientific evidence (Langemo et al., 2008). They also bring to light healthcare issues, influence public policy, and improve quality of care (Woolf, Grol, Hutchinson, Eccles, & Grimshaw, 1999). The original guidelines for PrU prevention and treatment were developed in 1992 by the Agency for Healthcare Policy and Research—now the Agency for Healthcare Research and Quality (AHRQ). These guidelines were re-endorsed in 2001 (Shekelle et al., 2001).

Most recently, members of the NPUAP and EPUAP, and the Pan Pacific Pressure Injury Alliance (PPPIA)—an alliance of wound care organizations across Australia, New Zealand, Hong Kong, and Singapore—collaborated to update the NPUAP and EPUAP 2009 Clinical Practice Guidelines for the Prevention and Treatment of Pressure Ulcers. These guidelines were released October of 2014 (NPUAP, EPUAP, PPPIA, 2014). Others have developed guidelines including the Wound Ostomy Continence Nurses (WOCN) Society guidelines that were first developed in 2003, and updated in 2010 (Ratliff, Tomaselli, & The Guideline Task Force, 2010). The NPUAP/EPUAP/PPPIA guidelines and the WOCN guidelines are well-known, easily accessible, and commonly accepted (Berlowitz et al., n. d.; Soban et al., 2011).

These PrU prevention guidelines focus on reducing risk for PrU development. The first step in prevention is gathering information on an individual's risk for PrU development and the status of their skin. Based on this information, specific PrU prevention interventions are selected

for implementation. These interventions include: (1) PrU risk assessment; (2) skin assessment; (3) reducing risk by minimizing shear and minimizing pressure through routine repositioning, use of a pressure-redistribution surface, and offloading pressure from the heels (elevating the lower legs to completely lift the heel from the bed); (4) moisture management and skin care; (5) nutritional support; and (6) patient and caregiver education (NPUAP/EPUAP/PPPIA, 2014; Ratliff et al., 2010).

Pressure Ulcer Risk Assessment and Skin Assessment

Pressure ulcer risk assessment is determining the degree that an individual is at risk for PrU development. A risk assessment involves identifying patient factors known to be associated with PrU development. A PrU risk assessment should be completed (and documented) using a reliable and valid tool within 8 hours of admission, at regular intervals, and after a change in patient status. These assessments guide which prevention interventions are to be performed. Additional risk factors (not on the structured PrU risk assessment tool) should be considered; such as advanced age, obesity, tobacco use, co-morbid conditions, and history of PrU. Risk status, therefore, should be determined by a mixture of clinical judgment and a risk assessment tool.

Skin assessment is a head-to-toe assessment of an individual's skin. It includes visual inspection of the entire body, as well as assessment of skin temperature, moisture, color, turgor, integrity (Ayello et al., 2009), edema, tissue density (firm or mushy), and localized pain. Skin assessments should be systematically performed and documented upon admission, at regular intervals, with each PrU risk assessment, and after a change in status. A skin assessment must be performed and documented as close as possible to the time of admission in order to accurately determine if a PrU is hospital-acquired or community-acquired (i.e. present-on-admission).

Pressure Ulcer Prevention Interventions

Minimizing shear involves keeping skin clean and dry, using lift sheets, maintaining the head of the bed at or below 30 degrees, lowering the head of the bed one hour after meals, use of overhead trapeze bars, avoiding vigorous massage over bony prominences, and applying transparent films or hydrocolloids to bony prominences.

Pressure is minimized by routine repositioning at a frequency determined by individual activity/mobility level and medical condition(s)—usually every 2 hours (while on a standard mattress) or every 4 hours (while on a pressure-redistribution mattress), and at least every 1 hour while seated. Those seated who can reposition themselves, however, should relieve pressure every 15 minutes, such as by a chair push-up or forward-leaning. Pressure-redistribution surfaces (foam, gel, or air mattresses/overlays/cushions) are recommended while in bed or a chair; as are frequent small position changes using pillows and wedges. Heels should be offloaded. Strategies to reduce PrUs among patients in the operating room include the use of a pressure-redistribution surface for surgeries lasting four hours or longer. Minimizing hypotensive and low body temperature episodes during surgery are also important as are proper body positioning to minimize pressure—including offloading of the heels.

Moisture management and skin care includes establishing a bowel/bladder retraining program. Skin must be kept clean and dry using gentle, pH-balanced cleansers and smooth disposable wipes. Other interventions are the use of skin protectants such as barrier creams/ointments/pastes, fecal management systems, and high-quality disposable diapers/pads.

Nutritional support equals offering the person at nutritional and PrU risk at least 30 – 35 kcal/kg/day that includes 1.25 – 1.5 g/kg/day of protein, and 1 ml of fluid/kcal/day. In

addition, these individuals should receive a dietary consult. In general, nutritional intake must be monitored and appropriate nutritional interventions provided through the appropriate route.

Patient and caregiver education on causes and risk factors for PrUs, as well as how to decrease these risks, should be presented. Specifically, patients and caregivers must be taught the importance of (and how to properly perform) routine repositioning, regular skin inspections, maintaining adequate nutritional and fluid status, preventing friction and shear, and keeping the skin clean and dry.

The NPUAP/EPUAP/PPPIA and the WOCN Society Guidelines provide strength of evidence ratings for each of their recommendations with Level A being the strongest, followed by Level B, and then Level C (the weakest). Within the WOCN Society Guidelines, a Level A rating indicates that the recommendation was supported by (a) at least two randomized controlled trials (RCTs) of at least 10 people with PrUs, (b) a meta-analysis of RCTs, or (c) a Cochran Review of RCTs. A Level B rating indicates that the recommendation was supported by (a) at least one controlled trial of at least 10 people with PrUs, or (b) at least two supporting non-randomized trials of at least 10 people with PrUs. The lowest level, a Level C rating, indicates that the recommendation was supported by (a) two supporting case series of at least 10 people with PrUs, or (b) expert opinion. Within the WOCN Guidelines, only three of 22 (12%) recommendations were supported by Level B evidence, while 22 of 25 (88%) recommendations were supported by Level C evidence (Ratliff et al., 2010). There were no Level A ratings.

The NPUAP/EPUAP/PPPIA definition of Level A differs slightly from the WOCN's definition. Level A is defined by NPUAP/EPUAP/PPPIA as "direct scientific evidence from properly designed and implemented controlled trials on pressure ulcers in humans (or humans at-risk for pressure ulcers); providing statistical results that consistently support the guideline

statement (Level 1 studies required)" (NPUAP, EPUAP, & PPPIA, 2014, p. 4). Level 1 studies are "Random trial(s) with clear-cut results and low risk of error OR systematic literature review or meta-analysis according to the Cochrane methodology or meeting at least 9 out of 11 quality criteria according to AMSTAR appraisal tool" (p. 4). Within the NPUAP/EPUAP/PPPIA Guidelines, none of the recommendations were supported by Level A evidence, 24% by Level B evidence, and 76% by Level C evidence (NPUAP/EPUAP/PPPIA, 2014).

The National Database for Nursing Quality Indicators®

The NDNQI was established by the American Nurses Association (ANA) in 1998 as part of its Safety and Quality Initiative (ANA, 2014; Montalvo, 2007). During the 1990s, healthcare saw a reduction in patient care provided by registered nurses (RNs), decreased length of patient stay, and other cost-cutting measures. As a result, the ANA sought to collect data to show that RNs affect quality of care, and to fill the need for collecting nurse-sensitive outcomes—as opposed to the medical-focused outcomes commonly reported at that time (ANA, 2014; Montalvo, 2007). The purpose of the NDNQI is to (1) collect data on quality indicators amenable to nursing care, (2) provide information to participating hospitals so they may compare their own performance over time and to similar hospitals and nursing units, and (3) to build evidence on the relationship between nurse staffing and patient outcomes (ANA, 2014). The NDNQI was managed by the University of Kansas Medical Center (KUMC), School of Nursing under contract with the American Nurses Association and with Press Ganey through December, 2014. The NDNQI is an accepted nurse registry that satisfies CMS voluntary reporting (ANA, 2013). It also meets Magnet application requirements (ANA, 2013).

From its beginning, the NDNQI began collecting data on their PrU outcome indicator, but *quarterly* HAPU data collection did not begin until the year 2000 – 2001 (Bergquist-

Beringer, 2011). Over the years, this data collection was expanded to include suspected deep tissue injury in 2008, and pediatric units in 2009. Also, inter-rater reliability studies were conducted on PrU identification, staging, and origin in 2004 through 2006; the online PrU Training program was developed in 2006, and has been updated routinely (2007, 2008, 2009, 2010, 2012, and 2014) (S. Bergquist-Beringer, personal communication, March 19, 2014). The NDNQI PrU advisory committee was formed in 2007.

The NDNQI Stage II and higher HAPU rate (HAPU2+) is a National Quality Forum (NQF) endorsed measure. For a measure to be endorsed by the NQF, it must go through a rigorous process and meet specific criteria. Endorsed measures are generated from (and based on) established guidelines and an expert panel review, which continues until the measure is retired. This endorsement process is described in Chapter II. The NDNQI uses surveillance data to determine HAPU rates. This means the skin of all patients is directly examined for PrUs on a single day (point in time). Those PrUs identified are classified as community-acquired or hospital-acquired and staged. A hospital-acquired PrU (HAPU) is a PrU that developed after admission to the hospital. Any PrU that is not documented in the patient record upon admission is considered to be a HAPU (NDNQI, 2013). The NDNQI counts all categories/stages of HAPUs.

Since 1998, the following *process measures* have been added to the suite of NDNQI pressure ulcer measures (a) *Skin assessment on admission*, (b) *PrU risk assessment on admission*, (c) *Time since last PrU risk assessment,* (d) *Risk assessment method,* (e) *Risk status*, (f) *Performance of PrU prevention (yes/no)*, (g) *Pressure-redistribution surface use*, (h) *Nutritional support*, (i) *Routine repositioning*, (j) *Skin assessment within the past 24 hours,* and (k) *Moisture management* (Bergquist-Beringer et al., 2013).

The accuracy of the NDNQI pressure ulcer data is supported by ongoing: (a) data checks by the NDNQI staff and participating hospitals, as well as error report alerts; (b) training of data collectors; and (c) assurance by facilities' site coordinators that data are collected in accordance with NDNQI's definitions and collection requirements (NDNQI, 2013). It is essential, however, that empirical support for the NDNQI PrU risk and prevention data's reliability be established, something that has not been done. This study hopes to begin establishing that support.

Measuring Quality of Care

Measuring quality is the first step to improve quality of care and typically involves measuring healthcare structures, processes, and patient outcomes. Structures include elements such as the physical characteristics, size, and type of facility; organizational structure; resources and equipment; qualifications and numbers of staff; and financing. Processes are the care itself, such as the activities that go on between the caregiver and the patient. These include characteristics of care such as if care is timely, appropriate, and comprehensive. Outcomes are measures of the end results of care such as morbidity, mortality, prevalence, incidence, readmissions, adverse events, cost of services provided, and patient satisfaction and quality of life (Donabedian, 1978). Collectively, these measures are called *quality indicators*. Data from the measurement of quality indicators are used to measure quality of care, and monitor and guide quality improvement efforts (de Vos, Graafmans, Kooistra, Meijboom, & Westert, 2009).

Sound measurement is essential if research findings are to be credible: it is the cornerstone of quality research. It is imperative that what is intended to be measured is actually measured, and the same or similar results are produced when the measurement process is repeated. Two fundamental components of any measurement tool or measuring method are

validity and reliability. Valid tools measure what they are intended to measure, while reliability is concerned with consistency of measures (DeVon et al., 2007; Salkind, 2006).

Establishing reliability is the *first step* in making sure an instrument is psychometrically sound (Fleiss, Levin, & Paik, 2003; Portney & Watkins, 2009; Salkind, 2006). While a measure can be reliable and not valid, it *cannot be valid without also being reliable*: It is meaningless to even ask what is associated with a variable when values given to it cannot be trusted (Fleiss et al., 2003; Salkind, 2006; Shrout, 1998).

While numerous studies have examined PrU processes and outcomes (Bergquist-Beringer et al., 2013; Niederhauser et al., 2012, Soban et al., 2011; Sullivan & Schoelles, 2013), the processes themselves—as well as how they were measured—vary across studies. Only one study used standardized measures of process indicators to evaluate PrU risk and prevention performance (Bergquist-Beringer et al., 2013). Especially noteworthy, however, is that there is a scarcity of evidence supporting the reliability of the data on these process of care indicators; and *no studies* have examined the reliability of data collected from the NDNQI PrU risk and prevention process care measures.

Statement of the Problem

Healthcare consumers and payers demand that care be effective, efficient, timely, accessible, and cost-effective. As a result, measuring and reporting performance indicators have become the norm. Measurements of quality indicators guide improvement efforts, inform consumers, and are increasingly used by public and private payers for payment.

While numerous research studies have been conducted on PrU prevention, recent literature reviews have shown that the majority of these studies have focused on patient outcomes (Niederhauser et al., 2012; Soban et al., 2011; Sullivan & Schoelles, 2013). Few have

evaluated process of care measures. In addition, how these process of care measures have been operationalized has varied. Most concerning, however, is the lack of evidence for these measures' reliability, even though reliability is crucial to ensure findings are credible.

Because nursing is the largest group of professionals in the healthcare workforce, their impact on quality is significant. It is crucial, therefore, that *reliable nurse-sensitive indicators* serve as a strong foundation for quality improvement efforts. Although the NDNQI performs rigorous cleaning of data submitted to them by participating hospitals (Klaus, Dunton, Gajewski, & Potter, 2013), verifying the reliability of data is essential to establish their accuracy. Credibility of these PrU risk and prevention indicators is suggested by the fact that (a) CMS recognizes NDNQI as an accepted nurse registry for CMS's reporting requirements, (b) submission of these indicators fulfill Magnet application requirements, and (c) select NDNQI indicators have been endorsed by the National Quality Forum (ANA, 2014). However, *no studies have evaluated the reliability of NDNQI pressure ulcer risk and prevention indicators*.

Significance

This study is significant in that it is the first study to examine the inter-rater reliability of PrU risk and prevention data from the NDNQI. Establishing reliability is necessary because the data are used to drive improvement efforts, yet their accuracy has not been established. Because of the nation's demand for quality patient care, the fact that nurses are the largest workforce in providing that care, and that the NDNQI collects nurse-sensitive process measures for PrU risk and prevention from more than 1,400 hospitals across the U.S., *there is widespread interest in the findings this data generates*. Those involved with the NDNQI, therefore, should ensure data are accurate.

Specific Aims and Research Questions

Purpose Statement: The purpose of this study was to examine the reliability of the NDNQI pressure ulcer risk and prevention measures. The specific aims of this study are:

Aim 1: Examine the reliability of the NDNQI pressure ulcer risk and prevention measures within and across NDNQI hospitals. For this aim, this study used Cohen's kappa, prevalence-adjusted kappa (PAK) (see discussion on p. 73 – 76), percent agreement, intraclass correlation coefficient (*ICC*), and agreement matrices/descriptive statistics.

Question 1: What is the agreement between expert participant ratings and non-expert participant ratings *for each* of the 11 NDNQI pressure ulcer risk and prevention measures *within* each hospital? This question was answered by hospital. A Cohen's kappa value for each of the expert to non-expert comparisons was calculated for every risk and prevention measure. The Cohen's kappa values obtained from these comparisons for each measure were averaged. This yielded one Cohen's kappa value for each of the 11 NDNQI risk and prevention measures per hospital. Prevalence-adjusted kappa (PAK) values and percent agreement were calculated in the same way to obtain one PAK value and one percent agreement value for each of the 11 NDNQI risk and prevention measures per hospital.

Question 2: What is the *overall* agreement between expert participant ratings and non-expert participant ratings for the NDNQI PrU *risk* measures per hospital and the *overall* agreement between expert participant ratings and non-expert participant ratings for the NDNQI PrU *prevention* measures per hospital? The average Cohen's kappa value obtained for each measure in Question 1 was used to answer Question 2. Specifically, the average Cohen's kappa value obtained at each hospital from the expert and non-expert comparisons for each of the five PrU *risk* measures (skin assessment on admission, risk assessment on admission, time since last

risk assessment, risk assessment scale, and risk status) was averaged to obtain one overall kappa value for pressure ulcer *risk* per hospital. Similarly, the average Cohen's kappa value obtained from the expert and non-expert comparisons for each of the six PrU *prevention* measures (any PrU prevention, skin assessment, pressure-redistribution surface use, routine repositioning, nutritional support, and moisture management) at each hospital was averaged to obtain one overall Cohen's kappa value for PrU *prevention* per hospital.

Using this same method, each hospital's average PAK value obtained in Question 1 for each measure was used to answer Question 2. Specifically, the average PAK value obtained from the expert and non-expert comparisons for each of the five PrU *risk* measures (skin assessment on admission, risk assessment on admission, time since last risk assessment, risk assessment scale, and risk status) at each hospital were averaged to obtain one overall PAK value for pressure ulcer *risk* per hospital. Similarly, the average PAK value obtained from the expert and non-expert comparisons for each of the six PrU *prevention* measures (any PrU prevention, skin assessment, pressure-redistribution surface use, routine repositioning, nutritional support, and moisture management) at each hospital were averaged to obtain one overall PAK value for PrU *prevention* per hospital.

Question 3: What is the *average* of the *within hospital* agreements between expert participant ratings and non-expert participant ratings for each of the 11 NDNQI pressure ulcer risk and prevention measures *across* hospitals? The average Cohen's kappa value that was obtained for each measure to answer Question 1 was used to answer Question 3. Specifically, the average Cohen's kappa value that was obtained for each pressure ulcer risk and prevention measure at each hospital was averaged across hospitals to obtain one Cohen's kappa value for each risk and prevention measure across hospitals. Similarly, the average PAK value that was

obtained for each pressure ulcer risk and prevention measure at each hospital was averaged across hospitals to obtain one PAK value for each risk and prevention measure across hospitals. In addition, the average percent agreement value that was obtained for each PrU risk and prevention measure per hospital was averaged to obtain the percent agreement for each risk and prevention measure across hospitals.

Question 4: What is the intraclass correlation coefficient (agreement) between expert participant ratings and non-expert participant ratings for each of the 11 NDNQI PrU risk and prevention measures *across* hospitals? To answer this question, the Cohen's kappa value obtained from *each* of the expert to non-expert comparisons per risk and prevention measure at each hospital, was used to calculate an ICC_{kappa} for each of the 11 NDNQI risk and prevention measures *across* hospitals. Similarly, the PAK value from *each* of the expert to non-expert comparisons per risk and prevention measure at each hospital was used to calculate an ICC_{PAK} for each of the 11 NDNQI risk and prevention measures across hospitals.

Question 5: Where is the lack of agreement between expert and non-expert participant ratings on the 11 NDNQI pressure ulcer risk and prevention measures occurring? For instance, is the lack of agreement most often between "No" and "Unnecessary for patient", or "No" and "Patient refused", and least often between "No" and "Yes"? To answer this question, agreement matrices and descriptive statistics were created.

Aim 2: Examine the methods and processes used by participant raters to gather data on the NDNQI PrU risk and prevention measures. Descriptive analysis was used to address Aim 2.

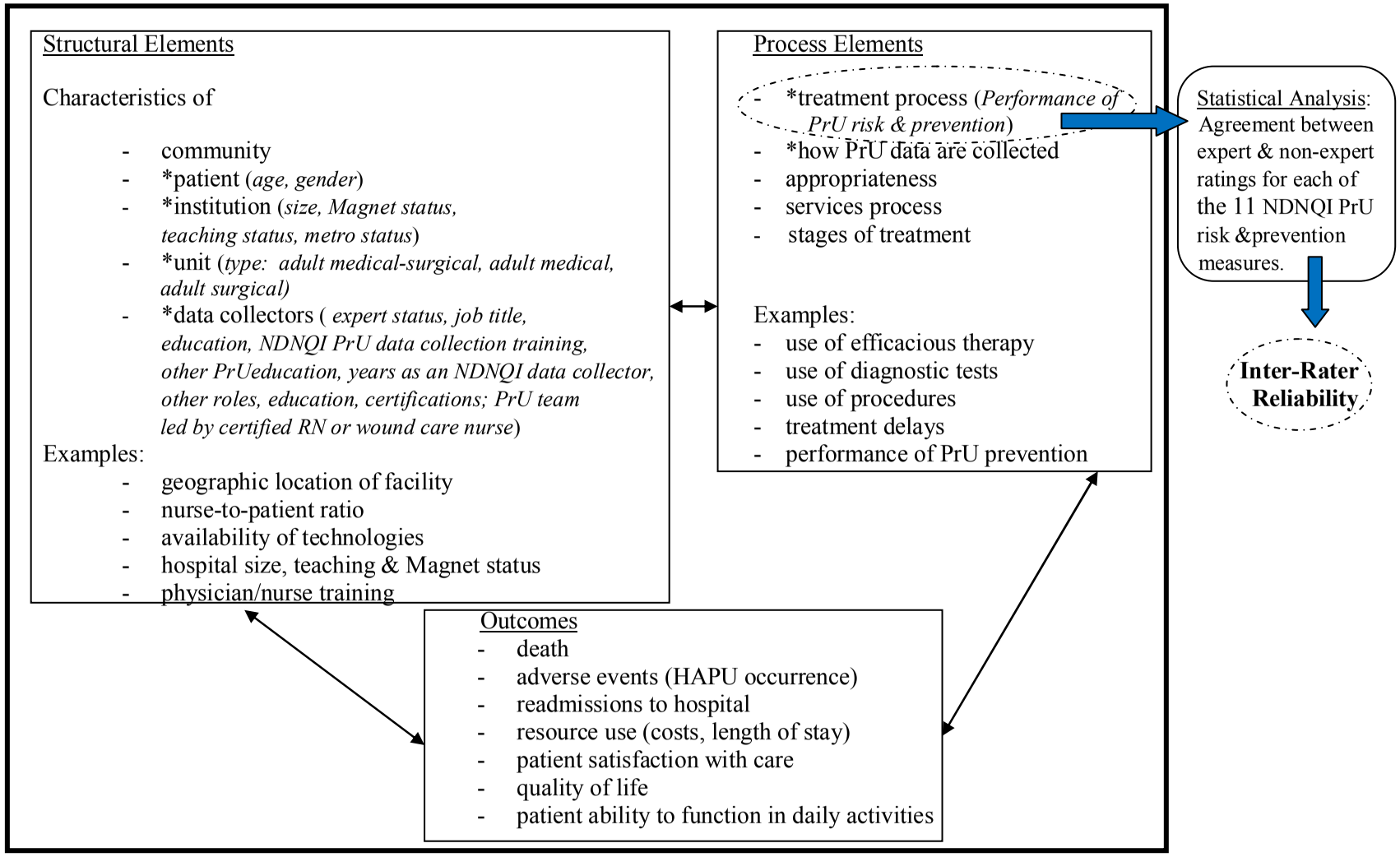
Conceptual Framework

Donabedian's model of measuring healthcare system performance (Donabedian, 1978) guided this study. Donabedian (1919 – 2000) was a physician and health services researcher. In

1966, while teaching at the University of Michigan, Donabedian first identified his structure-process-outcome model. This model posits that healthcare quality is influenced by structures, processes, and outcomes; each of which influences the others. Donabedian's purpose for building this model was to develop a sound method for evaluating healthcare quality because, at that time, he found the current methods of evaluating care to be inadequate (Donabedian, 2005). Donabedian's 1966 article, reprinted in 2005 (Donabedian, 2005), is considered a citation classic among healthcare systems research (Frenk, 2000). His early model was criticized for being too linear (Mitchell, Ferketich, & Jennings, 1998), and was modified to give details of the give-and-take relationships among "the system, interventions, client, and outcome components" (Duffy, 2009, p. 29). Provision of patient care, therefore, is mediated by characteristics of the patient and the system; and outcomes are influenced by all the variables in the model. The more current model, therefore, reflects multidirectional relationships between the concepts of structure, process, and outcome, and is used by the NDNQI; offering a comprehensive method for evaluating nursing-sensitive measures as they relate to healthcare quality (Montalvo, 2007).

Donabedian's modified model relevant to the NDNQI PrU outcome and process measures is presented in Figure 1. **Structures** include characteristics of the community, institution, patient, and provider. In general, structures are easily quantified and measured. For this study, structures are operationalized as: (a) data collection team members' characteristics that include expert status, job title, education level, review of the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers, completion of the NDNQI Pressure Ulcer Training program, other PrU data collection education, years as an NDNQI PrU data collector, other roles (besides chart reviewer) in NDNQI PrU data collection, wound care certification, and wound/skin care nurse status (b) the unit type that reflects the patient population in adult.

Figure 1. **Donabedian's Model Modified for Measuring Healthcare System Performance** (Donabedian, 1978; Mitchell et al., 1998) **& Reliability Assessment related to Pressure Ulcers**



* = Elements considered in this study

medical-surgical, medical, and surgical units; and (c) the hospital type of General Acute Care, number of staffed beds, teaching status, Magnet status, and metropolitan status

Processes include the care itself, such as the activities that go on between the caregiver and the patient. These patient care characteristics are to capture if the care is timely, appropriate, and comprehensive. Measuring these characteristics, therefore, is challenging. "Timely", "appropriate", and "comprehensive" must be defined, and those definitions vary among patient populations. Processes considered relevant to PrUs by the NDNQI are the performance of routine skin assessments and PrU risk assessments, identifying PrU risk status, and implementation of appropriate PrU prevention interventions for those "at risk" (skin assessment, pressure-redistribution surface use, repositioning, nutritional support, and moisture management). Pressure ulcer care process measures are rarely standardized measures, and data derived from them lack evidence of reliability.

Outcomes measure the end result of care and include morbidity, mortality, prevalence, incidence, readmission, and adverse event rates; cost of services provided; and patient satisfaction and quality of life (Donabedian, 1978). Outcomes considered relevant to PrUs by the NDNQI include HAPU rates, unit-acquired HAPU rates, and PrU stage/category (all stages/categories are included). No outcomes were measured in this study.

To test Donabedian's model—the influence of structure and process on outcome—one must assume that care processes are being accurately and reliably measured. Relevant to this study, the NDNQI uses Donabedian's model to understand the influence of PrU processes on PrU outcomes. The model implies that in order to understand the relationships between PrU processes and PrU outcomes, the reliability of the NDNQI PrU process measures have been established, however, this is something that has not been done. Specifically, different NDNQI

data collectors evaluating the same patient should come to the same (or similar) conclusions about the performance of PrU risk and prevention. This study addressed this gap.

Definition of Terms

Reliability

Conceptual definition. Reliability is "the degree to which observations or measures are consistent or stable" (Rosenthal & Rosnow, 2008, p. 757). Reliability of the observed score (True Variance /Observed Variance) is written as (Waltz, Strickland, & Lenz, 2010):

$$\frac{\text{True Variance}}{\text{Observed Variance}} = 1.0 - \frac{\text{Error Variance}}{\text{Observed Variance}}$$

Operational definition. Inter-rater reliability (IRR) is the degree to which two or more participant raters agree on their judgments of an outcome (Rosenthal & Rosnow, 2008). The IRR parameters for this study are Cohen's kappa, PAK, percent agreement, and intraclass correlation coefficients (*ICCs*).

Kappa. Kappa (k) is a statistical measure for inter-rater agreement and can be mathematically written as

$$k = (P_o - P_{e)} / (1 - P_e)$$

where P_o is the proportion of observed total agreement, and P_e is the proportion of agreement that is expected by chance (Feinstein & Cicchetti, 1990, p. 544). Kappa, therefore, is the proportion of agreement greater than what is expected by chance alone (Sim & Wright, 2005).

Prevalence-adjusted kappa. Kappa can be adjusted for high or low category prevalence by using average cell counts of agreement (cells on the diagonal) in place of actual cell counts of agreement, while calculating kappa. This is called prevalence-adjusted kappa (PAK) and is discussed in detail in Chapter II.

Percent agreement. Percent agreement is the proportion of *cases* that two raters agree in their ratings (Fleiss et al., 2003).

Intraclass correlation coefficient (ICC). The intraclass correlation coefficient (*ICC*) is the ratio of between-subjects variance to total variance. Theoretically, an *ICC* has a possible value of 0 to 1.0 with 1.0 representing perfect reliability (Portney & Watkins, 2009). However, an *ICC* will be negative when the average covariance within subjects is negative (Nichols, 1999). For the purpose of the study, *ICC* is defined as the ratio (Hart, Bergquist, Gajewski, & Dunton, 2006):

$$\frac{\text{(Variability of kappa's between hospitals)}}{\text{(Variability of kappa's between hospitals) + (Variability of kappa's within hospitals)}}$$

Therefore, an *ICC* near zero is desirable and suggests *within*-hospital variance is much greater than *between*-hospital variance (Hart et al., 2006).

Expert

Because an *expert* is considered to have a high degree of skill in PrU data collection, and a high level of knowledge about skin and wound care; expert is operationalized as the person identified by each site coordinator as the individual with the most experience and/or skill in chart abstraction on the PrU data collection team. The **non-expert rater** is defined as (a) not the expert, and (b) part of the PrU data collection team who usually reviews patient records for data on PrU risk and prevention.

Pressure Ulcer Risk and Prevention

Pressure ulcer **risk** is the degree to which a person is susceptible to develop a PrU. Risk status is determined by a risk assessment tool (i.e. Braden or Norton Scales) and by clinical judgment based on the presence of other factors (e.g. advanced age, low body weight, current or previous PrU, poor perfusion, and co-morbidities). Pressure ulcer risk is operationalized as the

performance of PrU risk assessment(s) and skin assessment(s), and determination of PrU risk status; as documented in the patient record. Pressure ulcer **prevention** is minimizing risk for PrU development, and is operationalized as performance of a documented skin assessment, pressure redistribution surface use, routine repositioning as prescribed, nutritional support, or moisture management; within the 24 hours prior to data collection for patients "at risk" for PrU development.

Participant Rater, Unit, and Hospital Characteristics

Participant rater characteristics are the characteristics of the chart abstractors and PrU team leaders. The NDNQI data collection guidelines state PrU team leaders should be: (a) certified in wound care; or (b) have received formal education in skin assessment, PrU risk assessment, PrU identification/staging, and PrU prevention (Bergquist-Beringer & Davidson, 2014). Team members may be wound care certified, but before data collection they must receive training in PrU risk, identification, staging, prevention, data collection, and be skilled at reading patient health records. Participant characteristics is operationalized as expert status, and self-reported information on job title, highest level of education, review of the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers, other PrU training or review, years collecting NDNQI PrU data, other roles in NDNQI PrU data collection (besides chart abstractor), and wound care certification status. The data collection team characteristic of interest is whether or not the PrU team is led by a certified wound nurse or wound/skin care nurse, as reported by participant raters at the time of participation in the study.

Unit characteristics are the characteristics of the nursing unit from which data are collected. This is operationalized as the unit identification (ID) code as recorded on the *Data Collection Form* by participant raters at the time of data collection. Unit ID identifies unit type.

Unit type included adult medical-surgical, adult medical, and adult surgical units. Unit ID was the unit identifier code that NDNQI uses to identify each unit within that hospital.

Hospital characteristics are characteristics of the hospital from which data are collected, and is operationalized as staffed bed size, hospital type (General Acute Care), Magnet status, teaching status, and metropolitan status; as found on record with the NDNQI. All variables are discussed in detail in Chapter III.

Assumptions

1. Data collectors use multiple sources to gather PrU data, such as the electronic health record, paper patient records, direct observation, and input from patient care staff.
2. Participant raters rated patients independently of each other.
3. During the study, participant raters rated patients in a manner similar to how they rate patients during routine NDNQI PrU data collection.
4. Participant raters provided thoughtful and accurate responses.
5. "Expert" participant raters were accurately identified.
6. "Expert" participant raters accurately rated patients on the PrU risk and prevention items.
7. Documentation in the patient record accurately reflected performance of process measures.
8. *Intra*-rater reliability was assumed.

CHAPTER II: REVIEW OF THE LITERATURE

Pressure ulcers are considered avoidable and due to poor quality of care (Black et al., 2011). The first step in decreasing pressure ulcers is PrU prevention. Clinical Practice Guidelines have been published that translate what is known about interventions to prevent PrUs for healthcare use. The NDNQI® collects data on PrU risk and prevention for quality improvement purposes to decrease PrU rates. The reliability of this data, however, has not been reported. This chapter will review (a) national quality indicators on PrUs, (b) recommended PrU prevention interventions, (c) PrU prevention programs, (d) criteria for a sound measure, and (e) reliability studies of the NDNQI PrU measures.

National Quality Indicators on Pressure Ulcers

Quality is a complex and abstract term. In 2007—while Director of the Agency for Healthcare Research and Quality (AHRQ)—Carolyn Clancy, MD defined quality in healthcare as "the right care, for the right person, at the right time; the first time" (Clancy, 2008, slide #3). The Institute of Medicine (IOM) defined quality care as care that is safe, timely, effective, efficient, equitable, and patient-centered (IOM, 2001).

Established in 1970, the IOM is an independent, non-profit organization that provides recommendations to decision makers regarding the nation's healthcare system (IOM, 2013). In 1999, the IOM released its report "To Err is Human: Building a Safer Health System" (National Research Council, 2000). This report identified PrUs as a commonly occurring adverse event, and recommended strategies to reduce them. These recommendations included voluntary and mandatory reporting of adverse events, including PrUs. As a result, many state laws were passed in 2003 to 2005 mandating hospitals report PrUs and other adverse events. Mandatory reporting

within Minnesota found Stage III, Stage IV, and unstageable PrUs, were the most commonly occurring serious adverse health event in 2013 (Minnesota Department of Health, 2014).

Healthcare quality varies across settings and individuals, and this variability is not clearly associated with dollars spent (Clancy, 2008). As a result, Congress enacted the Medicare Modernization Act of 2003, which mandated that the IOM examine healthcare quality (Library of Congress, 2003). With the goal of improving quality and safety by reducing the number of adverse events, The Patient Safety and Quality Improvement Act of 2005 (Patient Safety Act) established a system for voluntary reporting of adverse events (including Stage III, Stage IV, and unstageable PrUs) and medical errors (Department of Health and Human Services, n. d.). To encourage participation, the confidentiality of this patient safety information is federally protected. Concomitant efforts to improve quality have been put forth by governmental agencies such as the Agency for Healthcare Research and Quality (AHRQ) and the Centers for Medicare and Medicaid Services (CMS); and private organizations such as the National Quality Forum (NQF), and the Joint Commission (IOM, 2005). These efforts include strategies to reduce PrU occurrence.

The Agency for Healthcare Research and Quality (AHRQ)

The AHRQ is a federal agency within the Department of Health and Human Services (AHRQ, 2012a). Established in 1989 in response to the Omnibus Budget Reconciliation Act of 1987, the AHRQ focuses on research to improve the quality, safety, efficiency, and effectiveness of healthcare. To achieve its goal, the AHRQ has developed indicators to measure hospital healthcare quality from available inpatient data, including PrUs.

Patient Safety Indicator #3 and Pediatric Quality Indicator #2. The AHRQ quality indicators on PrUs include Patient Safety Indicator (PSI) #3 and Pediatric Quality Indicator

(PDI) #2. These quality measures are derived from claims data; specifically the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes, diagnostic related groups (DRGs), and patient age and/or gender if included in the claims data (AHRQ, 2012b). The PSI #3 is defined as the number of medical/surgical discharges with a secondary diagnosis (not present on admission) of a Stage III, Stage IV, or unstageable PrU per 1,000 discharges, for those adults (18 years of age and older) hospitalized at least 5 days (AHRQ, 2013b & 2013c; Zhan & Miller, 2003). The PDI #2 is almost identical to the PSI #3, except the PDI #2 is for those 17 years of age or younger, excluding neonates (AHRQ, 2012b). Because PSI #3 and PDI #2 count only Stage III, IV, and unstageable PrUs, they underestimate the rate of all hospital-acquired PrUs by as much as 70% (Bergquist-Beringer et al., 2013; VanGilder et al., 2013). Further underestimation is likely because PSI #3 and PDI #2 do not count PrUs among those who: (a) are hospitalized less than 5 days; (b) have paralysis; and (c) are transfers from skilled nursing facilities, intermediate care, or other hospitals.

The reliability and validity of PSI #3 and PDI #2 were established during their development (Fabian & Geppert, 2011). The Agency for Healthcare Research and Quality (AHRQ) indicator development takes approximately 20 months, and implementation 1½ months. Development begins with an extensive literature review. An expert panel is involved from the early stages of development, and continues until the measure is retired. Technical specifications, importance, scientific acceptability, and usability were established. The feasibility of the PSI #3 and PDI #2 (as well as the full set of indicators) was pilot tested. Ongoing reliability and validity is empirically examined, though only measures that have support for their validity and reliability receive National Quality Forum endorsement (NQF, 2013d). The PDI #2 is a NQF-endorsed measure; the PSI #3 is not.

The reliability of the PSI #3 was recently supported by Schone, Hubbard, and Jones (2011). The purpose of their analysis was to provide hospitals with information on the reliability and minimum case thresholds for the CMS's outcome measures proposed for use in the 2014 Hospital Value-Based Purchasing Program. Discharge/claims data that were submitted between March 2010 through September 2010 to measure value-based purchasing performance were used to estimate the reliability of the PSI #3 measure at 6, 12, 18 and 24 months. A minimum median acceptable reliability level of .4 was established by Schone and colleagues. The study determined that the median reliability at 6, 12, 18, and 24 months was .82, .90, .93, and .95 respectively. For 6 months of data, the minimum number of cases at which reliability reached .4 was a reasonable $N = 44$ PrUs.

The PSI #3's *validity*, however, was questioned when the validity of administrative/claims data—such as diagnostic codes used by the AHRQ and CMS to calculate HAPU rates—was questioned. In 2009, from a sample of 196 hospitals in California, Meddings et al. (2013) conducted a retrospective analysis of nearly 2 million all-payer administrative (discharge) records and 96,355 patients' surveillance data. Hospital-acquired Stage II or higher PrU rates were determined from administrative (discharge/HAPU *incidence*) data or surveillance (patients examined/*point prevalence*) data. Meddings and colleagues expected administrative/claims data to return lower HAPU2+ rates than surveillance data because "point-prevalence assessments collect more data from patients with longer lengths of stay" (p. 508), which is a consequence of PrU. One could argue, however, that the opposite is true because point-prevalence data may miss patients who develop a HAPU after direct examination of the patient. When *administrative* data were used to calculate HAPU2+ rate, the mean rate was 0.15%, 95% CI [0.13% - 0.17%], compared to a 2.0% rate, 95% CI [1.8% - 2.2%] calculated

from *surveillance* data. The significance of this difference is even more alarming because three of the 49 hospitals found to be in the highest (worst) quartile of HAPU2+ rates (as determined by administrative data); were ranked as "superior" when HAPU2+ rates were calculated from surveillance data. Also, 14 other hospitals labeled "worst hospitals" by administrative data, were considered "above average" when surveillance data were used. Certainly, this questions if administrative data alone should be used to compare hospitals' HAPU2+ rates.

The reliability of PDI #2 has also been questioned. Although Scanlon, Harris II, Levy, and Sedman (2008) found PDI #2 had the highest positive predictive value (51.4%) among all PDIs; a later study by Bardach, Chien, and Dudley (2010) reported that only 4 out of 353 (1.1%) hospitals in California had the minimum number of 5,956 PDI #2 cases (over 3 years) to allow for identification of hospitals with pediatric HAPU rates twice the state's average. This two-fold "poorer performance" was determined a priori in order to identify "what could be considered a moderate deficiency" in providing quality patient care (p. 267).

Patient Safety Organization – Common Format for pressure ulcers. Besides the PSI #3 and the PDI #2, the AHRQ also collects data on PrUs through the Patient Safety Organization (PSO) (AHRQ, n.d.a). The PSO was sanctioned by the Patient Safety Act of 2005, with the purpose of allowing healthcare providers to report quality and patient safety data without fear of legal discovery. The PSO is comprised of public, private, for-profit, and not-for-profit organizations; to which hospitals voluntarily submit data on patient safety events. These organizations are regulated and managed by the AHRQ, and allow for confidential collection, aggregation, and analysis of patient event data. Although hospitals voluntarily submit data on *all* PrU stages/categories, the PSO reports only Stage III, Stage IV, and unstageable PrUs. These same stages are included in the PSI #3 and PDI #2; but the PSI and PDI are from

administrative/claims data, their submission is not voluntary, and they are used for reimbursement purposes.

For AHRQ PSO data on PrUs, information is collected using the Pressure Ulcer Common Format (AHRQ, n. d.a). The PrU *process* and *structure* measures included in the Pressure Ulcer Common Format are: (a) performance of a PrU risk assessment (when performed, what scale was used, and risk status); (b) PrU prevention (use of a pressure-redistribution surface, repositioning, hydration/nutritional support, and skin care practices); (c) risk factors (devices or appliances); and (d) contributing factors (not specified) (AHRQ, 2013a). The first two align with the NDNQI pressure ulcer risk and prevention process measures. Unlike the PSI #3, the PDI #2, and the PSO; the NDNQI counts and reports *all* PrU categories/stages, and includes patients hospitalized less than five days, patients with paralysis, and patients transferred from other facilities.

The Centers for Medicare and Medicaid Services (CMS)

Medicare and Medicaid were enacted into law in 1965 (CMS, 2013b). The Deficit Reduction Act of 2005 required the Secretary of the Department of Health and Human Services to identify hospital-acquired conditions (HACs) that (1) are high-cost or high-volume, (2) result in a diagnosis-related group (DRG) with a larger payment as a secondary diagnosis than as a primary diagnosis, and (3) are preventable. Currently there are eight HACs—one is Stage III or Stage IV HAPU. Effective October 2008, CMS no longer reimburses providers for treatment of hospital-acquired Stage III or Stage IV PrUs (QualityNet, n. d.). Private insurers followed suit (O'Reilly, 2008). Blue Cross and Blue Shield of Connecticut stopped reimbursing providers for treatment of any current or future CMS-identified HAC in March of 2012 (Blue Cross & Blue Shield, 2012).

Hospital-acquired conditions data come from discharge (administrative/claims) data, data with questionable accuracy (Coomer & McCall, 2012; Meddings et al., 2013; Schone et al., 2011). Coomer and McCall (2012) examined Medicare claims data (for the years 2009 and 2010) to determine the accuracy of PrU coding. These codes are used by CMS to determine the HAC of hospital-acquired Stage III or Stage IV PrU rates. Pressure ulcers are coded by a *site* code and a *stage* code. Because CMS does not reimburse for treatment of Stage III or Stage IV PrUs, hospitals have reason to miscode PrU stage. For instance; a Stage III HAPU may be miscoded as a Stage II, and a Stage II present-on-admission PrU may be miscoded as a Stage III. Another coding concern is that only the first eight secondary-diagnoses are recorded in the Medicare data system, but hospitals are not required to list a PrU *stage* code among the first eight. In other words, while a PrU *site* code will likely appear among the first eight secondary diagnoses, the PrU *stage* code is often omitted. Coomer and McCall (2012) found that in 2009 only 54% of claims with a secondary diagnosis for PrU *site* had a secondary diagnosis for PrU *stage*. This increased to 61% in 2010. When only the first eight secondary-diagnoses were used, Stage III or IV HAPU rates were 30% to 62% lower than rates calculated using *all* secondary diagnoses (i.e. beyond the first eight). The larger the hospital, the less likely a PrU stage code accompanied a PrU site code. Because of these issues, the CMS has now revised their data system to include 25 diagnosis codes and 25 procedure codes per claim (Coomer & McCall, 2012).

Others have confirmed these doubts about the reliability of the CMS HAC measure. Schone and colleagues (2011) examined the reliability estimates and minimum case thresholds for the 2014 Hospital Value-Based Purchasing Program measures. The median reliability estimates for 6, 12, 18, and 24 months of PrU data were R = .28, .40, .50, and .53 respectively.

At least 12 months of data, therefore, are required for CMS's PrU data to be reliable at the minimum level of $R = .4$. In addition, the minimum number of cases at which R reached .4 was 2,195 Stage III or IV PrUs. This is much larger than the 44 PrUs needed for AHRQ's PSI #3 to reach $R = .4$. These investigators concluded the measure is unreliable in the short term and with smaller Stage III and Stage IV PrU rates. Because of these issues, starting in 2015, CMS will use the PSI Composite #90 (which includes PSI#3) as its HAC measure (CMS, 2013a).

From 2001 to 2007, CMS led the development and implementation of the Medicare Patient Safety Monitoring System (MPSMS) (Qualidigm, 2013a & 2013b). This was done in collaboration with the AHRQ, the Centers for Disease Control and Prevention, the Food and Drug Administration, the Office of the National Coordinator for Health Information Technology, and the Veterans Health Administration. Collectively these organizations are known as the Federal Agency Work Group (Qualidigm, 2013a). In 2009, the AHRQ took over the primary responsibility for the MPSMS in partnership with CMS. Using 21 measures, this national surveillance program has the goal of measuring adverse events among hospitalized patients, and includes all stages/categories of HAPUs (Hunt et al., 2005; Lyder et al., 2012). The MPSMS data are collected from a subset of inpatient records *and* administrative/claims data (versus PSI #3, PDI #2, and CMS's HAPU rates, which are based solely on claims data). Each year a random sample of more than 40,000 hospital records are sent to one of two Clinical Data Abstraction Centers. Until 2009, only Medicare recipients' were included in the MPSMS. Now this surveillance program also includes hospitalized patients 18 years of age or older; who were admitted for heart failure, acute myocardial infarction, pneumonia, or a subset of major surgeries. Patients included in MPSMS account for 26% of the nation's annual hospitalizations (Qualidigm, 2013a).

The reliability of the MPSMS is evaluated monthly: 40 randomly selected abstracted charts are exchanged between the two abstraction centers (Hunt et al., 2005). Percent agreement between the two abstraction centers was 96.66%. Accuracy of the raters was reported as 98.03%; which is the agreement between the aggregated agreement of the two abstractors and the gold standard of the expert abstractor (Hunt et al., 2005). Reporting only percent agreement, however, limits the usefulness of this reliability estimate because agreement due to chance alone is not accounted for (Fleiss et al., 2003).

The Joint Commission

The Joint Commission is an independent, non-profit organization that was established in 1951. Its mission is "to continuously improve healthcare for the public, in collaboration with other stakeholders, by evaluating healthcare organizations and inspiring them to excel in providing safe and effective care of the highest quality and value" (Joint Commission, 2013a, "Our Mission"). The Joint Commission accredits more than 20,000 healthcare organizations (Joint Commission, 2013b).

Each year, the Joint Commission publishes National Patient Safety Goals to promote improvements in patient safety (Joint Commission, 2013c). *Goal #14: Prevent health care-associated pressure ulcers (decubitus ulcers)*, applies only to long-term care settings (Joint Commission, 2013c). Meeting *Goal #14* includes (1) creating a written plan for identifying PrU risk and prevention, (2) performing a PrU risk assessment upon admission, (3) performing a systematic PrU risk assessment using a validated tool, (4) reassessing risk at intervals determined by the facility, (5) implementing PrU prevention measures for identified risks, and (6) educating staff on PrU risk and prevention (Joint Commission, 2013c). These align with the

NPUAP/EPUAP/PPPIA guidelines (NPUAP, EPUAP, & PPPIA, 2014) and—with the exception of staff education—are similar to the NDNQI pressure ulcer risk and prevention measures.

The Robert Wood Johnson Foundation provided funding to The Joint Commission for the testing of NQF-endorsed measures, including the PrU measure; *NQF 0201*. The *NQF 0201* is HAPU2+ prevalence among acute care patients. This measure is the same as the HAPU2+ measure from the NDNQI. The Joint Commission evaluated the reliability of PrU data by conducting on-site reliability testing in 19 of 20 randomly selected sites from April through August 2008 (Joint Commission, 2010). Based in part on the results of these reliability tests, the HAPU2+ measure was subsequently endorsed by the NQF. However, the Joint Commission wants to retire their stewardship of this measure because they lack the ability to maintain it (Munthali & Morsell, 2013). This makes sense because the commission does not collect data on *NQF 0201*. Others will need to take over stewardship of this measure if it is to continue (S. Bergquist-Beringer, personal communication, December 18, 2013).

The National Quality Forum (NQF)

The NQF is a nonprofit, consensus standards-setting organization established in 1999 with the mission of improving the nation's healthcare (NQF, 2003& 2013a). It *endorses* standardized performance measures through a rigorous process that relies on stakeholder input and empirical support: It does not *develop* performance measures (NQF, 2013c). Since 2004, the NQF has endorsed hundreds of quality of care indicators aligned with organizational (Hospital Quality Alliance, CMS, Joint Commission) and legislative (Deficit Reduction Act) measurement requirements (A Crosswalk, n. d.; NQF, 2013c).

The *NQF-endorsed measures* specific to PrUs, the **target population**, and the stewards for each include (NQF, 2013b):

Outcome measures. (1) *NQF 0181 Increase in number of PrUs (Stages I through IV) among **home health** patients*, uses Outcome and Assessment Information Set data (OASIS-C form), CMS is the steward; (2) *NQF 0201 HAPU2+ point-prevalence among **all inpatient settings***, Joint Commission is the steward; (3) *NQF 0337 HAPU2+ rate among **hospital/acute care patients < 18 years of age**, excluding neonates* [AHRQ's PDI #2], the steward is the AHRQ; (4) *NQF 0678 Percentage of residents or patients with a Stage II through IV PrU that are new or worsened among **short-stay (≤ 100 days) nursing home/rehab/long-term care patients***, the steward is CMS; and (5) *NQF 0679 Percentage of high risk residents with a Stage II through IV PrU among **long-stay (> 100 days) hospital/acute care patients***, CMS is the steward.

Process measure. (1) *NQF 0538 PrU prevention and care among **home health patients***—which includes performance of PrU risk assessment, PrU prevention included in the plan of care, and PrU prevention implemented during short stays; CMS is the steward. It is noteworthy that there is only one endorsed process measure, and it does not apply to hospitalized patients or long-term care residents. This underscores the need for reliable PrU risk and prevention measures for use among acute-care patients. Addressing this need is the purpose of this study.

The National Quality Forum (NQF) criteria for evaluating the quality of measures for endorsement were first published in September of 2003 (Joint Commission, 2014). Since then, numerous revisions have taken place. Their latest version was published in October, 2013 (NQF, 2013e). Measure endorsement is a rigorous process (see Appendix A) and NQF-endorsed measures are held as the gold standard of performance measures (NQF, 2013d). Candidate measures submitted for endorsement are examined by a broad and varied group of experts who use a consensus-based process of providing input that evaluates: (1) the importance of what is

being measured (clinical evidence, performance gap, and priority); (2) if the proposed measure meets scientific standards (reliability and validity); (3) feasibility; (4) usability (accountability/transparency, improvement); and (5) related or competing measures (NQF, 2013d & 2013e). While candidate measures will meet each of these criterions to varying degrees, the measure must meet the minimum acceptable level for the first two—which includes reliability, or it will not be evaluated on the other three criteria.

Importance pertains to if (a) the measure assesses a healthcare or patient *outcome* that is related to a process or structure, (b) there is opportunity for improvement in the problem, and (c) the measure addresses a high-priority or high-impact aspect of healthcare. If the measure evaluates a process or structure, versus an outcome (such as the NDNQI process measures that were evaluated in this study), it can still be endorsed if it is based on a systematic review and evidence that corresponds with what is being measured. It appears from these initial evaluation criteria that the NQF is more interested in outcome measures, because the endorsement process is more streamlined for outcome measures, than process or structure measures (see Algorithm 1 in Appendix A). Scientific acceptability indicates the measure's reliability and validity has been empirically and appropriately supported. Percent agreement does not fulfill the reliability requirement (see Algorithm 2 in Appendix A). Feasibility represents that the data are readily available, and data collection does not impose undue burden. Usability reflects that those who want the data have access to them, and the degree to which using the measure leads to improved quality of healthcare. Finally, a measure should not compete with other endorsed or related measures, or it must be clearly superior to such measures in order to justify its use. EMeasures are evaluated using the same criteria as non-emeasures (NQF, 2013e).

Recommended Pressure Ulcer Prevention Interventions: Literature Review

The etiology of PrUs is complex and poorly understood (Ratliff et al., 2010). Theories of PrU development include (a) tissue ischemia due to capillary occlusion leading to tissue anoxia and death, (b) tissue injury due to reperfusion of ischemic tissue, (c) lymphatic dysfunction resulting in an accumulation of metabolic wastes, and (d) mechanical bending/twisting of tissues (Kottner, Balzer, Dassen, & Heinze, 2009; Peirce, Skalak, & Rodeheaver, 2000). These theories are the basis for recommended interventions that minimize pressure and shear. In order to implement appropriate PrU prevention interventions, first an individual's risk for PrU development must be identified.

Pressure Ulcer Risk Assessment

Tools to assess PrU risk have been available and used for decades. There are more than 40 PrU risk assessment scales (Kottner & Balzer, 2010). The most commonly used, however, are the Norton Scale, the Waterlow Scale, and the Braden Scale (Anthony, Papanikolaou, Parboteeah, & Saleh, 2010). The Norton Scale has been in use the longest, 48 years. The Waterlow Scale is ever-present in the United Kingdom (Anthony et al., 2010), while the Braden Scale is the most commonly used tool in the U.S. (Armstrong et al., 2008). All have subscale scores that are totaled to measure overall risk, and are based on factors related to PrU development. Because more than 100 factors have been associated with PrU risk (Lyder & Ayello, 2008), it is not feasible for risk assessment tools to include them all; but tools should include factors known to have the greatest impact on risk. This highlights the importance of using clinical judgment in conjunction with validated assessment tools while determining risk.

Many studies have validated the use of the Braden Scale to identify patients at PrU risk (Bergstrom & Braden, 2002; Bergstrom, Braden, Kemp, Champagne, & Ruby, 1998; Bergstrom, Braden, Laguzza, & Holman, 1987; Kring, 2007). There is debate, however, as to what

sensitivity and specificity are needed, and the problem of decreasing risk once identified. Also, a risk assessment scale may have high sensitivity and specificity; but if effective PrU prevention is provided (and a PrU is avoided), sensitivity and specificity are decreased. One could reason, therefore, that sensitivity and specificity appropriate for a screening tool are acceptable for these scales. While it is difficult to have high sensitivity *and* high specificity, cut-off scores should be where the balance between the two is the most desirable. A cut-off score is the score at which a person is classified "at risk", and alerts care providers to implement appropriate PrU prevention interventions. Commonly accepted cut-off scores for each scale are Norton (≤ 15 – 16), Modified Norton (≤ 20 – 21), Braden (≤ 18) (Källman & Lindgren, 2014), and Waterlow (≥ 17) (Serpa, Santos, Peres, Cavicchioli, & Hermida, 2011).

Using a cross-sectional descriptive design, Källman and Lindgren (2014) examined the predictive value of four PrU risk assessment scales (Norton, Modified Norton, Braden, and their own Risk Assessment Pressure Sore). Using the risk assessment scales and the Swedish version of the EPUAP minimum data set, data were collected during one day in November 2009 on 346 adult inpatients (62% were > 70 years of age) at a hospital in Sweden. Sensitivity, specificity, positive predictive values (PPV), and negative predictive values (NPV) were calculated for five different cut-off scores per scale. The investigators found strong support for the use of the Risk Assessment Pressure Sore Scale and the Braden Scale. For the Braden Scale, the best balance between sensitivity and specificity was at a cut-off score of ≤ 18 (sensitivity = 74.5%; specificity = 73.7%; PPV = 35%). Others have confirmed similar sensitivity and specificity values for the Braden Scale (Bergstrom et al., 1998; Bergstrom et al., 1987; Bergquist & Frantz, 2001; Braden & Bergstrom, 1994; Defloor & Grypdonck, 2005; Tannen, Balzer, Kottner, Dassen, Halfens, & Mertens, 2010). These values demonstrate the usefulness of the Braden Scale in screening for

PrU risk; and despite limitations, risk assessment tools are considered accurate predictors of PrU risk (AHRQ, 2011b; Anthony et al., 2010; Ayello & Braden, 2002; Chou et al., 2013).

Few studies have linked performance of a PrU risk assessment with decreased PrU development (Anthony et al., 2010; Ratliff et al., 2010). In 2010, Moore and Cowman (2010) updated a 2008 Cochran Review with the purpose of determining if use of a PrU risk assessment tool decreased PrU incidence. The original (2008) review was unable to find a single study that met the inclusion criteria of (1) an RCT comparing the use of a structured PrU risk assessment tool to no structured PrU risk assessment, or to unaided judgment; or (2) an RCT comparing various structured PrU risk assessment tools to each other. The updated (2010) review revealed that only one randomized study (Saleh, Anthony, & Parboteeah, 2009) met these criteria. In this study, nurses ($N = 256$) from a military hospital in Saudi Arabia were randomized (by clusters of three nursing units per cluster) into one of three groups that received: (1) a mandatory wound care management study day, PrU prevention training with specific instruction on the Braden Scale, and then staff were required to implement the Braden Scale ($n = 74$); (2) identical to the first group, but staff were not required to implement the Braden Scale ($n = 76$); or (3) only the mandatory wound care study day ($n = 106$). Interestingly there was a reduction in PrU occurrence within all groups. No statistically significant difference in PrU incidence between the groups, however, was found. The reviewers (Moore & Cowman, 2010), however, reported Saleh et al.'s (2009) study was underpowered. Anthony et al. (2010) concluded that these (Saleh et al., 2009) results suggest: (a) education may be the relevant factor, not the Braden Scale; and (b) being made aware of a problem improves it. It is likely the routine performance of assessing PrU risk, not the scale used, is the relevant factor.

Nonetheless, risk assessment tools are widely used in hospitals to identify those at risk and determine appropriate prevention interventions (AHRQ, 2011b; Anthony et al., 2010; Ayello & Braden, 2002; Chou et al., 2013). Anthony and colleagues (2010) reviewed research conducted on PrU risk scales—much of which was conducted by them over the past 25 years in an attempt to improve these scales. They concluded that while risk assessment scales predict risk, there is no evidence that they reduce PrU development; and that this information is rarely sought in studies. Use of a risk assessment tool, however, has been linked with improved performance of PrU prevention interventions. Specifically, nurses performed prevention interventions more frequently and appropriately after training and implementation of the Norton Scale (Hodge, Mounter, Gardner, & Rowley, 1990). Clinical guidelines advocate the use of risk assessment tools because they formalize and make routine the assessment of PrU risk (NPUAP/EPUAP, 2009; Ratliff et al., 2010). No one tool, however, has been clearly identified as the most accurate PrU risk predictor (Chou et al., 2013; Ratliff et al., 2010; NPUAP/EPUAP, 2009), and prevention guidelines state clinical judgment should be part of a PrU risk assessment (NPUAP/EPUAP/PPPIA, 2014; Ratliff et al., 2010).

The exception to the findings of other researchers is a single study analyzing NDNQI data for the year 2010 (Bergquist-Beringer et al., 2013). This study is presented in detail later in this chapter. Specifically, after statistically controlling for hospital characteristics and nurse staffing, decreased odds of HAPU development were reported for those whose interventions included: (a) a PrU risk assessment on admission, OR = 0.82, 95% CI [0.73 – 0.92], $p < .001$; and (b) a PrU risk assessment performed within the 24-hour period before the data were collected, OR = 0.87, 95% CI [0.81 – 0.92], $p < .001$. Although the purpose of this study was to determine the process

measures that were associated with HAPU rates, results suggest that performance of risk assessment on admission sensitized staff to those who were at risk.

Skin Assessment

Early identification of erythema is important to prevent PrUs, and skin status has been identified as an independent risk factor for Stage II and higher PrU development (Nixon, Cranny, & Bond, 2007). Among persons with a spinal cord injury; the association between routine, systematic, skin inspections—which involves looking *and* touching—and PrU occurrence is unclear. Garber, Rintala, Hart, and Fuhrer (2000) performed a longitudinal cohort study of 118 men with spinal cord injury who were living at home. Researchers found that participants who self-reported having a Stage II or higher PrU within the year prior to data collection ($n = 37$, 31%) were *more* likely to perform daily skin checks than those who had not developed a Stage II or higher PrU (92% versus 74%), $\chi^2 = 4.98$, df = 1, $p = .03$. This is surprising because commonsense says frequent visual skin assessment would be especially important to prevent PrU development among those with loss of sensation. These findings do not suggest regular performance of a skin assessment is associated with higher PrU rates, but rather that performing a skin assessment may not affect PrU rate among those with a spinal cord injury. It is important to note this study did not directly measure completion of skin assessments, but rather participants retrospectively self-reported if they routinely performed daily skin inspections. Current PrU prevention recommendations for those with a spinal cord injury include at least twice daily total body skin inspections (Schubart, & Hilgart, n. d.).

Bergquist-Beringer et al. (2013), however, found different results when they analyzed NDNQI pressure ulcer data (collected in 2010) from more than 1,400 hospitals and 710,626 patients. Hierarchical logistic regression revealed a decreased risk of HAPU (all stages) for

those who: (a) received a skin assessment on admission compared to those who did not, OR = 0.76, 95% CI [0.67 – 0.87], $p < .001$; and (b) received a daily skin assessment compared to those who did not, OR = 0.82, 95% CI [0.72 – 0.94], $p = .003$. Maybe nurses who are more likely to perform (and document) routine skin assessments are more likely to engage in activities to prevent PrUs.

The AHRQ (AHRQ, 2011b), the WOCN Society guidelines (Ratliff et al., 2010), and the NPUAP/EPUAP/PPPIA guidelines (NPUAP/EPUAP/PPPIA, 2014), recommend daily skin assessments. Skin assessments may lead to earlier detection of PrUs than without these assessments. More evidence, however, is needed to support that skin assessments prevent PrU development.

Minimizing Shear

This intervention used to be minimizing *friction* and shear. The 2009 NPUAP/EPUAP pressure ulcer prevention and treatment guidelines, however, removed reference to friction as a factor of PrU development (NPUAP/EPUAP, 2009). Friction is a concern because it may lead to harmful shear and strain on tissues. To include shear *and* friction in the definition (of PrU) would be unnecessary as well as confusing because friction may cause blisters, which are not PrUs. Shear, therefore, remains the primary causative factor and friction has been eliminated from the definition (Antokal et al., 2012). Shear is "the mechanical force that is parallel rather than perpendicular to an area" (Ratliff et al., 2010, p. 48).

A randomized control trial was performed in Japan by Nakagami et al. (2007) to examine the relationship of a specialized shear-reducing dressing on HAPU rates. The sample consisted of 37 bed bound, elderly ($M = 86.4$ years of age) hospitalized patients at risk for PrU development. The dressing was randomly placed on either the right or left greater trochanter for

three weeks, while the other trochanter was left without a dressing. No HAPUs developed. Persistent erythema (defined as a *blanchable* reddened area that persists for at least 20 minutes), however, occurred less often on the trochanters with the dressing than the trochanters without the dressing, RR = 0.18, 95% CI [0.05 – 0.73], $p = 0.007$.

Weng (2008) conducted a quasi-experimental study to compare the effectiveness of two dressings in preventing facial PrUs caused by (the pressure and shear exerted by) oxygen masks. These snug-fitting masks are commonly used in the ICU to maintain positive airway pressures. Ninety ICU patients ($M = 76.4$ years of age) in Taiwan received either no dressing ($n = 30$), Tegaderm ($n = 30$), or Tegasorb ($n = 30$). Skin condition was assessed every 30 minutes—it was not stated how long these assessments continued and it was assumed assessments continued until a PrU developed or the mask was discontinued. Frequent skin assessments were possible because the dressings were transparent. Those who had a dressing (either one) had: (a) fewer face mask PrUs than the controls [no dressing (29/30, 96.7%), Tegaderm (16/30, 53.3%), and Tegasorb (12/30, 40%); $p < .01$]; and (b) a longer duration until PrU occurrence [no dressing (1,111 minutes), Tegaderm (2,628 minutes), and Tegasorb (3,272 minutes); $p < .01$]. Interestingly, Weng did not find a relationship between PrU development and sensory perception, moisture, nutrition provided, nutritional condition, and serum albumin levels. Perhaps this is because only mask ulcers developed and were the focus of this study.

Multivariate logistic regression was performed by Schindler et al. (2011) to identify prevention interventions associated with decreased HAPUs among 5,346 pediatric ICU patients; 545 (10.2%) of whom developed a HAPU. Data were collected via retrospective chart review on all pediatric ICU admissions during the 6-month study period. The HAPUs were: Stage I (63.34%), Stage II (32.07%), Stage III (3.68%), and Stage IV (0.9%). After controlling for age

and risk of mortality, lower odds of PrU development was associated with interventions that reduced shear such as: (a) keeping the head of the bed at or below 30º compared to more than 30º, OR = 0.150, 95% CI [0.117 – 0.193], $p < .001$; (b) using a draw sheet compared to no draw sheet, OR = 0.575, 95% CI [0.403 – 0.820], $p = .002$; and (c) use of a breathable waterproof transparent dressing—not stated but assumed it was placed on the sacral area—compared to no dressing, OR = 0.713, 95% CI [0.516 – 0.985], $p = .04$. Unfortunately, the generalization of findings is limited to at-risk pediatric ICU patients.

Different results were reported by Brindle and Wegelin's (2012) RCT of 85 cardiac ICU patients from a single CCU ($M = 61.8$ years of age). The purpose of this study was to find out if use of a foam dressing would decrease sacral PrU incidence. Participants received either standard PrU preventive care ($n = 35$), or standard preventive care with the application of a self-adherent silicone border foam dressing (Mepilex®) over the sacrum ($n = 50$). All participants were considered to be at "high" risk for PrU development. Twenty-one covariates (such as HAPU rate, hours in the operating room, serum albumin levels, incontinence, and use of vasopressors) were measured. No variable (including HAPU rate) differed significantly between the control and intervention groups. The number of participants (by group) who developed at least one PrU was 4/35 (11%) for the controls, and 1/50 (2%) for the intervention group, $p = .154$. Although researchers concluded a lack of statistical power, the difference between an 11% HAPU rate and a 2% HAPU rate *would be significant* to those who developed a HAPU, and those trying to prevent them.

More recently, support for Mepilex® dressing use to prevent PrUs was presented in a literature review conducted by PrU prevention experts from Australia, Portugal, the U.K, and the U.S. (Black et al., 2014). The purpose of this review was to present recommendations for wound

dressing use in PrU prevention. Twenty-nine studies met inclusion criteria in that they considered wound dressing use in PrU prevention, were written in English or Spanish, and were published from 1998 to 2011. Black and colleagues concluded that when used in conjunction with usual PrU prevention interventions, Mepilex® decreased PrU incidence; and should be applied (as early as upon arrival to the emergency department) to the sacrum, buttocks, and heels of those at risk for PrU development.

Although Black and colleagues identified a specific dressing, no one dressing has been identified as the best dressing to decrease shear and HAPU rates. Despite limitations, evidence supports minimizing shear in order to prevent PrU occurrence among all ages for those at risk for PrU development. Pressure ulcer prevention experts recommend further research on wound dressings and shear (Black et al., 2014; NPUAP, 2013).

Minimizing Pressure

Pressure's role in PrU development has been well-documented. Kosiak's 1959 article described the link between skin breakdown in dogs with the magnitude and time of exposure to pressure. As early as 1930, mean arteriolar capillary pressure was determined to be 32 mm Hg (Landis, 1930). Later, in 1976, Roaf (2006) reported this pressure to be the threshold below which tissues tolerate. More recent research has focused on specific strategies to decrease pressure; such as body positioning, frequency of repositioning, and pressure-redistribution surfaces.

Body position. Defloor (2000) assessed interface pressure of 10 body positions on a standard mattress and a polyethylene-urethane pressure-redistribution mattress. Interface pressure is the pressure on skin and underlying tissues due to the surface the patient is on. Sixty-two healthy adults volunteered for this study. Defloor reported pressure on the sacrum and heels

was lowest in a semi-Fowler's position—the head and the foot of the bed elevated 30º, while the lower legs are horizontal—relative to any supine position (the head of the bed flat; or raised 30º, 60º, or 90º) on both mattresses. In addition, a 30º (tilt) side lying position allowed for lower pressures than a 90º side lying position, which had the highest pressures of any position on either mattress. A small ($N = 46$) RCT by Young (2004), however, found no difference in HAPU rates between a 30º tilt (3/23 = 13%) and 90º side-lying position (2/23 = 8.7%) among hospitalized elderly at risk for PrU development. All PrUs ($N = 5$) were Stage I in the Young (2004) study; which was likely underpowered.

A cross-sectional study by Moody, Gonzales, and Cureton (2004) was performed to measure interface pressures of four patient positions: when the head of the bed was elevated to 45º, 60º, and 65º; and a 30º side lying with the head of the bed elevated 30º. In addition, two mattresses were used; a polyethylene-urethane pressure-redistribution mattress and a dynamic air flow mattress. Interface pressure was measured by the X-Sensor Pressure Mapping System. This system consists of a pad with pressure sensors throughout that is placed on the bed and displays pressure values along the body. The sample included 20 adults with quadriplegia. Pressures increased—body location not specified but assumed bony prominences such as the sacral area—as the degrees of elevation of the head of the bed increased. The 30º side lying position had the lowest pressures of all positions, on both mattresses. This study supported Defloor's (2000) findings that keeping the head of the bed as low as possible keeps pressure on tissues to a minimum.

Heels are a common site of PrU development. Fowler, Scott-Williams, and McGuire (2008) conducted a review of studies on interventions to prevent heel PrUs. Among the eight studies included, all were published in 2007. Key points were: (1) while evidence supports use

of pillows to offload heels, it is important pillows are used properly and actually keep pressure off heels; (2) heel offloading is the first step to prevent heel PrUs; (3) heel offloading decreased heel PrU incidence; and (4) heel padding devices are not the same as heel offloading devices, and therefore do not relieve heel pressure.

Lyman (2009) reported on the effectiveness of a heel offloading device to decrease heel PrUs among 550 long-term care residents at risk for PrU development. During the 6 months prior to the intervention, 39 facility-acquired heel PrUs occurred; with an incidence rate of 2.1% to 5% per month (3 to 13 heel PrUs/month). The first 3 months after the intervention, the incidence of facility-acquired heel PrUs was 0% to 3.2% (0 to 15 heel PrUs/month). No heel PrUs developed during the 4th, 5th, or 6th month post-intervention. No information on statistical significance was reported.

Evidence supports that offloading heels is associated with a decrease in heel PrU development. The problem is doing it correctly. If pillows are used, it is important to make sure they are firm enough and remain in place to properly offload heel pressure. Care providers need to understand the difference between heel padding and offloading devices.

Repositioning. The effect of patient repositioning on PrU rates was examined by Defloor, De Bacquer, and Grypdonck (2005). These investigators conducted an RCT that examined four different repositioning schedules among 838 long-term care patients, within 11 facilities in Belgium. The four schedules for repositioning were (a) every 2 hours on a standard mattress, (b) every 3 hours on a standard mattress, (c) every 4 hours on a pressure-redistribution (viscoelastic foam) mattress, and (d) every 6 hours on a pressure-redistribution (viscoelastic foam) mattress. Within these facilities, 32 units were randomly assigned to perform one of the four repositioning schedules. Each unit implemented the repositioning schedule for 4 weeks.

Then the randomization process was repeated and each unit implemented a different repositioning schedule for another 4 weeks. Defloor and colleagues (2005) reported repositioning patients every 4 hours combined with the use of a viscoelastic foam mattress, significantly reduced HAPU2+ incidence, OR = 0.12, 95% CI [0.03 – 0.48], $p = 0.003$, when compared to the other repositioning schedules. These data are difficult to interpret because the study design combined different repositioning schedules with different support surfaces.

An RCT was conducted by Vanderwee, Grypdonck, De Bacquer, and Defloor (2007) from a sample of 235 nursing home residents among 84 units within 16 Belgian facilities. All residents had viscoelastic foam (pressure-redistribution) mattresses. Randomization was by ward. The experimental group ($n = 122$) was alternated between a 2-hour lateral position, and a 4-hour supine position. The controls ($n = 113$) were repositioned every 4 hours. The study continued for 21 months. No statistically significant difference was found between the groups with regard to Stage II and higher PrU incidence (16.4% and 21.2% in the experimental and control groups respectively, $p = .40$), PrU severity ($p = .65$), and time to developing a PrU ($p = .29$). With these high HAPU rates, one might argue that neither repositioning schedule was often enough. In addition, many patients did not stay in the side-lying position, so results are difficult to interpret.

Moore, Cowman, and Conroy (2011) conducted a cluster-RCT (by facility) to examine the effect of nighttime repositioning on PrU incidence among long-term care residents ($N = 213$) in 12 facilities in Ireland. Study patients were repositioned from 8 p.m. to 8 a.m. in one of two ways: (1) using the 30º tilt (left, back, right, back) every 3 hours (experimental group, $n = 99$); or (2) using the 90º side lying position every 6 hours (control group, $n = 114$). The routine daytime care of repositioning every 2 to 3 hours was maintained for all participants. All study patients

were at risk for PrU, had no PrU at the start of the study, and were at least 65 years of age—53% were 81 to 90 years of age, and 13% were 91 to 100 years of age. The researchers reported those in the experimental group (every 3 hours, 30º tilt) had lower odds of developing a PrU than those in the control group (every 6 hours, 90º side lying); OR = 0.234, 95% CI [0.067 – 0.879], p = .034. All PrUs were Stage 1 (n = 7; 44%) or Stage II (n = 9; 56%) [total of 16 PrUs among 16 patients]. Pressure ulcers developed in 3/99 (3%) participants repositioned every 3 hours using the 30º tilt group, and in 13/114 (11%) participants repositioned every 6 hours using the 90º side lying position; p = .035 95% CI [0.031 – 0.038]. Because this study randomized by facility (cluster randomization), an intra-class correlation (*ICC* = 0.001) by cluster was presented. An *ICC* near zero is desirable and suggests in-cluster variance is greater than between-cluster variance. In other words, variance in PrU rates can be attributed to the repositioning schedule, not the effect of being in a particular facility. Unfortunately, findings do not reveal if the time between turning or the degree of tilt was responsible for improved HAPU rates.

A secondary data analysis of data from a prospective cohort study (Baumgarten et al., 2009) of 658 patients, from nine hospitals within Maryland and Pennsylvania, 65 years of age and older, and who underwent hip surgery, was performed by Rich et al. (2011a). The purpose of this study was to examine the association between repositioning frequency and HAPU2+ incidence. Secondary analysis included patients (n = 269) from the parent study who had their skin examined for PrUs at least once during the first five days of hospitalization, and were bed-bound according to their Braden subscale score. In the parent study, repositioning data were collected from the medical record, which counted how many times a patient was repositioned per day for the initial 5 days of hospitalization. Also during the parent study, PrU status was assessed upon admission and on alternating days for 21 days—follow-up continued after

discharge from the hospital. The analysis found no association between repositioning frequency and HAPU2+ rates. This study is difficult to interpret because PrU incidence was determined over 21 days, while repositioning was assessed only the first 5 days of hospitalization. In addition, the *number of times* someone is repositioned each day may not be associated with HAPU rates because *time between* repositioning is what matters. For instance, someone may be repositioned 10 times over 3 hours, then left in the same position the rest of the day.

Bergstrom et al. (2013) conducted a multi-site RCT to identify the best repositioning frequency for nursing home residents while on high-density foam (pressure-redistribution) mattresses. The sample ($N = 942$) included residents 65 years of age and older ($M = 85.1$ years), from among 20 U.S. and 7 Canadian nursing homes, who were at moderate ($n = 617$) or high ($n = 325$; Braden Scale score ≤ 13) risk for PrU. Participants were randomly assigned to a 2-hour ($n = 321$), 3-hour ($n = 326$), or 4-hour ($n = 295$) repositioning schedule for 3 weeks. This random assignment was stratified according to the patient's degree of risk for PrU (moderate or high risk). The coccyx/sacrum, trochanters, and heels were examined weekly by blinded assessors. Twenty-one PrUs (2 Stage I and 19 Stage II) on 19 participants developed. There was no statistically significant difference in PrU incidence rates between the groups: 2-hour, 8/321 (2.5%); 3-hour, 2/326 (0.6%); and 4-hour, 9/295 (3.1%); $p = .68$.

The evidence suggests that repositioning every 4 hours while on a pressure-redistribution surface may be enough to prevent PrUs, but every 2-hour repositioning is typically considered the standard of care for those who cannot reposition themselves (AHRQ, n. d.b; Black et al., 2011; Institute for Healthcare Improvement, 2011; NDNQI, 2013). Routine repositioning is a mainstay of PrU prevention (NPUAP, EPUAP, & PPPIA, 2014). The NPUAP acknowledges the need for further research on repositioning schedules (NPUAP, 2013).

Pressure-redistribution surface. Pressure-redistribution surfaces (such as specialty mattresses or overlays that are filled with air, water, foam, gel, or a combination of these; dynamic surfaces such as alternating-pressure mattresses, and alternating-pressure overlays; and chair cushions) have been well-studied and found to reduce interface pressure (Chou et al., 2013; Reddy, Gill, & Rochon, 2006). In fact, the large number of studies is "staggering and confusing" (Clancy, 2013, p. 58). Randomized controlled trials on pressure-redistribution surfaces, however, are rare.

Among the RCTs conducted, Nixon (2006) performed one to compare PrU development among those on an alternating-pressure *mattress*, to those on an alternating-pressure *overlay*. The sample (N = 1,972) included hospitalized persons at least 55 years of age (M = 75.2 years). Critical care admits and people considerably over or under weight were excluded. No difference was found in HAPU rates between the two pressure-redistribution surfaces (10.3% and 10.7% for the mattress and overlay respectively). Nixon (2006) found the only surfaces that consistently relieved sacral pressure were the low-air-loss and fluid-filled mattresses (not the overlays). Unfortunately, the population most at risk for PrU, the critically ill and the underweight, were excluded from this study.

Comfort's (2008) meta-analysis began with a literature review to find studies that cited hospitals implementing the Braden Scale and subsequent use of a pressure-redistribution surface for those identified "at risk" for PrU. Nine such hospitals/studies were found. These studies were published from 1996 to 2000. Each hospital measured HAPU prevalence before and after implementing the Braden Scale followed by best practices. Although which prevention interventions were initiated varied across hospitals, all hospitals included some type of pressure-redistribution surface in their PrU prevention protocol. All hospitals reported improvement in

HAPU prevalence for patients who received the intervention (n = 2,593) compared to those who did not (n = 2,589), but the degree of improvement varied, OR = 0.335, 95% CI [0.22 – 0.508]. The reviewers concluded that providing support surfaces *after* the appearance of a Stage I or II PrU was associated with failure of the surface to prevent PrUs advancing to a higher stage. Pressure-redistribution surfaces, therefore, should be provided to all at risk patients *upon admission*.

A Cochran Review (McInnes, Jammali-Blasi, Bell-Syer, Dumville, & Cullum, 2012) of 53 RCTs and quasi-randomized trials examined pressure-redistribution surfaces. Reviewers concluded that: (1) foam alternative (pressure-redistribution) mattresses were associated with lower PrU occurrence than standard hospital mattresses, RR = 0.40, 95% CI [0.21 – 0.74]; (2) evidence is not clear as to the value of alternating- and constant-low pressure devices; (3) overlay use in the operating room decreased PrU occurrence for high-risk surgical patients and those with surgeries lasting > 3 hours; and (4) sheepskin use decreased PrU occurrence. The reviewers recommended that those at risk for PrU development should be placed on *any* pressure-redistribution surface. In addition, alternating-pressure mattresses may be more cost effective than alternating-pressure overlays.

Using multivariate logistic regression, Schindler et al. (2011) identified prevention interventions associated with decreased HAPUs among 5,346 pediatric ICU patients; 545 (10.2%) of who developed a PrU during hospitalization. After controlling for age and risk of mortality, patients who were on a specialty bed had lower HAPU rates than patients not on a specialty bed, OR = 0.226, 95% CI [0.167 – 0.306], $p < .001$. The type of specialty bed was not defined. Also, PrU development was less likely for patients on a sheepskin, compared to patients not on a sheepskin, OR = 0.448, 95% CI [0.325 – 0.618], $p < .001$.

Different results were reported from Rich et al.'s (2011b) secondary analysis of data from a sample of 658 hospitalized persons who were at least 65 years of age (46.4% were ≥ 85 years) and had undergone hip fracture surgery at one of nine Baltimore area hospitals. The purpose of this study was to determine the association between use of a pressure-redistributing surface and HAPU2+ incidence. Upon admission and every other day for 21 days, patients were assessed for PrU presence, pressure-redistribution surface use, and bedbound status. Because some patients' hospital stay did not last 21 days, patient follow-up continued in the 105 post-acute care facilities to which these patients were discharged.

The rate of HAPU2+ was compared between patients who had: (1) no pressure-redistribution surface; (2) a powered pressure-redistribution surface (alternating-pressure mattresses, low-air-flow mattresses, or alternating-pressure overlays); and (3) a non-powered pressure-redistribution surface (high-density foam, static air, gel mattress, or overlay). Because participants' use of a pressure-redistribution surface was not necessarily consistent throughout the 21-day data collection period, HAPU2+ rates were presented per follow-up visit (versus per total patients). In other words, a Stage II or higher PrU was found among (1) 195 of the 4,632 visits (4.2% of visits) during which no pressure-redistribution surface was in use, (b) 28 of the 623 visits (4.5% of visits) during which a powered pressure-redistribution surface was in use, and (c) 54 of the 1,496 visits (3.6% of visits) during which a non-powered pressure-redistribution surface was in use. Analysis revealed no statistically significant difference in these rates (p = .52). This data is difficult to interpret because it was cross-sectional and perhaps patients were put on a pressure-redistribution surface *after* they developed a PrU.

A 2012 review of RCTs (N = 45) published from 2000 to 2010 was conducted by French investigators to determine which support surfaces should be used for PrU prevention and

treatment (Colin et al., 2012). These RCTs compared different support surfaces to each other and/or to the standard hospital mattress. The types of support surfaces examined varied widely across the studies. Reviewers concluded (1) structured foam surfaces were more effective in preventing PrUs than standard hospital mattresses, (2) alternating-pressure mattresses were more effective in preventing PrUs than viscoelastic mattresses, (3) low-air loss mattresses were more effective at preventing heel PrUs than mixed-pulsating air mattresses, (4) some sheepskins decreased sacral PrU occurrence in orthopedic patients, and (5) the use of overlays in operating rooms decreased PrU occurrence among surgical patients.

The AHRQ reviewed the evidence on the effectiveness of prevention interventions in decreasing PrU incidence or PrU severity (Chou et al., 2013). Studies from 1946 to 2012 were searched for RCTs or cohort studies that examined the effect of prevention interventions on clinical outcomes ($N = 120$). This review confirmed that use of a more advanced mattress or overlay is associated with fewer PrUs compared to standard mattress use; and that there is no clear evidence as to which pressure-redistribution surface is best to decrease PrU rates. Other findings from this AHRQ efficacy study are presented within each intervention's review that follows. Similar conclusions regarding support surfaces were made by the NPUAP/EPUAP and the WOCN Society (NPUAP/EPUAP, 2009; Ratliff et al., 2010). As clinicians and researchers, we can agree the standard hospital mattress should be replaced with one considered to redistribute pressure.

Moisture Management and Skin Care

An RCT with blinded assessors was conducted by Bates-Jensen, Alessi, Al-Samarrai, and Schnelle (2003a) to examine skin health outcomes (such as wetness, blanchable erythema, and PrUs) after an exercise and incontinence program. The sample included 144 incontinent nursing

home residents (M = 87.5 years of age) among four nursing homes. The intervention group (n = 70) received exercise and incontinence care every 2 hours from 0800 to 1630, 5 days a week, for 32 weeks. The control group (n = 74) received usual care. Twenty-two residents (15.3%) developed a Stage I or Stage II PrU: No PrUs higher than a Stage II occurred. While those who received the intervention had better functional, incontinence, and skin wetness outcomes than the controls; there was no difference in PrU incidence rates between the intervention group (14.3%) and the controls (16.2%).

Through retrospective chart review, Schindler et al. (2011) examined moisture management interventions associated with decreased HAPUs among 5,346 pediatric ICU patients; 545 of who developed a PrU during hospitalization. After controlling for age and risk of mortality, decreased HAPU rates were associated with nursing care interventions to manage moisture that included: (a) use of dry-weave diapers, OR = 0.286, 95% CI [0.218 – 0.375], p < .001; (b) urinary catheter, OR = 0.441, 95% CI [0.347 – 0.559], p < .001; (c) use of disposable under pads, OR = 0.345, 95% CI [0.252 – 0.473], p < .001; and (d) application of body lotion, OR = 0.655, 95% CI [0.478 – 0.897], p = .008. These two studies present different findings for two very different samples; the incontinent-elderly nursing home resident, and the critically-ill hospitalized pediatric patient. Incongruent findings are likely due to the differences in sample characteristics, or that Schindler et al.'s (2011) study was so large even a small effect would be statistically significant.

In addition to protection from excessive moisture, emollients should be used to hydrate dry skin (NPUAP, EPUAP, & PPPIA, 2014). In a study of 286 hospitalized patients, multivariable Cox regression analysis revealed that those with dry skin were more likely to develop a HAPU than those without dry skin, RR = 2.31, 95% CI [1.02 – 5.21] (Allman, Goode,

Patrick, Burst, & Bartolucci, 1995). The best emollient to prevent PrUs, however, has not been determined (NPUAP/EPUAP, 2009). The literature review performed by AHRQ (to examine the efficacy of PrU prevention interventions in reducing PrU rates) concluded evidence was meager and insufficient to form reliable conclusions as to the effectiveness of lotions, creams, and cleansers to prevent PrUs, compared to standard care (Chou et al., 2013).

Nutritional Support

For decades, low serum protein levels have been associated with poor wound healing (Thompson, Ravdin, & Frank, 1938) and pressure ulcer development (Mulholland, Tui, Wright, Vinci, & Shafiroff, 1943). Current research confirms serum albumin is associated with PrU development. A retrospective analysis of data from 9,409 persons 14 years of age and older, who were patients at a United Kingdom hospital (from April 2006 through November 2007), was conducted to compare subscale scores on the Waterlow Scale with serum albumin levels in predicting HAPU development (Anthony, Rafter, Reynolds, & Aljezawi, 2011). One hundred and thirty-nine participants (1.5%) developed a HAPU. Pair-wise testing (with a Bonferroni correction and revised p value of .05/25 = .002 for 25 pair-wise tests) determined that serum albumin levels < 3.2 g/dl (binary variable) were associated with increased PrU development, $\chi^2 = 47.8$, df = 1, $p < .001$.

Iizaka, Okuwa, Sugama, and Sanada (2010) conducted a case-control study to investigate the effect of nutritional status on PrU development and PrU severity in the home care setting. Two hundred and seven home health offices in Japan were randomly selected and stratified by region. Participants with a PrU ($n = 290$) were matched to those without a current or previous PrU ($n = 456$). Eleven nutritional risk factors for PrU development were measured. After adjusting for non-nutritional risk factors: (a) the odds for PrU development were higher among

those with malnutrition than those without malnutrition, OR = 2.29, 95% CI [1.53 – 3.44]; and (b) malnutrition was associated with more severe PrUs, OR = 1.88, 95% CI [1.03 – 3.45]. Malnutrition was widely defined as having (a) a body mass index ≤ 18.5, (b) serum albumin ≤ 3.0 g/dl, (c) hemoglobin ≤ 11.0 g/dl, (d) weight loss, (e) presence of edema, or (f) inadequate nutritional intake.

Other studies have examined nutritional *supplements* and HAPU occurrence. Researchers in France conducted a multi-site (19 wards) RCT that included 672 critically ill elderly (M = 83.3 years of age) patients (Bourdel-Marchasson et al., 2000). The nutritional intervention group (n = 295) received two daily oral supplements of 200 kcal each, for 15 days, in addition to the standard hospital diet of 1800 kcal/day. The control group (n = 377) received the hospital diet only. At day 15, 40.6% of the intervention group developed a PrU, compared to 47.2% of the controls. Ninety percent (90%) of the PrUs were Stage I. Multivariate analysis revealed: (a) those in the control group had a higher risk of developing a PrU than those receiving the supplements, RR = 1.57, 95% CI [1.03 – 2.38], p = .04; and (b) for every 1g/L decrease in serum albumin level on admission, there was a 5% increased risk of PrU development, RR = 1.05, 95% CI [1.02 – 1.07], p < .001.

Houwing et al. (2003) conducted a double-blind RCT to investigate the effect of nutritional supplements on (new) PrU development. The sample included 103 hip fracture patients (M = 81 years of age) among three facilities in the Netherlands. Participants were randomized to receive either: (1) 400 ml of a supplement with protein, arginine, zinc, and antioxidants (n = 51); or (2) a zero-calorie water-based placebo (n = 52). Pressure ulcer presence was assessed daily for 28 days or until discharge (median = 10 days). There was no difference in HAPU rate between the intervention (55%) and placebo (59%) groups. However, those

receiving the supplement had a longer time before PrU onset (3.6 +/- 0.9 days) relative to those who received the placebo (1.6 +/- 0.9 days). No Stage III or higher PrUs developed. These HAPU rates, however, are considerably higher than current U.S rates of 7.9% for those at risk for PrU reported by Bergquist-Beringer et al. (2013). Certainly these participants were at risk for PrU due to advanced age and hip fracture, but overall body mass index ($M \approx 24$ BMI) did not suggest malnutrition among either group; so perhaps participants had adequate nutritional intake without the supplement. The study did not report information on changes in weight or other nutritional markers (NPUAP/EPUAP, 2009).

A Cochran Review (Langer, Knerr, Kuss, Behrens, & Schlomer, 2008) to evaluate the effectiveness of parenteral and enteral nutrition on PrU prevention and treatment examined eight RCTs. Reviewers reported difficulty in forming conclusions due to small sample sizes, high attrition rates, and variable interventions. Implications for practice included only that "elderly people recovering from acute illness appear to develop fewer pressure ulcers when given two daily supplement drinks" (p. 8).

Finally, from a sample of 5,346 critical care pediatric patients, those who received a nutritional consult had lower HAPU rates than those who did not have a consult (after controlling for age and risk of mortality), OR = 0.206, 95% CI [0.156 – 0.272], $p < .001$ (Schindler et al., 2011). It is likely the consult led to nutritional support not measured in the study as no other nutritional measures were collected.

Studies on nutritional interventions typically have short follow-up periods, making it difficult to identify effects of the intervention on PrU development (NPUAP/EPUAP, 2009). Despite numerous studies, the NPUAP and the WOCN Society recognize a gap in knowledge with regard to the association between nutritional interventions and PrU development, and

recommend further research (NPUAP, 2013; Ratliff et al., 2010). This is confirmed by the AHRQ's recent efficacy review which concluded there is insufficient evidence to come to any reliable conclusions as to the effectiveness of nutritional support in preventing PrUs, when compared to standard care (Chou et al., 2013).

Education

Pressure ulcer prevention includes education. Caregivers, the person at risk for PrU development and their family, and all levels of healthcare providers should be instructed about PrU prevention (Bryant & Rolstad, 2001). The AHRQ recommends that PrU teaching focuses on risk factors, pathophysiology, and risk reduction (AHRQ, 2011b).

Caregiver education. While education has been shown to improve nurses' knowledge of PU prevention, improved knowledge has not always been linked with improved care—let alone decreased HAPU rates (Armstrong et al., 2008; Athlin, Idvall, Jernfält, & Johansson, 2009; Levine, Ayello, Zulkowski, & Fogel, 2012; Moore, 2010; Pieper & Mattern, 1997; Zulkowski, Ayello, & Wexler, 2007). Beeckman, Defloor, Schoonhoven, and Vanderwee (2011) conducted a study in Belgium to examine the relationships between nurses' knowledge of PrU prevention, attitudes towards PrU prevention, and performance of PrU prevention. This was a cross-sectional study of 553 nurses and 625 patients at risk for PrU development among 14 hospitals. The participants and hospitals were randomly selected. Researchers reported that education was positively correlated with knowledge scores ($p = .003$), but knowledge scores were not correlated with performance of prevention interventions ($p = .198$).

Abel et al. (2005) measured PrU prevention process of care measures and PrU incidence, before and after an education intervention. The sample of convenience included 34 nursing homes in Texas. The intervention was a change in the process of care systems. This change was

introduced via tools and intense education of care providers. Findings showed there was a statistically significant improvement in eight out of ten quality indicators (risk assessment within 2 days of admission, appropriate plan of care that triggered interventions, timely skin assessments, appropriate ulcer description, use of a pressure-redistribution surface, and orders/care plans reflect wound care policy) after the intervention. However, even though facility-acquired PrU occurrence decreased from 13.6% to 10% after the intervention; this decrease was not statistically significant, $\chi^2_{MH} = 3.66$, $p = .06$ [there is a 60% chance that the difference in HAPU incidence was from chance alone (Chi-square test, n. d.)].

Patient education. An RCT was conducted by Rintala, Garber, Friedman, and Holmes (2008) to determine if patient education and follow-up after PrU surgery decreased PrU recurrence. The sample consisted of 41 veterans (all male) with spinal cord injury ($n = 39$) or multiple sclerosis ($n = 2$). Participants were randomized into three groups (1) individualized PrU education and monthly phone follow-up ($n = 20$), (2) *monthly* mail or phone follow-up without the education portion ($n = 11$), and (3) *quarterly* mail or phone follow-up without the education portion ($n = 10$). All follow-up continued for 2 years, or PrU recurrence, or death. Patients who received the education (Group 1) had a lower rate of PrU recurrence relative to patients who did not (Groups 2 and 3) [33%, 60%, and 90%; $p = .007$]. Also, there was a longer time before PrU recurrence among those that received the education (Group 1) relative to those that did (Groups 2 and 3) [19.6, 10.1, and 10.3 months; $p = 002$]. Limitations to this study include a small sample size. Moreover, the groups differed with regard to type of PrU surgery, level of spinal cord injury, time since last PrU surgery for those who had a history of this, and self-reported health status at baseline. Despite limitations, this study supports the use of education to decrease PrU recurrence. While the literature differs as to the effect education has on PrU occurrence,

commonsense says one must have the knowledge before they can be expected to properly perform PrU prevention. In other words, education is necessary but not sufficient to decrease PrU occurrence. The NPUAP recommends further research on patient literacy and teaching (NPUAP, 2013).

In summary, because numerous variables are associated with PrU development, it is challenging and often not feasible to apply the rigorous controls necessary to determine causality. Evidence supports that while individually each intervention may be necessary to prevent PrUs, they may not be sufficient. An interesting descriptive-cohort study was performed by Bates-Jensen et al. (2003b) to determine whether nursing homes (in California) that ranked among the extreme quartiles of PrU prevalence, provided different PrU prevention interventions than nursing homes in the other extreme quartile. From among a total of 45 nursing homes in these upper and lower quartiles, 16 participated (upper quartile, $n = 10$; lower quartile, $n = 6$). Pressure ulcer prevalence rates were those reported as part of the Minimum Data Set PrU Indicator. Data on 16 PrU care processes were collected from medical records, direct patient observation, interviews, and wireless thigh movement monitors to measure repositioning. Nine of these indicators measured aspects of PrU *prevention* such as risk assessment, every 2-hour repositioning, pressure reduction, nutritional support, and moisture management; and seven indicators measured aspects of PrU *treatment*. The only significant difference in PrU prevention between nursing homes with low PrU prevalence rates and high PrU prevalence rates, was that homes with low PrU prevalence were *less* likely to use pressure-redistribution surfaces (M = 52%, SD = 38.7) than the high PrU prevalence homes (M = 68%, SD = 33.1); $p < .001$. This is contrary to what was expected. Knowing their performance of PrU prevention would be examined, perhaps only nursing homes that performed well on PrU prevention interventions

agreed to participate. Or perhaps what matters is *what combination* of prevention interventions is applied, such as with the combination of interventions implemented with a PrU prevention program. There is still much to learn about PrU prevention.

Pressure Ulcer Prevention Programs

Pressure ulcer prevention programs use an interdisciplinary approach to implement a combination of PrU prevention strategies (AHRQ, 2011b). Numerous studies have described the implementation of PrU prevention programs. Three recent literature reviews have evaluated these PrU prevention studies. Each review focused on *programs* of prevention; which is important because that is typically how PrU prevention is implemented, and *immediate* adoption of a multifaceted intervention program is among the top ten patient safety strategies recommended by an international panel of patient safety and quality care experts (Shekelle et al., 2013). Many of the studies reviewed, implemented a combination of similar PrU prevention strategies. However, which strategies were included, how they were implemented, and how processes and outcomes were measured as they related to the intervention; varied widely among the studies.

A meta-analysis was performed by Soban et al. (2011) to examine studies that implemented PrU prevention programs. Of the 39 studies published from 1990 to 2009 that met inclusion criteria, 22 were U.S. studies. Most used a pre-, post-intervention design at a single site. Interventions implemented varied widely across the studies. The number of studies in the review, and the interventions studied, were: (a) protocol developed/implemented, 29/39; (b) staff education, 28/39; (c) risk assessment tool, 21/39; (d) performance monitoring at least three times during the study, 20/39; (e) team assembled, 19/39; (f) beds/support surfaces, 14/39; (g) guideline implemented based on published guidelines specified in the study, 11/39; (h) feedback to nursing staff and/or nurse managers, 10/39; and (i) link/resource nurse, 9/39. No standardized

measure of these interventions was available, however, and evidence as to the reliability of the measures was lacking.

Most of the studies included in the meta-analysis found their PrU prevention initiatives were associated with improvement in at least one nursing care process or patient outcome measure. Nearly all (37 out of 39) reported the effect of implementation on patient *outcome* measures. From this, Soban and colleagues computed an overall risk ratio of -0.07; 95% CI [-0.0976 to -0.418]; $p < .0001$, reflecting a decreased risk of developing a HAPU after implementation of at least one nurse-focused intervention. However, only eight of the 39 studies reported the effect of implementation on nursing care processes. Moreover, how these nursing care processes were measured, varied across studies. Due to the lack of evidence for the reliability of data from these process measures, the credibility of the study findings must be questioned.

A literature review was performed by Niederhauser et al. (2012) to examine the evidence for "comprehensive" PrU prevention programs in acute or long-term care settings**.** Studies published from 1995 to 2010 were included in the review ($N = 24$) if their program had two or more interventions and involved a multidisciplinary team. Similar to Soban et al.'s (2011) meta-analysis, most of the studies (21/24) reported the effect of the prevention program on patient outcome measures, while only five studies considered how the program affected care processes (i.e. improved performance of repositioning).

Overall, 17 of the 20 studies that reported the intervention's effect on care process or patient outcome measures found improvement after the intervention. Five of the six studies that measured facility-acquired PrU rates found a decrease in these rates after the intervention. Researchers reported post-intervention improvement in some care processes, but not others.

Also, interventions included in the prevention programs varied across studies, from repositioning and use of pressure-redistribution mattresses, to mentoring, initiation of a physician communication form, and organization of supplies.

Results from this review suggest that a multifaceted program of prevention that includes repositioning, skin assessment, PrU risk assessment, heel protectors, and use of pressure-redistribution surfaces such as new mattresses, are important to decreasing PrU rates. This is significant because PrU prevention is usually delivered through a multifaceted approach. Unfortunately, due to the wide-variety (and various combinations) of interventions implemented, and the lack of standardized and reliable process measures, the relationship between a PrU prevention program and improvement in patient outcome measures is difficult to understand and imprecise.

The third and most recent review was conducted by Sullivan and Schoelles (2013) to examine "multicomponent" PrU prevention programs. Multicomponent was broadly defined as programs with more than one component, such as a skin champion and education. Only studies that reported PrU rates as an outcome variable for at least 6 months after initiating the program were included. Twenty-six studies, 18 acute care and 8 long-term care settings, published from 2000 to 2012 met study criteria. Components of the PrU prevention programs, and the number of studies including them, were: (a) risk assessment, 24/26; (b) repositioning, 23/26; (c) moisture management, 21/26; (d) support surfaces, 21/26; (e) nutrition, 17/26; (f) skin examinations, 13/26; (g) reduce friction and shear, 12/26; (h) patient and family education, 2/26; and (i) use of the Agency for Health Care Policy and Research Clinical Practice Guidelines for PrU Prediction and Prevention, 1/26.

Findings from this review were similar to the previous literature reviews (Soban et al., 2011; Niederhauser et al., 2012) in that 24 out of 26 studies reported improved PrU rates. Eleven of those 24 studies had a statistically significant improvement, and out of the 13 studies not reaching statistical significance, five reported improvement in process of care measures (and PrU rates). Once again, the care processes that were measured before and after initiation of the prevention program, and how they were measured, varied across studies: They were not standardized measures. Also concerning is that no evidence to support the reliability of these process measures was presented. Similar findings are not surprising because many studies were included in more than one review. Only 11 of the 26 Sullivan and Schoelles (2013) studies had not been included in either the Niederhauser et al. (2012) or Soban et al. (2011) reviews: Nine studies were included in all three reviews.

Sullivan and Schoelles' (2013) review was unique because it reported on: (a) the models or theories that guided the research—most of which were quality improvement models such as Plan-Do-Study-Act, and failure mode and effects analysis; (b) the details of how initiatives were implemented—some very unique such as staff sitting on bedpans for 30 minutes (Young, Ernsting, Kehoe, & Holmes, 2010); (c) solutions to perceived barriers to PrU prevention intervention initiatives; and (d) cost-savings. Four of the 26 studies that evaluated cost, reported savings, with two studies reporting an annual savings of $2.4 million (Courtney, Ruppman, & Cooper, 2006) and $6.7 million (McInerney, 2008) per facility.

These three reviews, the NPUAP/ EPUAP/PPPIA (2014), and the WOCN Society (Ratliff et al., 2010); recommend further research be performed to better understand the effect of interventions on PrU occurrence. But in order to link prevention interventions with decreased PrUs, we as nurse researchers must have standardized measures with empirically supported

reliability. Only one study was found that used standardized measures of PrU prevention interventions to examine their association with HAPU rates (Bergquist-Beringer et al., 2013). The sample included 710,626 adult patients on the adult critical care, step-down, medical, surgical, and medical/surgical units who were surveyed for PrUs among 1,419 U.S. hospitals that participated in the NDNQI. Data on PrU risk assessment and PrU prevention were gathered by chart review, and data on PrU identification and staging were gathered by direct patient examination.

Of the 710,626 patients, 282,500 (40%) were considered "at risk" for PrU development. Among patients at risk for PrU, 7.9% (n = 22,317 patients) developed a PrU after hospital admission. After controlling for hospital characteristics and nurse staffing, hierarchical logistic regression revealed a decreased odds of HAPU for those who: (a) received a skin assessment on admission, OR = 0.76, 95% CI [0.67 – 0.87], $p < .001$; (b) received a PrU risk assessment on admission, OR = 0.82, 95% CI [0.73 – 0.92], $p < .001$; (c) were reassessed for PrU risk during the 24 hours prior to the NDNQI PrU survey, OR = 0.87, 95% CI [0.81 – 0.92], $p < .001$; (d) received a daily skin assessment, OR = 0.82, 95% CI [0.72 – 0.94], $p = .003$; or (e) were repositioned as prescribed, OR = 0.86, 95% CI [0.81 – 0.92], $p < .001$. Higher HAPU rates were reported for patients who: (a) received nutritional support, OR = 1.58, 95% CI [1.51 – 1.66], $p < .001$, compared to those who did not; (b) daily skin assessment was contraindicated, OR = 2.84, 95% CI [1.64 – 4.95], $p < .001$, compared to those for whom it was not contraindicated; or (c) refused repositioning, OR = 1.78, 95% CI [1.43 – 2.20], $p < .001$, compared to those who did not refuse.

Overall, findings support the use of PrU prevention interventions to decrease HAPU development. However, a fundamental problem remains: the reliability of these PrU prevention

process measures was not reported. The fact that a large number of hospitals use the same NDNQI structure, process, and outcome measures, makes NDNQI research especially newsworthy to any acute care facility's administration and nursing staff. This emphasizes the pressing need for establishing the reliability of the PrU prevention measures used in NDNQI research. The literature highlights the need for, and value of, using standardized national quality indicators to improve the quality of patient care across the nation.

Criteria for a Sound Measure

While reliability is the focus of this proposed research, validity is a fundamental requirement of a psychometrically sound measure and therefore warrants a brief discussion.

Validity

Validity is "the degree to which the measures or observations are appropriate or meaningful in the way they claim to be" (Rosenthal & Rosnow, 2008, p. 763). Systematic error is the central threat to validity (Salkind, 2006). Systematic error is predictable; it does not fluctuate from one measuring condition to another (Waltz et al., 2010). Validity assessment falls into three categories (1) content validity, (2) construct validity, and (3) criterion validity (Rosenthal & Rosnow, 2008).

Content validity is the extent to which a test measures the content of the concept being tested, and is usually evaluated during instrument development through literature review, expert opinion, and qualitative research. After the tool is developed, content validity may be supported by content validity indices calculated from experts' opinions (DeVon et al., 2007). This involves rating each item for relevance. Face validity is the subjective evaluation that the tool measures what it claims to measure (DeVon et al., 2007).

The overarching validity of a measure is construct validity (DeVon et al., 2007). Construct validity refers to whether a measure accurately reflects the concepts as operationally defined (Waltz et al., 2010). Methods to assess construct validity include contrasted groups, hypothesis testing, a multitrait-multimethod approach, and factor analysis (Waltz et al., 2010).

Criterion validity refers to the relationship between the instrument and its performance on some other variable. Three types of criterion validity exist: (a) predictive validity, which is the extent scores predict performance on a future measure; (b) concurrent validity, which reflects the amount scores are correlated to a related criterion at the same point in time; and (c) convergent and discriminant validity. Convergence reflects a high correlation between theoretically similar constructs, while discriminant validity represents an instrument's ability to distinguish between theoretically similar and different constructs (DeVon et al., 2007; Rosenthal & Rosnow, 2008).

Reliability

Reliability is "the degree to which observations or measures are consistent or stable" (Rosenthal & Rosnow, 2008, p. 757). A reliable measure—whether it be a quantitative survey, an observation, or a subjective judgment—allows for replication of the same or similar results. Random error is the main threat to reliability (Salkind, 2006). According to Classical Measurement Theory, every observed score is comprised of a true score and the influence random error has on that true score (Salkind, 2006; Waltz et al., 2010):

$$\text{Observed Score} = \text{True Score} + \text{Error Score}$$

It is not possible to measure the true score, as the true score is a theoretical representation of the actual amount of the trait being measured: A true score is 100% accurate (Salkind, 2006).

Random error, therefore, is present (to some extent) in all measurement; but it too remains unknown, only the observed score is known (Waltz et al., 2010). Random error can be

classified as trait error or method error (Salkind, 2006). Trait error is error caused by the individual, such as filling out a survey while ill, being distracted, and motivation. Method error is caused by the testing situation, such as unclear instructions or an uncomfortable testing environment. In addition, random error can originate after data collection, such as with data entry mistakes (Wambach, 2012b). Keeping random error to a minimum, therefore, is essential if a measure is to be reliable. Random error results in higher *and* lower than expected scores. If the measurement process is repeated enough times, therefore, random error cancels itself out and average scores will represent true values (Waltz et al., 2010).

Reliability is mathematically defined as the proportion of variance in the distribution of the observed scores that is caused from true score differences (versus caused by random error). Waltz et al. (2010, p. 73) statistically defined reliability with the following steps:

$$\text{Reliability} = \frac{\text{True Variance}}{\text{Observed Variance}}$$

Because *unreliability* is the result of random error causing variance in the observed score, unreliability can be statistically defined as

$$\text{Unreliability} = \frac{\text{Error Variance}}{\text{Observed Variance}}$$

Applying the Classical Measurement Theory, variance (commonly displayed as a 95% confidence interval) is written as

$$\text{Observed Variance} = \text{True Variance} + \text{Error Variance}$$

Dividing both sides of the equation by the Observed Variance results in the equation

$$1.0 = \frac{\text{True Variance}}{\text{Observed Variance}} + \frac{\text{Error Variance}}{\text{Observed Variance}}$$

Reliability of the observed score (True Variance /Observed Variance), therefore, is written as

$$\frac{\text{True Variance}}{\text{Observed Variance}} = 1.0 - \frac{\text{Error Variance}}{\text{Observed Variance}}$$

Types of reliability include stability, internal consistency, and equivalence (Waltz et al., 2010).

Stability reliability. Stability assesses if the instrument measures the concept consistently over time. It is evaluated by repeating the measure twice within an appropriate specified time. This is called test-retest reliability (Portney & Watkins, 2009; Waltz et al., 2010).

Internal consistency reliability. Internal consistency reflects the homogeneity of the items in an instrument. For example, internal consistency reflects the logic behind combining the items together to measure a single attribute (Salkind, 2006). Internal consistency is commonly estimated by Cronbach's alpha (α), which is based on correlations between an individual's score on one item and their total score (Salkind, 2006), and therefore reflects the extent that performance on any single item predicts performance on any other item in the measure (Waltz et al., 2010).

Equivalence reliability. Equivalence means similar results are obtained from two equivalent forms of a test (DeVon et al., 2007). When measurement involves observation, equivalence is assessed by inter-rater and intra-rater reliability. In*tra*-rater reliability is the stability of the data when multiple observations by the same person, separated by a short interval, produce the same results. The number of repeated observations depends on how much variability is expected (Portney & Watkins, 2009). Intra-rater reliability is an assumption for inter-rater reliability testing (Gwet, 2012), and can be estimated by Pearson's *r* correlations (Wambach, 2012a).

Inter-rater reliability is the degree to which two or more observers (raters) agree on their judgments of an outcome (Rosenthal & Rosnow, 2008). The best method for estimating inter-rater reliability depends on the number of raters (two, or more than two), and whether all raters score the same participants or if multiple ratings per subject are done by different raters (Fleiss et

al., 2003). Method selection also depends on if the ratings are dichotomous, nominal, ordinal, or ratio. Finally, for inter-rater reliability, issues of prevalence, bias, and interdependence must be considered throughout analysis and interpretation of results (Sim & Wright, 2005). These issues are discussed fully, later in this chapter. Statistical tests to determine inter-rater reliability include percent agreement, kappa, intraclass correlation coefficient (*ICC*), and Pearson's *r* or its non-parametric counterpart Spearman rho (Portney & Watkins, 2009).

Percent agreement is the simplest agreement index, and is used when measurements are categorical. It measures how often raters agree on the rating they give someone on a particular attribute, or how often test-retest scores agree. Percent agreement is the proportion of observations on which raters agree and is written as (Portney & Watkins, 2009, p. 598):

$$\% \text{ Agreement} = \frac{\text{Number of Agreements}}{\text{Number of Possible Agreements}}$$

Its main limitation is that it does not account for agreement due to chance. Although taken alone it has limitations, percent agreement is an important commonsense descriptive statistic (Uebersax, 2009).

Kappa was first proposed by Cohen (1960). Kappa takes into consideration that some agreement will occur by chance alone and corrects for this chance agreement (Fleiss et al., 2003). This correction, however, assumes statistical independence of raters. Lack of independence will inflate kappa (Sim & Wright, 2005; Uebersax, 2010a). For nominal and ordinal ratings, inter-rater reliability is best estimated by Cohen's kappa when there are two raters, and Fleiss' kappa is best for more than two raters (Gwet, 2012; Salkind, 2006). Fleiss' kappa also allows for the number of raters to be more than the number of raters per subject (i.e. not every rater needs to evaluate every subject). What matters is that each subject is evaluated the same number of times (Zaiontz, 2013). Fleiss' kappa, therefore, measures the overall agreement between all raters

(Uebersax, 2010b), and is based on the proportion of agreeing rater-rater pairs compared to the total number of rater-rater pairs (Zaiontz, 2013). Kappa (*k*) can be mathematically written as

$$k = (P_o - P_{e)} / (1 - P_e)$$

where P_o is the proportion of observed total agreement (i.e. percent agreement), and P_e is the proportion of agreement that is expected by chance (Feinstein & Cicchetti, 1990, p. 544).

When ratings have more than two choices, kappa can be "weighted". Weighted kappa considers the degree of agreement and will be higher for raters who are closer to agreement than those who are not (Fleiss et al., 2003). For instance, raters who differ between *completely agree* and *mostly agree* will have a higher weighted kappa value than raters who differ between *completely agree* and *completely disagree*. Different methods of weighting exist and will result in different kappa values (Sim & Wright, 2005).

Kappa equals 1.0 when there is complete agreement. If the observed agreement is greater than or equal to the agreement due to chance alone, kappa is ≥ 0. If observed agreement is less than agreement expected from chance alone—which rarely occurs in clinical contexts (Sim & Wright, 2005)—then kappa is < 0 (Fleiss et al., 2003). Researchers have categorized kappa's strength of agreement as:

Landis and Koch (1977, p. 165)		Shrout (1998, p. 308)		Fleiss et al. (2003, p. 604)	
< 0.00	Poor	0.00 to 0.10	Virtually none	< 0.40	Poor
0.00 to 0.20	Slight	0.11 to 0.40	Slight	0.40 to 0.75	Good
0.21 to 0.40	Fair	0.41 to 0.60	Fair	> 0.75	Excellent
0.41 to 0.60	Moderate	0.61 to 0.80	Moderate		
0.61 to 0.80	Substantial	0.81 to 1.00	Substantial		
0.81 to 1.00	Almost Perfect				

The classifications by Landis and Koch (1977) are most commonly used (Bergquist-Beringer, Gajewski, Dunton, & Klaus, 2011; Fleiss et al., 2003; Hart, Bergquist, Gajewski, & Dunton, 2006). What is considered the acceptable level of kappa, however, depends on the situation.

Nonetheless, for the current study, the recommended reliability level is ≥ .610, which is similar to Polit and Beck's (2012) recommendation of > .60, but much lower than Salkind's (2006) recommendation of .90. Commonsense says inter-rater reliability should be better than "moderate" agreement.

Prevalence-adjusted kappa (PAK) is an additional statistic to measure observer agreement, and is useful to adjust kappa values when prevalence may affect kappa values (Sim & Wright, 2005). Prevalence is the difference between counts in cells of agreement (cells on the diagonal, or cells *a* and *d* in Table 1 below). Similarly, bias-adjusted kappa (BAK) is useful to adjust kappa values when bias my affect kappa values (Sim & Wright, 2005). Bias is the difference between counts in cells of disagreement (cells off the diagonal, or cells *c* and *b* in Table 1).

Table 1**.**

Illustration of Prevalence in a Distribution of 100 Subjects by Rater and Response Category

		Rater 1: Yes	Rater 1: No	
Rater 2	Yes	a 98	b 1	$g_1 = 99$
	No	c 1	d 0	$g_2 = 1$
		$f_1 = 99$	$f_2 = 1$	$n = 100$

The effect of prevalence on kappa can be explained by referring to the equation for kappa:

$$k = (P_o - P_e)/ (1 - P_e)$$

If the proportion of expected by chance agreement (P_e) is large, even a large proportion of observed agreement (P_o) can be reflected as a low kappa. A prevalence index can be calculated

to determine if prevalence may influence the kappa values. The equation for calculating prevalence index is (Sim & Wright, 2005, p. 261):

$$\text{Prevalence index} = \frac{|a - d|}{n}$$

The prevalence index has a possible value of 0 to 1.0. A value at or near zero would reflect equiprobable (positive and negative) cases. For the hypothetical data in Table 1 (above) showing a 98% "Yes" rating, agreement expected by chance alone is:

$$P_e = \frac{\left(\frac{f1 \; x \; g1}{n}\right) + \left(\frac{f2 \; x \; g2}{n}\right)}{n} = \frac{\left(\frac{99 \text{ x } 99}{100}\right) + \left(\frac{1 \text{ x } 1}{100}\right)}{100} = .98$$

Observed agreement is:

$$P_o = (a + d)/n = (98 + 0)/100 = .98$$

and $k = 0$ even though the raters disagreed on only two out of 100 cases.

$$k = \frac{(P_o - P_e)}{1 - P_e} = \frac{(.98 - .98)}{1 - .98} = 0/.02 = 0$$

Kappa values, therefore, are difficult to interpret unless prevalence is considered.

In addition to prevalence, bias may be present and affect kappa. Bias is the degree to which raters disagree on the *proportion* of positive and negative cases, and is reflected as unbalanced marginal totals. Marginal homogeneity and/or a bias index can be used to identify and evaluate the degree of imbalance among marginal totals (Sim & Wright, 2005). The bias index has a possible value of 0 to 1.0:

$$\text{Bias index} = \frac{|b - c|}{n}$$

Table 2 (from Feinstein & Cicchetti, 1990, p. 545) is an example of the effect unbalanced marginal totals may have on kappa. The proportion of observed (P_o) agreement is the same for both tables [$P_{o \text{ Table 2A}} = (45 + 15)/100 = 60\%$; $P_{o \text{ Table 2B}} = (25 + 35)/100 = 60\%$]; yet the

proportion of expected agreement (P_e) causes different kappa values [$P_{e\ Table\ 2A}$ = (0.70)(0.60) + (0.30)(0.40) = 0.54, so k = (0.60 - 0.54)/(1 - 0.54) = 0.13; $P_{e\ Table\ 2B}$ = (0.30)(0.60) + (0.70)(0.40) = 0.46, so k = (0.60 - 0.46)/(1 - 0.46) = 0.26].

Table 2.

Illustration of the Influence of Bias on Kappa Value

2A

		Rater 2		
		Yes	No	Total
Rater 1	Yes	45	15	60
	No	25	15	40
	Total	70	30	100

2B

		Rater 2		
		Yes	No	Total
Rater 1	Yes	25	35	60
	No	5	35	40
	Total	30	70	100

Compared to balanced marginal totals (in Table 2A), unbalanced totals (in Table 2B) will produce a higher kappa because P_e is smaller for unbalanced totals (Feinstein & Cicchetti, 1990).

Kappa can be adjusted for prevalence (PAK), bias (BAK), or both (PABAK). These adjustments use *average* cell counts instead of *actual* cell counts of agreement (prevalence-adjustment), and *average* cell counts instead of *actual* cell counts of disagreement (bias-adjustment), between raters (Sim & Wright, 2005). An example of PAK is presented in Table 3 below. In this example, the actual cell counts in Table 3A produce a k of .52; and average cell counts of agreement in Table 3B produce a PAK of .72.

Table 3.

Example of Calculating Prevalence-Adjusted Kappa and Change in Kappa Value

A

		Rater 2			
		Yes	No	Pending	Total
Rater 1		**90**	3	4	97
		7	**10**	2	19
		3	5	**5**	13
		100	18	11	129

B

		Rater 2			
		Yes	No	Pending	Total
Rater 1		**35**	3	4	42
		7	**35**	2	44
		3	5	**35**	43
		45	43	41	129

While some believe PAK should not be used because the kappa calculated from average cell counts do not convey the circumstances in which the initial ratings were made (Hoehler, 2000), others suggest reporting kappa with and without PAK in order to gain insight into how prevalence may have affected kappa values (J. Sim, personal communication, May 5, 2014; Sim & Wright, 2005). Because prevalence was a concern for this study, PAK values were presented to better understand the effect high prevalence may have had on the Cohen's kappa values (Sim & Wright, 2005).

Intraclass correlation coefficients *(ICCs)* are widely used to estimate inter-rater reliability, and are appropriate for quantitative (continuous level) ratings regardless of the number of raters (Fleiss et al., 2003; Gwet, 2012; Salkind, 2006; Shrout, 1998; Uebersax, 2010b). When the same raters are used for the same subjects, *ICC*s are identical to kappa values (Fleiss et al., 2003). The *ICC* is the ratio of between-subjects variance to total variance, and theoretically has a possible value of 0 to 1.0 with 1.0 representing perfect reliability (Portney & Watkins, 2009). However, an *ICC* will be negative when the average covariance within subjects is negative (Nichols, 1999).

There are six forms of *ICC*s, each giving different results with the same data. Selecting the appropriate *ICC* model depends on if the subjects are rated by the same or different sets of raters, and if one wants to generalize findings to a larger population of raters (Portney & Watkins, 2009; Shrout & Fleiss, 1979). According to Shrout and Fleiss (1979) three *ICC* models pertain to raters. In the model that was used in this study (Model 1) participants are judged by different groups of raters. The raters are considered randomly selected. For this model, a one-way analysis of variance (ANOVA) estimates the variance between participants, as well as error.

The error is the variation within a participant across raters; and is made up of true changes, rater error, and unexplained error. The type of error, however, is not differentiated.

In the inter-rater reliability of PrU identification, staging, and origin study by Hart et al. (2006), Model 1 was used to calculate the *ICC* to compare variability of responses within and between hospitals (B. Gajewski, personal communication, May 15, 2014). Specifically, in Hart et al.'s (2006) study the *ICC* represented the ratio:

$$\frac{\text{(Variability of kappa's between hospitals)}}{\text{(Variability of kappa's between hospitals) + (Variability of kappa's within hospitals)}}$$

An *ICC* near zero, therefore, would reveal a large proportion of the variability on an item is within hospitals. This equates to hospitals having similar inter-rater reliabilities. Put another way, a hospital's given data collectors identified, staged, and classified PrUs with the same amount of error as other hospitals' data collectors. This use of *ICC* was employed in this study to estimate inter-rater reliability *between* hospitals.

Signal-to-noise ratio is another measure of reliability. The National Quality Forum (NQF) recommended that measures used for value-based purchasing, such as CMS's measure of Stage III or Stage IV PrUs, are evaluated using a signal-to-noise ratio (NQF, 2013d). Signal-to-noise ratio represents the ratio of useful information (signal) to irrelevant information (noise); and is a measure of precision (Adams, Mehrotra, Thomas, & McGlynn, 2010). For value-based purchasing measures, this is the ratio of real variation in performance (signal), to variation due to imprecision of the measure (noise). It is used for value-based payments to classify the provider's level of performance in relation to another provider's performance (Adams et al., 2010). One can understand the value of assessing the signal-to-noise ratio because CMS payments are not determined in isolation, but instead relative to other providers' level of performance (Tompkins, Higgins, & Ritter, 2009).

Reliability Studies of the NDNQI Measures

A number of reliability studies on the NDNQI measures have been published (Table 4). They include: (a) PrU identification, staging, and origin (Bergquist-Beringer et al., 2011; Gajewski, Hart, Bergquist-Beringer, & Dunton, 2007; Hart et al., 2006); (b) Fall classification (Simon, Klaus, Gajewski, & Dunton, 2013); (c) Patient day (Simon, Yankovskyy, & Dunton, 2010; Simon, Yankovskyy, Klaus, Gajewski, & Dunton, 2011); (d) RN Satisfaction survey (Taunton et al., 2004; Elliott, 2006); and (e) Nursing care hour (Klaus et al., 2013). The purpose of reviewing these studies was to identify the research designs and methods that might be used in this study. Three of these nine reliability studies (Bergquist-Beringer et al., 2011; Gajewski et al., 2007; Hart et al., 2006) were deemed relevant to this study, and therefore are discussed below. These studies are closely related and (although somewhat dated) support the reliability of these PrU outcome measures.

Hart, Bergquist, Gajewski, and Dunton (2006) investigated the inter-rater reliability of PrU identification, PrU staging, and PrU origination (nosocomial or community-acquired). A random selection of 48 hospitals (87% response rate) from among NDNQI hospitals participated in the study. Each hospital was randomly assigned to take one of two versions of a criterion-referenced online test. Both versions of the test included 20 photographs of PrUs and 5 photographs of other wounds, but only one version included a short narrative describing the wound in the photograph. Two hundred, fifty-six RN raters (69% response rate) [1 to 21 raters (M = 5.3) per hospital] completed the test.

Cohen's kappa values were computed by comparing responses from each hospital rater to the "correct" response as determined by an expert panel. The overall kappa value for wound identification (yes PrU/no PrU) was 0.56. The overall kappa value for PrU staging was 0.65 and

Table 4.

Review of Literature: Reliability of NDNQI Measures

Reference & Purpose	Sample & Setting	Methods	Analysis	Findings
		Pressure Ulcer Identification, Staging, and Origin		
1. Hart et al. (2006) Assess IRR of NDNQI PrU identification, staging, & origin.	48 NDNQI Hospitals 256 RN raters (U.S.)	Random assignment to 1 of 2 versions of on-line test (25 photos w/ & w/o descriptors)	IRR: *k*, *ICC* HLM	PrU (yes/no) identification $k = 0.84$; PrU staging $k = 0.65$; PrU origin $k = .80$ *ICC*s near 0 revealed most variation was within hospitals, not between hospitals. No association between rater characteristics, hospital characteristics, or test version with classifying PrU as nosocomial or hospital-acquired.
2. Bergquist-Beringer et al. (2011) Evaluate reliability of PrU identification, staging, & origin	31 NDNQI Hospitals 162 Raters (M = 5.2/hospital) (U.S.)	Raters collectively evaluated 591 PrUs directly and were randomized to 1 of 2 online tests (w/ & w/o wound descriptors) (Triangulation approach)	Direct observation: *k* between rater & expert; HLM Web-based: *k* between each rater & experts HLM	PrU Staging by Direct Observation: Mean $k = 0.60$. HLM: Teams led by WOCN had higher *k* values Web-Based Test: 1. Staging Overall $k = 0.69$ (substantial) With descriptors $k = 0.81$ (near perfect) Without descriptors $k = 0.18$ (slight) 2. Wound Identification (yes a PrU/not a PrU) Overall $k = 0.83$ (near perfect) 3. PrU Origin Overall $k = 0.79$ (substantial) 4. HLM: No difference in PrU staging *k* values for teams led by WOCN compared to those not led by WOCNs. WOCNs higher agreement on PrU identification and PrU origin compared with non-WOCNs.

Table 4 (continued).

Review of Literature: Reliability of NDNQI Measures

Reference & Purpose	Sample & Setting	Methods	Analysis	Findings
		Pressure Ulcer Identification, Staging, and Origin (continued)		
3. Gajewski et al. (2007) Assess IRR of PrU identification, staging, and origin from direct observation and web-based photographs.	20 NDNQI Hospitals 140 raters (*M* = 7 raters/hospital) Total of 347 PrUs (U.S.)	Described the method for estimating IRR	Bayesian hierarchical model (BHM)	*ICC* = 0.57 (0.06 to 1.0) BMH more accurate estimate of IRR than non-Bayesian methods
		Fall Classification		
4. Simon et al. (2013) Investigate how experts & hospital staff classify falls compared to NQF's definition	Phase I: 24 experts Phase II: 6,342 hospital staff (78% RNs) from 362 units & 170 hospitals (U.S.)	Both phases used an online video survey with 20 fall scenarios	Bayesian determined cutoff of 70.2% agreement *ICC*s	Experts did not agree (≤ 70.2% agreement) on fall classification for 6 of the 20 scenarios. Experts and staff did not agree on 4 of the 20 scenarios. NQF's definition of fall needs refinement.

Table 4 (continued).

Review of Literature: Reliability of NDNQI Measures

Reference & Purpose	Sample & Setting	Methods	Analysis	Findings
	Patient Day			
5. Simon et al. (2010) Determine the most accurate method to collect patient day data.	Hypothetical "average" surgical unit	Simulated inpatient census data varying % of short stay patients from 0% - 50% (10,000 iterations) Compared data collection methods to gold standard of actual hours of inpatients. and short stay patients.	Descriptive	Actual hours of inpatient and short stay patients is best method for collecting patient day data Bias present for all other NDNQI-endorsed methods and increased as percent of short stay patients increased. Second best methods are (1) every 6 hour census, (2) noon/midnight census.
6. Simon et al. (2011) Assess reliability of 5 measures of patient patients. day to the "gold standard" of every 2-hour census.	255 out of 282 units from 54 hospitals (U.S.)	Compared 5 approved methods to gold standard of q 2-hour census; on 7 random days in Sept. 2008.	*ICC*s; Bayesian regression analysis	Overall agreement between 5 methods and gold standard of q 2-hour census was excellent, *ICC* = 0.967, 95% CI [0.958-0.974]. 2 methods underestimated patient day. 1 method underestimated patient day (by 7.6%) for units with lowest degree of short stay patients, and overestimates (by 7.7%) for units with highest degree of short stay Most all NDNQI hospitals use reliable patient day measurement methods, however, units with large or small degrees of short stay patients should consider multiple census collection methods.

Table 4 (continued).

Review of Literature: Reliability of NDNQI Measures

Reference & Purpose	Sample & Setting	Methods	Analysis	Findings
	RN Satisfaction Survey			
7. Taunton et al. (2004) Use psychometric evaluation to adapt the Stamps' Index of Work Satisfaction for use by NDNQI as a measure of RN job satisfaction.	Study I 10 NDNQI Hospitals - 918 RNs Study II 11 NDNQI Hospitals -2277 RNs (U.S.)	Mail-in Survey Mail-in or online Survey	Factor Analysis Concurrent Validity	The NDNQI-Adapted Index of Work Satisfaction is a reliable and valid measure or RN work satisfaction. Online survey acceptable alternative to mail-in form
8. Elliott (2006) Psychometric evaluation of aggregated NDNQI-RN Satisfaction Survey data from the individual level to the unit & hospital levels.	206 NDNQI Hospitals 50% response 55,516 RNs 40% response 66,160 RNs (U.S.)	Secondary data analysis of 2004 data Psychometric evaluation	Validity: hypothesis testing Reliability: alphas inter-item *r* ANOVA *ICC*s	NDNQI-RN Satisfaction Survey is reliable and valid measure of job satisfaction even after aggregating individual data to unit and hospital level. 40% response is likely acceptable if the data set is large (as opposed to the required 50% response rate).

Table 4 (continued).

Review of Literature: Reliability of NDNQI Measures

Reference & Purpose	Sample & Setting	Methods	Analysis	Findings
		Nursing Care Hour		
9. Klaus et al. (2013) Examine processes to generate nursing care hour data; evaluate IRR	714 NDNQI Site Coordinators for descriptive Part I 11 SCs for IRR Part II (U.S.)	Part I: Multiple choice online survey Part II: Compare "time-in/time-out" clock times to usual NDNQI extracted data. (Rater-to-Standard Design) Interview	Descriptive statistics; *ICC*s	15% & 17% incorrectly counted monitor techs and unit secretaries respectively. Only 26% correctly counted new RNs. IRR "high" with overall *ICC* = 0.97.

ANOVA = analysis of variance
BHM = Bayesian hierarchical model
HLM = hierarchical linear model
ICC = intraclass correlation coefficient
IRR = inter-rater reliability
k = kappa
NQF = National Quality Forum
WOCN = Wound Ostomy Continence Nurse
SC = site coordinator

overall kappa value for PrU origin was 0.80. These values demonstrated moderate to almost perfect reliability (Landis & Koch, 1977). The *ICC* analysis revealed hospitals were equally reliable in identifying wound type, identifying PrUs (yes/no), staging PrUs, and determining PrU origin; *ICC* values = .19, .09, .16, and .21, respectively. This study will calculate *ICC*s in order to examine rater agreement across hospitals.

Bergquist-Beringer et al. (2011) expanded the previous study to evaluate the reliability of PrU staging by direct observation. The sample of convenience included 31 NDNQI participating hospitals. Two to ten raters per hospital (M = 5.3, Total = 180 raters) rounded on patients with a PrU. Collectively, 591 PrUs were independently staged from I to IV or classified as unstageable (M = 3.3 PrUs per rater). Among the raters at each hospital was one person who was the hospital expert on PrUs. After PrU staging by direct observation was completed, 162 of the 180 raters (90% response rate) completed the 3-part criterion-referenced web-based test used in the Hart et al. (2006) study. Each rater was randomly assigned to one of two versions of the online test; one with and one without the narrative descriptors.

The kappa agreement among raters from direct observation of the wounds was determined by comparing the ratings of each rater to the expert rater. For the online test, Cohen's kappa values were computed by comparing responses from each hospital rater to the "correct" response determined by experts. Finally, hierarchical linear modeling was used to examine the relationships between participant characteristics, hospital characteristics, and kappa values.

Pressure ulcer staging by *direct observation* revealed moderate reliability (k = 0.60, SD = 0.29). The average kappa for PrU staging from direct observation of PrUs was significantly higher (p = .027) for hospitals whose raters where led by a wound certified nurse (k = 0.68, SD =

0.68) compared to those teams without a certified nurse (k = 0.57, SD = 0.22). Overall agreement for PrU staging *from photos* was substantial (k = 0.69, SD = 0.20). The reliability of PrU staging was significantly higher ($p \leq .001$) for the test version with wound descriptors (k = 0.81, SD = 0.16) compared to the version without (k = 0.59, SD = 0.18). The overall kappa value for PrU identification (yes a PrU/not a PrU) using the online test was 0.83; and overall kappa for PrU origin (community- or hospital-acquired) was 0.79. There was no difference in kappa values between *certified* and *noncertified* nurses (p = .241) after adjusting for test version. Kappa values for *wound/skin care* nurses were higher relative to nurses not recognized as wound/skin care nurses, but results were not statistically significant (p = .063).

This study is similar to the Hart et al. (2006) study in that raters directly rated patients on the PrU risk and prevention measures. Also, a reliability coefficient, Cohen's kappa, was calculated in the same manner as the Bergquist-Beringer et al. (2011) and Hart et al. (2006) studies. In these studies, the expert's rating was the criterion and all other raters were compared to the expert.

Gajewski et al. (2007) investigated the inter-rater reliability of PrU staging by secondary analysis of data on wound staging from direct observation of these wounds (Bergquist-Beringer et al., 2011) using a subsample of 20 NDNQI participating hospitals. The overall purpose of Gajewski et al.'s (2007) study was to address the shortcomings of the traditional statistical approach typically used to estimate inter-rater reliability of PrU staging (a one-way Analysis of Variance random-effects model). In addition, this study focused on dealing with unstageable PrUs during statistical analysis. Specifically, whether or not unstageable PrUs should be classified as Stage III or Stage IV PrUs was considered because these ulcers are typically dropped from analysis. Other shortcomings addressed were that PrU stage is incorrectly treated

as continuous level data, when they are actually ordinal level data, and with small samples, within-hospital variance in PrU classification may be estimated as zero, leading to an *ICC* of 1.0, 95% CI [1.0 – 1.0], which is "scientifically unreasonable" (p. 4606).

Three to nine raters (*M* = 7) per hospital independently staged PrUs within their facility. Collectively, 347 PrU had been staged, 6 to 108 PrUs per hospital. Findings supported the inter-rater reliability of PrU staging from direct observation (mean *ICC* = 0.57), however the Bayesian method presented the most accurate reliability estimates when compared to traditional analysis.

These PrU incidence, staging, and origin studies (Gajewski et al., 2007; Bergquist-Beringer et al., 2013; and Hart et al., 2006) provide support for the reliability of the NDNQI outcome measure on PrU. Data on the reliability of PrU risk and prevention measures, however, is still needed to examine the reliability of the NDNQI pressure ulcer risk and prevention measures.

CHAPTER III: METHODOLOGY

The purpose of this study was to examine the reliability of the NDNQI® pressure ulcer risk and prevention measures. Although more than 1,400 hospitals collect the NDNQI pressure ulcer data, no studies have evaluated the reliability of this database's PrU risk and prevention measures. In Chapter III, the research design is presented, followed by a discussion of the setting and sample, study variables and their measures, and data collection procedures. Human Subject Committee review, data analysis, and the pilot study are described. Lastly, the study timeline is presented.

Research Design

A descriptive design was used to accomplish the aims of this study. The study is a multi-site, two-part, inter-rater reliability study. In Part 1, participants collected NDNQI PrU risk and prevention data from patient records. In Part 2, these same participants completed an online survey asking them about the methods and processes they use to rate patients on these measures. Participants completed Part 1 before beginning Part 2. Advantages of the study design include (a) data collection procedures reflected actual PrU data collection procedures that are performed in hospitals or clinical settings, (b) data collection from a large geographical area, (c) a multi-faceted approach of using chart reviews with data extraction, and online survey provided more insight than a single approach, (d) methods were useful for identifying future areas of research in areas not well studied, and (e) data were easily quantified in that they did not require complex statistical analysis (Grimes & Schultz, 2002; Polit & Beck, 2012). Disadvantages of this design include (a) causal relationships cannot be determined, (b) lack of control over study variables, and (c) the survey response options may not have accurately captured how respondents wanted to

respond (Grimes & Schultz, 2002; Polit & Beck, 2012). Study limitations are discussed in Chapter V.

Setting and Sample

Participation in the NDNQI is voluntary, and participating hospitals are not required to collect data on all the indicators, although they are encouraged to do so. While not all participating hospitals submit PrU data, most do: As of January 2014, more than 1,400 out of the 1,986 hospitals that participated in the NDNQI submitted PrU data (NDNQI, 2014b). Hospitals are encouraged to submit PrU data quarterly, but can elect to submit less frequently.

Sample of Hospitals

The NDNQI classifies hospital type as General Acute Care, Critical Access, Non-USA, and Specialized (NDNQI, 2013). In order to eliminate the bias that may occur if hospital types vary widely (among each other), only General Acute Care Hospitals were included in this study, which by definition are located in the U.S. A General Acute Care Hospital "primarily offers services for medical-surgical patients and may or may not also include services for obstetrics, pediatrics, rehabilitation and psychiatry" (NDNQI, 2013, p. 11).

Following usual NDNQI recruitment procedures for participation in research studies, 750 randomly selected General Acute Care Hospitals that (1) submitted PrU data at least two of the last four quarters, and (2) had at least two adult medical-surgical, medical, or surgical units, were invited to participate in the study. Considering the NDNQI has had hospital response rates of 10% to 15% for reliability studies (S. Bergquist-Beringer, personal communication, March 24, 2014), it was anticipated that approximately 75 to 112 hospitals would agree to participate.

Sample of Units and Patients

The NDNQI classifies unit type based on patient population characteristics (i.e. acuity level, age, and type of care provided). For this study, it was planned that data on PrU risk and prevention would be obtained from the records of 50 patients who were located on adult medical-surgical, adult medical, and /or adult surgical unit types in the hospital. Much of the NDNQI data on PrUs is submitted by these unit types. Of the 15,400 units that submitted PrU data for the 2nd Quarter of 2014, 7,471 (49%) units were classified as medical-surgical (n = 1,774), medical (n = 2,468), or surgical (n = 3,229) (NDNQI Statistical Analyst, personal communication, November 21, 2014). At least 90% of the patients on these units receive the level of care appropriate for the unit type (NDNQI, 2013).

Critical care units were excluded from the study for consistency in the type of documentation system from which data were retrieved. For example, critical care units sometimes have significantly different flow sheets for documentation than non-critical care units. Moreover, the rapid turnover of patients in critical care units could have confounded study procedures.

Study Variables and Measures

NDNQI Pressure Ulcer Risk and Prevention Variables and Measures

NDNQI pressure ulcer risk and skin assessment variables. There are 11 NDNQI PrU risk and prevention measures. These measures include (1) *Skin assessment within 24 hours of admission*, (2) *PrU risk assessment within 24 hours of admission*, (3) *Time since last PrU risk assessment*, (4) *Last PrU risk assessment scale and score*, (5) *PrU risk status*, (6) *PrU prevention use within the last 24 hours*, (7) *Skin assessment within the last 24 hours*, (8) *Pressure-redistribution surface use within the last 24 hours*, (9) *Routine repositioning as*

prescribed within the last 24 hours, (10) *Nutritional support within the last 24 hours*, and (11) *Moisture management within the last 24 hours* (NDNQI, 2013).

The NDNQI requires hospitals follow specific guidelines to collect data on PrUs (NDNQI, 2013). Specifically, data on PrU risk and prevention are collected by staff at each hospital who have received training in PrUs—the NDNQI Pressure Ulcer Survey Team. According to the NDNQI guidelines, the NDNQI Pressure Ulcer Survey Team should be led by a person who is certified in wound/skin care, or has received additional training in PrU identification and staging, and is knowledgeable of the NDNQI PrU data collection guidelines. Training for new PrU team members includes: (a) the NDNQI guidelines on PrU prevalence indicators; (b) skin assessment; (c) NPUAP pressure ulcer staging definitions; (d) PrU stage appearance; (e) other wound types; (f) ability to differentiate between community-, hospital-, and unit-acquired PrUs; and (g) data extraction from the patient record (Bergquist-Beringer & Davidson, 2014). Prior to each subsequent skin survey, team members should review data collection guidelines, skin assessment, and PrU identification and staging. Chart reviewers are members of the NDNQI Pressure Ulcer Survey Team who collect data on the PrU risk and prevention measures, and should demonstrate competence in the performance of skin and risk assessments and data abstraction from the patient record. Team members are supposed to be assigned to units other than the ones where they usually work in order to discourage biased reporting.

For this study, data on PrU risk and prevention were recorded by participant raters on the *Data Collection Form* (Appendix B). Each variable is discussed as follows, and the conceptual and operational definitions and level of analyses are listed in Table 5.

Table 5.

Major Study Variables: NDNQI Pressure Ulcer Risk and Prevention Measures

		Recorded by Participant Raters on the *Data Collection Form* (Part 1)	
	Conceptual Definition	Operational Definition	Response Options & Analysis Level
1. Pressure Ulcer Risk Assessment	Evaluation of the patient's risk for PrU development	**1. Pressure ulcer risk assessment documented w/in 24 hours of admission?**	Yes; No; Pending (admitted w/in the last 24 hours) Categorical
		2. How long ago was the last pressure ulcer risk assessment performed?	> 0 – 12 hours; > 12 – 24 hours; > 24 – 48 hours; > 48 – 72 hours; > 72 hours to 1 week; > 1 week; Never assessed for risk Categorical
		3. Last risk assessment scale & score?	Braden scale; Norton scale; Other; Clinical factors Categorical Last Score _____ Continuous
		4. Based on last assessment, is patient "at risk" for pressure ulcers?	Yes-based on risk assessment score; Yes-based on clinical factors; No Categorical

Table 5 (continued).

Major Study Variables: NDNQI Pressure Ulcer Risk and Prevention Measures

		Recorded by Participant Raters on the *Data Collection Form* (Part 1)	
	Conceptual Definition	Operational Definition	Response Options & Analysis Level
2. Skin Assessment	A head-to-toe evaluation of the the patient's skin focusing on bony prominences and other at risk areas.	**1. Skin assessment documented w/in 24 hours of admission?**	Yes; No; Pending (admitted w/in the last 24 hrs) Categorical
		2. Skin assessment documented (within the past 24 hours for "at risk" patient).	Yes; No; Documented contraindication Categorical
3. PrU Prevention Use	The performance of any PrU prevention strategy for "at risk" patients.	**Pressure ulcer prevention in use w/in the past 24 hrs for "at risk" patient?**	Yes; No; Pending (admitted w/in the last 24 hrs) Categorical
4. Pressure-Redistribution Surface Use	The use of a special support surface that redistributes pressure on skin and subcutaneous tissue.	**Pressure-redistribution surface use (within the past 24 hours for "at risk" patient).**	Yes; No; Documented contraindication; Unnecessary for pt.; Pt. refused Categorical
5. Repositioning	Routine repositioning as determined by patient characteristics	**Routine repositioning as prescribed (within the past 24 hours for "at risk" patient).**	Yes; No; Documented contraindication; Unnecessary for pt.; Pt. refused Categorical

Table 5 (continued).

Major Study Variables: NDNQI Pressure Ulcer Risk and Prevention Measures

	Recorded by Participant Raters on the *Data Collection Form* (Part 1)		
	Conceptual Definition	Operational Definition	Response Options & Analysis Level
6. Nutritional Support	Adequate nutritional intake.	**Nutritional support (within the past 24 hours for "at risk" patient).**	Yes; No; Documented contraindication; Unnecessary for pt.; Pt. refused Categorical
7. Moisture Management	Keeping the patient's skin clean and dry**.**	**Moisture management (within the past 24 hours for "at risk" patient).**	Yes; No; Documented contraindication; Unnecessary for pt.; Pt. refused Categorical

Pressure ulcer risk assessment. Pressure ulcer risk assessment is the evaluation of the patient's risk for PrU development. Use of a validated scale to determine risk status is recommended by AHRQ (AHRQ, 2011b) and the Wound Ostomy and Continence Nurses Society (WOCN) (Ratliff et al., 2010). The Institute for Healthcare Improvement (IHI) recommends PrU risk be assessed within 4 hours of admission, and reassessed at least daily (NDNQI, 2013). For the NDNQI, members of the PrU data collection team review each patient record to determine if a PrU risk assessment was documented within 24 hours of admission. "Pending" is selected if the patient was admitted within the last 24 hours and risk assessment is in process. According to 2010 NDNQI data, "Pending" was selected for only 0.7% of patients (Bergquist-Beringer et al., 2013). Documentation in the patient record is also reviewed to determine how long before the NDNQI PrU Survey the last risk assessment was performed. Risk assessments performed at the time of the survey should not be counted. The method used to assess PrU risk is identified as well as the score if risk was determined by the Braden, Norton, Braden Q, or Neonatal Skin Risk Assessment Scales. Risk status is then identified.

Skin assessment. Skin assessment is the head-to-toe evaluation of the patient's skin, focusing on bony prominences and other at risk areas of the body. The goal of the skin assessment is to assess skin integrity and early identification of PrUs. A skin assessment should be done on admission to detect and document any PrUs present on admission. Ideally the skin of those at risk for PrU is inspected frequently, such as every shift or each time the patient is turned (NDNQI, 2013). For the NDNQI, members of the PrU data collection team review patient records to determine if a skin assessment was documented within 24 hours of admission. "Pending" is selected if the patient was admitted within the last 24 hours and skin assessment is

in process. According to 2010 NDNQI data, "Pending" was selected for only 0.5% of patients (Bergquist-Beringer, 2011).

Documentation in the patient record is also reviewed to determine how long before the NDNQI PrU survey the last skin assessment was performed. If no documentation for skin assessment is found, the chart should be reviewed for documentation of a reason why a skin assessment is contraindicated. Documented contraindication can be any reason recorded in the patient record as to why a particular PrU prevention intervention is contraindicated, such as hemodynamic instability, spinal cord injury, or increased intracranial pressure.

NDNQI pressure ulcer prevention defined. For NDNQI, PrU prevention is defined as the performance of any PrU prevention intervention for "at risk" patients within the 24-hour period before the NDNQI Pressure Ulcer Survey (NDNQI, 2013). Prevention interventions include skin assessment, pressure redistribution surface use, routine repositioning, nutritional support, and moisture management.

Pressure ulcer prevention use. If the patient is considered to be "at risk", their patient record is reviewed for documentation of the performance of PrU prevention within the previous 24 hours. Documentation of only one PrU prevention intervention is sufficient to select "Yes" to the indicator of *Pressure ulcer prevention within the last 24 hours*. "Pending" is selected if the patient was admitted within the last 24 hours and intervention implementation is in process. If pending is selected, then no interventions will have been implemented. According to 2010 NDNQI data, "Pending" was selected for only 0.7% of patients (Bergquist-Beringer, 2011).

Pressure-redistribution surface use. Pressure-redistribution surface use is the use of a special support surface that redistributes pressure on skin and underlying subcutaneous tissue or other parts of the body exposed to pressure. These surfaces can be special mattresses or seat

cushions, such as overlay mattresses or surfaces with air, gel, water, or high density foam. Pressure-redistribution also includes pressure relief of heels, such as suspending heels off the bed with pillows (NDNQI, 2013). For the NDNQI, members of the PrU data collection team review each patient record to determine if any pressure-redistribution surface was in use within the 24 hours prior to PrU data collection. Direct observation of pressure-redistribution surface use is sufficient evidence of use—documentation of use in this case is not required if the patient is observed on the surface.

Routine repositioning. Routine repositioning reduces the duration and amount of pressure on tissues. Repositioning frequency should be determined by individual patient characteristics (e.g. degree of immobility, body mass index, age, disease process) and the support surface being used. The commonly accepted frequency of repositioning is every 2 hours while in bed; and every 1 hour while in a chair (NDNQI, 2015). Routine repositioning is determined by reviewing the patient's record for documentation of *prescribed* frequency of repositioning. Prescribed frequency is the frequency of repositioning prescribed in the patient's plan of care. For the NDNQI, members of the PrU data collection team review each patient record to determine if routine repositioning was documented for the 24 hours prior to PrU data collection.

Nutritional support. Adequate nutritional intake helps the skin tolerate pressure. The recommended intake to prevent PrUs is ≥ 30 kcal/kg of body weight per day, and 1.25 to 1.5 grams/kg/day of protein. Nutritional support may be achieved by providing assistance with meals, nutritional consults, nutritional supplements, enteral or parenteral nutrition, and adequate hydration (NDNQI, 2013). For the NDNQI, members of the PrU data collection team review each patient record to determine if nutritional support was provided within the 24 hours prior to PrU data collection. Dietary consults not yet performed are adequate documentation of

nutritional support only for a patient who has been admitted for less than 24 hours. Direct observation of the patient receiving parenteral or enteral feedings is sufficient evidence of nutritional support—documentation of nutritional support is not required if the patient is observed receiving parenteral or enteral feedings.

Moisture management. Because excessive moisture may lead to skin breakdown when exposed to pressure, keeping patient skin clean and dry prevents PrU development. Use of absorbent pads or moisture barrier creams; management of urinary incontinency, fecal incontinence, and draining wounds; are examples of moisture management (NDNQI, 2013). Evidence of moisture management includes documentation of such interventions in the patient's record. Incontinence products or barrier creams at the bedside are not adequate evidence of their use. For the NDNQI, members of the PrU data collection team review each patient record to determine if moisture management was documented within the 24 hours prior to PrU data collection.

Pressure ulcer risk and prevention data collection form. The student investigator developed a data collection form, (the *Data Collection Form*), that was derived from forms used by NDNQI to capture the ratings on PrU risk and prevention for its PrU indicator (Appendix B). The investigator-developed form captures these data in the same order and was worded like the data collection form typically used during the NDNQI PrU Skin Survey (NDNQI, 2013). The investigator-developed form also included the date and time of data collection, participant rater number and corresponding *Survey ID*, patient room/bed number, patient age and gender, and unit ID. As before, data were collected only from nursing units that the NDNQI has classified as an adult medical-surgical, adult medical, or adult surgical unit.

The *Data Collection Form* was created in Excel. Prior to the study, face validity was established by two NDNQI staff members. Feasibility of use was determined during pilot study work by the participant raters and the skin survey team leader at the pilot site (see Pilot Study).

The *Pressure Ulcer Risk and Prevention Reliability Survey*

An online survey was developed by this student researcher to collect data about participant rater characteristics and information about the methods and processes used to rate the NDNQI pressure ulcer risk and prevention measures (Appendix C). This survey, the *Pressure Ulcer Risk and Prevention Reliability Survey*, is a survey that was formatted into *REDCap* (http://project-redcap.org/). *REDCap* allows for forward and backward navigation. Alerts were in place to notify participants of missing or incomplete answers, but allowed respondents to submit incomplete surveys. The survey could be accessed from any computer with internet access. The University of Kansas Medical Center supports the use of *REDCap*, and information collected on *REDCap* is considered secure.

Survey items #1 - #8 – Participant rater characteristics. The first eight items of the survey collected self-reported participant rater characteristics. Specifically, there are eight forced-choice questions on participant rater characteristics: (1) job title, (2) highest level of education, (3) review of the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers, and/or completion of the NDNQI Pressure Ulcer Training program, (4) education on PrU data collection other than the NDNQI Guidelines for Data Collection and Submission on Pressure Ulcers or the NDNQI Pressure Ulcer Training program, (5) years collecting NDNQI PrU data, (6) other role(s) in the NDNQI PrU survey, (7) who leads the PrU team, and (8) certifications. These characteristics reflect those collected by NDNQI in the past. Each participant rater characteristic is discussed below (Table 6) and the conceptual

Table 6.

Participant Rater Characteristics of Interest

	Self—Reported on the *Pressure Ulcer Risk and Prevention Reliability Survey*	
Conceptual Definition	Operational Definition	Response Options and Analysis Level
1. Job title that best reflects the position participant raters were hired for and are paid to perform at this hospital.	**What is your job title? (Select one.)**	Staff Nurse; Clinical Nurse Specialist; Advanced Practice Nurse; Nurse Manager; Nursing Administrator; Quality Improvement; Wound/Skin care Nurse; NDNQI Site Coordinator; Other (Please describe) Categorical
2. Highest level of education obtained.	**RN Nursing Education? (Select one.)**	Yes: Associate's Degree Nursing; Bachelor's Degree Nursing; Master's Degree Nursing; Doctorate Nursing No: High School Graduate/GED; Associate's Degree non-nursing; Bachelor's Degree non-nursing; Master's Degree non-nursing; Doctorate non-nursing] Categorical
3. Education on data collection for the NDNQI PrU risk and prevention measures.	**Please select what you have completed within the last 12 months. (Select all that apply.)**	I have reviewed the NDNQI Guidelines for Data Collection and Submission on Pressure Ulcers; I have completed all 4 modules of the NDNQI Pressure Ulcer Training program; I have completed some (not all) of the 4 NDNQI Pressure Ulcer Training program modules; None of the above Categorical

Table 6 (continued).

Participant Rater Characteristics of Interest

	Self—Reported on the *Pressure Ulcer Risk and Prevention Reliability Survey*	
Conceptual Definition	Operational Definition	Response Options and Analysis Level
4. Any training for data collection on PrUs that the participant rater received other than review of the NDNQI Guidelines for Data Collection and Submission or the NDNQI PrU Training program.	**Have you received education for data collection in pressure ulcers other than reviewing the NDNQI Guidelines for Data Collection and Submission on Pressure Ulcers, or completing the NDNQI Pressure Ulcer Training Program? (Select one.)**	Yes; No Dichotomous
5. Years collecting NDNQI PrU data reflects experience in PrU data collection.	**How many years have you collected NDNQI pressure ulcer data? (Select one.)**	< 1 year; 1 year, 2 years; 3 years, 4 years, 5 years, 6 years, 7 years, 8 years, 9 years, 10 years, 11 years, 12 years, 13 years, 14 years, 15 years, > 15 years Categorical
6. NDNQI PrU data collection role(s) reflects experience/expertise in PrU data collection.	**In addition to the chart abstractor, what other role(s) in the NDNQI Pressure Ulcer Survey have you had? (Select all that apply.)**	Site Coordinator; Patient skin inspection—rounding on all patients; Patient skin inspection—rounding on selected patients to confirm pressure ulcer presence or stage; Training of the pressure ulcer team; Data entry; None; Other (please describe) Categorical

Table 6 (continued).

Participant Rater Characteristics of Interest

	Self—Reported on the *Pressure Ulcer Risk and Prevention Reliability Survey*	
Conceptual Definition	Operational Definition	Response Options and Analysis Level
7. PrU team leader certification status reflects education and/or skill in wound/skin care.	**Who usually leads your NDNQI Pressure Ulcer Survey data collection team? (Select one.)**	Someone certified in wound care; Someone who is not certified in wound care but is the wound/skin care nurse; Neither of the above I don't know their certification status Categorical
8. Certification reflects specialized education and/or skill in wound/skin care.	**Which of the following active certification do you hold? (Select all that apply.)**	CWOCN--Certified Wound, Ostomy, Continence Nurse CWCN—Certified Wound Care Nurse COCN—Certified Ostomy Care Nurse CCCN—Certified Continence Care Nurse CWON—Certified Wound Ostomy Nurse CWS—Certified Wound Specialist WCC—Wound Care Certified No certification in wound care Other (please describe) Categorical

and operational definitions and level of analyses are listed.

Job title reflects the position participants fulfill, are paid for, and that they were hired to perform. Because some participant raters may not be nurses, highest education level included nursing and non-nursing degrees, as well as high school graduate/GED. The NDNQI Guidelines for Data Collection and Submission on Pressure Ulcers provides specific instructions to participating hospitals on how to collect data for the PrU measures (NDNQI, 2015). Participant raters were asked if they reviewed these guidelines within the last 12 months. The NDNQI Pressure Ulcer Training program has been shown to be an effective and easily accessible education method (Bergquist-Beringer et al., 2009). This online program has four modules that provide a comprehensive overview of PrU staging and origin, other wound types, PrU prevention, and guidelines to data collection (Bergquist-Beringer & Davidson, 2014).

Participant raters were asked if they completed all or some of the four modules of the NDNQI Pressure Ulcer Training program. Ongoing pressure ulcer training and review helps ensure data integrity. Participant raters were asked if they received any education for data collection on PrUs other than reviewing the NDNQI Guidelines for Data Collections and Submission on Pressure Ulcers or completing the NDNQI Pressure Ulcer Training program. The NDNQI data collection guidelines state team members should receive PrU training or review prior to each skin survey (NDNQI, 2015).

Years as an NDNQI pressure ulcer chart abstractor and other roles in the NDNQI Pressure Ulcer Survey helped capture the level of experience participant raters have in NDNQI PrU data collection. Participant raters were asked how many years that they have been collecting NDNQI pressure ulcer risk and prevention data, and other roles they have had in NDNQI PrU data collection. Wound and skin care expertise identifies participant raters who have advanced

knowledge and expertise in the area of wound and skin care management. Participant raters were asked about current certifications. Pressure ulcer data collection team characteristics were the wound/skin care certification status and the wound/skin care expertise of the leader of the PrU data collection team.

Survey items #9 - #35 – Methods used to collect NDNQI pressure ulcer data. The remaining survey items included 27 Likert-type questions (categorical-level) about the methods and processes used to gather information needed in order to rate the NDNQI PrU risk and prevention measures. This information is helpful to better understand how data are collected, and if reported methods reflect the NDNQI data collection guidelines. Prior to the study, face validity of the survey was established by two NDNQI staff members. Feasibility was verified by the participant raters and the skin survey team leader at the pilot site during pilot work (see Pilot Study).

Hospital Characteristics

Hospital characteristics of interest to this study included hospital bed size, Magnet status, teaching status, and metropolitan status. Hospitals self-classify this information using the NDNQI Hospital Demographic Summary (NDNQI, 2013). Characteristics are determined before hospitals are enrolled in the NDNQI and verified by the site coordinator each quarter. Although no validity studies have been conducted on these hospital variables, quarterly verification lends support to their accuracy. Information on these hospital characteristics were provided by the NDNQI staff for the 2nd Quarter of 2014, which is the quarter before data collection took place.

Hospital size. Hospital size is based on the number of staffed beds (AHRQ, 2011a), and is categorized as < 50, 50 – 74, 75 – 99, 100 – 199, 200 – 299, 300 – 399, 400 – 499, and 500 or

more beds. Staffed beds include those that are occupied and those that are vacant and available, regardless of if they are included in NDNQI reporting (NDNQI, 2013).

Magnet status. Magnet status is recognition given to hospitals by the American Nurses Credentialing Center (ANCC), and signifies the hospital has achieved excellence in nursing (American Nurses Credentialing Center, 2013). The NDNQI categorizes Magnet status as (1) have applied, (2) intend to apply, (3) Magnet recognition, (4) no plans to apply, and (5) unsuccessful application. However, for this study, Magnet status available to the student researcher included "Magnet" or "Non-Magnet". Magnet recognition is verified by the NDNQI staff using information available on the ANCC website. Hospitals that have applied for Magnet status but not yet received the result verify this information by sending the NDNQI (via facsimile) their ANCC application letter.

Hospital teaching status. Hospitals can identify themselves as an academic medical center, a teaching hospital, or a non-teaching hospital. An academic medical center is the primary clinical site for a medical school; a teaching hospital has medical interns or residents; and a non-teaching hospital does not have interns or residents (NDNQI, 2013).

Metropolitan status. Metropolitan status is based on the U.S. Census Bureau's definitions (United States Census Bureau, 2013) and refers to the location of the hospital in a metropolitan area, a micropolitan area, or neither a metropolitan area or micropolitan area. A metropolitan area is a single county or group of adjacent counties that has a core urban population of at least 50,000 residents. A micropolitan area is a county or group of adjacent counties that has a core urban population of greater than 10,000 but less than 50,000 residents. A non-metropolitan area is defined as a county that is not metropolitan or micropolitan (NDNQI, 2014a).

Procedures

Recruitment and Enrollment

Before the study, approval by the Human Subjects Committee (HSC) at the University of Kansas Medical Center (KUMC) was obtained. On August 1, 2014, 750 randomly selected General Acute Care Hospitals that submitted PrU data at least two of the last four quarters, and had at least two adult medical-surgical, medical, or surgical units; were invited to participate in the study via email (*Invitation to Participate* presented in Appendix D). The study's original recruitment procedure was to exclude hospitals with < 100 staffed beds. This exclusion was inadvertently not put in place during random selection of the hospitals. Subsequently, hospitals with < 100 staffed beds were sent an *Invitation to Participate*. This email invitation was reviewed by the NDNQI staff, and was sent from NDNQI. The invitation described the study, identified each hospital's eligible units, and included a link to a *REDCap* survey for site coordinator response. Site coordinators (n = 124) replied to the *Invitation to Participate* by accessing the *Reply to Invitation to Participate* (Yes; No; or I'm not sure, I need more information) and provided their NDNQI hospital code and name and email address (Appendix E). The deadline to reply to the invitation was August 19, 2014. However, a few site coordinators asked to participate after the deadline, and the survey was reopened until August 25.

Before Data Collection

Study researchers mainly communicated with each hospital through the hospital's NDNQI site coordinator. On occasion, hospital personnel other than the site coordinator (but still persons who were involved in the study) initiated and established communication with study researchers. Each site coordinator, however, served as the contact person, distributed study-related materials, and managed the study at their hospital. This reflects usual NDNQI practice in

that site coordinators are the primary point of contact between the hospital and the NDNQI (NDNQI, 2013). Site coordinators are responsible for quarterly NDNQI data collection and submission. Site coordinators are also those with whom NDNQI communicates for optional research projects.

Study materials. Site coordinators of hospitals who indicated their interest in study participation received study materials via email on August 27, 2014. These materials included: an *Overview of the Pressure Ulcer Risk and Prevention Reliability Study,* (Appendix F), *Site Coordinator Instructions* (Appendix G), *Participant Rater Instructions* (Appendix H), the *Data Collection Form* (Appendix B), list of their eligible units (*Unit List*), and the *Teleconference Instructions* (Appendix I). The email that accompanied the study materials is presented in Appendix I. The *Overview of the Pressure Ulcer Risk and Prevention Reliability Study* was distributed to site coordinators and participant raters, and provided both with an overview of the study procedures.

The *Site Coordinator Instructions* provided step-by-step instructions for the study including selection of the participant raters, data collection procedures, access to and completion of the online survey, and data submission. The *Participant Rater Instructions* provided step-by-step directions for the study including collecting data (Part 1 and Part 2), accessing and completing the online survey, and submitting the completed *Data Collection Form*.

The *Data Collection Form* was used by participant raters to record data on patient PrU risk and prevention measures. Prior to the study, this Excel file form was reviewed and approved by NDNQI staff.

A list of each hospital's eligible medical-surgical, medical, and surgical units from which study data were to be collected, (the *Unit List*), was provided as a separate document among

initial study materials. These units were listed in the order from which data were to be collected. The *Unit List* included unit name, and each unit's corresponding unit type and unit identification number.

The student researcher hosted a 1-hour teleconference session on September 9, 2014 for site coordinators and participant raters from hospitals who agreed to participate in the study. The purpose of the teleconference was to explain the study and data collection procedures, and answer questions. The *Teleconference Instructions* provided information about the teleconference's time and date, how to attend, and how to obtain an audio-recording of the teleconference. During the conference, the study instructions were reviewed and attendees were able to ask (and have answered) their questions. One week before the teleconference, site coordinators received an email from the student researcher reminding them of the teleconference date, time, and access information. The audio-recording was emailed to those who requested it. Thereafter, the student researcher was available within NDNQI office space at routine intervals to answer pre-study and during-study questions.

Participant raters. It was planned that each site coordinator would identify one expert and two or three non-expert raters (participant raters). At one hospital, however, one expert and *four* non-experts were identified to collect data. Site coordinators selected participant raters from their hospital's established PrU data collection team, i.e. team of staff who usually collect these data for the quarterly NDNQI Pressure Ulcer Survey. Site coordinators first selected the expert rater, defined as the PrU team member with the most experience and/or skill in patient record review for the quarterly PrU survey. Site coordinators then selected non-expert raters, defined as any PrU team members (other than the identified expert) who usually review patient records. The number of expert and non-expert raters was selected because it is reasonable to expect that

hospitals would have at least three persons who usually review patient records. The number of two *or three* non-expert raters was selected to better capture variability between the expert and non-expert ratings than if there were only two non-expert raters per hospital.

Site coordinators distributed the *Participant Rater Instructions* (Appendix H) and a print-copy the *Data Collection Form* (Appendix B) to each participant rater. The site coordinator assigned each participant rater a rater number. The site coordinator recorded this number on each *Data Collection Form* prior to distributing the form to the participant rater. The "expert" was "Rater 1", and the non-experts were "Rater 2" and "Rater 3". If a third non-expert rater was available, this rater was assigned to be "Rater 4". The hospital with a fourth non-expert rater identified this rater as "Rater 5". Each rater number had a corresponding *Survey ID.* The *Survey ID* is a unique anonymous code comprised of the hospital's NDNQI alphanumeric identifier code (hospital code) followed by the number "1", "2", "3", or "4". Specifically, the expert ("Rater 1") was assigned the *Survey ID* that was their hospital code followed by the number "1". Each of the non-experts was assigned a *Survey ID* that was their hospital code followed by the number "2" or "3". If a third non-expert was available, their *Survey ID* was their hospital code followed by the number "4". The *Survey ID* for the fourth non-expert rater (from the single hospital with four non-experts) was their hospital code followed by the number "5". Site coordinators also distributed the *Participant Rater Instructions* (Appendix H) to each participant rater.

Study day. It was planned that the data collection period for the study would be September 29 to October 13, 2014. The site coordinator and participant raters chose what day during the study data collection period they would conduct the study. Specifically, the site coordinator and participant raters were to agree on a single day within this period to collect the study data. The goal across hospitals was to have data *collected* within 14 days of the study

period start, and have data *submitted* within 14 days after the data collection period ended. However, the data *collection* period was extended to the data *submission* deadline of October 27, 2014 in order to accommodate hospital schedules.

Participant raters were instructed to collect the data on the study day during a 5-hour time period beginning at 7:00 a.m. and ending at 12 noon. They were directed to review the patient record for care provided in the 24 hour period between 7:00 a.m. the day before data collection to 7:00 a.m. the day of data collection. Although participant raters collected data on the same morning, they were instructed to collect data independently.

On the morning of data collection, each participant rater was to receive a list of patients (from the site coordinator) that included patient name and the room/bed numbers of those who were on each eligible unit at 7:00 a.m. (the *Patient List*). Participant raters were asked to collect PrU risk and PrU prevention data only for patients on this list that were still in the room/bed number at the time of data collection. Participant raters were directed not to seek or find the patients who had moved.

Data Collection

Part 1. Participant raters were asked to *independently review* patient records and *independently rate* (the same) 50 patients on the NDNQI pressure ulcer risk and prevention measures. They were asked not to collect data together, not to ask others for help, and not to share or compare their data with the others after collecting it. A 15 to 30 minute staggered data collection start time was suggested in the *Site Coordinator Instructions* and the *Participant Rater Instructions* to minimize the risk of participant raters sharing information with each other. Participant raters were to begin data collection on the first unit identified in the *Unit List*, and

then proceed by unit in the order listed. Units were listed starting with adult medical-surgical unit(s), then adult medical unit(s), and lastly adult surgical unit(s).

Working from the 7:00 a.m. *Patient Lists* generated for each unit, participant raters were to start collecting patient data from the patient in the lowest room/bed number on the first unit in the *Unit List*. For example, if the first unit on the *Unit List* was 6-East, data collection was to begin on 6-East; and the *Patient List* for 6-East was to be used to guide data collection on this unit. Data collection was to continue in chronological order to the patient in the highest room/bed number on that unit. Next, participant raters were to move to the next unit on the *Unit List* and collect data in the same manner; starting with the patient in the lowest room/bed number on that unit, and continuing in chronological order to the patient in the highest room/bed number. Data collection was to continue in this manner until the records of 50 patients were rated. However, some smaller hospitals reported that they may not have 50 patients. These hospitals were instructed to choose their data collection day to be a day when they expected to have the highest patient census, and then collect data on however many patients were available.

If the patient in the room/bed was different than the patient reported to be in that room/bed (according to the 7:00 a.m. *Patient List*), then data were not to be collected from this patient. This was necessary to ensure that participant raters collected data on the same patients, and because some NDNQI PrU measures may be rated by direct observation. This is consistent with usual NDNQI PrU data collection in that data are not collected on discharged patients. While collecting data, participant raters were to include only patient data for the 24-hour period beginning 7:00 a.m. the day *before* data collection to 7:00 a.m. the day *of* data collection. It was not necessary for the patient to have been in that room/bed for the entire 24-hour period.

Participant raters used the *Data Collection Form* (Appendix B) to: (a) record the date and time data collection began; (b) record each patient's room/bed number, age, and gender; and (c) record the unit ID. As data were collected, participant raters could enter patient data onto their *print-copy* of the *Data Collection Form*, or enter data directly into their *electronic* version of the *Data Collection Form*. Those who recorded data onto a print-copy then transcribed their data into their electronic *Data Collection Form* after data collection was completed. Participant raters were to personally enter the data they collected into the electronic *Data Collection Form* in order to minimize the risk of results being changed to reflect agreement among participant raters. Participant raters were to check the accuracy of their data entry.

Part 2: Online *REDCap* survey. After data collection, participant raters logged into *REDCap* (http://project-redcap.org/) in order to (a) upload their completed *Data Collection Form* Excel file (Appendix B), and (b) complete the *Pressure Ulcer Risk and Prevention Reliability Survey* (Appendix C). Only participant raters whose patient data were included in kappa value calculations, (Part I of the study), were included in the analysis of survey data ($N = 120$ participant raters). The *Participant Rater Instructions* (Appendix H) provided participant raters with the link to the *REDCap* survey. The deadline for uploading completed *Data Collection Forms,* and completing the survey was October 27, 2014. However, participant raters from three hospitals submitted their *Data Collection Forms* and completed the survey after the deadline (October 28 through October 31).

Access to the online survey was by a common URL address. The *Data Collection Form* Excel file upload process was presented as a survey item, with a direct upload (single click) link. *REDCap* provided step-by-step instructions to complete the file upload. In addition, *REDCap* automatically tracked and recorded user activities, and did not allow for changes to the survey

after data collection began. In essence, *REDCap* kept its own audit trail, code key, and data log. Site coordinators were responsible to make sure all *Data Collection Forms* were submitted, and to encourage participant raters to complete the online survey. The student researcher sent emails to site coordinators notifying them of the status of their participant raters' data submission.

NDNQI personnel extracted the hospital characteristics of interest (size, Magnet status, teaching status, and metropolitan status) for the 2nd Quarter of 2014: the quarter before the study data collection period. The student researcher received this information from NDNQI.

Human Subject Review

Approval for NDNQI activities was granted by the KUMC Institutional Review Board (IRB). For this study (including the pilot study), approval was obtained from the Human Subjects Committee (HSC) at KUMC. Specifically, a change in protocol to the NDNQI study was submitted. The student researcher completed KUMC Human Subjects Protection training and Health Insurance Portability and Accountability Act (HIPAA) training, signed a conflict-of-interest form, completed the NDNQI Confidentiality Form, and was added as student personnel to the NDNQI study. The student conducted the study within assigned NDNQI office space.

Data on hospitals within the NDNQI dataset received for this study were de-identified. Patient data recorded on the *Data Collection Form* by participant raters was de-identified. *REDCap* is a secure web-based application that provided a secure link to the survey. Respondent anonymity was maintained by a de-identified number that was assigned to them. Only members of the investigation team had access to the responses. De-identified datasets were stored on a password-protected computer or secured KUMC drive. All data are presented without identifiers.

Data Analysis

Preparation of the Data for Analysis

Hospital characteristics were extracted by NDNQI staff for the 2nd Quarter of 2014: the quarter before the study data collection period. These data were sent to the student researcher who transferred the data into the Statistical Package for the Social Services (SPSS) version 22 for analysis. After participant raters uploaded their completed *Data Collection Form* onto *REDCap*, the student researcher also transferred these files into (SPSS). Because each hospital had one expert and two, three, or four non-expert participant raters; three, four, or five *Data Collection Forms* were submitted per hospital. Data from the online survey (*Pressure Ulcer Risk and Prevention Reliability Survey*) was also retrieved and electronically transferred into SPSS. All KUMC email is considered "secure" within KUMC, including *REDCap*. The student researcher maintained a code book, data log, and audit trail throughout data analysis.

Data cleaning and data exploration. Data collected from patient record review (the *Data Collection Forms*), and data collected on *REDCap* (*Pressure Ulcer Risk and Prevention Reliability Survey*) was examined for errors. Specifically, data entry errors on the *Data Collection Forms* (i.e. patient data) were identified and site coordinators were contacted as needed. Most *Data Collection Forms* were submitted to and retrieved from *REDCap* the day patient data were collected. This meant that site coordinators were contacted regarding data entry errors soon after data were collected. Most of the time, therefore, the student researcher was successful in resolving data entry errors.

Review of data missing from the *Data Collection Forms* suggested the patient was transferred, discharged, or died between the time the list was generated and data collection was completed by the three (four, or five) participant raters. Data on patient age and gender were

examined by hospital to make sure that patient room/bed number, age and gender matched across participant raters within that hospital. Only patients who had data collected on them by all of the participant raters at that hospital were included in the analysis. Data were also examined to identify outliers. The only outlier identified was a patient age of "5 years", and this patient's data was not included in the analysis because it did not match patient age recorded by the other participant raters.

The data from *The Pressure Ulcer Risk and Prevention Reliability Survey* were examined for missing data. The number of missing cases is presented by item in the survey data analysis results.

Descriptive statistics. Descriptive statistical analysis was performed to summarize **patients** by age and gender, the **hospitals** (size, Magnet status, teaching status, and metropolitan status), and the **units** (type). For continuous level data (patient age), the mean and standard deviation (SD) were calculated. For categorical level data (all other variables), frequencies and percentages were obtained.

Descriptive statistics were also used to describe the **participant raters** (expert status, job title, highest level of education, review of the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers within the last 12 months, completion of the NDNQI Pressure Ulcer Training program within the last 12 months, education for data collection on PrUs other than review of the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers or completion of the NDNQI Pressure Ulcer Training program, years collecting NDNQI PrU data, other role(s) in the NDNQI PrU survey, certification status of the PrU data collection team leader, and certifications). Frequencies and percentages were obtained. Responses to the option "Other (please specify)", were grouped into conceptually similar categories. These items and

their item number were (#1) "What is your job title?" (#6) "In addition to chart abstractor, what other role(s) in the NDNQI Pressure Ulcer Survey do you hold?" and (#8) "Which of the following active certifications do you hold?" The frequencies and percentages of responses for each of these categories were determined.

Data Analysis by Study Aim and Research Question

Aim 1: Examine the reliability of the NDNQI pressure ulcer risk and prevention measures *within* and *across* NDNQI hospitals. For this aim, Cohen's kappa, prevalence-adjusted kappa (PAK), percent agreement, intraclass correlation coefficient (*ICC*), and agreement matrices/descriptive statistics were performed.

Question 1: **What is the agreement between expert participant ratings and non-expert participant ratings *for each* of the 11 NDNQI pressure ulcer risk and prevention measures *within* each hospital?** Agreement was assessed by Cohen's kappa, PAK, and percent agreement in order to obtain one Cohen's kappa value for each of the 11 NDNQI risk and prevention measures per hospital, one PAK value for each of the 11 NDNQI risk and prevention measures per hospital, and a single percent agreement value for each of the 11 NDNQI risk and prevention measures per hospital. It was planned that SPSS would be used to compute Cohen's kappa values, and that Excel would be used to calculate PAK values. However, two issues became apparent: (1) the frequency tables were very sparse (i.e. many cells had very small counts or values of zero), and (2) many participant raters' ratings were considered constants (i.e. many participant raters rated all 50 patients as having received the intervention). These two issues meant that neither SPSS nor Excel would compute a large percentage of Cohen's *k* values. In order to produce Cohen's *k* values, it was necessary to add 0.0001 to each cell. This

adjustment was made to *every cell* regardless of its value, and it was kept in place during the PAK calculations. Therefore, Excel was used to calculate Cohen's kappa (and PAK) values.

A frequency table was formulated in Excel to calculate all Cohen's *k* and PAK values. Percent agreement was calculated by hand from each frequency table. Accuracy of this Excel formula table was verified by hand calculations of Cohen's *k* and PAK values for use as a 3 x 3 table, 4 x 4 table, 5 x 5 table, and 7 x 7 table (i.e. all possible frequency tables that were used in this study).

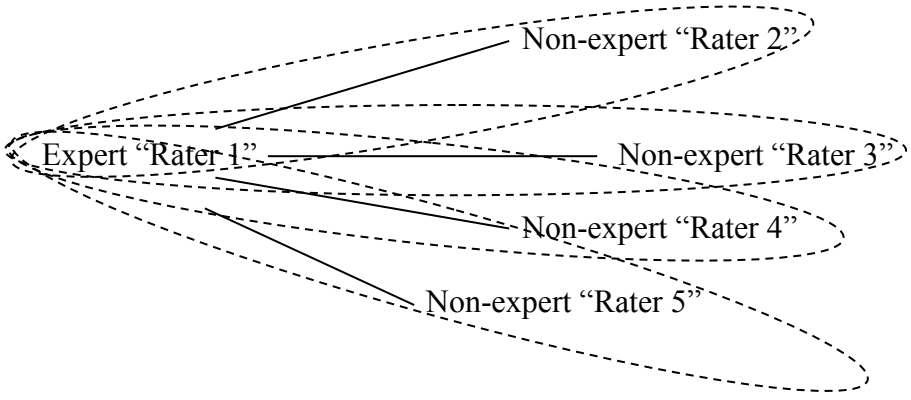
When SPSS is used to calculate Cohen's kappa, a frequency table representing agreement (and disagreement) is constructed and presented in the output. Each of the 11 measures had 84 expert to non-expert pairs, across 36 hospitals. Therefore, for each measure, 84 frequency tables were constructed (in SPSS) representing agreement (and disagreement) between *each* expert to non-expert pair (Total = 924 frequency tables). Cell counts from each frequency table (plus 0.0001) were then entered (by this student researcher), into an Excel frequency formula table. For each table, Excel calculated a Cohen's *k* value and a PAK value for that expert/non-expert pair. In other words, for each measure, 84 Cohen's kappa values were calculated to estimate the agreement between *each* of the 84 expert to non-expert pairs, and 84 PAK values were calculated to estimate the agreement between *each* of the 84 expert to non-expert pairs. These *calculations* were completed in Excel, but *the cell counts* (that were used for these calculations) were identified in SPSS.

Risk for error while entering the cell counts into the Excel frequency table was minimized by: (1) reconciling the marginal totals as identified in the SPSS table with the marginal totals identified in Excel table; (2) ensuring the Cohen's *k* value calculated in the Excel table equaled the Cohen's *k* value calculated in SPSS, when available; and (3) verifying that,

within each hospital, the expert's margin totals were identical across all expert/non-expert pairs for that measure (because the expert's ratings for each measure should be the same across all expert/non-expert pairs within that hospital). No weighting of Cohen's *k* or PAK values was performed. This was because the measures had categorical-level response options, and the loss of information that would have resulted from weighting kappa.

Cohen's kappa. For every hospital, ratings of each NDNQI PrU *risk* measure (*Skin assessment on admission*, *Risk assessment on admission*, *Time since last risk assessment*, *Risk assessment scale*, and *Risk status*) for the 23 – 50 patients, made by the 2 – 4 non-expert raters; were compared to the expert's ratings. Similarly, (for every hospital), ratings of each NDNQI PrU *prevention* measure (*Any PrU prevention*, *Skin assessment*, *Pressure-redistribution surface use*, *Routine repositioning*, *Nutritional support*, and *Moisture management*) for the 5 – 37 "at risk" patients, made by the 2 – 4 non-expert raters; were compared to the expert's ratings. In other words, at each hospital, there was two, three, or four expert to non-expert comparisons (kappa values) per patient (Figure 2) for each measure.

Figure 2. Expert to Non-Expert Pairs



One hundred twenty raters (36 "experts" and 84 "non-experts", *M* = 3.3 raters/hospital) rated a total of 1,637 patients (*M* = 45.5 patients/hospital, SD = 6.4). Of these 1,637 patients, 553 (33.8%) were identified by the "expert" to be "at risk" for PrU. However, not all non-experts

agreed with the expert on risk status, and 528 to 530 patients were rated on each of the prevention measures (5 – 37 "at risk" patients per hospital, $M = 14.7$, SD = 7.0). To answer Question 1, the 2 – 4 kappa coefficients for each measure were averaged to yield a single kappa value for each risk and prevention measure per hospital. Corresponding standard deviations were computed for each hospital's per measure kappa value.

Comparing pairs of expert/non-expert agreement evaluated the "correctness" (or validity) of the ratings; it is assumed the expert correctly rated the study PrU risk and prevention measures. The equation for Cohen's kappa is (Sim & Wright, 2005):

$$k = \frac{\text{(observed agreement)} - \text{(chance agreement)}}{1 - \text{(chance agreement)}}$$

Table 7 below illustrates how observed agreement and chance agreement were calculated for *Routine repositioning as prescribed.* The response options for *Routine repositioning as prescribed* include "Yes", "No", "Documented contraindication", "Unnecessary for pt.", and "Pt. refused".

Observed agreement (P_o) is:

$$P_o = \frac{(a + f + l + r + x)}{n}$$

Chance agreement (P_e) is:

$$P_e = \frac{\left(\frac{f1 \; x \; g1}{n}\right) + \left(\frac{f2 \; x \; g2}{n}\right) + \left(\frac{f3 \; x \; g3}{n}\right) + \left(\frac{f4 \; x \; g4}{n}\right) + \left(\frac{f5 \; x \; g5}{n}\right)}{n}$$

Table 7.

Illustration for Calculating Observed Agreement and Expected Agreement

		Non-Expert Rater					
		Yes	No	Contrain.	Unnec.	Pt. Refused	
Expert Rater	Yes	*a*	*b*	*c*	*d*	*e*	*g1*
	No	*e*	*f*	*g*	*h*	*i*	*g2*
	Contraindicated	*j*	*k*	*l*	*m*	*n*	*g3*
	Unnecessary	*o*	*p*	*q*	*r*	*s*	*g4*
	Pt. Refused	*t*	*u*	*v*	*w*	*x*	*g5*
		f1	*f2*	*f3*	*f4*	*f5*	***n***

Note. Contrain. = Contraindicated; Unnec. and Unnecessary = Unnecessary for the patient; Pt. = Patient

The main diagonal cells (in bold outline) represent agreement, and the off-diagonal cells represent disagreement. Landis and Koch's (1977) classifications (< 0 = "poor", 0 – .20 = "slight", .21 – .40 = "fair", .41 – .60 = "moderate", .61 – .80 = "substantial", and .81 – 1.0 = "almost perfect") were used to categorize kappa values' strength of agreement. For this study, the recommended reliability level for the study measures is ≥ .610.

Prevalence-adjusted kappa (PAK). For the NDNQI PrU risk and prevention measures, a large proportion of patients receive appropriate PrU risk and prevention interventions (e.g. 98.4% of "at risk" patients received a skin assessment within 24 hours of admission in Bergquist-Beringer, 2011). Therefore, the proportion of agreement by chance was expected to be large (i.e. high prevalence of "Yes" responses were expected). Cohen's kappa values decrease as chance for agreement increases.

In order to evaluate the effect of prevalence on the Cohen's kappa values, prevalence-adjusted kappa (PAK) values were computed. High (or low) prevalence *decreases* estimated kappa values (Sim & Wright, 2005). Prevalence-adjusted kappa values were calculated from average cell counts of agreement (mean of cell counts on the diagonal). Table 8 below illustrates how cell counts were calculated for **PAK** from cell counts for *Routine repositioning as prescribed*.

Table 8.

*Illustration of Cell Counts used to calculate **PAK** from Actual Cell Counts*

		Non-Expert Rater					
		Yes	No	Contrain.	Unnec.	Pt. Refused	
Expert Rater	Yes	*a* *100.0001* **24.0001**	*b* *10.0001*	*c* *0.0001*	*d* *10.0001*	*e* *2.0001*	*g1= 122.0005* **46.0005**
	No	*e* *10.0001*	*f* *5.0001* **24.0001**	*g* *5.0001*	*h* *0.0001*	*i* 0.0001	*g2= 20.0005* **39.0005**
	Contraindicated	*j* *0.0001*	*k* *5.0001*	*l* *10.0001* **24.0001**	*m* *10.0001*	*n* *3.0001*	*g3= 28.0005* ***42.0005***
	Unnecessary	*o* *8.0001*	*p* *2.0001*	*q* *6.0001*	*r* *5.0001* **24.0001**	*s* *0.0001*	*g4= 21.0005* **40.0005**
	Pt. Refused	*t* *2.0001*	*u* *3.0001*	*v* *4.0001*	*w* *0.0001*	*x* *0.0001* **24.0001**	*g5= 9.0005* **33.0005**
		f1=120.0005 **44.0005**	*f2=25.0005* **44.0005**	*f3=25.0005* **39.0005**	*f4=25.0005* **44.0005**	*f5=5.0005* **29.0005**	***n** = 200.0025*

Note. Contrain. = Contraindicated; Unnec. and Unnecessary = Unnecessary for the patient; Pt. = Patient

Using the same formula that was used to calculate Cohen's kappa values—only this time using the revised cell counts—a PAK value (and corresponding standard deviation) was calculated for each of the 11 NDNQI PrU risk and prevention measures per hospital. Just as with the

calculations for Cohen's kappa values, 0.0001 was added to each cell (i.e. nothing changed except averaging the cells of agreement). Landis and Koch's (1977) classifications (< 0 = "poor", 0 – .20 = "slight", .21 – .40 = "fair", .41 – .60 = "moderate", .61 – .80 = "substantial", and .81 – 1.0 = "near perfect") were used to categorize PAK values' strength of agreement. For this study, the recommended reliability level for the study measures is ≥ .610.

Percent agreement. Using the 2 – 4 frequency tables that were presented by SPSS, percent agreement was calculated for each expert/non-expert pair, per measure. These are the same frequency tables that were used to calculate each expert/non-expert pair's Cohen's *k* and PAK values, only this time without the addition of 0.0001 to each cell. These 2 – 4 percent agreement values for each measure were averaged to yield a single percent agreement for each risk and prevention measure per hospital. Corresponding standard deviations were computed for each hospital's per measure percent agreement value.

Question 2: What is the *overall* agreement between expert participant ratings and non-expert participant ratings for the five NDNQI PrU *risk* measures per hospital, and the *overall* agreement between expert participant ratings and non-expert participant ratings for the six NDNQI PrU *prevention* measures per hospital? For each hospital, the average Cohen's kappa value obtained for each measure in Question 1 was used to answer Question 2. For each hospital, the average PAK value obtained in Question 1 for each measure, was used to answer Question 2.

Cohen's kappa. The average Cohen's kappa value obtained from the expert and non-expert comparisons for each PrU *risk* measure (*Skin assessment on admission*, *Risk assessment on admission*, *Time since last risk assessment*, *Risk assessment scale*, and *Risk status*) was averaged to obtain one overall Cohen's kappa value for PrU *risk* per hospital. Similarly, the

average Cohen's kappa value obtained from the expert and non-expert comparisons for each PrU *prevention* measure (*Any PrU prevention*, *Skin assessment*, *Pressure-redistribution surface use*, *Routine repositioning*, *Nutritional support*, and *Moisture management*) was averaged to obtain one overall Cohen's kappa value for PrU *prevention* per hospital. Corresponding standard deviations were computed for each of these per hospital Cohen's *k* values for the PrU *risk* measures, and similarly for the per hospital Cohen's *k* values for the PrU *prevention* measures.

Prevalence-adjusted kappa. The average PAK value obtained from the expert and non-expert comparisons for each PrU *risk* measure (*Skin assessment on admission*, *Risk assessment on admission*, *Time since last risk assessment*, *Risk assessment scale*, and *Risk status*) was averaged to obtain one overall PAK value for PrU *risk* per hospital. Similarly, the average PAK value obtained from the expert and non-expert comparisons for each PrU *prevention* measure (*Any PrU prevention*, *Skin assessment*, *Pressure-redistribution surface use*, *Routine repositioning*, *Nutritional support*, and *Moisture management*) was averaged to obtain one overall PAK value for PrU *prevention* per hospital. Corresponding standard deviations were computed for each of these per hospital PAK values for the PrU *risk* measures, and similarly for the per hospital PAK values for the PrU *prevention* measures.

Question 3: What is the *average* of the *within hospital* agreement between expert participant ratings and non-expert participant ratings for each of the 11 NDNQI pressure ulcer risk and prevention measures *across* hospitals? Each hospital's average Cohen's kappa value that was obtained for each measure to answer Question 1 was used to answer Question 3. Similarly, each hospital's average PAK value and percent agreement that was obtained for each measure in Question 1, was used to answer Question 3.

Cohen's kappa. The average Cohen's kappa value that was computed for each pressure ulcer risk and prevention measure per hospital (to answer Question 1) was used to calculate an average Cohen's kappa value for each pressure ulcer risk and prevention measure across hospitals. For example, the average Cohen's kappa value for *Risk status* that was obtained at each hospital was used to compute an average Cohen's kappa value for *Risk status across* hospitals. This was repeated for each of the other 10 NDNQI measures. Therefore, there are 11 across hospital Cohen's kappa values (one for each of the 11 measures). Corresponding 95% CIs were computed for each of these 11 Cohen's kappa values.

Prevalence-adjusted kappa (PAK). In the same manner, the average PAK value that was computed for each pressure ulcer risk and prevention measure per hospital (to answer Question 1) was used to calculate an average PAK value for each pressure ulcer risk and prevention measure across hospitals. For example, the average PAK value for *Risk status* that was obtained at each hospital was used to compute an average PAK value for *Risk status across* hospitals. This was repeated for each of the other 10 NDNQI measures. Therefore, there are 11 across hospital PAK values (one for each of the 11 measures). Corresponding 95% CIs were computed for each of these 11 PAK values.

Percent agreement. In the same manner, the average percent agreement that was computed for each pressure ulcer risk and prevention measure per hospital (to answer Question 1) was used to calculate an average percent agreement for each pressure ulcer risk and prevention measure across hospitals. For example, the average percent agreement for *Risk status* that was obtained at each hospital was used to compute an average percent agreement for *Risk status across* hospitals. This was repeated for each of the other 10 NDNQI measures. Therefore,

there are 11 across hospital percent agreements (one for each of the 11 measures). Corresponding 95% CIs were computed for each of these 11 percent agreements.

Question 4: What is the *ICC* (agreement) between expert participant ratings and non-expert participant ratings for each of the 11 NDNQI PrU risk and prevention measures *across* hospitals? To answer this question, the Cohen's kappa value obtained for each of the first two expert to non-expert comparisons (expert-to-Rater 1, and expert-to-Rater 2) per risk and prevention measure at each hospital, were used to calculate an ICC_{kappa} for each of the 11 NDNQI risk and prevention measures *across* hospitals. Similarly, the PAK value from each of the first two expert to non-expert comparisons (expert-to-Rater 1, and expert-to-Rater 2) per risk and prevention measure at each hospital were used to calculate an ICC_{PAK} for each of the 11 NDNQI risk and prevention measures *across* hospitals.

The choice for two rater pairs (expert to Rater 1 and expert to Rater 2) was because not all hospitals had three or more rater pairs. Inclusion of all rater pairs in the analysis resulted in the listwise exclusion of all hospitals but one. This is because the analysis method used requires that each hospital has the same number of raters, or data are excluded. Therefore, when all rater pairs were included in the analysis ($N = 84$ kappa values from $N = 36$ hospitals), the analysis listwise excluded all but the single hospital that had 4 rater pairs. Likewise, when only hospitals with 2 or 3 rater pairs ($n = 35$ hospitals) were included, the analysis listwise excluded 25 (69.4%) hospitals. An alternative statistical approach to calculating *ICC*s was considered, such as linear mixed model analysis, but this method did not allow for the calculation of 95% CIs. Although the *ICC* calculations did not include all 84 rater pairs, such as were used to calculate agreement estimates for the NDNQI risk and prevention measures (k, PAK, and percent agreement), the number of values that were included is comparable ($N = 72$). Using two rater pairs per hospital

(to calculate *ICC*s) may be justified because the Cohen's kappa values, PAK values, and percent agreements that were calculated for each measure are conceptually different than the *ICC*s. The *ICC*s examined if participant raters *across* hospitals rated the study items with similar reliability. In contrast, the Cohen's kappa values, PAK values, and percent agreements examined agreement *within* hospitals.

Intraclass correlation coefficient (ICC). Intraclass-correlation coefficients (*ICC*s) were used to examine inter-rater reliability *across* hospitals. The sample size for *ICC*s was two kappa values per hospital, across 36 hospitals. This translates to 72 kappa values per measure that were used to calculate *ICC*s.

Intraclass correlation coefficients (*ICC*s) were calculated as in Hart et al. (2006):

$$\frac{\text{(Variability of kappa's between hospitals)}}{\text{(Variability of kappa's between hospitals) + (Variability of kappa's within hospitals)}}$$

Specifying Model 1 (Shrout & Fleiss, 1979), a one-way random ANOVA, 11 $ICCs_{kappa}$—one for each PrU risk and prevention measure—were calculated across hospitals. The *ICC* reported is the "Single Measure" value. This is because in practice, data on the study measures are collected by a single rater (B. Gajewski, personal communication, May 15, 2014). Corresponding 95% CIs are reported for each ICC_{kappa}.

The choice for a one-way random model to calculate *ICC* values was based on previous research (Hart et al., 2006) that used this analysis to examine the reliability of PrU staging (B. Gajewski, personal communication, May 15, 2014). The assumptions for a one-way ANOVA were addressed. The assumptions include (1) the dependent variable is normally distributed for the population, (2) variances of the dependent variables are the same for all populations, and (3) the cases represent random samples from the populations and the scores are independent of each other (Green & Salkind, 2008). According to Green and Salkind (2008), violation of the

normality assumption may have been tempered by requesting that an equal number of patients be rated per hospital, and that a random sample of hospitals meeting inclusion criteria were invited to participate in the study.

An ICC_{PAK} for each measure was calculated in the same way using the same formula from PAK values. Specifying Model 1, a one-way random ANOVA, 11 $ICCs_{PAK}$—one for each PrU risk and prevention measure—was calculated across hospitals. The "Single Measure" *ICC* was reported. Corresponding 95% CIs were calculated for each ICC_{PAK}.

An *ICC* close to zero (i.e. < .22) is desirable and suggests *within*-hospital variance is much greater than *between*-hospital variance (Hart et al., 2006); an ICC < .50 suggests within-hospital variability was greater than between-hospital variability. This means that the variance among the agreement between the expert and non-experts *within* hospitals, was greater than the variance in kappa values (agreement) *across* hospitals. In other words, an *ICC* near zero suggests participant raters across hospitals rated the study items with similar reliability.

Question 5: Where is the lack of agreement between expert and non-expert participant ratings on the 11 NDNQI pressure ulcer risk and prevention measures occurring? (Agreement matrices/Descriptive statistics)

Descriptive statistics. Descriptive statistics were used to evaluate where the lack of agreement between experts' ratings and non-experts' ratings on each of the study measures occurred. This included constructing "confusion" or agreement matrices in order to evaluate where the lack of agreement occurred (Bakeman & Gottman, 1997; Lloyd, n. d.). These matrices are identical to the distribution table (Table 8) above in that cell counts on the diagonal indicate agreement between two raters, while cell counts off the diagonal identify disagreement. The number of columns/rows equals the number of response categories. Data used to construct these

matrices were the same as that were used to calculate Cohen's kappa values (frequencies of agreement and disagreement between expert and non-expert participant ratings); data were presented in a matrix in order to visually identify areas of "confusion". Data were aggregated to present one matrix for each NDNQI PrU risk and prevention measure across hospitals. These matrices give important insight into where the disagreement occurred in order to target efforts at improving reliability (as needed).

Table 9 below is an example of an agreement matrix with hypothetical data. From this table, areas of disagreement can be identified. For instance, most disagreement occurred when one rater believed routine repositioning was performed ("Yes"), and the other rater thought repositioning was "Unnecessary".

Table 9.

Illustration of an Agreement Matrix for Hypothetical Data: Routine Repositioning

Expert Rater	Non-Expert Rater: Yes	No	Contrain.	Unnec.	Pt. Refused	
Yes	546	28	7	(73)	4	*g1*= 658
No	25	115	11	18	7	*g2* =176
Contraindicated	1	21	52	10	2	*g3* =86
Unnecessary	(53)	22	4	35	0	*g4* =114
Pt. Refused	1	30	1	2	18	*g5*=52
	f1=626	*f2=216*	*f3=75*	*f4=138*	*f5=31*	***n* = 1086**

Note. Contrain. = Contraindicated; Unnec. and Unnecessary = Unnecessary for the patient; Pt. = Patient

Aim 2: Examine the methods and processes used by participant raters to gather data on the NDNQI PrU risk and prevention measures. (Descriptive statistics)

Descriptive statistics. Descriptive statistics were used to describe the methods and processes used to collect information needed to rate the study indicators, as self-reported on the online *REDCap* survey, *Pressure Ulcer Risk and Prevention Reliability Survey* (Appendix C). Frequencies and percentages are presented for each item on the survey. For analysis of the open-ended responses describing participant rater characteristics [(#1) What is your job title?—Other (please specify); (#6) In addition to chart abstractor, what other role(s) in the NDNQI Pressure Ulcer Survey do you hold?—Other (please specify); (#8) Which of the following active certifications do you hold?—Other (please specify)], responses were grouped into conceptually similar categories. If a participant rater accessed the survey more than once, only survey data submitted during the first login was included in the data analysis. Frequencies and percentages of responses for each of these categories are reported.

Pilot Study

A pilot study using a single NDNQI participating hospital was performed to assess the feasibility of the study and of the *Data Collection Form*, and test *REDCap* use. Purposive sampling was used to identify a Midwestern hospital for the pilot study. Except for recruitment and pre-study teleconference, the study protocol was followed during the pilot study. Using the operational definitions specified in the study protocol, one expert and three non-experts were identified by the site coordinator. After IRB approval at the pilot hospital, data were collected on August 7, 2014. Participant raters did not submit their data and complete the online survey, however, until August 26. After this date, pilot study participant raters and the skin survey team leader were asked to provide feedback on the *Overview of the Pressure Ulcer Risk and*

Prevention Study, the *Site Coordinator Instructions*, the *Participant Rater Instructions,* the *Data Collection Form*, and the *Pressure Ulcer Risk and Prevention Reliability Survey*—including feedback on the process of uploading their Excel file, and any issues or difficulties they encountered during the pilot study. A *REDCap* survey was developed by this student researcher that was used to solicit anonymous feedback from these individuals (Appendix K). In addition, the pilot study gave the student researcher experience with *REDCap*.

Anonymous feedback on the study from pilot participants ($N = 4$) was received on August 27, 2014. This feedback (Appendix K) identified some technical issues with the *Data Collection Form*, which were addressed before the start of the study. No other issues were identified. Other feedback from pilot participants included (a) it took approximately 3 to 4 hours to rate 50 patients on these measures, and (b) survey completion and file upload took ≤ 15 to 30 minutes. No pilot data were included in the study. The hospital that participated in the pilot study was not eligible to participate in the study.

Study Timeline

Table 10.

Study Timeline

JUNE 2014	JULY 2014	AUG 2014	SEPT-OCT 2014	NOV-DEC 2014	JAN-MAR 2015
KUMC HSC & NDNQI approval	Study materials sent to pilot site. (July 27)	Sent *Invitation to Participate* (Aug. 1) & selected hospitals Pilot study (Aug. 7 – Aug. 27)	Teleconference (Sept. 9)	Data clarification & cleaning	Write *Results, Discussion, & Conclusion* Results to participating hospitals (Feb.)
Pilot hospital recruited	Pilot IRB approval (July 30)	Study materials sent to participating hospitals (Aug. 27)	Data collection & submission (Sept. 29 – Oct. 31)	Data analysis	

Note. KUMC = Kansas University Medical Center; HSC = Human Subjects Committee.

CHAPTER IV: RESULTS

The purpose of this study was to examine the reliability of the NDNQI® pressure ulcer risk and prevention measures. Specific study aims were to (1) **Aim 1**: examine the reliability of the NDNQI pressure ulcer risk and prevention measures within and across NDNQI hospitals, and (2) **Aim 2**: examine the methods and processes used by participant raters to gather data on the NDNQI PrU risk and prevention measures. In Chapter IV, the hospital, unit, patient, and participant rater characteristics are described. Next, results by study aim and question are presented.

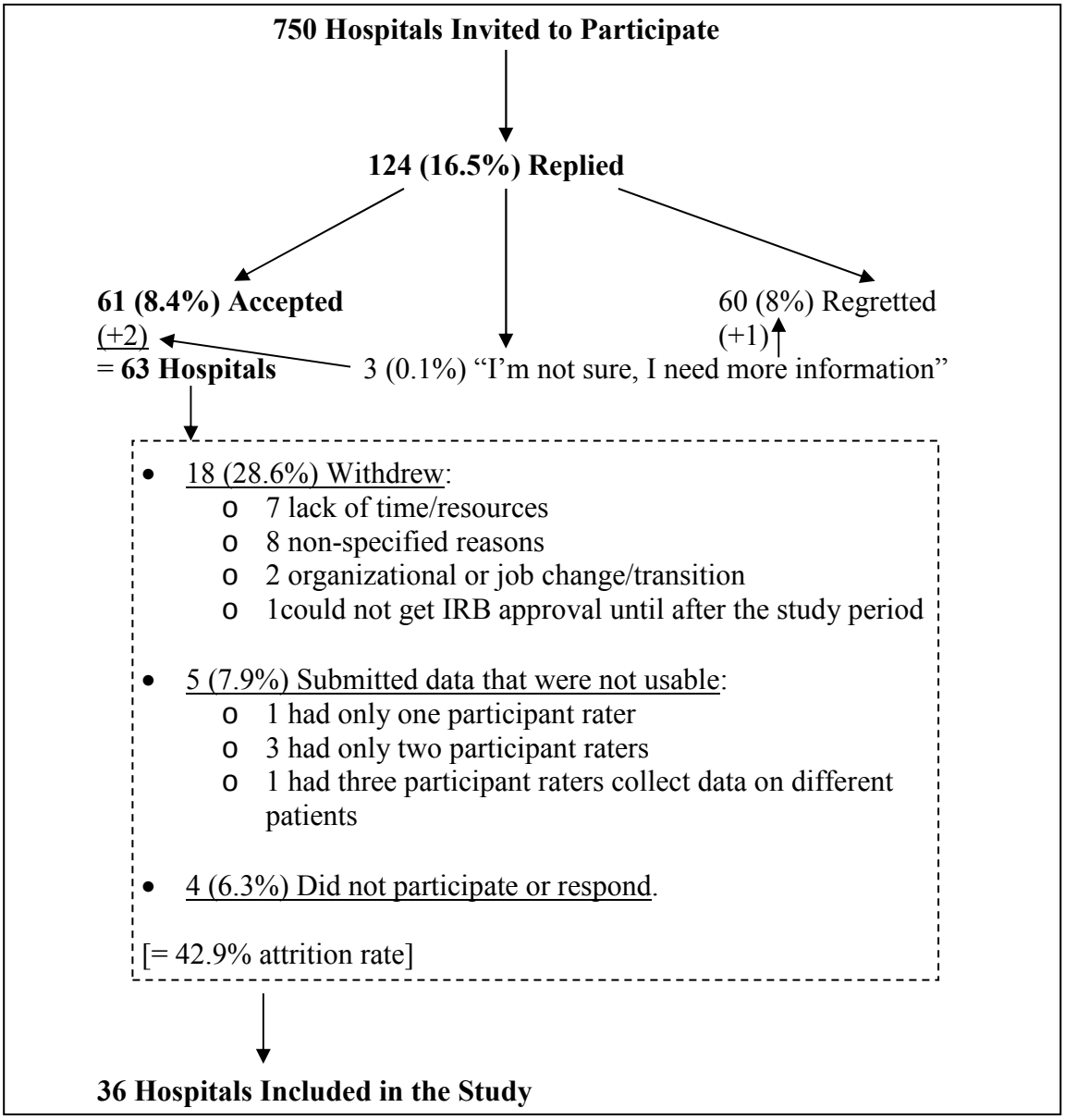
Descriptive Statistics

Participant raters from among 41 hospitals rated patients on the 11 NDNQI PrU risk and prevention measures. Of these 41 hospitals, data from five hospitals were removed from the analysis because (a) only a single participant rater scored patients (1 hospital), (b) only two participant raters scored patients (3 hospitals), and (c) three participant raters collected data on different patients (1 hospital), see Figure 3 below. Among the final sample of 36 hospitals, 120 participant raters collected data on 1,637 patients. These patients were located on 92 units; specifically on 62 (67%) adult medical-surgical units, 19 (21%) adult medical units, and 11 (12%) adult surgical units.

Hospitals ($N = 36$)

Most of the hospitals were a Non-Teaching Facility [$n = 21$ (58.3%)], located in a metropolitan area [$n = 31$ (86.1%)], and Non-Magnet [$n = 27$ (75%)]. The majority were 100 – 299 staffed beds in size [$n = 26$ (72.2%)]. Two hospitals were < 100 staffed beds. The decision to include these two smaller hospitals in the final analysis was based on visual (box-and-whisker

Figure 3. Derivation of the Hospital Sample



plots) and statistical inspection of the Cohen's kappa and PAK values that were generally within the range of all the other hospitals' kappa values. Additional support for including these small hospitals came from the fact that the number of patients rated at each of these two small hospitals was similar to the number of patients rated at the other hospitals. Hospital characteristics are presented in Table 11 below.

Table 11.

Hospital Characteristics (N = 36): Size, Teaching Status, Magnet Status, and Metropolitan Status

	n (%)
Number of Staffed Beds	
50 – 74	1 (2.8%)
75 – 99	1 (2.8%)
100 – 199	13 (36.1%)
200 – 299	13 (36.1%)
300 – 399	5 (13.9%)
400 – 499	3 (8.3%)
≥ 500	0
Teaching Status	
Academic Medical Center	3 (8.3%)
Teaching Facility	12 (33.3%)
Non-Teaching Facility	21 (58.3%)
Magnet Status	
Magnet	9 (25%)
Non-Magnet	27 (75%)
Metropolitan Status	
Non-Metropolitan/Non-Micropolitan	1 (2.8%)
Micropolitan	4 (11.1%)
Metropolitan	31 (86.1%)

Patients (*N* = 1,637)

Among the 36 hospitals, data were collected on a total of 1,737 patients. Of these 1,737 patients, data from 100 patients were removed from analysis because patient age, patient gender, and patient room number did not agree across raters within that particular hospital. The final sample has 1,637 patients, for an average of 45.5 patients per hospital (SD = 6.4). *All* 1,637 patients were rated on the PrU *risk* measures. Of these, 553 (33.8%) patients were considered by the expert rater to be "at risk" for PrU development. However, not all non-experts agreed with the expert on risk status, so the sample of patients for each of the prevention measures was 528 to 530. Therefore, 528 to 530 patients were rated on *both* the risk and the prevention measures. This is an average of 14.7 "at risk" patients per hospital (range = 5 to 37; SD = 7.0).

Patient age ranged from 17 – 102 years, with a mean age of 63.3 years (SD = 18.4). A majority of patients were female, *n* = 893 (55%). Of the 1,637 patients; 905 (55.3%) were on a medical-surgical unit, 619 (37.8%) were on a medical unit, and 113 (6.9%) were on a surgical unit. These frequencies are logical because data collection was to begin on medical-surgical units, and then proceed to medical units, followed by surgical units. Adult medical-surgical, medical, and surgical unit types represent the units from where nearly half (49%) of NDNQI PrU reporting is derived (NDNQI Statistical Analyst, personal communication, November 21, 2014).

Participant raters at 30 (83.3%) of the 36 hospitals, each collected data on 50 patients. Table 12 below presents patient age and gender, the number of participant raters, the total number of patients rated, the number of patients included in the statistical analysis, and the number of patients considered to be "at risk" by the "expert". For instance, Table 12 shows that Hospital 1 had three participant raters who each collected data on 50 patients. However, only 42

Table 12.

Patient Age and Gender, Number of Participant Raters (N = 120), and Number of Patients (N = 1,637) by Hospital (N = 36)

	Patient Age in Years	**Patient Gender**			**Number of Patients**		
Hospital	*Mean* (SD)	#F (%)	#M (%)	#of PRs	# Rated	# Included in Analysis	# "At Risk"(%)
1	54.5 (17.2)	22 (52)	20 (48)	3	50	42	12 (28.6)
2	63.3 (20.5)	28 (57)	21 (43)	3	50	49	11 (22.4)
3	62.5 (17.8)	24 (52)	22 (48)	3	50	46	10 (21.7)
4	62.4 (21.4)	26 (58)	19 (42)	4	50	45	17 (37.8)
5	61.2 (12.8)	28 (68)	13 (32)	3	42	41	7 (17.1)
6	70 (17.3)	29 (59)	20 (41)	3	50	49	17 (34.7)
7	62.4 (21.3)	22 (45)	27 (55)	4	50	49	27 (55.1)
8	56.9 (16.6)	19 (54)	16 (46)	4	36	35	5 (14.3)
9	65.2 (16.8)	24 (50)	24 (50)	3	50	48	16 (33.3)
10	56.4 (20.4)	11 (48)	12 (52)	3	24	23	6 (26.1)
11	65.3 (16.3)	25 (51)	24 (49)	3	50	49	15 (30.6)
12	46.8 (15.5)	20 (50)	20 (50)	4	50	40	7 (17.5)
13	65.8 (13.5)	15 (58)	11 (42)	4	26	26	7 (26.9)
14	60.1 (20.7)	27 (57)	20 (43)	3	50	47	22 (46.8)
15	64.6 (16.8)	24 (48)	26 (52)	3	50	50	37 (74)
16	62.4 (20)	26 (52)	24 (48)	3	50	50	20 (40)
17	64.4 (18.4)	25 (54)	21 (46)	3	50	46	21 (45.7)
18	66.7 (17.6)	32 (71)	13 (29)	4	50	45	16 (35.6)
19	69.8 (15.7)	30 (60)	20 (40)	3	50	50	11 (22)

Note. F = Female; M = Male; PRs = Participant Raters.

Table 12 (continued).

Patient Age and Gender, Number of Participant Raters (N = 120), and Number of Patients (N = 1,637) by Hospital (N = 36)

	Patient Age in Years	**Patient Gender**			**Number of Patients**		
Hospital	*Mean* (SD)	#F (%)	#M (%)	#of PRs	# Rated	# Included in Analysis	# "At Risk" (%)
20	64.2 (18.2)	25 (54)	21 (46)	3	50	46	21 (45.7)
21	62 (17)	20 (43)	27 (57)	3	50	47	25 (53.2)
22	65.4 (15.9)	36 (72)	12 (28)	3	50	50	19 (38)
23	64.6 (20.5)	30 (61)	19 (39)	4	50	49	15 (30.6)
24	62.8 (18.1)	28 (58)	20 (42)	5	55	48	15 (31.3)
25	69.2 (19.1)	26 (54)	22 (46)	3	50	48	14 (29.2)
26	62.6 (21.6)	27 (55)	22 (45)	4	50	49	10 (20.4)
27	65.1 (11.7)	27 (54)	23 (46)	3	54	50	20 (40)
28	66.5 (16.9)	15 (42)	21 (58)	4	50	36	6 (16.7)
29	69.2 (15.1)	25 (50)	25 (50)	3	50	50	11 (22)
30	58.6 (18.6)	35 (70)	15 (30)	3	50	50	5 (10)
31	69 (19)	21 (48)	23 (52)	4	50	44	21 (47.7)
32	61.5 (20.7)	33 (66)	17 (34)	3	50	50	18 (36)
33	72.4 (14.8)	24 (52)	22 (48)	3	50	46	24 (52.2)
34	53.9 (18.9)	17 (38)	28 (62)	3	50	45	6 (13.3)
35	65.7 (17.4)	25 (51)	24 (49)	3	50	49	19 (38.8)
36	60.8 (16.5)	22 (44)	28 (56)	3	50	50	20 (40)

Note. F = Female; M = Male; PRs = Participant Raters.

of those patients were rated by *all three* of these participant raters, and therefore, only these 42 patients were included in the statistical analysis of the NDNQI PrU *risk* measures. Of those 42 patients, 12 (28.6%) were considered to be "at risk" for PrU development by the "expert". Only these 12 patients, therefore, were included in the data analysis for the NDNQI PrU *prevention* measures.

Participant Raters (*N* = 120)

Among the 36 hospitals, 120 raters participated in the study (36 experts—one at each hospital—and 84 non-experts), which is on average 3.3 participant raters per hospital. Specifically, 25 (69%) hospitals had 3 participant raters (1 expert, 2 non-experts), 10 (28%) hospitals had 4 participant raters (1 expert, 3 non-experts), and 1 (3%) hospital had 5 participant raters (1 expert, 4 non-experts). Participant rater characteristics (job title, education, and certifications) are presented in Table 13.

Most of the experts reported being a Wound/Skin Care Nurse [*n* = 17 (47.2%)] or Staff RN [*n* = 8 (22.2%)], with a Bachelor's [*n* = 18 (50%)] or Master's [*n* = 12 (33.3%)] Degree in Nursing. Conversely, most of the non-expert participant raters reported being a "Staff RN" [*n* = 44 (52.8%)], with a Bachelor's [*n* = 37 (44%)] or Associate's [*n* = 21 (25%) Degree in Nursing. Fourteen (38.9%) experts reported having "no certifications in wound care", while 62 (73.8%) non-experts reported the same. Among the experts, 16 (44.4%) were Certified Wound, Ostomy, Continence Nurses (CWOCN), 4 (11.1%) were Wound Care Certified (WCC), 1 (2.8%) was a Certified Wound Care Nurse (CWCN), and 1 (2.8%) was a Certified Wound Ostomy Nurse (CWON). Overall, fewer non-experts were certified in wound care (8 [9.5%] CWOCNs, 2 [2.4%] WCC, 2 [2.4%] CWCNs, and 1 [1.2%] CWON). Five (13.9%) experts and 9 (10.7%) non-experts identified certification in areas other than wound care.

Table 13.

Participant Rater (N = 120) Characteristics: Job Title, Education, and Certifications

	Expert (*n* = 36) *n* (%)	Non-Expert (*n* = 84) *n* (%)	Total Participant Raters (*N* = 120) *n* (%)
What is your job title? [a]			
Staff RN	8 (22.2)	44 (52.4)	52 (43.3)
CNL	4 (11.1)	2 (2.4)	6 (5.0)
APN	1 (2.8)	2 (2.4)	3 (2.5)
Nurse Manager	3 (3.8)	3 (3.6)	6 (5.0)
Nursing Administrator	–	1 (1.2)	1 (0.8)
Quality Improvement	–	4 (4.8)	4 (3.3)
Wound/Skin Care Nurse	17 (47.2)	15 (17.9)	32 (26.7)
NDNQI Site Coordinator	1 (2.8)	4 (4.8)	5 (4.2)
Other, specify			
Clinical Educator	1 (2.8)	2 (2.4)	3 (2.5)
Infection Prevention	1 (2.8)	–	1 (0.8)
Certified Nurse Aid	–	2 (2.4)	2 (1.7)
Research Council	–	1 (1.2)	1 (0.8)
Dietitian	–	1 (1.2)	1 (0.8)
LPN	–	1 (1.2)	1 (0.8)
Clinical Supervisor	–	1 (1.2)	1 (0.8)
Registered Nurse (RN) nursing education. [b]			
Yes	33 (91.7)	72 (85.7)	105 (87.5)
No	3 (8.3)	11 (13.1)	14 (11.7)

[a] Missing values = 1 (0.8%)
[b] Missing values = 1 (0.8%)

Table13 (continued).

Participant Rater (N = 120) Characteristics: Job Title, Education, and Certifications

	Expert (*n* = 36) *n* (%)	Non-Expert (*n* = 84) *n* (%)	Total Participant Raters (*N* = 120) *n* (%)
Select highest RN education level. [c]			
Associate's Degree Nursing	3 (8.3)	21 (25)	24 (20)
Bachelor's Degree Nursing	18 (50)	37 (44)	55 (45.8)
Master's Degree Nursing	12 (33.3)	12 (14.3)	24 (20)
Doctorate Degree Nursing	–	–	–
Select highest non-RN education level. [d]			
High School Graduate/GED	–	2 (2.4)	2 (1.7)
Associate's Degree Non-nursing	2 (5.6)	4 (4.8)	6 (5.0)
Bachelor's Degree Non-nursing	1 (2.8)	4 (4.8)	5 (4.2)
Master's Degree Non-nursing	–	1 (1.2)	1 (0.8)
Doctorate Degree Non-Nursing	–	–	–

[c] Missing values = 2 (1.7%)
[d] Missing values = 0

Table 13 (continued).

Participant Rater (N = 120) Characteristics: Job Title, Education, and Certifications

	Expert (*n* = 36) *n* (%)	Non-Expert (*n* = 84) *n* (%)	Total Participant Raters (*N* = 120) *n* (%)
Which of the following active certifications do you hold? [e] (Select all that apply.)			
CWOCN	16 (44.4)	8 (9.5)	24 (20)
CWCN	1 (2.8)	2 (2.4)	3 (2.5)
COCN	–	–	–
CCCN	–	–	–
CWON	–	–	–
WCC	4 (11.1)	2 (2.4)	6 (5.0)
No certifications in wound care	14 (38.9)	62 (73.8)	76 (63.3)
Other Certifications (please describe)			
• CIC	1 (2.8)	–	1 (0.8)
• Adult Health CNS	1 (2.8)	–	1 (0.8)
• CCRN	1 (2.8)	1 (1.2)	2 (1.7)
• CNRN	1 (2.8)	1 (1.2)	2 (1.7)
• Medical-Surgical	1 (2.8)	3 (3.6)	4 (3.3)
• ONC	–	2 (2.4)	2 (1.7)
• CNL	–	1 (1.2)	1 (0.8)
• CWCA	–	1 (1.2)	1 (0.8)

[e] Totals do not equal the number of raters because respondents were to select "all that apply".

Note.
CWOCN = Certified Wound, Ostomy, Continence Nurse
CWCN = Certified Wound Care Nurse
COCN = Certified Ostomy Care Nurse
CCCN = Certified Continence Care Nurse
CWON = Certified Wound Ostomy Nurse
CWS = Certified Wound Specialist
WCC = Wound Care Certified
CIC = Certified Infection Control
CNS = Clinical Nurse Specialist
CCRN = Critical Care Registered Nurse
CNRN = Certified Neuroscience RN
ONC = Oncology/Chemotherapy Administrator
CNL = Clinical Nurse Leader
CWCA = Certified Wound Care Associate

Results by Study Aim

Aim 1

Aim 1 was to examine the reliability of the NDNQI pressure ulcer risk and prevention measures within and across NDNQI hospitals. Results are reported by question.

Question 1. What is the agreement between expert participant ratings and non-expert participant ratings *for each* of the 11 NDNQI pressure ulcer risk and prevention measures *within* each hospital? To address this question, each hospital's Cohen's kappa value, PAK value (see discussion p. 119 – 121), and percent agreement are presented by measure below. Altogether, 396 Cohen's kappa values, 396 PAK values, and 396 percent agreement values are reported (36 hospitals x 11 measures = 396). Overall, the level of agreement for Cohen's kappa values was more varied than the PAK values, and ranged from "poor" [*n* = 15 (3.8%)] to "near perfect" [*n* = 34 (8.6%)]. More than 40% of the Cohen's kappa values indicate "moderate" [*n* = 170 (42.9%)] agreement. Another 40% indicate "fair" [*n* = 74 (18.7%)] or "slight" agreement [*n* = 83 (21%)]. Only 84 of the 396 (21.2%) Cohen's kappa values reflect at least "substantial" agreement. Hospital PAK values across the 11 measures ranged from "poor" [*n* = 1 (0.3%)] to "near perfect" [*n* = 265 (66.9%)], with 309 of the 396 (78.0%) PAK values reflecting at least "substantial" agreement.

Risk measure – Skin assessment within 24 hours of admission. Cohen's kappa values for the PrU risk measure, *Skin assessment within 24 hours of admission* (Table 14), ranged from -.014 "poor" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 2 (5.6%) hospitals, Cohen's kappa values indicate "substantial" to "near perfect" agreement. In 17 (47.2%) hospitals, Cohen's kappa values indicated "moderate" agreement; in 11 (30.6%)

Table 14.

Aim 1 Question 1: ***"Skin Assessment within 24 hours of Admission"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36 Hospitals)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.236 (.356)	.911 (.025)	94.05 (1.68)
2	2	.500 (0)	1.00 (0)	100 (0)
3	2	.500 (0)	1.00 (0)	100 (0)
4	3	.334 (.288)	.989 (.019)	99.26 (1.28)
5	2	.001 (0.00)	.963 (0.00)	97.56 (0.00)
6	2	-.014 (.021)	.939 (.043)	95.92 (2.89)
7	3	.500 (0)	1.00 (0)	100 (0)
8	3	1.00 (0)	1.00 (0)	100 (0)
9	2	.500 (0)	1.00 (0)	100 (0)
10	2	.250 (.354)	.968 (.046)	97.83 (3.08)
11	2	.500 (0)	1.00 (0)	100 (0)
12	3	.772 (.197)	.975 (.022)	98.33 (1.44)
13	3	.500 (0)	1.00 (0)	100 (0)
14	2	.500 (0)	1.00 (0)	100 (0)
15	2	.500 (0)	1.00 (0)	100 (0)
16	2	.500 (0)	1.00 (0)	100 (0)
17	2	.154 (.218)	.854 (.023)	90.22 (1.53)
18	3	.334 (.288)	.989 (.019)	99.26 (1.28)
19	2	.251 (.353)	.985 (.021)	99.00 (1.41)
20	2	.219 (.310)	.919 (.023)	94.57 (1.53)
21	2	.001 (.001)	.952 (.023)	96.81 (1.51)
22	2	.500 (0)	1.00 (0)	100 (0)
23	3	.334 (.577)	.979 (.018)	98.61 (1.20)
24	4	.375 (.250)	.992 (.016)	99.48 (1.04)
25	2	.500 (0)	1.00 (0)	100 (0)
26	3	.334 (.288)	.990 (.018)	99.32 (1.18)
27	2	.500 (0)	1.00 (0)	100 (0)
28	3	.500 (0)	1.00 (0)	100 (0)
29	2	.250 (.354)	.882 (.168)	92.00 (11.31)
30	2	-.010 (.015)	.955 (.021)	97.00 (1.41)
31	3	.333 (.289)	.977 (.039)	98.48 (2.63)
32	2	.500 (0)	1.00 (0)	100 (0)
33	2	.001 (0.00)	.966 (0.00)	97.73 (0.00)
34	2	.500 (0)	1.00 (0)	100 (0)
35	2	.500 (0)	1.00 (0)	100 (0)
36	2	.500 (0)	1.00 (0)	100 (0)
Total Expert to Non-Expert pairs = 84				

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

hospitals there was "fair" agreement; in 4 (11.1%) hospitals, there was "slight" agreement; and in 2 (5.6%) hospitals, Cohen's kappa values indicate "poor" agreement between the expert and non-expert raters.

Prevalence-adjusted kappa (PAK) values ranged from .854 – 1.0, indicating "near perfect" agreement between expert and non-expert ratings on *Skin assessment within 24 hours of admission*. Percent agreement ranged from 90.22% – 100% across hospitals for this measure.

Risk measure – Risk assessment within 24 hours of admission. Cohen's kappa values for the PrU risk measure, *Risk assessment within 24 hours of admission* (Table 15), ranged from -.027 "poor" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 4 (11.1%) hospitals, Cohen's kappa values indicate "near perfect" agreement; and in 1 (2.8%) hospital, there was "substantial" agreement. In 22 (61.1%%) hospitals, Cohen's kappa values indicate "moderate" agreement; in 4 (11.1%) hospitals, there was "fair" agreement; in 4 (11.1%) hospitals, there was "slight" agreement; and in 1 (2.8%) hospital, the Cohen's kappa value indicates "poor" agreement between the expert and non-expert raters.

Prevalence-adjusted kappa (PAK) values ranged from .822 – 1.0, indicating "near perfect" agreement between expert and non-expert ratings on *Risk assessment within 24 hours of admission*. Percent agreement ranged from 87.76% – 100% across hospitals for this measure.

Risk measure – Time since last pressure ulcer risk assessment. Cohen's kappa values for the PrU risk measure, *Time since last pressure ulcer risk assessment* (Table 16), ranged from -.052 "poor" to .826 "near perfect" agreement between the expert and non-expert raters. For 1 (2.8%) hospital, the Cohen's kappa value indicates "near perfect" agreement; and in 3 (8.3%) hospitals, there was "substantial" agreement. In 13 (36.1%) hospitals, Cohen's kappa values indicate "moderate" agreement; in 7 (19.4%) hospitals, there was "fair" agreement; in 8 (22.2%)

Table 15.

Aim 1 Question 1: ***"Risk Assessment within 24 hours of Admission"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.251 (.353)	.982 (.025)	98.81 (1.68)
2	2	.500 (0)	1.00 (0)	100 (0)
3	2	.500 (0)	1.00 (0)	100 (0)
4	3	.500 (0)	1.00 (0)	100 (0)
5	2	.500 (0)	1.00 (0)	100 (0)
6	2	.250 (.354)	.822 (.252)	87.76 (17.32)
7	3	.500 (0)	1.00 (0)	100 (0)
8	3	.913 (.151)	.985 (.025)	99.02 (1.70)
9	2	.001 (0.00)	.969 (0.00)	97.92 (0.00)
10	2	.556 (.128)	.903 (.046)	93.48 (3.08)
11	2	.251 (.353)	.985 (.022)	98.98 (1.44)
12	3	.658 (.296)	.950 (.043)	96.67 (2.89)
13	3	.500 (0)	1.00 (0)	100 (0)
14	2	.500 (0)	1.00 (0)	100 (0)
15	2	.001 (0.00)	.970 (0.00)	98.00 (0.00)
16	2	.500 (0)	1.00 (0)	100 (0)
17	2	.151 (0.00)	.828 (0.00)	90.22 (1.53)
18	3	1.00 (0)	1.00 (0)	100 (0)
19	2	.500 (0)	1.00 (0)	100 (0)
20	2	.501 (.706)	.984 (.023)	98.92 (1.53)
21	2	-.027 (.008)	.904 (.045)	93.62 (3.01)
22	2	1.00 (0)	1.00 (0)	100 (0)
23	3	.914 (.149)	.990 (.018)	99.32 (1.18)
24	4	.581 (.501)	.977 (.030)	98.44 (2.00)
25	2	.500 (0)	1.00 (0)	100 (0)
26	3	.334 (.288)	.990 (.018)	99.32 (1.18)
27	2	.500 (0)	1.00 (0)	100 (0)
28	3	.500 (0)	1.00 (0)	100 (0)
29	2	.500 (0)	1.00 (0)	100 (0)
30	2	.500 (0)	1.00 (0)	100 (0)
31	3	.167 (.288)	.977 (.020)	98.49 (1.31)
32	2	.500 (0)	1.00 (0)	100 (0)
33	2	.500 (0)	1.00 (0)	100 (0)
34	2	.500 (0)	1.00 (0)	100 (0)
35	2	.500 (0)	1.00 (0)	100 (0)
36	2	.500 (0)	1.00 (0)	100 (0)

Total Expert to Non-Expert pairs = 84

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

Table 16.

Aim 1 Question 1: ***"Time since Last Pressure Ulcer Risk Assessment"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.001 (.001)	.791 (.055)	80.95 (3.37)
2	2	.002 (.041)	.658 (.382)	69.39 (34.63)
3	2	-.052 (.074)	.257 (.075)	35.87 (1.54)
4	3	.270 (.036)	.827 (.065)	85.18 (5.59)
5	2	.039 (.028)	.570 (.078)	64.64 (5.17)
6	2	.188 (.035)	.454 (.063)	54.08 (4.33)
7	3	.500 (0)	1.00 (0)	100 (0)
8	3	.426 (.328)	.717 (.163)	76.19 (13.51)
9	2	.299 (.493)	.902 (.069)	91.67 (5.89)
10	2	-.020 (.033)	.924 (.036)	93.48 (3.08)
11	2	.705 (.127)	.834 (.067)	85.72 (5.78)
12	3	.500 (0)	1.00 (0)	100 (0)
13	3	.315 (.595)	.548 (.415)	58.98 (38.72)
14	2	.252 (.351)	.988 (.018)	98.94 (1.51)
15	2	.004 (0)	.977 (0)	98.00 (0.00)
16	2	.729 (0)	.953 (0)	96.00 (0.00)
17	2	.305 (.412)	.500 (.491)	53.27 (47.66)
18	3	.664 (.364)	.871 (.112)	88.89 (9.69)
19	2	.430 (.308)	.849 (.147)	87.00 (12.73)
20	2	.455 (.643)	.899 (.107)	91.31 (9.23)
21	2	.389 (.548)	.913 (.052)	92.55 (4.51)
22	2	.826 (.016)	.953 (0)	91.00 (7.07)
23	3	.506 (.111)	.726 (.059)	76.98 (4.81)
24	4	.397 (.079)	.647 (.048)	70.22 (3.88)
25	2	.004 (0.00)	.976 (0.00)	97.92 (0.00)
26	3	.504 (.449)	.944 (.036)	95.24 (3.12)
27	2	.500 (0)	1.00 (0)	100 (0)
28	3	-.014 (.103)	.487 (.188)	57.41 (15.88)
29	2	.497 (.023)	.638 (.016)	69.00 (1.41)
30	2	.525 (0.00)	.766 (0.00)	80.00 (0.00)
31	3	.061 (.106)	.713 (.086)	75.17 (6.12)
32	2	.500 (0)	1.00 (0)	100 (0)
33	2	.470 (.274)	.668 (.206)	71.59 (17.68)
34	2	-.040 (.059)	.765 (.260)	80.00 (22.01)
35	2	.513 (.204)	.952 (.034)	95.92 (2.89)
36	2	.001 (.001)	.791 (.065)	82.00 (5.66)
Total Expert to Non-Expert pairs = 84				

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

hospitals, there was "slight" agreement; and in 4 (11.1%) hospitals, Cohen's kappa values indicate "poor" agreement between the expert and non-expert raters.

Prevalence-adjusted kappa (PAK) values ranged from .257 "fair" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 30 (83.3%) hospitals, PAK values indicate "substantial" to "near perfect" agreement. In 5 (13.9%) hospitals, PAK values indicate "moderate" agreement; and in 1 (2.8%) hospital, the PAK value indicates "fair" agreement between expert and non-expert ratings for the measure *Time since last pressure ulcer risk assessment*. Percent agreement ranged from 35.87% – 100% across hospitals for this measure.

Risk measure – Risk assessment scale. Cohen's kappa values for the PrU risk measure, *Risk assessment scale* (Table 17), ranged from .319 "fair" to .500 "moderate" agreement between the expert and non-expert raters. For 35 (97.2%) hospitals, Cohen's kappa values indicate "moderate" agreement; and for 1 (2.8%) hospital, the Cohen's kappa value indicates "fair" agreement between the expert and non-expert raters.

Prevalence-adjusted kappa (PAK) values ranged from .829 – 1.0, indicating "near perfect" agreement between expert and non-expert ratings on *Risk assessment scale*. Percent agreement ranged from 91.90% – 100% across hospitals for this measure.

Risk measure – Risk status. Cohen's kappa values for the last PrU risk measure, *Risk status* (Table 18), ranged from .314 "fair" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 23 (63.9%) hospitals, there was "near perfect" agreement; and in 9 (25.0%) hospitals, there was "substantial" agreement. In 3 (8.3%) hospitals, Cohen's kappa values indicate "moderate" agreement; and in 1 (2.8%) hospital, there was "fair" agreement between the expert and non-expert raters.

Table 17.

Aim 1 Question 1: ***"Risk Assessment Scale"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.500 (0)	1.00 (0)	100 (0)
2	2	.500 (0)	1.00 (0)	100 (0)
3	2	.500 (0)	1.00 (0)	100 (0)
4	3	.500 (0)	1.00 (0)	100 (0)
5	2	.500 (0)	1.00 (0)	100 (0)
6	2	.500 (0)	1.00 (0)	100 (0)
7	3	.500 (0)	1.00 (0)	100 (0)
8	3	.500 (0)	1.00 (0)	100 (0)
9	2	.500 (0)	1.00 (0)	100 (0)
10	2	.500 (0)	1.00 (0)	100 (0)
11	2	.500 (0)	1.00 (0)	100 (0)
12	3	.500 (0)	1.00 (0)	100 (0)
13	3	.500 (0)	1.00 (0)	100 (0)
14	2	.500 (0)	1.00 (0)	100 (0)
15	2	.500 (0)	1.00 (0)	100 (0)
16	2	.500 (0)	1.00 (0)	100 (0)
17	2	.500 (0)	1.00 (0)	100 (0)
18	3	.500 (0)	1.00 (0)	100 (0)
19	2	.319 (.451)	.892 (.040)	91.90 (2.97)
20	2	.500 (0)	1.00 (0)	100 (0)
21	2	.500 (0)	1.00 (0)	100 (0)
22	2	.500 (0)	1.00 (0)	100 (0)
23	3	.500 (0)	1.00 (0)	100 (0)
24	4	.500 (0)	1.00 (0)	100 (0)
25	2	.500 (0)	1.00 (0)	100 (0)
26	3	.500 (0)	1.00 (0)	100 (0)
27	2	.500 (0)	1.00 (0)	100 (0)
28	3	.500 (0)	1.00 (0)	100 (0)
29	2	.500 (0)	1.00 (0)	100 (0)
30	2	.500 (0)	1.00 (0)	100 (0)
31	3	.500 (0)	1.00 (0)	100 (0)
32	2	.500 (0)	1.00 (0)	100 (0)
33	2	.500 (0)	1.00 (0)	100 (0)
34	2	.500 (0)	1.00 (0)	100 (0)
35	2	.500 (0)	1.00 (0)	100 (0)
36	2	.500 (0)	1.00 (0)	100 (0)
Total Expert to Non-Expert pairs = 84				

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

Table 18.

Aim 1 Question 1: ***"Risk Status"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.914 (.037)	.947 (.025)	96.43 (1.68)
2	2	.970 (.043)	.985 (.022)	98.98 (1.44)
3	2	.905 (.047)	.951 (.023)	96.74 (1.54)
4	3	.659 (.217)	.746 (.169)	82.96 (11.41)
5	2	.429 (.035)	.711 (.053)	80.49 (3.45)
6	2	.788 (.015)	.849 (.002)	89.90 (0.14)
7	3	.609 (.066)	.633 (.062)	75.51 (4.08)
8	3	.961 (.067)	.985 (.026)	98.99 (1.75)
9	2	.314 (.039)	.443 (.030)	60.42 (2.95)
10	2	.947 (.076)	.968 (.046)	97.83 (3.08)
11	2	.947 (.008)	.969 (0.00)	97.94 (0.03)
12	3	.723 (0.00)	.884 (0.00)	92.31 (0.00)
13	3	.918 (.143)	.943 (.099)	96.15 (6.66)
14	2	.713 (.038)	.748 (.001)	82.98 (0.00)
15	2	.924 (.107)	.926 (.105)	95.00 (7.07)
16	2	.917 (0.00)	.940 (0.00)	96.00 (0.00)
17	2	1.00 (0)	1.00 (0)	100 (0)
18	3	.573 (.139)	.646 (.142)	75.56 (10.64)
19	2	.839 (.152)	.924 (.065)	94.92 (4.36)
20	2	.915 (.060)	.935 (.046)	95.66 (3.08)
21	2	.851 (.211)	.888 (.059)	92.50 (10.61)
22	2	.794 (.023)	.835 (.021)	89.00 (1.41)
23	3	.786 (.331)	.838 (.254)	88.65 (17.84)
24	4	.889 (.127)	.921 (.095)	94.68 (6.38)
25	2	.950 (.071)	.969 (.045)	100 (0)
26	3	.636 (.111)	.786 (.081)	85.72 (5.40)
27	2	.957 (0.00)	.969 (0.00)	97.96 (0.00)
28	3	.969 (.054)	.986 (.024)	99.07 (1.61)
29	2	.944 (0.00)	.970 (0.00)	98.00 (0.00)
30	2	.618 (.017)	.866 (.021)	91.00 (1.41)
31	3	.831 (.094)	.871 (.073)	91.42 (4.82)
32	2	.935 (.030)	.955 (.021)	97.00 (1.41)
33	2	.848 (.215)	.885 (.163)	92.39 (10.76)
34	2	.683 (.322)	.821 (.206)	87.78 (14.14)
35	2	.912 (0.00)	.939 (0.00)	95.92 (0.00)
36	2	.937 (.089)	.955 (.064)	97.00 (4.24)
Total Expert to Non-Expert pairs = 84				

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

Prevalence-adjusted kappa (PAK) values ranged from .443 "moderate" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 35 (97.2%) hospitals, PAK values indicate "substantial" to "near perfect" agreement. For 1 (2.8%) hospital, the PAK value indicates "moderate" agreement between expert and non-expert ratings on *Risk status*. Percent agreement ranged from 60.42% – 100% across hospitals for this measure.

Prevention measures – Any prevention within the last 24 hours. Cohen's kappa values for the PrU prevention measure, *Any prevention within the last 24 hours* (Table 19), ranged from -.133 "poor" to .855 "near perfect" agreement between the expert and non-expert raters. For 2 (5.6%) hospitals, Cohen's kappa values indicate "near perfect" agreement. In 20 (55.6%) hospitals, there was "moderate" agreement; in 6 (16.7%) hospitals, there was "fair" agreement; in 7 (19.4%) hospitals, there was "slight" agreement; and in 3 (8.3%) hospitals, Cohen's kappa values indicate "poor" agreement between the expert and non-expert raters.

Prevalence-adjusted kappa (PAK) values ranged from -.089 "poor" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 31 (86.1%) hospitals, PAK values indicate "substantial" to "near perfect" agreement. In 1 (2.8%) hospital, the PAK value indicates "moderate" agreement; in 2 (5.6%) hospitals, there was "fair" agreement; in 1 (2.8%) hospital there was "slight" agreement; and in 1 (2.8%) hospital, the PAK value indicates "poor" agreement between expert and non-expert ratings on *Any prevention within the last 24 hours*. Percent agreement ranged from 13.26% – 100% across hospitals for this measure.

Prevention measure – Skin assessment documented within the last 24 hours. Cohen's kappa values for the PrU prevention measure, *Skin assessment documented within the last 24 hours* (Table 20), ranged from .000 "slight" to .501 "moderate" agreement between the expert and non-expert raters. For 29 (80.6%) hospitals, Cohen's kappa values indicate "moderate"

Table 19.

Aim 1 Question 1: ***"Any Prevention within the Last 24 hours"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	-.133 (.046)	-.089 (.012)	13.26 (6.97)
2	2	.500 (0)	1.00 (0)	100 (0)
3	2	.500 (0)	1.00 (0)	100 (0)
4	3	.333 (.289)	.746 (.441)	80.00 (34.64)
5	2	.500 (0)	1.00 (0)	100 (0)
6	2	.500 (0)	1.00 (0)	100 (0)
7	3	.334 (.288)	.976 (.041)	97.41 (4.48)
8	3	.848 (.263)	.902 (.170)	93.33 (11.55)
9	2	.500 (0)	1.00 (0)	100 (0)
10	2	.500 (0)	1.00 (0)	100 (0)
11	2	.500 (0)	1.00 (0)	100 (0)
12	3	.500 (0)	1.00 (0)	100 (0)
13	3	.500 (0)	1.00 (0)	100 (0)
14	2	.500 (0)	1.00 (0)	100 (0)
15	2	.500 (0)	1.00 (0)	100 (0)
16	2	.500 (0)	1.00 (0)	100 (0)
17	2	.380 (.095)	.858 (.001)	90.48 (0.00)
18	3	.063 (.188)	.303 (.190)	54.86 (9.84)
19	2	.855 (.205)	.932 (.096)	95.46 (6.43)
20	2	0 (0)	.777 (0.00)	85.00 (0.00)
21	2	.171 (.168)	.369 (.165)	54.17 (17.68)
22	2	.500 (0)	1.00 (0)	100 (0)
23	3	.500 (0)	1.00 (0)	100 (0)
24	4	.500 (0)	1.00 (0)	100 (0)
25	2	.500 (0)	1.00 (0)	100 (0)
26	3	.500 (0)	1.00 (0)	100 (0)
27	2	.001 (0.00)	.917 (0.00)	94.44 (0.00)
28	3	.167 (.289)	.835 (.143)	88.89 (9.62)
29	2	.500 (0)	1.00 (0)	100 (0)
30	2	0 (0)	.180 (.081)	32.50 (10.61)
31	3	.107 (.189)	.524 (.180)	67.46 (13.35)
32	2	.250 (.354)	.956 (.062)	97.06 (4.16)
33	2	.500 (0)	1.00 (0)	100 (0)
34	2	.250 (.354)	.763 (.335)	83.34 (23.57)
35	2	.319 (.451)	.869 (.062)	91.18 (4.16)
36	2	.500 (0)	1.00 (0)	100 (0)
Total Expert to Non-Expert pairs = 84				

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

Table 20.

Aim 1 Question 1: ***"Skin Assessment Documented within the Last 24 hours"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.250 (.354)	.644 (.503)	72.73 (38.57)
2	2	.500 (0)	1.00 (0)	100 (0)
3	2	.500 (0)	1.00 (0)	100 (0)
4	3	.334 (.288)	.944 (.096)	97.92 (3.61)
5	2	.500 (0)	1.00 (0)	100 (0)
6	2	.250 (.354)	.886 (.161)	92.31 (10.88)
7	3	.500 (0)	1.00 (0)	100 (0)
8	3	.500 (0)	1.00 (0)	100 (0)
9	2	.500 (0)	1.00 (0)	100 (0)
10	2	.500 (0)	1.00 (0)	100 (0)
11	2	.500 (0)	1.00 (0)	100 (0)
12	3	.500 (0)	1.00 (0)	100 (0)
13	3	.500 (0)	1.00 (0)	100 (0)
14	2	.500 (0)	1.00 (0)	100 (0)
15	2	.500 (0)	1.00 (0)	100 (0)
16	2	.500 (0)	1.00 (0)	100 (0)
17	2	.501 (.706)	.965 (.050)	97.62 (3.37)
18	3	.500 (0)	1.00 (0)	100 (0)
19	2	.500 (0)	1.00 (0)	100 (0)
20	2	.500 (0)	1.00 (0)	100 (0)
21	2	.251 (.353)	.970 (.042)	98.00 (2.83)
22	2	.500 (0)	1.00 (0)	100 (0)
23	3	.500 (0)	1.00 (0)	100 (0)
24	4	.500 (0)	1.00 (0)	100 (0)
25	2	.500 (0)	1.00 (0)	100 (0)
26	3	.500 (0)	1.00 (0)	100 (0)
27	2	.500 (0)	1.00 (0)	100 (0)
28	3	.333 (.289)	.918 (.143)	94.44 (9.62)
29	2	.500 (0)	1.00 (0)	100 (0)
30	2	0 (0)	.107 (.022)	22.50 (3.54)
31	3	.334 (.288)	.972 (.048)	98.15 (3.21)
32	2	.500 (0)	1.00 (0)	100 (0)
33	2	.500 (0)	1.00 (0)	100 (0)
34	2	.500 (0)	1.00 (0)	100 (0)
35	2	.500 (0)	1.00 (0)	100 (0)
36	2	.500 (0)	1.00 (0)	100 (0)
Total Expert to Non-Expert pairs = 84				

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

agreement. In 6 (16.7%) hospitals, Cohen's kappa values indicate "fair" agreement; and in 1 (2.8%) hospital, there was "slight" agreement between the expert and non-expert raters.

Prevalence-adjusted kappa (PAK) values ranged from .107 "slight" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 35 (97.2%) hospitals, PAK values indicate "substantial" to "near perfect" agreement; and in 1 (2.8%) hospital, the PAK value indicates "slight" agreement between expert and non-expert ratings on *Skin assessment documented within the last 24 hours*. Percent agreement ranged from 22.50% – 100% across hospitals for this measure.

Prevention measure - Pressure-redistribution surface use within the last 24 hours. Cohen's kappa values for the PrU prevention measure, *Pressure-redistribution surface use within the last 24 hours* (Table 21), ranged from .002 "slight" to .828 "near perfect" agreement between the expert and non-expert raters. For 1 (2.8%) hospital, the Cohen's kappa value indicates "near perfect" agreement. In 17 (47.2%) hospitals, Cohen's kappa values indicate "moderate" agreement; in 8 (22.2%) hospitals, there was "fair" agreement; and in 10 (27.8%) hospitals, Cohen's kappa values indicate "slight" agreement between expert and non-expert ratings on *Pressure-redistribution surface use within the last 24 hours*.

Prevalence-adjusted kappa (PAK) values ranged from .252 "fair" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 28 (77.8%) hospitals, PAK values indicate "substantial" to "near perfect" agreement. In 5 (13.9%) hospitals, PAK values indicate "moderate" agreement; and in 3 (8.3%) hospitals, there was "fair" agreement between expert and non-expert ratings on *Pressure-redistribution surface use within the last 24 hours*. Percent agreement ranged from 29.37% – 100% across hospitals for this measure.

Table 21.

Aim 1 Question 1: ***"Pressure-Redistribution Surface Use within the Last 24 hours"****—Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.500 (0)	1.00 (0)	100 (0)
2	2	.500 (0)	1.00 (0)	100 (0)
3	2	.500 (0)	1.00 (0)	100 (0)
4	3	.500 (0)	1.00 (0)	97.92 (3.61)
5	2	.500 (0)	1.00 (0)	100 (0)
6	2	.002 (0.00)	.904 (0.00)	92.31 (0.00)
7	3	.167 (.288)	.795 (.234)	83.13 (19.42)
8	3	.500 (0)	1.00 (0)	100 (0)
9	2	.251 (.353)	.905 (.135)	92.31 (10.88)
10	2	.040 (.053)	.522 (.384)	58.33 (35.36)
11	2	.251 (.353)	.917 (.117)	93.34 (9.43)
12	3	.500 (0)	1.00 (0)	100 (0)
13	3	.500 (0)	1.00 (0)	100 (0)
14	2	.500 (0)	1.00 (0)	100 (0)
15	2	.376 (.040)	.517 (.062)	60.81 (5.73)
16	2	.500 (0)	1.00 (0)	100 (0)
17	2	.251 (.352)	.970 (.042)	97.62 (3.37)
18	3	.500 (0)	1.00 (0)	100 (0)
19	2	.390 (.031)	.607 (.018)	68.34 (2.35)
20	2	.368 (.237)	.624 (.097)	70.00 (7.07)
21	2	.190 (.203)	.284 (.206)	36.00 (22.63)
22	2	.500 (0)	1.00 (0)	100 (0)
23	3	.500 (0)	1.00 (0)	100 (0)
24	4	.006 (.007)	.501 (.475)	53.22 (46.08)
25	2	.500 (0)	1.00 (0)	100 (0)
26	3	.167 (.289)	.511 (.426)	56.67 (37.86)
27	2	.002 (0.00)	.934 (0.00)	94.74 (0.00)
28	3	.168 (.288)	.862 (.120)	88.89 (9.62)
29	2	.500 (0)	1.00 (0)	100 (0)
30	2	.159 (.058)	.362 (.071)	45.00 (7.07)
31	3	.103 (.091)	.252 (.345)	29.37 (36.53)
32	2	.251 (.353)	.891 (.155)	91.08 (12.47)
33	2	.251 (.352)	.974 (.037)	97.92 (2.95)
34	2	.500 (0)	1.00 (0)	100 (0)
35	2	.828 (.082)	.890 (.052)	91.18 (4.16)
36	2	.500 (0)	1.00 (0)	100 (0)

Total Expert to Non-Expert pairs = 84

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

Prevention measure – Routine repositioning as prescribed within the last 24 hours. Cohen's kappa values for the PrU prevention measure, *Routine repositioning as prescribed within the last 24 hours* (Table 22), ranged from -.242 "poor" to .750 "substantial" agreement between the expert and non-expert raters. For 4 (11.1%) hospitals, Cohen's kappa values indicate "substantial" agreement. In 4 (11.1%) hospitals, Cohen's kappa values indicate "moderate" agreement; in 6 (16.7%) hospitals, there was "fair" agreement; in 19 (52.8%) hospitals, there was "slight" agreement; and in 3 (8.3%) hospitals, Cohen's kappa values indicate "poor" agreement between the expert and non-expert raters.

Prevalence-adjusted kappa (PAK) values ranged from .069 "slight" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 16 (44.4%) hospitals, PAK values indicate "substantial" to "near perfect" agreement. In 10 (27.8%) hospitals, PAK values indicate "moderate" agreement; in 8 (22.2%) hospitals, there was "fair" agreement; and in 2 (5.6%) hospitals, PAK values indicate "slight" agreement between expert and non-expert ratings on *Routine repositioning as prescribed within the last 24 hours*. Percent agreement ranged from 10.00% – 100% across hospitals for this measure.

Prevention measure – Nutritional support within the last 24 hours. Cohen's kappa values for the PrU prevention measure, *Nutritional support within the last 24 hours* (Table 23), ranged from -.050 "poor" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 1 (2.8%) hospital, the Cohen's kappa value indicates "near perfect" agreement; and in 2 (5.6%) hospitals, there was "substantial" agreement. In 4 (11.1%) hospitals, Cohen's kappa values indicate "moderate" agreement; in 11 (30.6%) hospitals, there was "fair" agreement; in 15 (41.7%) hospitals, there was "slight" agreement between expert and non-expert ratings.

Table 22.

Aim 1 Question 1: ***"Routine Repositioning as Prescribed within the Last 24 hours"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.047 (.167)	.412 (.102)	51.89 (9.11)
2	2	.001 (.001)	.644 (.327)	70.00 (28.28)
3	2	.197 (.076)	.375 (.088)	47.78 (11.00)
4	3	.188 (.099)	.464 (.073)	56.79 (4.98)
5	2	.001 (0.00)	.650 (0.00)	71.43 (0.00)
6	2	.010 (.185)	.122 (.160)	30.77 (10.88)
7	3	.051 (.087)	.574 (.148)	64.87 (13.45)
8	3	.681 (.276)	.815 (.163)	85.00 (13.23)
9	2	.002 (.001)	.874 (.042)	99.91 (3.40)
10	2	-.242 (.061)	.444 (.171)	58.34 (11.79)
11	2	.377 (.457)	.672 (.230)	73.34 (18.86)
12	3	0 (0)	.254 (.101)	33.33 (11.55)
13	3	-.188 (.184)	.244 (.195)	42.86 (14.29)
14	2	.501 (.705)	.963 (.052)	97.06 (4.16)
15	2	.130 (.031)	.290 (.029)	40.54 (0.00)
16	2	.002 (0.00)	.930 (0.00)	94.44 (0.00)
17	2	.380 (.095)	.881 (.001)	90.48 (0.00)
18	3	.296 (.273)	.522 (.270)	59.72 (25.71)
19	2	.664 (.131)	.834 (.058)	86.67 (4.72)
20	2	.264 (.136)	.596 (.220)	67.50 (17.68)
21	2	.669 (.289)	.723 (.236)	78.00 (19.80)
22	2	0 (0)	.709 (.045)	76.32 (3.73)
23	3	.500 (0)	1.00 (0)	100 (0)
24	4	.058 (.102)	.384 (.154)	47.26 (15.23)
25	2	.750 (.353)	1.00 (0)	100 (0)
26	3	.167 (.288)	.835 (.143)	86.67 (11.55)
27	2	.061 (.024)	.390 (.107)	52.63 (7.44)
28	3	.168 (.288)	.862 (.120)	88.89 (9.62)
29	2	.427 (.226)	.700 (.063)	81.82 (12.86)
30	2	.084 (.118)	.069 (.098)	10.00 (14.14)
31	3	.192 (.330)	.363 (.354)	49.21 (28.05)
32	2	.334 (.055)	.412 (.044)	50.00 (4.16)
33	2	.379 (.172)	.459 (.170)	56.25 (14.74)
34	2	.529 (.313)	.593 (.284)	66.67 (23.57)
35	2	-.021 (.029)	.425 (.115)	52.94 (8.32)
36	2	.107 (.067)	.300 (.054)	43.56 (2.04)
Total Expert to Non-Expert pairs = 84				

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

Table 23.

Aim 1 Question 1: ***"Nutritional Support within the Last 24 hours"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.015 (.021)	.135 (.061)	21.59 (4.82)
2	2	0 (0)	.266 (.065)	35 (7.07)
3	2	.250 (.354)	.569 (.610)	60.00 (56.57)
4	3	.020 (.059)	.152 (.110)	22.81 (12.04)
5	2	.001 (.001)	.515 (.434)	57.14 (40.40)
6	2	.340 (.265)	.511 (.161)	61.54 (10.88)
7	3	.347 (.264)	.459 (.333)	52.57 (33.31)
8	3	.388 (.235)	.585 (.239)	65.00 (21.79)
9	2	.434 (.482)	.525 (.451)	59.14 (40.11)
10	2	.001 (.001)	.297 (.420)	33.34 (47.14)
11	2	.015 (.021)	.341 (.358)	40.00 (37.72)
12	3	.333 (.289)	.771 (.397)	80.00 (34.64)
13	3	-.040 (.169)	.072 (.173)	23.81 (16.49)
14	2	.001 (0.00)	.842 (.016)	87.30 (1.33)
15	2	.271 (.106)	.454 (.105)	55.41 (9.55)
16	2	.100 (0.00)	.595 (0.00)	66.67 (0.00)
17	2	.393 (.141)	.645 (.081)	71.43 (6.73)
18	3	.362 (.324)	.685 (.087)	74.31 (7.31)
19	2	1.00 (.001)	1.00 (0)	100 (0)
20	2	.260 (.342)	.502 (.269)	60.00 (21.21)
21	2	-.009 (.096)	.079 (.054)	28.00 (0.00)
22	2	.001 (.001)	.773 (.136)	81.58 (11.17)
23	3	.179 (.104)	.354 (.082)	45.24 (8.25)
24	4	0 (0)	.361 (.165)	44.17 (16.50)
25	2	.566 (.093)	.956 (.063)	96.43 (5.05)
26	3	.159 (.273)	.737 (.228)	89.63 (0.64)
27	2	.522 (.004)	.600 (.001)	68.42 (0.00)
28	3	.616 (.375)	.735 (.298)	77.78 (25.46)
29	2	.666 (.471)	.780 (.311)	81.82 (25.71)
30	2	.383 (.244)	.502 (.269)	57.50 (24.75)
31	3	.119 (.063)	.343 (.091)	46.48 (8.43)
32	2	-.050 (.103)	.040 (.053)	14.71 (4.16)
33	2	.211 (.124)	.620 (.151)	69.74 (12.54)
34	2	.139 (.087)	.199 (.038)	33.33 (0.00)
35	2	.017 (.033)	.362 (.081)	47.06 (8.32)
36	2	.449 (.320)	.628 (.154)	64.21 (5.95)
Total Expert to Non-Expert pairs = 84				

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

Prevalence-adjusted kappa (PAK) values ranged from .040 "slight" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 12 (33.3%) hospitals, PAK values indicate "substantial" to "near perfect" agreement. In 11 (30.6%) hospitals, PAK values indicate "moderate" agreement; in 7 (19.4%) hospitals, there was "fair" agreement; and in 6 (16.7%) hospitals, PAK values indicate "slight" agreement between expert and non-expert ratings on *Nutritional support within the last 24 hours*. Percent agreement ranged from 14.71% – 100% across hospitals for this measure.

Prevention measure – Moisture management within the last 24 hours. Cohen's kappa values for the last PrU prevention measure, *Moisture management within the last 24 hours* (Table 24), ranged from -.111 "poor" to .848 "near perfect" agreement between the expert and non-expert raters. For 1 (2.8%) hospital, the Cohen's kappa value indicates "near perfect" agreement. In 6 (16.7%) hospitals, Cohen's kappa values indicate "moderate" agreement; in 12 (33.3%) hospitals, there was "fair" agreement; in 16 (44.4%) hospitals, there was "slight" agreement; and in 1 (2.8%) hospital, the Cohen's kappa value indicates "poor" agreement between the expert and non-expert raters.

Prevalence-adjusted kappa (PAK) values ranged from .000 "slight" to 1.0 "near perfect" agreement between the expert and non-expert raters. For 14 (38.9%) hospitals, PAK values indicate "substantial" to "near perfect" agreement. In 14 (38.9%) hospitals, PAK values indicate "moderate" agreement; in 3 (8.3%) hospitals, there was "fair" agreement; and in 5 (13.9%) hospitals, PAK values indicate "slight" agreement between expert and non-expert ratings on *Moisture management within the last 24 hours*. Percent agreement ranged from 0.00% – 100% across hospitals for this measure.

Table 24.

Aim 1 Question 1: ***"Moisture Management within the Last 24 hours"—****Number of Expert/Non-Expert Pairs, Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement by Hospital (N = 36)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.069 (.097)	.155 (.060)	21.97 (7.50)
2	2	0 (0)	.473 (.227)	55.00 (21.21)
3	2	.096 (.206)	.223 (.258)	36.11 (19.64)
4	3	0 (0)	.192 (.041)	26.65 (4.94)
5	2	.001 (.001)	.581 (.342)	64.29 (30.30)
6	2	0 (0)	.250 (0.00)	33.33 (0.00)
7	3	.275 (.344)	.758 (.144)	80.60 (11.52)
8	3	.848 (.262)	.917 (.143)	93.33 (11.55)
9	2	.013 (.226)	.588 (.250)	67.55 (19.37)
10	2	.357 (.505)	.588 (.332)	66.67 (23.57)
11	2	.125 (.177)	.671 (.002)	73.33 (0.00)
12	3	.500 (0)	1.00 (0)	100 (0)
13	3	.382 (.544)	.496 (.446)	57.14 (37.80)
14	2	.247 (.347)	.841 (.018)	87.30 (1.33)
15	2	.143 (.243)	.424 (.169)	51.35 (15.29)
16	2	.500 (0)	1.00 (0)	100 (0)
17	2	.280 (.042)	.537 (.082)	61.91 (6.74)
18	3	.429 (.245)	.638 (.156)	71.53 (11.85)
19	2	.573 (.002)	.771 (.031)	81.67 (2.35)
20	2	.078 (.134)	.179 (.153)	32.50 (10.61)
21	2	.253 (.121)	.405 (.182)	52.00 (16.97)
22	2	.002 (.001)	.902 (.046)	92.11 (3.73)
23	3	.334 (.288)	.970 (.051)	97.62 (4.12)
24	4	.176 (.108)	.451 (.258)	52.26 (26.17)
25	2	.470 (.042)	.911 (.126)	92.86 (10.10)
26	3	.057 (.095)	.683 (.321)	72.96 (28.55)
27	2	.248 (.169)	.476 (.093)	57.90 (7.45)
28	3	.168 (.288)	.862 (.120)	88.89 (9.62)
29	2	-.111 (.158)	.471 (.429)	59.09 (32.15)
30	2	0 (0)	0 (0)	0 (0)
31	3	.268 (.266)	.458 (.372)	53.97 (34.03)
32	2	.316 (.126)	.400 (.129)	50.00 (12.47)
33	2	.461 (.083)	.633 (.075)	70.84 (5.89)
34	2	.143 (0.00)	.181 (.098)	25.00 (11.78)
35	2	.308 (.033)	.444 (.039)	55.88 (4.16)
36	2	.313 (.250)	.522 (.211)	61.84 (16.74)
Total Expert to Non-Expert pairs = 84				

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

Question 2. What is the *overall* agreement between expert participant ratings and non-expert participant ratings for the NDNQI PrU *risk* measures per hospital, and the *overall* agreement between expert participant ratings and non-expert participant ratings for the NDNQI PrU *prevention* measures per hospital? Average Cohen's kappa value for the 5 *risk* measures (*Skin assessment within 24 hours of admission, Risk assessment within 24 hours of admission, Time since last PrU risk assessment, Risk assessment scale, and Risk status*) are presented by hospital in Table 25, as are the average PAK value and the average percent agreement for these *risk* measures. Likewise, the average Cohen's kappa value for all 6 *prevention* measures (*Any prevention, Skin assessment documented, Pressure-redistribution surface use, Routine repositioning as prescribed, Nutritional support, and Moisture management*) are presented in Table 26 by hospital, as are the average PAK value and the average percent agreement for these *prevention* measures.

Overall agreement for the (five) PrU risk measures. Cohen's kappa values representing the overall agreement between the expert and non-expert raters on the five NDNQI PrU *risk measures* (Table 25) ranged from "fair" (.294) to "substantial" (.760) agreement, (M = .498, 95% CI [.463 – .533]). Specifically, for 6 (16.7%) hospitals, Cohen's kappa values ranged from .608 to .760, indicating "substantial" agreement; for 23 (63.9%) hospitals, Cohen's kappa values ranged from .422 to .591, indicating "moderate" agreement; and for 7 (19.4%) hospitals, Cohen's kappa values ranged from .294 to .387, indicating "fair" agreement.

Prevalence-adjusted kappa (PAK) values representing the overall agreement on the five NDNQI PrU *risk* measures ranged from .812 – .994, (M = .924, 95% CI [.910 – .939]), indicating "near perfect" agreement between expert and non-expert ratings on this *overall risk*

Table 25.

*Aim 1 Question 2: Overall Agreement for the Five Pressure Ulcer **Risk Measures** by Hospital (N = 36 Hospitals)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.380 (.347)	.926 (.083)	94.05 (7.67)
2	2	.494 (.342)	.923 (.151)	93.67 (13.58)
3	2	.471 (.341)	.842 (.327)	86.52 (28.35)
4	3	.453 (.154)	.912 (.119)	93.48 (8.63)
5	2	.294 (.252)	.849 (.197)	88.54 (15.65)
6	2	.342 (.309)	.812 (.213)	85.53 (18.24)
7	3	.522 (.049)	.927 (.164)	95.10 (10.95)
8	3	.760 (.274)	.937 (.124)	94.84 (10.44)
9	2	.323(.204)	.863 (.238)	90.00 (16.89)
10	2	.445 (.361)	.952 (.039)	96.52 (2.92)
11	2	.580 (.260)	.957 (.070)	96.53 (6.10)
12	3	.631 (.126)	.962 (.048)	97.46 (3.19)
13	3	.546 (.222)	.898 (.197)	91.03 (17.99)
14	2	.493 (.163)	.947 (.112)	96.38 (7.51)
15	2	.386 (.390)	.975 (.031)	98.20 (2.05)
16	2	.629 (.189)	.979 (.030)	98.40 (2.19)
17	2	.422 (.353)	.838 (.204)	86.74 (19.34)
18	3	.614 (.247)	.901 (.153)	92.74 (10.70)
19	2	.468 (.229)	.930 (.063)	94.56 (5.33)
20	2	.518 (.251)	.947 (.043)	96.09 (3.49)
21	2	.343 (.367)	.931 (.045)	95.09 (3.26)
22	2	.724 (.219)	.958 (.071)	96.00 (5.52)
23	3	.608 (.236)	.907 (.121)	92.71 (9.94)
24	4	.548 (.207)	.907 (.149)	92.56 (12.66)
25	2	.491 (.335)	.989 (.015)	99.58 (0.93)
26	3	.461 (.129)	.942 (.090)	95.95 (6.01)
27	2	.591 (.204)	.994 (.014)	99.59 (0.91)
28	3	.491 (.348)	.895 (.228)	91.30 (18.95)
29	2	.538 (.251)	.898 (.153)	91.80 (13.16)
30	2	.427 (.249)	.917 (.101)	93.60 (8.44)
31	3	.379 (.303)	.908 (.120)	92.71 (10.36)
32	2	.587 (.194)	.991 (.020)	99.40 (1.34)
33	2	.464 (.302)	.904 (.140)	92.34 (12.01)
34	2	.429 (.273)	.917 (.115)	93.56 (9.24)
35	2	.585 (.183)	.978 (.030)	98.37 (2.23)
36	2	.488 (.332)	.949 (.091)	95.8 (7.82)

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

Table 26.

*Aim 1 Question 2: Overall Agreement for the Six Pressure Ulcer **Prevention Measures** by Hospital (N = 36 Hospitals)*

Hospital	# Of Expert to Non-Expert Pairs	Cohen's kappa mean (SD)	PAK mean (SD)	Percent Agreement mean (SD)
1	2	.125 (.221)	.376 (.396)	46.91 (34.37)
2	2	.250 (.274)	.730 (.319)	76.67 (27.87)
3	2	.340 (.182)	.694 (.352)	73.98 (29.49)
4	3	.229 (.197)	.583 (.370)	63.68 (33.76)
5	2	.251 (.273)	.791 (.233)	82.14 (20.08)
6	2	.184 (.213)	.612 (.372)	68.38 (31.10)
7	3	.279 (.155)	.760 (.215)	79.77 (18.41)
8	3	.628 (.195)	.870 (.156)	89.44 (13.19)
9	2	.283 (.232)	.815 (.208)	84.82 (17.33)
10	2	.193 (.305)	.637 (.295)	69.44 (26.18)
11	2	.295 (.200)	.767 (.257)	80.00 (23.09)
12	3	.389 (.202)	.837 (.300)	85.56 (26.81)
13	3	.276 (.309)	.635 (.422)	70.64 (33.86)
14	2	.375 (.209)	.941 (.079)	95.28 (6.28)
15	2	.320 (.166)	.614 (.308)	68.02 (25.65)
16	2	.350 (.234)	.921 (.162)	93.51 (13.34)
17	2	.364 (.089)	.809 (.178)	84.92 (14.81)
18	3	.358 (.165)	.691 (.273)	76.74 (19.41)
19	2	.663 (.228)	.857 (.153)	88.69 (12.40)
20	2	.245 (.184)	.613 (.275)	69.17 (22.95)
21	2	.254 (.225)	.472 (.321)	57.70 (26.21)
22	2	.250 (.273)	.897 (.129)	91.67 (10.45)
23	3	.419 (.135)	.887 (.262)	90.48 (22.18)
24	4	.207 (.236)	.616 (.301)	66.15 (26.43)
25	2	.548 (.104)	.978 (.037)	98.21 (2.99)
26	3	.258 (.192)	.794 (.191)	84.32 (16.85)
27	2	.222 (.241)	.719 (.263)	78.02 (20.85)
28	3	.270 (.182)	.846 (.606)	87.96 (5.46)
29	2	.414 (.269)	.825 (.217)	87.12 (16.37)
30	2	.104 (.151)	.203 (.192)	27.92 (21.53)
31	3	.187 (.096)	.485 (.256)	57.44 (23.43)
32	2	.267 (.180)	.616 (.390)	67.16 (34.32)
33	2	.384 (.127)	.781 (.239)	82.46 (19.17)
34	2	.343 (.187)	.623 (.369)	68.06 (32.67)
35	2	.325 (.316)	.665 (.284)	73.04 (23.49)
36	2	.395 (.159)	.742 (.302)	78.27 (24.86)

Note. SD = standard deviation; PAK = prevalence-adjusted kappa

measure. Percent agreement ranged from 85.5% - 99.6%, (*M* = 94.1%, 95% CI [92.9% - 95.3%]) across hospitals on this measure.

Overall agreement for the (six) PrU prevention measures. Whether considering Cohen's kappa, PAK, or percent agreement; overall agreement for the *prevention* measures was uniformly lower than the overall agreement for the *risk* measures. Cohen's kappa values representing the overall agreement between the expert and non-expert raters on the six NDNQI PrU *prevention* measures (Table 26) ranged from "slight" (.104) to "substantial" (.628), (*M* = .312, 95% CI [.273 – .352]). Specifically, for 2 (5.6%) hospitals, Cohen's kappa values ranged from .628 to .633 indicating "substantial" agreement; for 3 (8.3%) hospitals, Cohen's kappa values ranged from .414 to .548, indicating "moderate" agreement; for 26 (72.2%%) hospitals, Cohen's kappa values ranged from .207 to .395, indicating "fair" agreement; and for 5 (13.9%) hospitals, Cohen's kappa values are .104 or .187, indicating "slight" agreement between expert and non-expert ratings on this *overall prevention* measure.

Prevalence-adjusted kappa (PAK) values for the overall agreement on the six NDNQI PrU *prevention* measures ranged from "slight" (.203) to "near perfect" (.978) agreement between expert and non-expert ratings on this *overall prevention* measure, (*M* = .714, 95% CI [.661 – .767]). Specifically, for 12 (33.3%) hospitals, PAK ranged from .809 to .978, indicating "near perfect" agreement; for 19 (52.8%) hospitals, PAK ranged from .612 to .794, indicating "substantial" agreement; for 3 (8.3%) hospitals, PAK ranged from .472 to .583, indicating "moderate" agreement; for 1 (2.8%) hospital, the PAK was .376, indicating "fair" agreement; and for 1 (2.8%) hospital, the PAK was .203, indicating "slight" agreement between expert and non-expert ratings on this *overall prevention* measure. Percent agreement ranged from 27.9% - 98.2%, (*M* = 76.2%, 95% CI [71.5% - 80.9%]) across hospitals for this measure.

Question 3. What is the *average* of the *within hospital* agreements between expert participant ratings and non-expert participant ratings for each of the 11 NDNQI pressure ulcer risk and prevention measures *across* hospitals? The average Cohen's kappa value, the average PAK value, and the average percent agreement across hospitals for each of the risk and prevention measures are presented in Table 27. The average Cohen's kappa value ranged from .216 for *Routine Repositioning as Prescribed within the Last 24 hours*, indicating "fair" agreement, to .819 for *Risk Status*, indicating "near perfect" agreement. For 10 of the 11 measures, however, Cohen's kappa values indicate only "fair", $n = 7$ (63.6%), or "moderate", $n = 3$ (27.3%) agreement.

The average PAK value ranged from .500 for *Nutritional Support within the Last 24 hours*, indicating "moderate" agreement, to .997 for *Risk Assessment Scale Used*, indicating "near perfect" agreement. Specifically, for 7 (63.6%) of the NDNQI PrU risk and prevention measures, PAK values ranged from .839 to .997, indicating "near perfect" agreement; for 1 (9.1%) measure, the PAK value was .790, indicating "substantial" agreement; and for 3 (27.3%%) measures, PAK values ranged from .500 to .577, indicating "moderate" agreement between expert and non-expert ratings on each of these risk and prevention measure. The mean percent agreement for each measure across hospitals ranged from 57.6% to 99.8%.

Question 4. What is the intraclass correlation coefficient (agreement) between expert participant ratings and non-expert participant ratings for each of the 11 NDNQI PrU risk and prevention measures *across* hospitals? The *ICC*s for Cohen's kappa values across hospitals, and the *ICC*s for PAK values across hospitals, are presented in Table 28. For this study, *ICC* = the variability of kappa values between hospitals, divided by the sum of the variability of kappa values between hospitals and the variability of kappa values within hospitals.

Table 27.

*Aim 1 Question 3: Mean Cohen's Kappa, Prevalence-Adjusted Kappa, and Percent Agreement for Each **Pressure Ulcer Risk and Prevention** Measure (N = 11 Measures) across Hospitals (N = 36 Hospitals)*

	Cohen's kappa [95% CI]	PAK [95% CI]	Percent Agreement [95% CI]
Risk Measures			
Skin Assessment within 24 hours of Admission	.379 [.308 – .451]	.977 [.966 – .989]	98.48 [97.70 – 99.27]
Risk Assessment within 24 hours of Admission	.472 [.393 - .552]	.978 [.964 - .993]	98.58 [97.65 – 99.52]
Time since Last PrU Risk Assessment	.324 [.241 - .406]	.790 [.729 - .852]	81.79 [76.52 – 87.06]
Risk Assessment Scale	.495 [.485 - .505]	.997 [.991 – 1.00]	99.78 [99.33 – 100]
Risk Status	.819 [.766 - .873]	.877 [.838 - .917]	91.75 [88.98 – 94.51]
Prevention Measures (in Use within the Last 24 hours)			
Any Prevention	.387 [.315 - .460]	.856 [.769 - .943]	89.41 [82.75 – 96.07]
Skin Assessment Documented	.451 [.415 - .488]	.956 [.904 – 1.00]	96.49 [92.07 – 100]
Pressure-Redistribution Surface Use	.353 [.289 - .417]	.839 [.763 - .916]	86.06 [79.27 – 92.85]
Routine Repositioning as Prescribed	.216 [.134 - .298]	.577 [.494 - .661]	65.36 [58.21 – 72.51]
Nutritional Support	.235 [.155 - .315]	.500 [.418 - .581]	57.59 [50.35 – 64.83]
Moisture Management	.231 [.164 - .298]	.556 [.469 - .643]	62.37 [54.44 – 70.31]

Note. *k* = kappa; PAK = prevalence-adjusted kappa

Table 28.

Aim 1 Question 4: Intraclass Correlation Coefficients for Cohen's Kappa and Prevalence-Adjusted Kappa Values across Hospitals for ***Pressure Ulcer Risk and Prevention*** *Measures (N = 72 Cohen's Kappa Values; N = 72 PAK Values)*

	ICC [95% CI] from Cohen's *k*	*ICC* [95% CI] from PAK
Risk Measures		
Skin Assessment within 24 hours of Admission	.373 [.059 – .622]	.399 [.089 – .640]
Risk Assessment within 24 hours of Admission	.573 [.309 – .756]	.295 [-.030 – .564]
Time since Last PrU Risk Assessment	.298 [-.027 – .566]	.422 [.155 – .655]
Risk Assessment Scale	-.513 [-.716 – -.228]	.877 [.773 – .935]
Risk Status	.510 [.226 – .716]	.511 [.226 – .716]
Prevention Measures (in Use within the Last 24 hours)		
Any Prevention	.539 [.263 – .734]	.800 [.644 – .892]
Skin Assessment Documented	-.052 [-.368 – .276]	.718 [.516 – .845]
Pressure-Redistribution Surface Use	.396 [.085 – .637]	.610 [.359 – .779]
Routine Repositioning as Prescribed	.453 [.153 – .677]	.713 [.508 – .842]
Nutritional Support	.525 [.245 – .725]	.426 [.121 – .659]
Moisture Management	.259 [-.068 – .537]	.620 [.373 – .786]

Note. *ICC* = Intraclass correlation coefficient; *k* = kappa; PAK = prevalence-adjusted kappa

This formula suggests that an *ICC* near zero (< .22) indicates that the *within*-hospital variance in rater agreement is much greater than the *between*-hospital variance in rater agreement. Support for this conclusion is found in Hart et al. (2006) who also defined "Near zero" as < .22. This aligns with the level of agreement for "poor" to "slight" (< 0 to .20) by Landis and Koch (1977). For this study, an *ICC* near zero means each hospital rated the NDNQI PrU risk and prevention measures with similar levels of agreement as the other hospitals rated the measures.

Intraclass correlation coefficients (*ICC*s) calculated from Cohen's kappa values ranged from -.513 to .573. Only 2 of the 11 *ICC*s (*Risk assessment scale*, and *Skin assessment documented within the last 24 hours*) were near zero; i.e. < .22. Intraclass correlation coefficients (*ICC*s) calculated from PAK values ranged from .295 to .800. None of the measures' *ICC*s (calculated from PAK values) are considered to be at or near zero.

Question 5. Where is the lack of agreement between expert and non-expert participant ratings on the 11 NDNQI pressure ulcer risk and prevention measures occurring? Data were aggregated to present one matrix (i.e. frequency table) for each NDNQI PrU risk and prevention measure. Cell counts along the diagonal indicate agreement between the expert and the non-expert rater, while cell counts off the diagonal represent disagreement. To help identify areas of confusion, the cell(s) of disagreement with the largest counts in the table have been circled (Table 29).

Some of the cells in these frequency tables are sparsely populated. While logical because most patients received the intervention, this sparseness provided support for the addition of 0.0001 to each cell count, in order to compute Cohen's kappa and PAK values. Five of the 11 tables have at least 40% of their cells with a cell count of zero. For example, the matrix for *Time*

Table 29.

Aim 1 Question 5: Agreement Matrices—Agreement between the Expert (N = 36) and Non-Expert (N = 84) Raters for Each of the 11 NDNQI ***Pressure Ulcer Risk and Prevention*** *Measures*

Risk Measures

Skin Assessment within 24 hours of Admission

		Non-Expert			
		Yes	No	Pending	
Expert	Yes	3712	20	4	3736
	No	22	3	2	27
	Pending	4	0	6	10
		3738	23	12	= 3773

Risk Assessment within 24 hours of Admission

		Non-Expert			
		Yes	No	Pending	
Expert	Yes	3703	16	3	3722
	No	15	12	4	31
	Pending	1	1	11	13
		3719	29	18	= 3766

Table 29 (continued).

Aim 1 Question 5: Agreement Matrices—Agreement between the Expert (N = 36) and Non-Expert (N = 84) Raters for Each of the 11 NDNQI ***Pressure Ulcer Risk and Prevention*** *Measures*

Time since Last PrU Risk Assessment

		Non-Expert							
		>0-12	>12-24	>24-48	>48-72	>72-1wk	>1 wk	Never	
Expert	>0-12	2689	290	8	0	0	0	11	2998
	>12-24	285	370	8	0	0	0	5	668
	>24-48	11	43	29	1	0	0	0	84
	>48-72	0	0	2	2	0	0	0	4
	>72-1wk	0	0	0	0	0	0	0	0
	>1wk	0	0	0	0	0	0	0	0
	Never	7	0	0	0	0	0	14	21
		2992	703	47	3	0	0	30	= 3775

Risk Assessment Scale

		Non-Expert				
		Braden	Norton	Other	Clinical Factors	
Expert	Braden	3741	0	0	0	3741
	Norton	0	0	0	0	0
	Other	0	0	0	0	0
	Clinical Factors	8	0	0	3	11
		3749	0	0	3	= 3752

Table 29 (continued).

Aim 1 Question 5: Agreement Matrices—Agreement between the Expert (N = 36) and Non-Expert (N = 84) Raters for Each of the 11 NDNQI ***Pressure Ulcer Risk and Prevention*** *Measures*

Risk Status

		Non-Expert: YRA	Non-Expert: YCL	Non-Expert: No	
Expert: YRA		985	49	52	1086
Expert: YCL		57	73	46	176
Expert: No		56	70	2353	2479
		1098	192	2451	= 3741

Prevention Measures

Any Prevention provided within the Last 24 hours

		Non-Expert: Yes	Non-Expert: No	Non-Expert: Pending	
Expert: Yes		1012	42	2	1056
Expert: No		69	33	5	107
Expert: Pending		0	1	0	1
		1081	76	7	= 1164

Skin Assessment Documented within the Last 24 hours

		Non-Expert: Yes	Non-Expert: No	Non-Expert: Contra.	
Expert: Yes		1137	12	0	1149
Expert: No		1	1	0	2
Expert: Contra.		7	0	0	7
		1145	13	0	= 1158

Note. YRA = Yes risk assessment; YCL = Yes clinical factors; Contra. = Contraindicated

Table 29 (continued).

Aim 1 Question 5: Agreement Matrices—Agreement between the Expert (N = 36) and Non-Expert (N = 84) Raters for Each of the 11 NDNQI ***Pressure Ulcer Risk and Prevention*** *Measures*

Pressure-Redistribution Surface Use within the Last 24 hours

		Non-Expert					
		Yes	No	Contra.	Unnecessary	Pt. Refused	
Expert	Yes	858	47	0	26	0	931
	No	28	75	0	41	0	144
	Contraindicated	2	0	0	0	0	2
	Unnecessary	7	39	0	36	0	82
	Pt. Refused	4	1	0	1	0	6
		899	162	0	104	0	= 1165

Routine Repositioning as Prescribed within the Last 24 hours

		Non-Expert					
		Yes	No	Contra.	Unnecessary	Pt. Refused	
Expert	Yes	537	80	1	65	6	689
	No	116	123	1	26	0	266
	Contraindicated	7	0	0	1	1	9
	Unnecessary	64	43	0	84	0	191
	Pt. Refused	8	1	0	0	1	10
		732	247	2	176	8	= 1165

Note. Contra. = Contraindicated; Pt. = Patient; Unnecessary = Unnecessary for the patient

Table 29 (continued).

Aim 1 Question 5: Agreement Matrices—Agreement between the Expert (N = 36) and Non-Expert (N = 84) Raters for Each of the 11 NDNQI ***Pressure Ulcer Risk and Prevention*** *Measures*

Nutritional Support within the Last 24 hours

		Non-Expert					
		Yes	No	Contra.	Unnecessary	Pt. Refused	
Expert	Yes	438	121	5	90	0	654
	No	55	94	8	41	0	198
	Contraindicated	36	21	9	11	0	77
	Unnecessary	44	65	1	121	1	232
	Pt. Refused	0	1	0	1	0	2
		573	302	23	264	1	= 1163

Moisture Management within the Last 24 hours

		Non-Expert					
		Yes	No	Contra.	Unnecessary	Pt. Refused	
Expert	Yes	588	74	0	89	0	751
	No	64	46	0	30	0	140
	Contraindicated	19	1	1	4	0	25
	Unnecessary	94	48	1	99	1	243
	Pt. Refused	0	0	0	0	0	0
		765	169	2	222	1	= 1159

Note. Contra. = Contraindicated; Pt. = Patient; Unnecessary = Unnecessary for the patient

since last PrU risk assessment reveals that while collectively rating 1,637 patients, 36 experts and 84 non-experts *never selected* the options "> 72 – 1 week", or "> 1 week".

Aim 2

Aim 2 was to examine the methods and processes used by participant raters to gather data on the NDNQI PrU risk and prevention measures. Responses to the *Pressure Ulcer Risk and Prevention Reliability Survey* (*n* = 35 items) were used to answer this question. Only the 120 participant raters (*n* = 36 experts, and *n* = 84 non-experts), whose patient data were included in the study, had their responses to the survey included in the data analysis. Results on the first three items of the survey: job title, nursing education, and certifications are described under aim 1 in Table 13 (p. 137). Findings from the remaining five survey items on participant training and education in PrU data collection, years collecting NDNQI PrU data, other roles during PrU data collection, and team leader certification status are described below. The other survey items (*n* = 27) asked participant raters how they collect NDNQI PrU risk and prevention data.

Participant Rater Characteristics (Training and education in PrU data collection, years collecting NDNQI PrU data, roles during PrU data collection, and skin team leader certification status). The majority of experts [*n* = 24 (66.7%)] and non-experts [*n* = 60 (71.4%)] reported reviewing the NDNQI Guidelines for Data Collection and Submission of PrUs within the last 12 months (Table 30). Twenty-six (72.2%) experts reported completing all four modules of the NDNQI PrU Training program within the last 12 months. Less than half of the non-experts [*n* = 40 (47.6%)] reported the same. Surprisingly, 3 (8.3%) experts and 11 (13.1%) non-experts reported that within the last 12 months they had not completed *any* of the four training modules, *nor* had they reviewed the NDNQI Guidelines for Data Collection and Submission of PrUs. These three experts (and four of these 11 non-experts) also answered "No" when asked if

Table 30.

Aim 2: Participant Rater Characteristics—Training and Experience in Collecting Pressure Ulcer Data

	Expert (*n* = 36) *n* (%)	Non-Expert (*n* = 84) *n* (%)	Total Participant Raters (*N* = 120) *n* (%)
Please select what you have completed within the last 12 months.[a] (Select all that apply.)			
I have reviewed the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers.	24 (66.7)	60 (71.4)	84 (70.0)
I have completed all 4 modules of the NDNQI Pressure Ulcer Training program.	26 (72.2)	40 (47.6)	66 (55.0)
I have completed only some (not all) of the 4 NDNQI Pressure Ulcer Training program modules.	5 (13.9)	12 (14.3)	17 (14.2)
None of the above.	3 (8.3)	11 (13.1)	14 (11.7)
Have you received education for data collection on pressure ulcers other than reviewing the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers or completing the NDNQI Pressure Ulcer Training program?[b]			
Yes	23 (63.9)	49 (58.3)	72 (60)
No	13 (36.1)	33 (39.3)	46 (38.3)

[a] Totals do not equal the number of raters because respondents were to select "all that apply".
[b] Missing values = 2 (1.7%)

Table 30 (continued).

Aim 2: Participant Rater Characteristics—Training and Experience in Collecting Pressure Ulcer Data

	Expert (n = 36) n (%)	Non-Expert (n = 84) n (%)	Total Participant Raters (N = 120) n (%)
How many years have you collected NDNQI pressure ulcer data?[c]			
< 1 year	4 (11.1)	15 (17.9)	19 (15.8)
1 year	2 (5.6)	7 (8.3)	9 (7.5)
2 years	3 (8.3)	11 (13.1)	14 (11.7)
3 years	3 (8.3)	10 (11.9)	13 (10.8)
4 years	6 (16.7)	10 (11.9)	16 (13.3)
5 years	4 (11.1)	12 (14.3)	16 (13.3)
6 years	3 (8.3)	3 (3.6)	6 (5.0)
7 years	3 (8.3)	4 (4.8)	7 (5.8)
8 years	3 (8.3)	3 (3.6)	6 (5.0)
9 years	–	1 (1.2)	1 (0.8)
10 years	2 (5.6)	3 (3.6)	5 (4.2)
11 years	1 (2.8)	3 (3.6)	4 (3.3)
12 years	–	1 (1.2)	1 (0.8)
13 years	–	–	–
14 years	1 (2.8)	–	1 (0.8)
15 years	–	–	–
> 15 years	1 (2.8)	–	1 (0.8)

[c] Missing values = 1 (0.8%)

Table 30 (continued).

Aim 2: Participant Rater Characteristics—Training and Experience in Collecting Pressure Ulcer Data

	Expert (*n* = 36) *n* (%)	Non-Expert (*n* = 84) *n* (%)	Total Participant Raters (*N* = 120) *n* (%)
In addition to chart abstractor, what other role(s) in the NDNQI Pressure Ulcer Survey have you had?[a] (Select all that apply.)			
Site Coordinator	9 (25.0)	9 (10.7)	18 (15.0)
Patient skin inspection—Rounding on all patients	26 (72.2)	56 (66.7)	82 (68.3)
Patient skin inspection—Rounding on selected patients to confirm pressure ulcer presence or stage	22 (61.1)	38 (45.2)	60 (50.0)
Training of pressure ulcer team	25 (69.4)	28 (33.3)	53 (44.2)
Data entry	22 (61.1)	29 (34.5)	51 (42.5)
No other roles	–	8 (9.5)	8 (6.7)
Other (please describe) Skin Team Leader	1 (2.8)	–	1 (0.8)
Who usually leads your NDNQI Pressure Ulcer Survey data collection team?[d]			
Someone certified in wound care	27 (75.0)	65 (77.4)	92 (76.7)
Someone not certified but the wound/skin care nurse	5 (13.9)	12 (14.3)	17 (14.2)
Neither of the above	2 (5.6)	4 (4.8)	6 (5.0)
I don't know their certification status	1 (2.8)	2 (2.4)	3 (2.5)

[d] Missing values = 2 (1.7%)

they had received other education for data collection on PrUs.

Using the value of 0.5 year for the participant raters who reported having collected NDNQI PrU data for < 1 year, and the value of 16 years for those who reported having collected NDNQI PrU data for > 15 years; the mean number of years collecting data was 5.2 (SD = 3.7) for the expert raters, and 3.9 (SD = 3.0) for the non-expert raters.

All 36 experts, and most of the non-experts, *n* = 77 (91.7%), reported having at least one other role in the NDNQI PrU Survey in addition to chart abstractor. Specifically, among the experts; 6 (16.7%) had one other role, 7 (19.4%) had two other roles, 11 (30.6%) had three other roles, 9 (25%) had four other roles, 2 (5.6%) had five other roles, and 1 (2.8%) expert rater had *six other roles* in the NDNQI PrU survey in addition to chart abstractor. Among the non-experts; 27 (32.1%) had one other role, 23 (27.6%) had two other roles, 21 (25%) had three, and 6 (7.1%) had four other roles in addition to chart abstractor. A single participant rater (0.8%) reported having a role other than those listed as a response option. This expert rater specified "Skin Team Leader". The majority of experts and non-experts reported that their NDNQI PrU Survey data collection team was led by "someone certified in wound care" [*n* = 27 (75%) experts, and *n* = 65 (77.4%) non-experts].

Methods and processes used to collect pressure ulcer risk and prevention data. All 36 experts reported that they are "almost always" or "always" comfortable with their skills to review patient records and collect NDNQI PrU risk and prevention data (Table 31). Seventy-four (88.1%) non-experts reported the same. Among all the respondents (i.e. experts and non-experts), 75 (62.5%) reported that their hospital "almost always" or "always" provides training or review prior to each PrU survey. Most respondents reported that RNs "almost always" or "always" collect PrU risk and prevention data, *n* = 111 (92.5%). One (0.8%) rater reported that

Table 31.

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

With regards to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:							
	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
1. I am comfortable with my skills to review patient records and collect NDNQI pressure ulcer risk and prevention data.							
Experts	27 (75.0)	9 (25.0)	–	–	–	–	[0]
Non-Experts	45 (53.6)	29 (34.5)	7 (8.3)	2 (2.4)	–	–	[1 (1.2)]
All Raters	72 (60.0)	38 (31.7)	7 (5.8)	2 (1.7)	–	–	[1 (0.8)]
2. Prior to each time that we collect NDNQI pressure ulcer data, our hospital provides training or reviews.							
Experts	20 (55.6)	8 (22.2)	2 (5.6)	4 (11.1)	2 (5.6)	–	[0]
Non-Experts	42 (50.0)	5 (6.0)	4 (4.8)	24 (28.6)	2 (2.4)	6 (7.1)	[1 (1.2)]
All Raters	62 (51.7)	13 (10.8)	6 (5.0)	28 (23.3)	4 (3.3)	6 (5.0)	[1 (0.8)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

With regards to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:							
	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
The people at our hospital who review patient records to collect NDNQI pressure ulcer risk and prevention data include:							
3. RNs							
Experts	34 (94.4)	1 (2.8)	1 (2.8)	–	–	–	[0]
Non-Experts	71 (84.5)	5 (6.0)	3 (3.6)	1 (1.2)	1 (1.2)	2 (2.4)	[1 (1.2)]
All Raters	105 (87.5)	6 (5.0)	4 (3.3)	1 (0.8)	1 (0.8)	2 (1.7)	[1 (0.8)]
4. LPNs							
Experts	2 (5.6)	2 (5.6)	–	4 (11.1)	15 (41.7)	2 (5.6)	[11 (30.6)]
Non-Experts	4 (4.8)	2 (2.4)	1 (1.2)	8 (9.5)	30 (35.7)	11 (13.1)	[28 (33.3)]
All Raters	6 (5.0)	4 (3.3)	1 (0.8)	12 (10.0)	45 (37.5)	13 (10.8)	[39 (32.5)]
5. Nurse aids							
Experts	5 (13.9)	2 (5.6)	–	2 (5.6)	16 (44.4)	2 (5.6)	[9 (25.0)]
Non-Experts	7 (8.3)	2 (2.4)	–	4 (2.8)	33 (39.3)	8 (9.5)	[30 (35.7)]
All Raters	12 (10.0)	4 (3.3)	–	6 (5.0)	49 (40.8)	10 (8.3)	[39 (32.5)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

With regards to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:	**Always** *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
The people at our hospital who review patient records to collect NDNQI pressure ulcer risk and prevention data include:							
6. Physical Therapists							
Experts	1 (2.8)	1 (2.8)	–	1 (2.8)	20 (55.6)	1 (2.8)	[12 (33.3)]
Non-Experts	2 (2.4)	3 (3.6)	–	2 (2.4)	38 (45.2)	7 (8.3)	[32 (38.1)]
All Raters	3 (2.5)	4 (3.3)	–	3 (2.5)	58 (48.3)	8 (6.7)	[44 (36.7)]
7. Respiratory Therapists							
Experts	–	–	–	–	21 (58.3)	1 (2.8)	[14 (38.9)]
Non-Experts	1 (1.2)	–	–	–	42 (50.0)	8 (9.5)	[33 (39.3)]
All Raters	1 (0.8)	–	–	–	63 (52.5)	9 (7.5)	[47 (39.2)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

With regards to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:							
	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
The people at our hospital who review patient records to collect NDNQI pressure ulcer risk and prevention data include:							
8. Clinical Secretaries							
Experts	–	–	–	2 (5.6)	18 (50.0)	3 (8.3)	[13 (36.1)]
Non-Experts	1 (1.2)	–	–	1 (1.2)	43 (51.2)	8 (9.5)	[31 (36.9)]
All Raters	1 (0.8)	–	–	3 (2.5)	61 (50.8)	11 (9.2)	[44 (36.7)]
9. Others than those listed above							
Experts	3 (8.3)	2 (5.6)	–	2 (5.6)	13 (36.1)	3 (8.3)	[13 (36.1)]
Non-Experts	8 (9.5)	2 (2.4)	–	–	27 (32.1)	9 (10.7)	[38 (45.2)]
All Raters	11 (9.2)	4 (3.3)	–	2 (1.7)	40 (33.3)	12 (10.0)	[51 (42.0)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

With regards to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:							
	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
10. At least once a year, we compare pressure ulcer team assessments of pressure ulcer risk and prevention to evaluate the reliability of this data.							
Expert	10 (27.8)	3 (8.3)	3 (8.3)	3 (8.3)	11 (30.6)	6 (16.7)	[0]
Non-Expert	25 (29.8)	7 (8.3)	2 (2.4)	7 (8.3)	23 (27.4)	16 (19.0)	[4 (4.8)]
All Raters	35 (29.2)	10 (8.3)	5 (4.2)	10 (8.3)	34 (28.3)	22 (18.3)	[4 (3.3)]
11. I collect data on the unit where I usually work.							
Expert	12 (33.3)	1 (2.8)	–	6 (16.7)	17 (47.2)	–	[0]
Non-Expert	33 (39.3)	8 (9.5)	6 (7.1)	12 (14.3)	22 (26.2)	1 (1.2)	[2 (2.4)]
All Raters	45 (37.5)	9 (7.5)	6 (5.0)	18 (15.0)	39 (32.5)	1 (0.8)	[2 (1.7)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

With regards to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:							
	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
12. I retrieve documentation on pressure ulcer risk and prevention from the patient's paper health record.							
Expert	8 (22.2)	2 (5.6)	1 (2.8)	3 (8.3)	21 (58.3)	–	[1 (2.8)]
Non-Expert	21 (25.0)	2 (2.4)	2 (2.4)	18 (21.4)	37 (44.0)	2 (2.4)	[2 (2.4)]
All Raters	29 (24.3)	4 (3.3)	3 (2.5)	21 (17.5)	58 (48.3)	2 (1.7)	[3 (2.5)]
13. I retrieve documentation of pressure ulcer risk and prevention from the patient's electronic health record.							
Expert	31 (86.1)	2 (5.6)	2 (5.6)	1 (2.8)	–	–	[0]
Non-Expert	68 (81.0)	6 (7.1)	1 (1.2)	5 (6.0)	1 (1.2)	1 (1.2)	[2 (2.4)]
All Raters	99 (82.5)	8 (6.7)	3 (2.5)	6 (5.0)	1 (0.8)	1 (0.8)	[2 (1.7)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

With regards to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:							
	Always ***n*** **(%)**	**Almost Always** ***n*** **(%)**	**Usually** ***n*** **(%)**	**Sometimes** ***n*** **(%)**	**Never** ***n*** **(%)**	**I Don't Know** ***n*** **(%)**	**[Missing Values]** ***n*** **(%)**
14. Persons from information technology extract data on pressure ulcer risk and prevention from the electronic health record for use in the pressure ulcer survey.							
Expert	5 (13.9)	–	2 (5.6)	1 (2.8)	20 (55.6)	8 (22.2)	[0]
Non-Expert	13 (15.5)	1 (1.2)	1 (1.2)	7 (8.3)	39 (46.4)	20 (23.8)	[3 (3.6)]
All Raters	18 (15.0)	1 (0.8)	3 (2.5)	8 (6.7)	59 (49.2)	28 (23.3)	[3 (2.5)]
15. If a patient was admitted to the hospital less than 24 hours before the pressure ulcer survey, I exclude them from the survey.							
Expert	3 (8.3)	1 (2.8)	–	3 (8.3)	29 (80.6)	–	[0]
Non-Expert	13 (15.5)	–	2 (2.4)	12 (14.3)	50 (59.5)	4 (4.8)	[3 (3.6)]
All Raters	16 (13.3)	1 (0.8)	2 (1.7)	15 (12.5)	79 (65.8)	4 (3.3)	[3 (2.5)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

With regards to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:							
	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
16. I decide if a patient is at risk for pressure ulcers by the score obtained on the risk assessment scale (i.e. Braden, Norton) at the time of the pressure ulcer survey.							
Expert	17 (47.2)	11 (30.6)	5 (13.9)	2 (5.6)	1 (2.8)	–	[0]
Non-Expert	50 (59.5)	20 (23.8)	2 (2.4)	6 (7.1)	2 (2.4)	1 (1.2)	[3 (3.6)]
All Raters	67 (55.8)	31 (25.8)	7 (5.8)	8 (6.7)	3 (2.5)	1 (0.8)	[3 (2.5)]
17. If a patient's risk assessment score (i.e. Braden or Norton Scale score) does not classify them as being "at risk for pressure ulcers", I look for documentation of other factors indicating the patient is at pressure ulcer risk.							
Expert	17 (47.2)	10 (27.8)	4 (11.1)	2 (5.6)	3 (8.3)	–	[0]
Non-Expert	51 (60.7)	10 (11.9)	3 (3.6)	11 (13.1)	6 (7.1)	–	[3 (3.6)]
All Raters	68 (56.7)	20 (16.7)	7 (5.8)	13 (10.8)	9 (7.5)	–	[3 (2.5)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

These last questions (18 – 27) apply to patients "at risk" for pressure ulcers. With regard to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:

	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
18. I directly observe patients in their rooms in order to rate pressure-redistribution surface use.							
Expert	23 (63.9)	8 (22.2)	1 (2.8)	1 (2.8)	3 (8.3)	–	[0]
Non-Expert	55 (65.5)	6 (7.1)	4 (4.8)	8 (9.5)	9 (10.7)	–	[2 (2.4)]
All Raters	78 (65.0)	14 (11.7)	5 (4.2)	9 (7.5)	12 (10.0)	–	[2 (1.7)]
19. I directly observe patients in their rooms in order to rate nutritional support.							
Expert	12 (33.3)	6 (16.7)	3 (8.3)	10 (27.8)	5 (13.9)	–	[0]
Non-Expert	31 (36.4)	8 (9.5)	4 (4.8)	20 (23.8)	18 (21.4)	–	[3 (3.63)]
All Raters	43 (35.8)	14 (11.7)	7 (5.8)	30 (25.0)	23 (19.2)	–	[3 (2.5)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

These last questions (18 – 27) apply to patients "at risk" for pressure ulcers. With regard to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:

	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
20. I directly observe patients in their rooms in order to rate moisture management.							
Expert	18 (50.0)	11 (30.6)	4 (11.1)	–	3 (8.3)	–	[0]
Non-Expert	44 (52.4)	11 (13.1)	4 (4.8)	12 (14.3)	10 (11.9)	–	[3 (3.6)]
All Raters	62 (51.7)	22 (18.3)	8 (6.7)	12 (10.0)	13 (10.8)	–	[3 (2.5)]
21. If the nurse tells me a patient refused to be repositioned, but there is no documentation of this in the patient record, I select "Patient refused".							
Expert	2 (5.6)	1 (2.8)	2 (5.6)	5 (13.9)	22 (61.1)	4 (11.1)	[0]
Non-Expert	5 (6.0)	1 (1.2)	2 (2.4)	12 (14.3)	57 (67.9)	5 (6.0)	[2 (2.4)]
All Raters	7 (5.8)	2 (1.7)	4 (3.3)	17 (14.2)	79 (65.8)	9 (7.5)	[2 (1.7)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

These last questions (18 – 27) apply to patients "at risk" for pressure ulcers. With regard to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:

	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
22. I ask patient care staff for information in order to determine how to rate Routine repositioning as prescribed.							
Expert	2 (5.6)	3 (8.3)	6 (16.7)	8 (22.2)	16 (44.4)	–	[1 (2.8)]
Non-Expert	10 (11.9)	5 (6.0)	6 (7.1)	23 (27.4)	34 (40.5)	3 (3.6)	[3 (3.6)]
All Raters	12 (10.0)	8 (6.7)	12 (10)	31 (25.8)	50 (41.7)	3 (2.5)	[4 (3.3)]
23. If repositioning is appropriate for a patient; I allow a 30 minute leeway past the time repositioning was to occur before I rate Routine repositioning as prescribed as "No".							
Expert	7 (19.4)	7 (19.4)	4 (11.1)	4 (11.1)	12 (33.3)	2 (5.6)	[0]
Non-Expert	11 (13.1)	14 (16.7)	9 (10.7)	13 (5.5)	20 (23.8)	14 (16.7)	[3 (3.6)]
All Raters	18 (15.0)	21 (17.5)	13 (10.8)	17 (14.2)	32 (26.7)	16 (13.3)	[3 (2.5)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

These last questions (18 – 27) apply to patients "at risk" for pressure ulcers. With regard to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:

	Always *n* **(%)**	**Almost Always** *n* **(%)**	**Usually** *n* **(%)**	**Sometimes** *n* **(%)**	**Never** *n* **(%)**	**I Don't Know** *n* **(%)**	**[Missing Values]** *n* **(%)**
24. To evaluate frequency of patient repositioning, I look in the patient record for physical therapy documentation of patient activity such as standing or ambulation.							
Expert	5 (13.9)	4 (11.1)	4 (11.1)	12 (33.3)	11 (30.6)	–	[0]
Non-Expert	17 (20.2)	11 (13.1)	5 (6.0)	17 (20.2)	28 (33.3)	3 (3.6)	[3 (3.6)]
All Raters	22 (18.3)	15 (12.5)	9 (7.5)	29 (24.2)	39 (32.5)	3 (2.5)	[3 (2.5)]
25. If a nutritional consult is ordered for a patient who has been hospitalized for 2 or more days with poor intake, but the consult has not yet been completed, I rate Nutritional support as "Yes".							
Expert	1 (2.8)	1 (2.8)	2 (5.6)	6 (16.7)	24 (66.7)	2 (5.6)	[0]
Non-Expert	6 (7.1)	6 (7.1)	2 (2.4)	10 (11.9)	50 (59.5)	8 (9.5)	[2 (2.4)]
All Raters	7 (5.8)	7 (5.8)	4 (3.3)	16 (13.3)	74 (61.7)	10 (8.3)	[2 (1.7)]

Table 31 (continued).

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data (n = 36 Experts, n = 84 Non-Experts)

These last questions (18 – 27) apply to patients "at risk" for pressure ulcers. With regard to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements:							
	Always ***n* (%)**	**Almost Always** ***n* (%)**	**Usually** ***n* (%)**	**Sometimes** ***n* (%)**	**Never** ***n* (%)**	**I Don't Know** ***n* (%)**	**[Missing Values]** ***n* (%)**
26. Barrier skin cream at the bedside is sufficient for me to rate Moisture management as "Yes".							
Expert	5 (13.9)	3 (8.3)	2 (5.6)	8 (22.2)	17 (47.2)	–	[1 (2.8)]
Non-Expert	10 (11.9)	6 (7.1)	6 (7.1)	18 (21.4)	38 (45.2)	3 (3.6)	[3 (3.6)]
All Raters	15 (12.5)	9 (7.5)	8 (6.7)	26 (21.7)	55 (45.8)	3 (2.5)	[4 (3.3)]
27. If I rate that a patient did not receive a particular pressure ulcer prevention intervention, someone else reviews the patient record to make sure I selected the correct rating.							
Expert	4 (11.1)	3 (8.3)	3 (8.3)	8 (22.2)	18 (50.0)	–	[0]
Non-Expert	10 (11.9)	6 (7.1)	3 (3.6)	18 (21.4)	32 (38.1)	12 (14.3)	[3 (3.6)]
All Raters	14 (11.7)	9 (7.5)	6 (5.0)	26 (21.7)	50 (41.7)	12 (10.0)	[3 (2.5)]

RNs "never" collect these data, but the three other raters at that hospital reported that RNs "always" do. Participant raters among 13 (36.1%) hospitals reported that RNs are the only ones who collect NDNQI PrU risk and prevention data at their hospital. Twenty-three (19.2%) raters among 12 (33.3%) hospitals reported that LPNs at least "sometimes" collect these data. Twenty-two (18.3%) raters among 15 (41.7%) hospitals reported that nurse aids at least "sometimes" collect these data. Ten (8.3%) raters among 7 (19.4%) hospitals reported that physical therapists at least "sometimes" collect these data. Four (3.3%) raters among 2 (5.6%) hospitals reported that clinical secretaries at least "sometimes" collect these data. Finally, 1 (0.8%) rater reported that respiratory therapists at least "sometimes" collect these data. Seventeen (14.2%) raters among 5 (13.9%) hospitals specified that others than those listed "almost always" or "always" collect these data. These included dieticians (2 hospitals), nursing students (2 hospitals), research assistants (1 hospital), data specialists (1 hospital), diabetic educators (1 hospital), and the site coordinator (1 hospital).

Forty-five (37.5%) participant raters reported that their hospital "almost always" or "always" evaluates the reliability of the PrU risk and prevention measures every year. Nearly half "almost always" or "always", $n = 54$ (45%) collect data on the floor where they usually work, while a third "never" do, $n = 39$ (32.5%). Respondents reported that their NDNQI PrU risk and prevention data is "almost always" or "always" obtained from review of the electronic health record, $n = 107$ (89.2%); compared to the patient's paper health record, $n = 33$ (27.5%). There was some confusion among participant raters with regards to whether or not persons from information technology retrieve any of their hospital's PrU data in that 28 (23.3%) replied "I don't know" to this survey item: Nearly half, however, replied "never", $n = 59$ (49.2%). This

item (about information technology extracting data) had the most "I don't know" responses than any other item, *n* = 28 (23.3%).

Most respondents reported they "never" exclude a patient who was admitted < 24 hours before the skin survey, *n* = 79 (65.8%). The PrU risk assessment scale score is often used to determine a patient's risk status: 98 (81.7%) participant raters reported that they "almost always" or "always" determine risk status by risk assessment scale score. In addition, 88 (73.3%) raters reported that they "almost always" or "always" consider other clinical factors while determining risk status.

Eighty-four (70%) participant raters reported they "almost always" or "always" rate *Moisture Management within the Last 24 hours* by direct observation. Likewise, 24 (20%) reported that barrier cream at the bedside is "almost always" or "always" sufficient to rate *Moisture Management within the Last 24 hours* as "Yes".

Responses to the survey items on repositioning highlight disagreement in rating *Routine repositioning as prescribed within the last 24 hours*. Thirty-two (26.7%) participant raters reported they "never" allow a 30-minute leeway from the time repositioning was to take place before rating this item as "No", while 39 (32.5%) reported they "almost always" or "always" do. More than half of the respondents reported they "never" rate *Routine repositioning as prescribed* as "Patient refused" without documentation in the patient record, regardless of what the nurse verbally reports, *n* = 79 (65.8%). At the same time, 20 (16.7%) respondents reported they "almost always" or "always" ask staff how to rate repositioning. Finally, 39 (32.5%) "never" look in the patient record for physical therapy documentation of patient activity to rate *Routine repositioning as prescribed*, while 37 (30.8%) "almost always" or "always" do.

Most participant raters reported they "never" rate *Nutritional support within the last 24 hours* as "Yes" for someone hospitalized for two or more days with poor intake and for whom the nutritional consult had not yet been completed, $n = 74$ (61.7%). Finally, less than a quarter reported that someone "almost always" or "always" verifies their ratings if they record that a patient did not receive a particular PrU prevention intervention; $n = 23$ (19.2%).

Chapter Summary

Cohen's kappa values varied widely within and across hospitals. Because most patients were assessed for PrU risk, and those at risk received prevention, the prevalence of a "Yes" response was high; suggesting PAK may be a better estimate of inter-rater reliability than Cohen's kappa. Prevalence-adjusted kappa (PAK) values for: (1) *Skin assessment*, PAK = .977, 95% CI [.966 – .989]; (2) *Risk assessment*, PAK = .978, 95% CI [.964 –.993]; (3) *Time since last risk assessment*, PAK = .790, 95% CI [.729 – .852]; (4) *Risk assessment scale*, PAK = .997, 95% CI [.991 – 1.0]; (5) *Risk status*, PAK = .877, 95% CI [.838 – .917]; (6) *Any prevention*, PAK = .856, 95% [.769 – .943]; (7) *Skin assessment documented*, PAK = .956, 95% CI [.904 – 1.0]; and (8) *Pressure-redistribution surface use*, PAK = .839, 95% CI [.763 – .916] indicated substantial to near perfect agreement. Prevalence-adjusted kappa values for: (9) *Routine repositioning*, PAK = .577, 95% CI [.494 – .661]; (10) *Nutritional support*, PAK = .500, 95% CI [.418 – .581]; and (11) *Moisture management*, PAK = .556, 95% CI [.469 – .643] indicated moderate agreement.

CHAPTER V: DISCUSSION, RECOMMENDATIONS, AND CONCLUSION

This is the first known study to examine the reliability of the 11 NDNQI® pressure ulcer risk and prevention measures. The specific aims were: (1) **Aim 1**: examine the reliability of the NDNQI pressure ulcer risk and prevention measures within and across NDNQI hospitals, and (2) **Aim 2**: examine the methods and processes used by participant raters to gather data on the NDNQI PrU risk and prevention measures.

The Sample

Hospital and Unit Sample

The **hospital** sample size ($N = 36$) was comparable to previous research that examined the reliability of PrU identification and staging: the study by Bergquist-Beringer et al. (2011) included 31 hospitals, Gajewski et al. (2007) had 20 hospitals, and 48 hospitals participated in the study by Hart et al. (2006). Additional NDNQI reliability studies, not specific to PrUs, included slightly more hospitals in their samples than the current study, with 48 (Choi et al., 2014) and 54 (Simon et al., 2011) hospitals respectively.

In this study, 25% ($n = 9$) of participating hospitals were Magnet hospitals. The percentage of hospitals participating in previous NDNQI research that were Magnet hospitals was 20.8% (Choi et al., 2014), 23% (Bergquist-Beringer et al., 2013), and 35% (Bergquist-Beringer et al., 2011). The proportion of hospitals in previous NDNQI research that were Teaching or Academic Medical Centers ranged from 44.5% (Bergquist-Beringer et al., 2013) to 65% (Staggs & He, 2013), compared to 41.6% for this study. Hospital size of < 300 staffed beds in previous NDNQI research was 72.1% (Choi et al., 2014) and 70.4% (Bergquist-Beringer et al., 2013), compared to 77.8% for this study. It is appropriate to state that the current study's

hospital sample was similar to the population of NDNQI hospitals that submit PrU data because the study by Bergquist-Beringer and colleagues (2013) included all 1,419 hospitals that submitted PrU data in 2010.

For this study, eligible **unit** type included adult medical-surgical, adult medical, and adult surgical units. Nearly half of the NDNQI data on PrUs is submitted by these unit types. Of the 15,400 units that submitted PrU data for the 2nd Quarter of 2014, 7,471 (49%) units were classified as medical-surgical (*n* = 1,774), medical (*n* = 2,468), or surgical (*n* = 3,229) (NDNQI Statistical Analyst, personal communication, November 21, 2014).

Participant Rater Sample

The study's sample of 120 participant raters (*n* = 36 experts, *n* = 84 non-experts) had fewer raters than the 140 raters (Gajewski et al., 2007), 162 raters (Bergquist-Beringer et al., 2011), and 256 raters (Hart et al., 2006) in previous NDNQI reliability studies; but more than 24 raters in the Simon et al. (2011) study. Similar to the current study, the raters in the Bergquist-Beringer et al. (2011) study included an "expert" at each hospital. Previous NDNQI researchers reported 46.3% (Bergquist-Beringer et al., 2011) to 47% (Hart et al., 2006) of the raters were Staff RNs, compared to 43.3% for this study; and that 54.5% (Choi et al., 2013) to 69.7% (Bergquist-Beringer et al., 2011) held a Bachelor's Degree or higher, compared to 70.8% of all raters (including non-nurses) for this study. In the current study, 63.3% (*n* = 76) of the participant raters reported they were *not* certified in wound care, compared to 76% in previous NDNQI PrU reliability research (Bergquist-Beringer et al., 2011).

Patient Sample

Data were extracted from 1,637 patient records. All 1,637 patients were rated on the (5) PrU *risk* measures (*M* = 45.5 patients per hospital). Of these patients, 553 (33.8%) were

considered by the expert to be "at risk" for PrU development. However, not all non-expert raters agreed with the expert on risk status. Therefore, 528 to 530 patients were also rated on the (6) *prevention* measures (M = 14.7 patients per hospital, SD = 7.0).

Patient sample size (***N* = 528 to 530)** was M = 14.7 patients per hospital. In previous research that examined the reliability of PrU identification and staging, Bergquist-Beringer et al. (2011) and Hart et al. (2006) presented 25 photographs of wounds for evaluation by each rater (***N* = 25**). Bergquist-Beringer et al. (2011) also had raters round on patients, and collectively they staged a total of ***N* = 591** PrUs, M = 19.1 PrUs per hospital; which is more than (but comparable to) the current study's sample size. Gajewski et al.'s (2007) sample (***N* = 347** PrUs) was smaller than the current study's overall sample. However, in the Gajewski et al. (2007) study, 6 – 108 PrUs (M = 17.4) per hospital were staged; which is more than the current study's 5 – 37 patients (M = 14.7) per hospital that were rated on the *prevention* measures. Based on previous studies, the sample size for this study (of 528 to 530 patients overall, and 14.7 patients per hospital) was near the range of previous NDNQI studies.

Similar to this study, previous research included slightly more female patients (52% and 54.9%) than male patients (Bergquist-Beringer et al., 2012; Choi et al., 2013). The percentage of patients found to be "at risk" for PrU development in this study (33.8%) is also comparable to previous research which found 37% of patients among medical-surgical, medical, and surgical unit types were "at risk" (S. Bergquist-Beringer, personal communication, March 11, 2014 in unpublished data from Bergquist-Beringer, 2011). Finally, the average age of patients in this study (M = 63 years) was similar to the average age of patients in Bergquist-Beringer (2011), (M = 62 years).

Aim 1: Calculations of Rater Agreement on the 11 NDNQI Pressure Ulcer Risk and Prevention Measures

In this study, there were 2 to 4 expert/non-expert pairs at each hospital (M = 2.33), yielding 2 to 4 kappa values for each measure per hospital (total of 84 kappa values across 36 hospitals). Because most patients were rated as having received the PrU risk and prevention interventions, rater response was often a constant. Consequently, SPSS and Excel would not compute Cohen's kappa values. Therefore, it was necessary to add a very small number (0.0001) to each cell in order for a Cohen's kappa value to be presented. Although this strategy has been suggested by others to address "constant" variables while calculating kappa (Weaver, 2007), the addition of 0.0001 to each cell distinguishes this study's data analysis from other studies presenting kappa values. Prevalence-adjusted kappa (PAK) values were also computed with the addition of 0.0001 to each cell (see Table 8, p. 120).

Aim 1, Question 1 – Reliability of Each Measure within Hospitals

Cohen's kappa values provide support for the reliability of *Risk status* measurement within hospitals. For 32 hospitals, Cohen's kappa values for *Risk status* were ≥ .610 to 1.00, indicating "substantial" to "near perfect" agreement between the expert and non-expert raters. For all the other measures, the majority of Cohen's kappa values were < .610. In contrast, PAK values provide support for the reliability of all 5 *risk* measures within hospitals, and 3 of 6 *prevention* measures within hospitals. Specifically, the majority of hospitals had PAK values ≥ .610 to 1.00 indicating "substantial" to "near perfect" agreement between the expert and non-expert raters for (1) *Skin assessment within 24 hours of admission*, (2) *Risk assessment within 24 hours of admission*, (3) *Time since last PrU risk assessment*, (4) *Risk assessment scale*, (5) *Risk*

status, (6) *Any prevention within the last 24 hours*, (7) *Skin assessment documented within the last 24 hours*, and (8) *Pressure-redistribution surface use within the last 24 hours*.

While considering within hospital Cohen's kappa and PAK values, it is important to remember that most (69.4%) of these values were calculated from two expert/non-expert pairs, and some hospitals had as few as 5 to 7 patients rated on the prevention measures. The low Cohen's kappa values were likely due to the large proportion of "Yes" responses for each measure (Sim & Wright, 2005). As expected, each PAK value was higher than its corresponding Cohen's kappa value, highlighting the presence and effect of high prevalence of "Yes" responses on estimates of rater agreement. Therefore, PAK values are likely better estimates of inter-rater reliability for the 11 NDNQI PrU risk and prevention measures, than Cohen's kappa values.

Fifteen (3.8%) Cohen's kappa values and 1 (0.3%) PAK value, were < 0; meaning agreement between the expert and the non-experts was less than expected by chance alone. A negative kappa value *close to zero* (i.e. > - .22 to < 0) is not surprising for this study's data, and suggests rater agreement was no better than chance alone—the most negative Cohen's kappa or PAK value was -.242. In addition, some negative kappa values are likely by chance alone because of the large number Cohen's kappa values (N = 396) and PAK values (N = 396) that were calculated. These negative Cohen's kappa and PAK values were found for 7 of the 11 NDNQI measures, and within 12 (33.3%) hospitals. They were present in hospitals that had both two and three expert/non-expert pairs.

Aim 1, Question 2 – *Overall* Agreement between Expert and Non-Expert Ratings for the NDNQI PrU *Risk* Measures, and the *Overall* Agreement between Expert and Non-Expert Ratings for the NDNQI PrU *Prevention* Measures

The average Cohen's kappa values for all 5 *risk* measures, and the average Cohen's kappa values for all 6 *prevention* measures, were < .610. Specifically, the overall Cohen's kappa value for the *risk* measures was .498, 95% CI [.463 – .533], indicating only "moderate" agreement between the expert and non-expert raters. The overall Cohen's kappa value for the *prevention* measures was even lower at .312, 95% CI [.273 – .352], indicating "fair" agreement. Cohen's kappa values leave in question the reliability of these measures overall.

In contrast, the average PAK value for all 5 *risk* measures and the average PAK value for all 6 *prevention* measures, were ≥ .610, and therefore, provide support for the reliability of these two overall measures. Specifically, the overall PAK value for the *risk* measures was .924, 95% CI [.910 – .939], indicating "near perfect" agreement between the expert and non-expert raters. The overall PAK value for the *prevention* measures was .714, 95% CI [.661 – .767], indicating "substantial" agreement. Prevalence-adjusted kappa (PAK) values may be the better measure of inter-rater reliability of these measures because of the large proportion of patients who received the PrU risk and prevention interventions (see discussion on p. 119 – 121).

Cohen's kappa and PAK values for the *prevention* measures were uniformly lower than Cohen's kappa and PAK values for the *risk* measures. This may be because fewer patients were rated on the *prevention* measures. Eight (22.2%) hospitals had as few as 5 to 7 patients who were rated on the *prevention* measures; while only 4 (11.1%) hospitals had less than 40 patients who were rated on the *risk* measures.

Aim 1, Question 3 – Average Agreement Estimates for Each Measure across 36 Hospitals

For each of the 11 PrU risk and prevention measures, a single Cohen's kappa value, a single PAK value, and a single percent agreement was presented. Most of these values had narrow 95% CIs, providing support for their accuracy.

The five NDNQI pressure ulcer *risk* measures.

***Cohen's kappa*.** Cohen's kappa values provide support for the reliability of the PrU *risk* measure, *Risk status*. The averaged (across hospitals) Cohen's kappa value indicated "near perfect" agreement (k = .819, 95% CI [.766 – .873]. Interestingly, the prevalence of agreement on *Risk status* was more evenly distributed among the cells of agreement compared to the other risk measures. This provides insight on the effect of high prevalence on Cohen's kappa values. Cohen's kappa values for the two risk measures, *Risk assessment within 24 hours of admission* (k = .472, 95% CI [.393 – .552]) and *Risk assessment scale* (k = .495, 95% CI [.485 – .505]), indicated "moderate" agreement between raters on average across hospitals. Agreement matrices show that each of these measures had a very large proportion of "Yes" responses, which decreased Cohen's kappa values. Among the four response options for *Risk assessment scale* ("Braden", "Norton", "Other", and "Clinical factors"); "Braden" was selected by each rater more than 99% of the time. All other response options combined were selected less than 1% of the time. Interestingly, the *Risk assessment scale* measure should be reliable because it essentially asks raters which risk assessment scale is used at their hospital; something one would expect all nurses at their hospital to know. Also, it is well-known that the Braden Scale is the most commonly used PrU risk assessment scale in the U.S. (Armstrong et al., 2008). Even so, the Cohen's kappa value reflected only "moderate" agreement for this measure.

Averaged Cohen's kappa values for the two other risk measures, *Skin assessment within 24 hours of admission* (k = .379, 95% CI [.308 – .451]), and *Time since last PrU risk assessment* (k = .324, 95% CI [.241 – .406]), indicated even lower ("fair") agreement between raters. Similar to other measures, agreement matrices show each these measures had a very large proportion of "Yes" responses, which lowered Cohen's kappa values. According to the

agreement matrix for *Skin assessment within 24 hours of admission*, most of the disagreement among raters was between the response options of "Yes" and "No", which is of concern, but the sum of the cell counts of disagreement accounted for only 1.5% of all cell counts combined. *Time since last PrU risk assessment* had the lowest average Cohen's kappa value, and the lowest average PAK value, among the risk measures. For this measure, high prevalence for the response "> 0 – 12" hours was apparent, but so was disagreement between this response and "> 12 – 24" hours. While interpreting agreement for *Time since last PrU risk assessment*, however, it is important to consider how study data were collected.

Participant raters were instructed to consider only patient data from 7:00 a.m. the day *before* the study, to 7:00 a.m. the day *of* the study. This is different than how data are collected for quarterly NDNQI PrU data submission: In practice, PrU data are collected by considering the 24-hour period *immediately prior to* the time of data collection. An example of this issue is as follows. Pressure ulcer risk assessments are typically performed at the beginning of each shift, so it is likely that most risk assessments were performed between 7:30 a.m. and 10:00 a.m. on the morning that study data were collected, as well as between 7:30 p.m. and 10:00 p.m. the evening before study data were collected. The participant raters, however, were not required to collect data within the same *hour* as other participant raters collected data at their hospital. Data collection was to occur only within the 5-hour data collection period of 7:00 a.m. to 12:00 p.m. In other words, if one of the raters collected study data on a particular patient at 8:00 a.m., and the last risk assessment was performed at 8:00 p.m., the correct rating for *Time since the last PrU risk assessment* is > 0 – 12 hours. However, if another rater collected study data on this same patient an hour later, then this rater may likely have recorded the *Time since the last PrU risk assessment* as > 12 – 24 hours. Both raters correctly used only data within the 24-hour period of

7:00 a.m. to 7:00 a.m., but how they "counted back" to determine the time since the last risk assessment was performed, may have influenced their rating.

Another issue with *Time since last PrU risk assessment* is that some participant raters may have incorrectly considered the risk assessment that was performed after 7:00 a.m. the day of the study, which is how NDNQI PrU data are collected in practice. This would result in disagreement among the raters because some would select > 0 – 12 hours, while others would select > 12 – 24 hours. Finally, in order to make sure participant raters were using the same time frame and the same risk assessment (as the other participant raters) they needed to score this measure as if they were collecting data at 7:00 a.m., something they were not instructed to do. For instance, if the last risk assessment was performed at 8:00 p.m. the evening before the day of the study, then unless participant raters either (a) counted back from 7:00 a.m. to arrive at > 0 – 12 hours, or (b) collected study data on a patient at the same time that the other participant raters collected data on that patient; disagreement between the raters was highly likely.

Nevertheless, the agreement matrix for *Time since last PrU risk assessment* (see Table 29 in Chapter IV) clearly shows that the overwhelming majority of disagreement was between the two response options of "> 0 – 12 hours" and "> 12 – 24 hours". When data for *Time since Last PrU Risk Assessment* were collapsed by combining the response options of "> 0 – 12" hours, and "> 12 – 24" hours, across hospital Cohen's kappa value increased from "fair" (.324) to "moderate" (.429); and the PAK value increased from "substantial" (.790) to "near perfect" (.899). Although, this "new" Cohen's kappa value was still < .610, these results suggest reliability would be improved by collapsing data for these two response options to "> 0 to 24" hours. Collapsing these data is supported by the Institute for Healthcare Improvement (IHI, 2015) and others (Stechmiller et al., 2008) who have established the standard of performing a

PrU risk assessment as every 24 hours in the acute care setting. In addition, even a recommendation of performing a risk assessment "every shift" would translate to > 0 – 24 hours (not > 0 – 12 hours). For example, for12-hour shifts, it is likely a risk assessment would be completed at 8:00 p.m. and again at 9:00 a.m. the next day, which is > 12 – 24 hours.

Prevalence-adjusted kappa (PAK). Prevalence-adjusted kappa (PAK) values provide support for the reliability of all (5) PrU *risk* measures. The averaged (across hospitals) PAK values for four of the five *risk* measures (*Skin assessment within 24 hours of admission*, *Risk assessment within 24 hours of admission*, *Risk assessment scale*, and *Risk status*) were .877 to .997, indicating "near perfect" agreement between the expert and non-expert raters. The averaged PAK value for the fifth risk measure, *Time since last PrU risk assessment*, was .790, 95% CI [.729 – .852], indicating "substantial" agreement.

Percent agreement. The average (across hospital) percent agreements were large (81.2% to 99.8%) for the PrU *risk* measures. These values were expected because of the large number of patients who receive the PrU risk and prevention interventions. It is important to remember that percent agreement does not consider agreement by chance alone. However, when considered with other agreement estimates, this commonsense statistic provides further insight. For instance, *Risk assessment scale* had a Cohen's kappa of .495, indicating only "moderate" agreement, but the raters agreed 99.8% of the time.

The six NDNQI pressure ulcer *prevention* measures.

As for NDNQI survey requirements, the PrU *prevention* measures in the study were rated only for persons who were determined to be "at risk" for PrU development. The expert identified 553 patients to be at risk. However, not all non-experts agreed with the expert on risk status. Therefore, 528 to 530 patients were rated on each of the prevention measures.

Cohen's kappa values. Average (across hospitals) Cohen's kappa values for the (6) prevention measures do not provide support for the reliability of these measures: None of these values were ≥ .610. The Cohen's kappa value for *Skin assessment documented within the last 24 hours* (k = .451, 95% CI [.415 – .488]), was the only average Cohen's kappa value among the prevention measures that indicated at least "moderate" agreement between the expert and non-expert raters. Average Cohen's kappa values for the other five prevention measures (*Any PrU prevention within the last 24 hours*, *Pressure-redistribution surface use within the last 24 hours*, *Routine repositioning as prescribed within the last 24 hours, Nutritional support within the last 24 hours*, and *Moisture management within the last 24 hours*) indicated "fair" agreement between the expert and non-expert raters (k = .216 to .387). Like the risk measures, high prevalence of "Yes" responses were expected and observed in the agreement matrices.

Prevalence-adjusted kappa values. Average (across hospitals) PAK values provide support for the reliability of three of the six prevention measures; (1) *Any prevention within the last 24 hours*, (2) *Skin assessment documented within the last 24 hours*, and (3) *Pressure-redistribution surface use within the last 24 hours*. These PAK values indicated "near perfect" agreement between the expert and non-expert raters (PAK = .839 – .956). The average PAK values for the remaining three prevention measures (*Routine repositioning as prescribed within the last 24 hours*, *Moisture management within the last 24 hours*, and *Nutritional support within the last 24 hours*) indicated "moderate" agreement (PAK = .500 – .577).

The agreement matrices show high prevalence of "Yes" responses; therefore, PAK values are likely better estimates of inter-rater reliability of these measures than Cohen's kappa values. For the prevention interventions *Routine repositioning as prescribed within the last 24 hours*, *Moisture management within the last 24 hours*, and *Nutritional support within the last 24 hours*,

the agreement matrices provide insight into where the disagreement occurred (see Table 29 in Chapter IV). "Unnecessary" is frequently among the cells of disagreement with the highest cell counts. Notably, this disagreement pattern was present for each of *the four* PrU prevention measures that have "Unnecessary" as a response option (*Pressure-redistribution surface use within the last 24 hours*, *Routine repositioning as prescribed within the last 24 hours*, *Nutritional support within the last 24 hours*, and *Moisture management within the last 24 hours*). There was similar disagreement counts between "Yes" and "Unnecessary", and "No" and "Unnecessary". This may reflect differences in hospital policy on PrU prevention, as well as individual differences in how raters score this item. Maybe some data collectors are using other factors besides the *total* Braden Scale score to rate patients on the prevention measures, while others are not. Perhaps some data collectors consider Braden *subscale* scores to rate these measures. For instance, some data collectors may rate *Nutritional support* as "Unnecessary" if the Braden subscale score for nutrition is greater than 3, while others look for documentation of *Nutritional support* for all patients "at risk" (i.e. total Braden score ≤ 18) regardless of the nutrition subscale score. This might also be happening with *Routine repositioning as prescribed*. Some raters may have chosen "Unnecessary" for patients who had a combined activity-mobility Braden subscale score ≥ 4; while other raters looked for documentation of repositioning for all "at risk" patients. In other words, some participant raters seemed to have based their rating on if repositioning *was necessary*, while others looked for documentation of *if it was done*. Others have reported that Braden subscale scores should be considered while determining the appropriate PrU prevention interventions (Braden, 2012). In addition, the NPUAP/EPUAP guidelines suggest risk assessment subscale scores are useful in determining appropriate prevention interventions (NPUAP/EPUAP, 2009).

Percent agreement. Average (across hospitals) percent agreement ranged from 57.6% to 96.5% for the *prevention* measures. Percent agreement was the lowest for, *Nutritional support within the last 24 hours* (57.6%, 95% CI [50.4% – 64.8%], *Moisture management within the last 24 hours* (62.4%, 95% CI [54.4% - 70.3%], and *Routine repositioning as prescribed within the last 24 hours* (65.4%, 95% CI [58.2% - 72.5%]. The lowest Cohen's kappa values and PAK values were also for these same measures.

Aim 1, Question 4 – Intraclass Correlation Coefficients from Kappa Values to Examine Variance in Rater Agreement across Hospitals

For this study, *ICC* was defined as the variance in kappa values *between* hospitals, divided by the sum of the variance in kappa values *between* hospitals and the variance in kappa values *within* hospitals. Following criteria established by Hart et al. (2006), an *ICC* near zero indicates *within*-hospital variability in rater agreement was much greater than *across*-hospital variability in rater agreement. This means that participant raters at a particular hospital scored the NDNQI PrU risk and prevention measures with similar inter-rater reliability as participant raters at the other hospitals that scored the measures.

Using a level of agreement similar to Landis and Koch's (1977) level of agreement categories for "poor" to "slight" agreement (i.e. < .20), Hart et al. (2006) identified that an *ICC* < .22 was considered to be "near zero", and suggests that *within*-hospital variability was much greater than *between*-hospital variability (an ICC < .50 suggests within-hospital variability was greater than between-hospital variability). The ICC_{kappa} values for two measures (*Risk Assessment Scale,* and *Skin Assessment Documented within the Last 24 hours*) were < .22.. However, these two *ICC*s were negative (ICC_{kappa} = -.513 and -.052, respectively). According to the *ICC* definition, (variability of kappa's between hospitals)/(total variability of kappa's), *ICC*

value is non-negative. However, negative *ICC* values are possible due to the way of separating between-subject variation and within-subject variation in SPSS. When the average covariance calculated from a dataset is negative, SPSS may generate negative *ICC* values (Nichols, 1999). Further examination of the Cohen's kappa values for these two measures reveals that for almost all the hospitals, Cohen's kappa values were .500 (and PAK values were 1.00). For *Risk assessment scale*, the Cohen's kappa value for 35 (97.2%) hospitals was .500 (see Table 17, p. 146). For *Skin assessment documented within the last 24 hours*, the Cohen's kappa value for 29 (80.6%) hospitals was .500 (see Table 20, p. 150). Kappa values for the other nine measures were more varied across hospitals. This was especially evident on each measure's box-and-whisker plot in that the two measures with negative *ICC*s had only a horizontal line at .500 (i.e. no "box" or "whiskers").

While the *ICC*s for the other nine measures were positive, they were not < .22 (ICC_{kappa} ranged from .259 to .573). In addition, in general the *ICC* 95% CIs were wide for all 11 measures; also likely due to the small sample size and that *ICC*s were calculated from only two kappa values per hospital.

When *ICC*s were calculated from PAK values, results revealed no measures with a negative ICC_{PAK}, or an ICC_{PAK} considered being near zero. However, a single measure's 95% CI (ICC_{PAK}) included zero: *Risk assessment within 24 hours of admission*, ICC_{PAK} = .295, 95% CI [-.03 - .564]. As before, the 95% CIs were wide, perhaps because *ICC*s were calculated from only two kappa values per hospital. Considering the above discussion, it is likely that the *ICC*s from the study (ICC_{kappa} and ICC_{PAK}) do not well represent the proportion of the total variability in kappa values that is due to between-hospital variability.

Differences in study results between this study and those by Hart et al. (2006) reflect differences in study methodology. In the study by Hart and colleagues (2006), *all raters across hospitals* scored a common set of photographs of wounds. In other words, the *ICC*s presented in Hart et al. (2006) estimated *across*-hospital inter-rater reliability. In this study, however, participants rated patients only at their hospital: *ICC*s in the current study *do not* examine *across*-hospital reliability. These *ICC*s merely evaluate if the participant raters at a particular hospital scored the NDNQI PrU risk and prevention measures with similar inter-rater reliability, as participant raters scored these measures at the other hospitals. This, and the large *ICC* confidence intervals, suggests the *ICC*s in the current study may not be very meaningful.

Aim 2: Methods and Processes Used to Collect NDNQI Pressure Ulcer Risk and Prevention Data

One hundred twenty participant raters completed the *Pressure Ulcer Risk and Prevention Reliability Survey.* Survey results suggest data collectors need training and review of the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers (NDNQI, 2013), which specify how these data should be collected.

Participant Rater Characteristics

Overall, participant raters were experienced in collecting NDNQI PrU data as reflected in the number of years they reported collecting data ($M = 4.3$ years), and the few raters, 8 (6.7%), who reported they had no other role in PrU data collection. However, 3 (8.3%) experts and 4 (4.8%) non-experts reported that within the last 12 months they had not reviewed or received *any* education or training for data collection on PrUs. This is concerning because the NDNQI (2013) recommends data collectors receive a review in PrU data collection prior to *each* skin survey.

Nevertheless, 75 (62.5%) participant raters reported that prior to each time they collect NDNQI PrU data, they "almost always" or "always" receive training or review.

Characteristics of the "expert" suggest these raters had greater expertise in wound/skin care and PrU data collection than the non-experts. Compared to the non-experts, a larger proportion of experts: (1) were wound/skin care nurses (47.2% to 17.9%); (2) had a Master's Degree in Nursing (33.3% to 14.3%); (3) completed all four modules of the NDNQI PrU Training program (72.2% to 47.6%); and (4) had other roles in PrU data collection, including training of the PrU data collection team (69.4% to 33.3%). In addition, 38.9% of experts reported having no certifications in wound care while 73.8% of non-experts reported the same.

On the other hand, responses to some survey items suggest the experts need review of the NDNQI PrU data collection guidelines as much (or more than) the non-experts. The NDNQI guidelines state moisture management must be documented in the patient record, but 80.6% of the experts "almost always" or "always" rate moisture management by direct observation, compared to 65.5% of the non-experts. Furthermore, 22.2% of the experts reported barrier cream at the bedside is "almost always" or "always" sufficient to rate moisture management as "Yes", compared to 19.0% of the non-experts. Finally, even though they are not supposed to do so, 36% of the experts "almost always" or "always" collect data on the unit where they usually work, as do 48.8% of the non-experts.

It is impressive that 76.7% of respondents reported that their skin survey team leader was certified in wound care, yet only 58.3% of the hospitals had at least 200 staffed beds. Results suggest that hospitals are sufficiently concerned about PrUs that they invest in credentialing these individuals.

The Methods Used to Collect NDNQI PrU Risk and Prevention Data

Respondents from no more than 13 (36.1%) hospitals reported that *only* RNs collect NDNQI PrU risk and prevention data. Even though there was a large number of missing data for items about who else besides RNs collect these data, responses still indicate that a large proportion of hospitals routinely use others such as nurses aids and LPNs to collect these data. The NDNQI (2013) identifies who should perform *skin inspections* during NDNQI PrU data collection (LPNs/Licensed Vocational Nurses and nurse aids should not perform skin inspections). However, specifications for *chart reviewers*—those who collect data on the PrU risk and prevention measures—are less specifically stated. The guidelines state these persons "need skill in reading documentation in patient records" (NDNQI, 2013, p. 104), but in the online teaching program, it is identified that skin survey team members should be trained on PrU data collection guidelines, skin and risk assessment, PrU identification and staging, PrU origin, and data extraction from the patient record (Bergquist-Beringer & Davidson, 2014). The majority of respondents (72.2%) reported completing the NDNQI PrU Training Program. It is difficult, however, to compare participant rater training with raters in earlier NDNQI PrU reliability studies because the NDNQI PrU training program did not exist during these previous studies. Chart reviewers should have received PrU team training and have demonstrated competence in performance of skin and risk assessments, and competence in data abstraction from the patient record (NDNQI, 2013). This requirement makes it difficult for LPNs, nurse aids, clinical secretaries, dieticians, nursing students, research assistants, or data specialists (e.g. those reported in the study as persons who usually collect these data) to meet the qualification to review patient records. Perhaps this affected the reliability of the study measures.

Only 45 (37.5%) respondents reported that their PrU data collection team "almost always" or "always" evaluates the reliability of this data at least once a year, as recommended by NDNQI. The NDNQI also recommends that data collectors collect data on units *other than* where they usually work in order to prevent bias reporting, but almost half of the respondents (*n* = 54, 45.0%) reported that they "almost always" or "always" *do* collect data on the unit where they usually work.

Most raters reported PrU risk and prevention data are "almost always" or "always" retrieved from the electronic patient record, *n* = 107 (89.2%). When asked if information technology (IT) extracts PrU risk and prevention data, 28 (23.3%) reported "I don't know". This item had the largest number of "I don't know" responses of all the items (with the second highest number being *n* = 16 (13.3%) for item #23 about allowing a 30-minute leeway for repositioning). There is no information as to the number of hospitals whose IT departments extract PrU risk and prevention data (S. Bergquist-Beringer, personal communication, February 9, 2015). Nevertheless, the location of data on PrU risk and prevention within the electronic patient record may have varied by unit and could have influenced study results.

Responses within survey items varied, highlighting patient scenarios where disagreement among these raters was (and is) likely to occur while rating patients on the PrU risk and prevention measures. All patients are to be included in NDNQI PrU data collection regardless of the time since their admission (NDNQI, 2013). Most respondents reported they "never" exclude a patient admitted less than 24 hours before the survey, *n* = 79 (65.8%), but 17 (14.2%) reported they "almost always" or "always" exclude these patients. A large majority reported they "almost always" or "always" determine risk status by the score obtained on the risk assessment scale, *n* = 98 (81.7%); while at the same time nearly the same number reported they "almost always" or

"always" look for other risk factors if the risk assessment scale score does not classify the patient at risk, *n* = 88 (73.3%). It is important that other risk factors are not overlooked because more than 100 factors have been associated with PrU development (Lyder & Ayello, 2008). In addition, PrU risk and prevention guidelines state other risk factors should be considered when determining risk status (NPUAP/EPUAP/PPPIA, 2014; Ratliff et al., 2010).

According to the NDNQI data collection guidelines (NDNQI, 2013) , only two PrU risk and prevention measures can be scored by direct observation of a patient in their room; *Pressure-redistribution surface use within the last 24 hours*, and *Nutritional support within the last 24 hours*. However, only 13 (10.8%) participant raters reported they "never" rate *Moisture management within the last 24 hours* by direct observation, while 84 (70%) reported they "almost always" or "always" do. In addition, less than half reported that barrier cream at the bedside is "never" sufficient to rate *Moisture management within the last 24 hours* as "Yes", *n* = 55 (45.8%). Evidence of moisture management includes documentation of such interventions in the patient's record. Incontinence products or barrier creams found in the patient's room are not adequate evidence of their use (NDNQI, 2013). Documentation of the use of these products must be found in the patient record for evidence that the intervention was performed (Bergquist-Beringer & Davidson, 2014).

There were four survey items specifically asking respondents about how they rate *Routine repositioning as prescribed within the last 24 hours*. Among all the measures, *Routine repositioning as prescribed within the last 24 hours* had the lowest across hospital Cohen's kappa value of .216, 95% CI [.124 – .298], and the third lowest PAK value of .577, 95% CI [.494 – 661]. Responses to the survey items provide insight into these values. Patient refusal of (or contraindications to) repositioning must be documented in the patient record (NDNQI, 2013).

Most respondents reported that if the nurse tells them that a patient refused repositioning, but there was no documentation of this refusal, then they "never" select "Patient refused", *n* = 79 (65.8%). Nevertheless, 9 (7.5%) respondents reported they "almost always" or "always" select "Patient refused" under this scenario. When asked if they allow a 30-minute leeway from when repositioning was to have occurred before they rate *Repositioning as prescribed* as "No", 32 (26.7%) respondents reported they "never" allow a 30-minute leeway, while nearly the same number of respondents, 39 (32.5%), reported they "almost always" or "always" do. Responses to this item varied *within* hospitals as only one hospital had participant raters who were unanimous in their responses to this item. The NDNQI (2013) guidelines do not specify whether or not data collectors should allow a leeway. In fact, the guidelines do not specify the time interval at which patients "at risk" for PrU should be repositioned other than "as prescribed" (p. 111). This is because many factors are to be considered while determining repositioning intervals (such as mattress type, patient age, and disease process). In addition, *each hospital's* repositioning procedure is to be followed when scoring this NDNQI measure; therefore, these procedures likely vary across hospitals. Perhaps each hospital needs to make their patient repositioning policy and documentation of its performance more apparent.

Responses also varied on where to look for information on repositioning. Thirty-nine (32.5%) raters reported they "never" look for physical therapy documentation of patient activity to rate repositioning, while 37 (30.8%) reported they "almost always" or "always" do. While the NDNQI PrU data collection guidelines do not specifically recommend looking for physical therapy documentation, responses to this survey item highlight the different methods data collectors use to score *Routine repositioning as prescribed*; methods that may affect inter-rater reliability. The NDNQI may want to include patient scenarios in their guidelines and the online

training program to further explain how data on repositioning should be collected. The survey results also suggest that inter-rater reliability may be better understood if survey data are examined by hospital.

Finally, most respondents, $n = 74$ (61.7%), reported they "never" rate *Nutritional support within the last 24 hours* as "Yes" if a nutritional consult is ordered for a patient who has been hospitalized for two or more days with poor intake, but the consult has not yet been completed. NDNQI guidelines state "If a pressure ulcer survey is being conducted early in the patient's hospital stay and a nutritional consult has been ordered, but not yet completed, this will count positively for nutritional support" (NDNQI, 2013, p. 103). Although not specifically stated, "early in a patient's hospital stay" is generally considered to be within the first 24 hours of admission.

Lastly, the inconsistency in the number of missing data among survey items was evident. Specifically, six items that asked who collects NDNQI PrU risk and prevention data had missing data rates of 32.5% to 39.2%. This was despite the response of "I don't know" being an option. In contrast, the rate of missing data for all the other survey items ranged from 0.8% to 3.3%.

Implications for the Conceptual Framework

Donabedian's model establishes the relationships between structures, processes, and outcomes. An assumption of the model is that the reliability of the NDNQI PrU process measures has been established (Donabedian, 2005). This study provides evidence to support the inter-rater reliability of eight NDNQI PrU risk and prevention process measures. They are (1) *Skin assessment within 24 hours of admission*, (2) *Risk assessment within 24 hours of admission*, (3) *Time since last PrU risk assessment*, (4) *Risk assessment scale*, (5) *Risk status*, (6) *Any prevention in use within the last 24 hours*, (7) *Skin assessment documented within the last 24*

hours, and (8) *Pressure-redistribution surface use within the last 24 hours*. Specifically, different NDNQI data collectors evaluated the same patients and in general, came to the same (or similar) conclusions about the performance of these PrU risk and prevention interventions. However, more research is needed on the inter-rater reliability of (1) *Routine repositioning as prescribed within the last 24 hours*, (2) *Nutritional support within the last 24 hours*, and (3) *Moisture management within the last 24 hours,* to provide support for the equivalence of these process measures in Donabedian's model.

Implications for Practice

Results from this study should be used to focus education efforts to clarify data collection procedures and improve reliability. Specifically, education efforts should target areas of disagreement identified in agreement matrices and in responses to the *Pressure Ulcer Risk and Prevention Reliability Survey*. Items from this survey, and similar items, might be integrated into the NDNQI PrU Training Program to clarify how PrU risk and prevention data are to be collected. Often people think they understand something until they are questioned about it, such as was done during survey completion. This lack of understanding was reflected in the "I don't know" responses. Clearly the results of the *Pressure Ulcer Risk and Prevention Reliability Survey* suggest confusion exists in how to rate patients within particular scenarios. Site coordinators need to clearly specify how data are to be collected, especially with regard to their hospital's policy on repositioning.

Data collectors must engage in ongoing training and review of the NDNQI Guidelines for Data Collection and Submission. In particular, education on when an intervention should be rated as "Unnecessary" (for a patient) or "No" (not received). Raters may be unfamiliar with the use of Braden *sub*scale scores to better identify appropriate PrU prevention interventions,

compared to use of the total Braden Scale score (Maklebust & Magnan, 2009). Education is needed at the hospital level on use of Braden *sub*scale scores to determine if an intervention is needed or "Unnecessary". On the other hand, even if a total Braden score is > 18, participant raters should know when to classify certain persons "at risk", such as anyone with a history of PrU or someone admitted after surgery lasting 5 hours or more. Another important consideration is that some hospitals may need to enhance the level of documentation in the patient health record to support a measure that asks if an intervention is "Unnecessary", "Contraindicated", or if the "Patient refused".

Even though NDNQI PrU risk and prevention data were collected on 1,637 patients by 120 raters, many response options were *never* selected. Whether or not these response options continue to be included in the NDNQI PrU risk and prevention measures for patients in the acute care setting should be considered. Elimination of responses rarely or never selected would make data collection less burdensome, and may improve reliability of these measures' data. However, this must be weighed against the reason for their presence.

Limitations

Despite sending 750 invitations, only 63 hospitals accepted the *Invitation to Participate* (8.4% response rate). Of these 63 hospitals; 18 withdrew, 4 never submitted data, and 5 submitted data that were not usable. Attrition rate was 42.9%. Consequently, it may be difficult to generalize the study's findings to other NDNQI-participating hospitals that submit PrU data. Moreover, the sample of hospitals, units, or participant raters may not represent the population of NDNQI participating hospitals, units, or data collectors that submit PrU data. In addition, the number of patients at each hospital who were rated on the six *prevention* measures was small (range = 5 to 37, M =14.7), and this likely contributed to the large standard deviations observed

for some hospital's Cohen's kappa and PAK values. The majority of hospitals (n = 25, 69.4%) had only two expert/non-expert pairs. So agreement estimates for the majority of hospitals were calculated from only two kappa values. The agreement estimates for another 10 (27.8%) hospitals were calculated from three kappa values, and the agreement estimates for 1 (2.8%) hospital was calculated from four kappa values. On the other hand, most averaged (across hospital) Cohen's kappa and PAK values (aim 1, Q3) had narrow 95% CIs, providing support for the accuracy of these average (across hospital) agreement estimates. Despite the low response rate and high attrition rate, the sample of hospitals and participant raters is similar to samples in comparable NDNQI research, and reflect the population of NDNQI hospitals that submit PrU data. Even so, generalizability of the survey results is limited by the > 15% of missing responses for the survey items on who collects quarterly NDNQI PrU risk and prevention data, suggesting that responses to these items may not reflect actual practice (Tabachnick & Fidell, 2007).

Prevalence-adjusted kappa (PAK) values were presented as estimates of inter-rater reliability for the study measures because of the high prevalence of a single response for the risk and prevention measures. However, PAK values do not convey the circumstances in which the initial ratings were made in that PAK values are calculated from the average cell counts of agreement (Hoehler, 2000). Nevertheless, PAK values reported, which adjusted for prevalence, provided insight into how prevalence affected the Cohen's kappa values (J. Sim, personal communication, May 5, 2014; Sim & Wright, 2005). The *ICC*s calculated in this study included only two kappa values per hospital (N = 72 kappa values), but average agreement estimates for each measure were determined from 2 to 4 kappa values per hospital (N = 84). Therefore, the *ICC*s may not have accurately captured the ratio of the between-hospital variance in kappa values to the total variance in kappa values.

Participant raters may not have followed study protocol and inappropriately shared information. Study instructions, however, stressed the importance of rating *independently* of others, and a staggered start for data collection was suggested in these instructions. Data collection procedure may have adversely influenced findings for the measure, *Time since last PrU risk assessment*. Also, the small number of "at risk" patients at some hospitals (i.e. those who were rated on the *prevention* measures), may have affected agreement estimates for this as well as the prevention measures.

Participant raters may have rated patients differently than how they usually rate patients just because they were part of a reliability study. The study created an artificial data collection situation where participants were not allowed to ask advice from other participants. In practice, data collectors work in teams and PrU data collection likely involves collaboration among data collectors, especially if they are uncertain of how to rate a patient. Participant raters may not have provided honest answers on the online survey because their responses may have revealed that they—or their hospital—do not follow the NDNQI data collection guidelines. Anonymity, however, was assured. Finally, *intra*-rater reliability is an assumption for *inter*-rater reliability testing (Gwet, 2012); but *intra*-rater reliability has not been established.

Future Research

Further research is needed to provide support for the NDNQI PrU risk and prevention measures. Specifically, inter-rater reliability studies with a larger patient sample than the current study's sample are needed. Regression analysis should also be conducted on this study's data to identify rater and hospital characteristics that influence reliability. In order to improve reliability of the NDNQI PrU risk and prevention measures, nurse researchers should conduct additional research on the methods and processes used to collect these data, especially the prevention

measures. In addition, *intra*-rater reliability must be established, as it is an assumption of *inter*-rater reliability. Responses of "sometimes", "usually", and "almost always" leads one to question *intra*-rater reliability. Finally, *across*-hospital reliability for the 11 NDNQI PrU risk and prevention measures must be examined and established, especially because performance of PrU risk and prevention interventions is compared across hospitals. This *across*-hospital inter-rater reliability research may be an online study presenting patient scenarios to raters across hospitals. For instance, actual patient scenarios could be randomly selected from existing electronic patient health records. In addition, for each of these scenarios, a corresponding photo of the patient room could be staged—this is because some measures can be (and are) rated by direct observation. These patient records and photos could then be presented to data collectors across the nation for them to rate each "patient" on the NDNQI PrU risk and prevention measures. If future studies replicate the reliability results on the prevention intervention measures found in this study, their value as quality of care measures might be reconsidered.

Conclusion

This is the first known study to examine the reliability of the NDNQI PrU risk and prevention measures. Overall findings provide support for the reliability of the five PrU *risk* measures (*Skin assessment within 24 hours of admission*, *Risk assessment within 24 hours of admission*, *Time since last PrU risk assessment*, *Risk assessment scale*, and *Risk Status*), and three of the six PrU *prevention* measures (*Any prevention in use within the last 24 hours*, *Skin assessment documented within the last 24 hours*, and *Pressure-redistribution surface use within the last 24 hours*). Cohen's kappa values were adjusted for prevalence because nearly all the patients received a PrU risk assessment, and most of those at risk received interventions to prevent them. The overall reliability of the PrU *risk* measures was higher than the overall

reliability of the PrU *prevention* measures. Education and further research are needed to improve the reliability of three of the six prevention measures (*Routine repositioning as prescribed*, *Nutritional support*, and *Moisture management*).

References

A crosswalk linking national quality indicators with national organizations, legislation, and clinical practice guidelines supporting use of the indicators (n. d). Retrieved from http://www.ashp.org/s_ashp/docs/files/QII_Crosswalk062907.pdf

Abel, R. L., Warren, K., Bean, G., Gabbard, B., Lyder, C., Bing, M., & McCauley, C. (2005). Quality improvement in nursing homes in Texas: Results from a pressure ulcer prevention project. *Journal of the American Medical Directors Association, 6*, 181-188. doi: 10.1016/j.jamda.2005.03.011

Adams, J. L., Mehrotra, A., Thomas, W., & McGlynn, E. A. (2010). Physician cost profiling – Reliability of risk misclassification. *New England Journal of Medicine, 362*, 1014-1021.

Agency for Healthcare Research & Quality. (n. d.a.). *Patient Safety Organizations: Common Formats*. Retrieved from http://www.pso.ahrq.gov/formats/commonfmt.htm

Agency for Healthcare Research & Quality (n. d.b). *Preventing pressure ulcers in hospitals.* Retrieved from https://www.premierinc.com/safety/topics/pressure-ulcer/pressure-ulcer-downloads/Preventing-Pressure-Ulcers-Hospitals.pdf

Agency for Healthcare Research & Quality. (2011a). AHRQ releases standardized hospital bed definitions. Retrieved from http://archive.ahrq.gov/research/havbed/definitions.htm

Agency for Healthcare Research & Quality. (2011b, April). *3. What are the best practices in pressure ulcer prevention that we want to use? Preventing pressure ulcers in hospitals: A toolkit for improving quality of care.* Retrieved from http://www.ahrq.gov/professionals/systems/long-term-care/resources/pressure-ulcers/pressureulcertoolkit/putool3.html

Agency for Healthcare Research & Quality. (2012a). *AHRQ at a glance*. Retrieved from http://www.ahrq.gov/about/mission/glance/index.html

Agency for Healthcare Research & Quality. (2012b). *Frequently asked questions: What are the quality indicators?* Retrieved from https://info.ahrq.gov/app/answers/detail/a_id/106/kw/PSI/related/1

Agency for Healthcare Research & Quality. (2013a, May). *National Quality Forum (NQF) endorsed individual and composite measures.* Retrieved from http://www.qualityindicators.ahrq.gov/Downloads/Modules/V45/Module%20NQF%20Endorsement%20V4.5.pdf

Agency for Healthcare Research & Quality. (2013b, May). *Pressure ulcer rate technical specification: Patient Safety Indicator #3 (PSI #3).* Retrieved from http://qualityindicators.ahrq.gov/Downloads/Modules/PSI/V45/TechSpecs/PSI%2003%20Pressure%20Ulcer%20Rate.pdf

Agency for Healthcare Research & Quality. (2013c, May). *Pressure ulcer rate technical specification: Pediatric Quality Indicator (PDI #2).* Retrieved from http://qualityindicators.ahrq.gov/Downloads/Modules/PDI/V45/TechSpecs/PDI%2002%20Pressure%20Ulcer%20Rate.pdf

Allman, R. M., Goode, P. S., Patrick, M. M., Burst, N., & Bartolucci, A. A. (1995). Pressure ulcer risk factors among hospitalized patients with activity limitation. *Journal of the American Medical Association, 273*, 865-870. doi: 10.1001/jama.1995.03520350047027

American Nurses Association. (2013). *The national database*. Retrieved from http://www.nursingworld.org/MainMenuCategories/ThePracticeofProfessionalNursing/PatientSafetyQuality/Research-Measurement/The-National-database.aspx

American Nurses Association. (2014). *Frequently asked questions about NDNQI*. Retrieved from http://www.nursingquality.org/FAQ#faq-measures

American Nurses Credentialing Center. (2013). *ANCC Magnet Recognition Program®*. Retrieved from http://www.nursecredentialing.org/Magnet.aspx

Anthony, D., Papanikolaou, P., Parboteeah, S., & Saleh, M. (2010). Do risk assessment scales for pressure ulcers work? *Journal of Tissue Viability, 19*(4), 132-136. doi: 10.1016/j.jtv.2009.11.006

Anthony, D., Rafter, L., Reynolds, T., & Aljezawi, M. (2011). An evaluation of serum albumin and the sub-scores of the Waterlow score in pressure ulcer risk assessment. *Journal of Tissue Viability, 20*(3), 89-99.

Antokal, S., Brienza, D., Bryan, N., Herbe, L., Logan, S., Maguire, J., . . . Siddiqui, A. (2012, November 20). *Friction induced skin injuries - Are they pressure ulcers?* A National Pressure Ulcer Advisory Panel White Paper. Retrieved from http://www.npuap.org/wp-content/uploads/2012/01/NPUAP-Friction-White-Paper.pdf

Armstrong, D. G., Ayello, E. A., Capitulo, K. L., Fowler, E., Krasner, D. L., Levine, J. M., . . . Smith, A. P. (2008). New opportunities to improve pressure ulcer prevention and treatment: Implications of the CMS inpatient hospital care Present on Admission (POA) indicators/hospital-acquired conditions (HAC) policy. A consensus paper from the International Expert Wound Care Advisory Panel. *Journal of Wound Ostomy Continence Nursing, 35*, 485-492. doi: 10.1097/01.won.0000335960.68113.82

Athlin, E., Idvall, E., Jernfält, M., & Johansson, I. (2009). Factors of importance to the development of pressure ulcers in the care trajectory: Perceptions of hospital and community care nurses. *Journal of Clinical Nursing, 19*, 2252-2258. doi: 10.1111/j.1365-2702.2009.02886.x

Ayello, E. A., & Braden, B. (2002). How and why to do pressure ulcer risk assessment? *Advances in Skin & Wound Care, 15*, 125-133.

Ayello, E. A., Capitulo, K. L., Fife, C. E., Fowler, E., Krasner, D. L., Mulder, G., . . . Yankowsky, K. W. (2009). Legal issues in the care of pressure ulcer patients: Key concepts for health care providers: A consensus paper from the International Expert Wound Care Advisory Panel. *Journal of Palliative Medicine, 12*, 995-1008. doi: 10.1089=jpm.2009.9939

Ayello, E. & Lyder, C. H. (2008). A new era of pressure ulcer accountability in acute care. *Advances in Skin & Wound Care, 21*, 134-140.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis (2nd ed.)*. Cambridge, UK: Cambridge University Press.

Bardach, N. S., Chien, A. T., & Dudley, A. (2010). Small numbers limit the use of the inpatient pediatric quality indicators for hospital comparison. *Academic Pediatrics, 10*, 266-273. doi: 10.1016/j.acap.2010.04.025

Bates-Jensen, B. M., Alessi, C. A., Al-Samarrai, N. R., & Schnelle, J. F. (2003a). The effects of an exercise and incontinence intervention on skin health outcomes in nursing home residents. *Journal of the American Geriatrics Society, 51*, 348-355. doi: 10.1046/j.1532-5415.2003.51108.x

Bates-Jensen, B. M., Cadogan, M., Osterweil, D., Levy-Storms, L., Jorge, J., Al-Samarrai, N., . . . Schnelle, J. F. (2003b). The Minimum Data Set Pressure Ulcer indicator: Does it reflect differences in care processes related to pressure ulcer prevention and treatment in nursing homes? *Journal of the American Geriatrics Society, 51*, 1203-1212. doi: 10.1046/j.1532-5415.2003.51403.x

Baumgarten, M., Margolis, D. J., Orwig, D. L., Shardell, M. D., Hawkes, W. G., Langenberg, P., . . . Magaziner, J. (2009). Pressure ulcers in elderly patients with hip fracture across the continuum of care. *Journal of the American Geriatric Society, 57*, 863-870.

Beeckman, D., Defloor, T., Schoonhoven, L., & Vanderwee, K. (2011). Knowledge and attitudes of nurses on pressure ulcer prevention: A cross-sectional multicenter study in Belgian hospitals. *Worldviews on Evidence-Based Nursing, 8*, 166-176. doi: 10.1111/j.1741-6787.2011.00217.x

Bergquist, S., & Frantz, R. (2001). Braden scale: Validity in community-based older adults receiving home health care. *Applied Nursing Research, 14*(1), 36-43. doi: 10.1053/apnr.2001.21079

Bergquist-Beringer, S. (2011, February 25). *National Database of Nursing Quality Indicators (NDNQI) update*. PowerPoint presentation at the NPUAP 12th National Biennial Conference: Emerging healthcare issues.

Bergquist-Beringer, S., & Davidson, J. (2014). *NDNQI Quality Improvement Solutions from ANA: Pressure Ulcer Training*. Retrieved from https://members.nursingquality.org/NDNQIPressureUlcerTraining/Default.aspx

Bergquist-Beringer, S., Davidson, J., Agosto, C., Linde, N. K., Abel, M., Spurling, K., . . . Christopher, A. (2009). Evaluation of the National Database of Nursing Quality Indicators (NDNQI) Training Program on Pressure Ulcers. *Journal of Continuing Education in Nursing, 40*, 252-258; quiz 259-260, 279.

Bergquist-Beringer, S., Dong, L., He, J., & Dunton, N. (2013). Pressure ulcers and prevention among acute care hospitals in the United States. *The Joint Commission Journal on Quality and Patient Safety, 39*, 404-414.

Bergquist-Beringer, S., Gajewski, B., & Davidson, J. (2012). Pressure ulcer prevalence and incidence: Report from the National Database of Nursing Quality Indicators (NDNQI). In B. Pieper (Ed.), *Pressure ulcers: Prevalence, incidence, and implications for the future* (pp. 175-187). Washington, DC: NPUAP.

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

Bergstrom, N., & Braden, B. J. (2002). Brief report. Predictive validity of the Braden Scale among black and white subjects. *Nursing Research, 51*, 398-403.

Bergstrom, N., Braden, B., Kemp, M., Champagne, M., & Ruby, E. (1998). Predicting Pressure Ulcer risk: a multisite study of the predictive validity of the Braden Scale. *Nursing Research, 47*, 261-269.

Bergstrom, N., Braden, B. J., Laguzza, A., & Holman, V. (1987). The Braden Scale for Predicting Pressure Sore Risk. *Nursing Research, 36*, 205-210.

Bergstrom, N., Horn, S. D., Rapp, M. P., Stern, A., Barrett, R., & Watkiss, M. (2013). Turning for ulcer reduction: A multisite randomized clinical trial in nursing homes. *Journal of the American Geriatrics Society, 61*, 1705-1713. doi: 10.1111/jgs.12440

Berlowitz, D., VanDeusen, C., Parker, V., Niederhauser, A., Silver, J., Logan, C., . . . Zulkowski, K. (n. d.). *Preventing pressure ulcers in hospitals*. Retrieved from https://www.premierinc.com/safety/topics/pressure-ulcer/pressure-ulcer-downloads/Preventing-Pressure-Ulcers-Hospitals.pdf

Black, J., Clark, M., Dealey, C., Brindle, C., Alves, P., Santamaria, N., & Call, E. (2014). Dressings as an adjunct to pressure ulcer prevention: Consensus panel recommendations. *International Wound Journal*. doi:10.1111/iwj.12197

Black, J. M., Edsberg, L. E., Baharestani, M. M., Langemo, D., Goldberg, M., McNichol, L., & Cuddigan, J. (2011). Pressure ulcers: Avoidable or unavoidable? Results of the National Pressure Ulcer Advisory Panel Consensus Conference. *Ostomy Wound Management, 57*(2), 24-37.

Blue Cross and Blue Shield of Connecticut. (2012, August 14). *Anthem Blue Cross and Blue Shield in Connecticut facility policy.* Retrieved from http://www.anthem.com/provider/noapplication/f1/s0/t0/pw_e170997.pdf?refer=ahpprovider

Bourdel-Marchasson, I., Barateau, M., Rondeau, V., Dequae-Merchadou, L., Salles-Montaudon, N., Emeriau, J.-P., . . . Dartigues, J.-F. (2000). A multi-center trial of the effects of oral nutritional supplementation in critically ill older inpatients. *Nutrition, 16*(1), 1-5. doi: http://dx.doi.org/10.1016/S0899-9007(99)00227-0

Braden, B. J. (2012). The Braden Scale for predicting pressure sore risk: Reflections after 25 years. *Advances in Skin & Wound Care, 25*, 61. doi: 10.1097/01.ASW.0000411403.11392.10

Braden, B. J., & Bergstrom, N. (1994). Predictive validity of the Braden Scale for pressure sore risk in a nursing home population. *Research in Nursing & Health, 17*, 459-470.

Brindle, C. T., & Wegelin, J. A. (2012). Prophylactic dressing application to reduce pressure ulcer formation in cardiac surgery patients. *Journal of Wound, Ostomy, and Continence Nursing, 39*, 133-142.

Bryant, R. A., & Rolstad, B. S. (2001). Utilizing a systems approach to implement pressure ulcer prediction and prevention. *Ostomy Wound Management, 47*(9), 26-36.

Centers for Medicare and Medicaid Services. (n. d.). *About the partnership*. Retrieved from http://partnershipforpatients.cms.gov/about-the-partnership/what-is-the-partnership-about/lpwhat-the-partnership-is-about.html

Centers for Medicare and Medicaid Services. (2013a, August 2). *Fact sheet: CMS final rule to improve quality of care during hospital inpatient stays*. Retrieved from http://www.cms.gov/newsroom/mediareleasedatabase/fact-sheets/2013-fact-sheets-items/2013-08-02-3.html

Centers for Medicare and Medicaid Services. (2013b, June 13). *History*. Retrieved from http://www.cms.gov/About-CMS/Agency-Information/History/index.html?redirect=/history/

Chan, B. C., Nanwa, N., Mittmann, N., Bryant, D., Coyte, P. C., & Houghton, P. E. (2012). The average cost of pressure ulcer management in a community dwelling spinal cord injury population. *The Journal of Spinal Cord Medicine, 35*, 473-474. doi: 10.1111/j.1742-481X.2012.01002.x

Chi-square test. (n. d.). Retrieved from http://www2.lv.psu.edu/jxm57/irp/chisquar.html

Choi, J., Bergquist-Beringer, S., & Staggs, V. (2013). Linking RN workgroup job satisfaction to pressure ulcers among older adults on acute care hospital units. *Research in Nursing & Health, 36*, 181-190. doi: 10.1002/nur.21531

Choi, J., Boyle, D. K., & Dunton, N. (2014). A standardized measure: NDNQI nursing care hours indicator. *Western Journal of Nursing Research, 36*(1), 105-116. doi: 10.1177/0193945913501723

Chou, R., Dana, T., Bougatsos, C., Blazina, I., Starmer, A., Reitel, K., & Buckely, D. (2013, May). *Pressure ulcer risk assessment and prevention: Comparative effectiveness. Comparative effectiveness review no. 87 (Prepared by Oregon Evidence-based Practice Center under Contract no. 290-2007-10057-I.* Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Clancy, C. (Producer). (2008, May). *Measuring health care quality* (AHRQ Narrated PowerPoint Presentation). Retrieved from http://www.kaiseredu.org/tutorials/quality/player.html

Clancy, M. J. (2013). Pressure redistribution devices: What works, at what cost and what's next? *Journal of Tissue Viability, 22*(3), 57-62. doi: 10.1016/j.jtv.2013.04.002

Clinical practice guideline for pressure ulcer prevention and treatment. (2013, August 19). *Wounds*. Retrieved from http://www.woundsresearch.com/news/clinical-practice-guideline-pressure-ulcer-prevention-and-treatment

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-47.

Colin, D., Rochet, J.-M., Ribinik, P., Barrois, B., Passadori, Y., & Michel, J.-M. (2012). What is the best support surface in prevention and treatment, as of 2012, for a patient at risk and/or suffering from pressure ulcer sore? Developing French guidelines for clinical practice. *Annals of Physical and Rehabilitation Medicine, 55*, 466-481. doi: 10.1016/j.rehab.2012.08.002

Comfort, E. H. (2008). Reducing pressure ulcer incidence through Braden Scale risk assessment and support surface use. *Advances in Skin & Wound Care, 21*, 330-334. doi: 10.1097/01.ASW.0000323519.08306.ea

Coomer, N. M., & McCall, N. T. (2012, April). *Examination of the accuracy of coding pressure ulcer stages: Final report* (CMS Contract No. HHSM-500-2005-00029I). Prepared for Centers for Medicare & Medicaid Services.

Courtney, B. A., Ruppman, J., B., & Cooper, H. M. (2006). Save our skin: Initiative cuts pressure ulcer incidence in half. *Nursing Management, 37*(4), 36-38, 40.

Defloor, T. (2000). The effect of position and mattress on interface pressure. *Applied Nursing Research, 13*(1), 2-11.

Defloor, T., De Bacquer, D., & Grypdonck, M. H. (2005). The effect of various combinations of turning and pressure reducing devices on the incidence of pressure ulcers. *International Journal of Nursing Studies, 42*, 37-46. doi: 10.1016/j.ijnurstu.2004.05.013

Defloor, T., & Grypdonck, M. H. (2005). Pressure ulcers: Validation of two risk assessment scales. *Journal of Clinical Nursing, 14*, 373-382

Department of Health & Human Services. (n. d.). *Health Information Privacy: Patient Safety and Quality Improvement Act of 2005 Statue and Rule*. Retrieved from http://www.hhs.gov/ocr/privacy/psa/regulation

Department of Health & Human Services. (2007). *Healthy People 2020 summary objectives: Older adults*. Retrieved from http://www.healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicId=31

de Vos, M., Graafmans, W., Kooistra, M., Meijboom, B., & Westert, G. (2009). Using quality indicators to improve hospital care: A review of the literature. *International Journal for Quality in Health Care, 21*(2), 119-129. doi: 10.1093/intqhc/mzn059

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., . . .Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship, 39*, 155-164.

Duffy, J. (2009). *Quality caring in nursing: Applying theory to clinical practice, education, and leadership*. NY: Springer.

Donabedian, A. (1978). The quality of medical care: Methods for assessing and monitoring the quality of care for research and for quality assurance programs. *Science, 200*, 856-864.

Donabedian, A. (2005). Evaluating the quality of medical care. 1966. *The Milbank Quarterly, 83*, 691-729.

Duncan, K. D. (2007). 5 Million Lives Campaign: Preventing pressure ulcers: The goal is zero. *Joint Commission Journal on Quality and Patient Safety, 33*, 605-610.

Elliott, C. G. S. (2006). *Using aggregated micro-level data as measures of macro-level phenomena: The case of the NDNQI-RN Satisfaction Survey.* Ph.D., University of Kansas. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=rzh&AN=2009711617&site=ehost-live Available from EBSCOhost rzh database

Fabian, L.A. & Geppert, J. (2011, May). *Quality Indicator Measure Development, Implementation, Maintenance, and Retirement Summary* (Prepared by Battelle, under Contract No. 290-04-0020). Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www.qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/QI%20Measure%20Development%20Implementation%20Maintenance%20Retirement%20Summary%2005-03-11.pdf

Feinstein, A. R., & Cicchetti, D. (1990). High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology, 43*, 543-549.

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). Hoboken, NJ: Wiley-Interscience.

Fowler, E., Scott-Williams, S., & McGuire, J. B. (2008). Practice recommendations for preventing heel pressure ulcers. *Ostomy Wound Management, 54*(10), 42-48, 50-42, 54-47.

Frenk, J. (2000). Obituary Avedis Donabedian. *Bulletin of the World Health Organization, 78*, 1475.

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

Garber, S. L., Rintala, D. H., Hart, K. A., & Fuhrer, M. J. (2000). Pressure ulcer risk in spinal cord injury: Predictors of ulcer status over 3 years. *Archives of Physical Medicine and Rehabilitation, 81*, 465-471.

Gorecki, C., Brown, J. M., Nelson, E. A., Briggs, M., Schoonhoven, L., Dealey, C., . . . Nixon, J. (2009). Impact of pressure ulcers on quality of life in older patients: A systematic review. *Journal of the American Geriatrics Society, 57*, 1175-1183. doi: 10.1111/j.1532-5415.2009.02307.x

Green, S. B., & Salkind, N. J. (2008). Lesson 25: One-way analysis of variance. *Using SPSS for Windows and Macintosh Analyzing and Understanding Data (5th ed.)*. Upper Saddle River, NJ: Pearson Prentice Hall.

Grimes, D. A., & Schultz, K. F. (2002). Descriptive studies: What they can and cannot do. *The Lancet, 359*, 145-149.

Gwet, K. L. (2012). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among multiple raters* (3rd ed.). Gaithersburg, MD: Advanced Analytics, LLC.

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

He, J., Staggs, V. S., Bergquist-Beringer, S., & Dunton, N. (2013). Unit-level time trends and seasonality in the rate of hospital-acquired pressure ulcers in US acute care hospitals. *Research in Nursing & Health, 36*, 171-180. doi: 10.1002/nur.21527

HealthGrades, Inc. (2013). *Variation in patient safety outcomes and the importance of being informed: Announcing the HealthGrades 2013 Patient Award Recipients.* Retrieved from https://d2dcgio3q2u5fb.cloudfront.net/54/b3/4e421ea847c1ba28d357e76b982e/2013-variation-in-patient-safety-outcomes-and-the-importance-of-being-informed.pdf

Hodge, J., Mounter, J., Gardner, G., & Rowley, G. (1990). Clinical trial of the Norton Scale in acute care settings. *Australian Journal of Advanced Nursing, 8*(1), 39-46.

Hoehler, F. K. (2000). Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. *Journal of Clinical Epidemiology, 53*, 499-503. doi: http://dx.doi.org/10.1016/S0895-4356(99)00174-2

Houwing, R. H., Rozendaal, M., Wouters-Wesseling, W., Beulens, J. W. J., Buskens, E., & Haalboom, J. R. (2003). A randomised, double-blind assessment of the effect of

nutritional supplementation on the prevention of pressure ulcers in hip-fracture patients. *Clinical Nutrition (Edinburgh, Scotland), 22*, 401-405.

Hunt, D. R., Verzier, N., Abend, S. L., Lyder, C., Jaser, L. J., Safer, N., & Davern, P. (2005). Fundamental of Medicare patient safety surveillance: Intent, relevance, and transparency. *Advances in Patient Safety, 2*, 105-117.

Iizaka, S., Okuwa, M., Sugama, J., & Sanada, H. (2010). The impact of malnutrition and nutrition-related factors on the development and severity of pressure ulcers in older patients receiving home care. *Clinical Nutrition, 29*, 47-53. doi: http://dx.doi.org/10.1016/j.clnu.2009.05.018

Institute for Healthcare Improvement (2011). *How to guide: Prevent pressure ulcers*. Cambridge, MA: Institute for Healthcare Improvement. Retrieved from http://www.ihi.org/knowledge/Knowledge%20Center%20Assets/Tools%20-%20How-toGuidePreventPressureUlcers_0a06721b-7d9b-42f6-8d40-53347fa21b65/HowtoGuidePreventPressureUlcers.pdf

Institute for Healthcare Improvement (2015). *Measures: Percent of patients receiving daily pressure ulcer risk reassessment*. Retrieved from http://www.ihi.org/resources/Pages/Measures/PercentPatientsDailyPressureUlcerRiskReassessment.aspx

Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st Century*. Washington, DC: National Academy Press.

Institute of Medicine. (2005). *Performance measurement: Accelerating improvement*. Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=11517&page=R1

Institute of Medicine. (2013). *About the IOM*. Retrieved from http://www.iom.edu/About-IOM.aspx

Joint Commission. (2010). *Implementation guide for the NQF endorsed nursing-sensitive care measure set*. Oakbrook Terrace, IL: Author

Joint Commission. (2013a). *About the Joint Commission*. Retrieved from http://www.jointcommission.org/about_us/about_the_joint_commission_main.aspx

Joint Commission. (2013b). *History of the Joint Commission*. Retrieved from http://www.jointcommission.org/about_us/history.aspx

Joint Commission. (2013c). *National Patient Safety Goals Effective January 1, 2014: Long term care accreditation program*. Retrieved from http://www.jointcommission.org/assets/1/6/LT2_NPSG_Chapter_2014.pdf

Joint Commission. (2014). *2014 National Patient Safety Goals*. Retrieved from http://www.jointcommission.org/standards_information/npsgs.aspx

Källman, U., & Lindgren, M. (2014). Predictive validity of 4 risk assessment scales for prediction of pressure ulcer development in a hospital setting. *Advances in Skin & Wound Care, 27*, 70-76. doi: 10.1097/01.ASW.0000439059.72199.41

Klaus, S. F., Dunton, N., Gajewski, B., & Potter, C. (2013). Reliability of the nursing care hour measure: A descriptive study. *International Journal of Nursing Studies, 50*, 924-932. doi: 10.1016/j.ijnurstu.2012.07.012

Kosiak, M. (1959). Etiology and pathology of ischemic ulcers. *Archives of Physical Medicine and Rehabilitation, 40*(2), 62-69.

Kottner, J., & Balzer, K. (2010). Do pressure ulcer risk assessment scales improve clinical practice? *Journal of Multidisciplinary Healthcare, 3*, 103-111.

Kottner, J., Balzer, K., Dassen, T., & Heinze, S. (2009). Pressure ulcers: A critical review of definitions and classifications. *Ostomy Wound Management, 55*(9), 22-29.

Kring, D. L. (2007). Reliability and validity of the Braden Scale for predicting pressure ulcer risk. *Journal of Wound, Ostomy & Continence Nursing, 34*, 399-406.

Landis, E. M. (1930, May). Micro-injection studies of capillary blood pressure in human skin. *Heart, 15*, 209-228

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Langemo, D., Cuddigan, J., Baharestani, M., Ratliff, C., Posthauer, M. E., Black, J., & Garber, S. (2008). Pressure ulcer guidelines: "Minding the gaps" when developing new guidelines. *Advances in Skin & Wound Care, 21*, 213-217. doi: 10.1097/01.ASW.0000305445.41317.56

Langer, G., Knerr, A., Kuss, O., Behrens, J., & Schlömer, G. J. (2008). Nutritional interventions for preventing and treating pressure ulcers. *Cochrane Database of Systematic Reviews,* (4). doi:10.1002/14651858.CD003216 Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003216/abstract;jsessionid=8D772A1D7BF40FA302894B2306DD5AFF.f02t02

Levine, J. M., Ayello, E. A., Zulkowski, K. M., & Fogel, J. (2012). Pressure ulcer knowledge in medical residents: An opportunity for improvement. *Journal of Advanced Skin & Wound Care, 25*(3), 115-117. doi: 10.1097/01.asw.0000412908.43335.46

Library of Congress. (2003). *Bill summary & status 108th Congress (2003 - 2004) H.R.1 all information*. Retrieved from http://thomas.loc.gov/cgi bin/bdquery/z?d108:HR00001:@@@L&summ2=m&

Lloyd, J. W. (n. d.). Assessing agreement: The confusion matrix. from http://faculty.virginia.edu/johnlloyd/edlf7330/resources/ConfusionMatrices.pdf

Lyder, C. H., & Ayello, E. A. (2008). Pressure ulcers: A patient safety issue. In: R. G. Hughes (Eds.), *Patient safety and quality: An evidence-based handbook for nurses* (pp. 1-269-299). Rockville (MD): Agency for Healthcare Research and Quality. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK2650/

Lyder, C. H., Wang, Y., Metersky, M., Curry, M., Kliman, R., Verzier, N., & Hunt, D. R. (2012). Hospital-acquired pressure ulcers: Results from the national Medicare Patient Safety Monitoring System study. *Journal of the American Geriatrics Society, 60*, 1603-1608. doi: 10.1111/j.1532-5415.2012.04106.x

Lyman, V. (2009). Successful heel pressure ulcer prevention program in a long-term care setting. *Journal of Wound, Ostomy, and Continence Nursing, 36*, 616-621. doi: 10.1097/WON.0b013e3181bd813e

McInerney, J. A. (2008). Reducing hospital-acquired pressure ulcer prevalence through a focused prevention program. *Advances in Skin & Wound Care, 21*, 75-78. doi: 10.1097/01.ASW.0000305410.58350.34

McInnes, E., Jammali-Blasi, A., Bell-Syer, S., Dumville, J., & Cullum, N. (2012). Preventing pressure ulcers—Are pressure-redistributing support surfaces effective? A Cochrane systematic review and meta-analysis. *International Journal of Nursing Studies, 49*, 345-359. doi: 10.1016/j.ijnurstu.2011.10.014

Meddings, J. A., Reichert, H., Hofer, T., & McMahon Jr., L. F. (2013). Hospital report cards for hospital-acquired pressure ulcers: How good are the grades? *Annals of Internal Medicine, 159*, 505-513. doi: 10.7326/0003-4819-159-8-201310150-00003

Minnesota Department of Health. (2014). *Adverse events in Minnesota: Tenth annual public report/January 2014*. Retrieved from http://www.health.state.mn.us/patientsafety/ae/2014ahereport.pdf

Mitchell, P. H., Ferketich, S., & Jennings, B. M. (1998). Quality health outcomes model. American Academy of Nursing expert panel on quality health care. *Image: Journal of Nursing Scholarship, 30*, 43-46.

Montalvo, I. (2007). The National Database of Nursing Quality Indicators (NDNQI). *The Online Journal of Issues in Nursing, 12*(3), 1-13. Manuscript 2. Retrieved from http://www.nursingworld.org/MainMenuCategories/ANAMarketplace/ANAPeriodicals/OJIN/TableofContents/Volume122007/No3Sept07/NursingQualityIndicators.html doi:10.3912/OJIN.Vol12No03Man02

Moody, P., Gonzales, I., & Cureton, V. Y. (2004). The effect of body position and mattress type on interface pressure in quadriplegic adults: A pilot study. *Dermatology Nursing, 16*, 507-512.

Moore, Z. (2010). Bridging the theory-practice gap in pressure ulcer prevention. *British Journal of Nursing, 19*, S15-18.

Moore, Z. E. H., & Cowman, S. (2010). Risk assessment tools for the prevention of pressure ulcers. *Cochrane Database of Systematic Reviews,* (3). Retrieved from http://www.mrw.interscience.wiley.com/cochrane/clsysrev/articles/CD006471/frame.html doi:10.1002/14651858.CD006471.pub2

Moore, Z., Cowman, S., & Conroy, R. M. (2011). A randomised controlled clinical trial of repositioning, using the 30 degrees tilt, for the prevention of pressure ulcers. *Journal of Clinical Nursing, 20*, 2633-2644. doi: 10.1111/j.1365-2702.2011.03736.x

Mulholland, J. H., Tui, C., Wright, A. M., Vinci, V., & Shafiroff, B. (1943). Protein metabolism and bed sores. *Annals of Surgery, 118*, 1015-1023.

Munthali, E., & Morsell, A. (2013, October 8). *Letter to the Consensus Standards Approval Committee regarding measure retirement request*: National Quality Forum.

Nakagami, G., Sanada, H., Konya, C., Kitagawa, A., Tadaka, E., & Matsuyama, Y. (2007). Evaluation of a new pressure ulcer preventive dressing containing ceramide 2 with low frictional outer layer. *Journal of Advanced Nursing, 59*, 520-529. doi: 10.1111/j.1365-2648.2007.04334.x

The National Database of Nursing Quality Indicators® (March 2014a). *Glossary & Reference Guide to Clinical Indicators®*

National Database of Nursing Quality Indicators. (2014b). *Hospital sites - January 2014*. Retrieved from http://www.nursingworld.org/MainMenuCategories/ThePracticeofProfessionalNursing/PatientSafetyQuality/NDNQI-Participating-Hospitals.pdf

National Database of Nursing Quality Indicators Staff. (2013, May). *NDNQI National Database of Nursing Quality Indicators guidelines for data collection and submission on quarterly indicators Version 10.0*. American Nurses Association.

National Pressure Ulcer Advisory Panel. (2012). NPUAP pressure ulcer Stages/Categories. Retrieved from http://www.npuap.org/resources/educational-and-clinical-resources/npuap-pressure-ulcer-stagescategories/

National Pressure Ulcer Advisory Panel. (2013, June 10). *Research priorities identified for pressure ulcer prevention, treatment & policy*. Retrieved from

http://www.npuap.org/research-priorities-identified-for-pressure-ulcer-prevention-treatment-policy/

National Pressure Ulcer Advisory Panel and European Pressure Ulcer Advisory Panel. (2009). *Prevention and treatment of pressure ulcers: Clinical practice guideline*. Washington DC: National Pressure Ulcer Advisory Panel.

National Pressure Ulcer Advisory Panel, European Pressure Ulcer Advisory Panel and Pan Pacific Pressure Injury Alliance. (2014). *Prevention and Treatment of Pressure Ulcers: Quick Reference Guide.* Emily Haesler (Ed.). Cambridge Media: Osborne Park, Western Australia. Retrieved from http://www.npuap.org/wp-content/uploads/2014/08/Updated-10-16-14-Quick-Reference-Guide-DIGITAL-NPUAP-EPUAP-PPPIA-16Oct2014.pdf

National Quality Forum. (2003). *Safe practices for better healthcare: A consensus report*. Retrieved from http://www.ahrq.gov/professionals/quality-patient-safety/patient-safety-resources/resources/nqfpract.pdf

National Quality Forum. (2013a). *About NQF*. Retrieved from http://www.qualityforum.org/About_NQF/About_NQF.aspx

National Quality Forum. (2013b). *Find measures; NQF-endorsed; pressure ulcer*. Retrieved from http://www.qualityforum.org/QPS/QPSTool.aspx#qpsPageState=%7B%22TabType%22%3A1,%22TabContentType%22%3A1,%22SearchCriteriaForStandard%22%3A%7B%22TaxonomyIDs%22%3A%5B%5D,%22SelectedTypeAheadFilterOption%22%3A%7B%22ID%22%3A14368,%22FilterOptionLabel%22%3A%22pressure+ulcer%22,%22TypeOfTypeAheadFilterOption%22%3A1,%22TaxonomyId%22%3A0%7D,%22Keyword%22%3A%22pressure+ulcer%22,%22PageSize%22%3A%2225%22,%22OrderType%22

%3A3,%22OrderBy%22%3A%22ASC%22,%22PageNo%22%3A1,%22IsExactMatch%22%3Atrue,%22QueryStringType%22%3A%22%22,%22ProjectActivityId%22%3A%220%22,%22FederalProgramYear%22%3A%220%22,%22FederalFiscalYear%22%3A%220%22,%22FilterTypes%22%3A2%7D,%22SearchCriteriaForForPortfolio%22%3A%7B%22Tags%22%3A%5B%5D,%22FilterTypes%22%3A0,%22PageStartIndex%22%3A1,%22PageEndIndex%22%3A25,%22PageNumber%22%3Anull,%22PageSize%22%3A%2225%22,%22SortBy%22%3A%22Title%22,%22SortOrder%22%3A%22ASC%22,%22SearchTerm%22%3A%22%22%7D,%22ItemsToCompare%22%3A%5B%5D,%22SelectedStandardIdList%22%3A%5B%5D%7D

National Quality Forum. (2013c). *Frequently asked questions*. Retrieved from http://www.qualityforum.org/Field_Guide/NQF_Help.aspx

National Quality Forum. (2013d). *Measure evaluation criteria: Effective July 2013*. Retrieved from http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx

National Quality Forum. (2013e, October 8). *Review and update of guidance for evaluating evidence and measure testing - technical report.* Retrieved from http://www.qualityforum.org/Publications/2013/10/Review_and_Update_of_Guidance_for_Evaluating_Evidence_and_Measure_Testing_-_Technical_Report.aspx

National Research Council. (2000). *To Err Is Human: Building a Safer Health System*. Washington, DC: The National Academies Press.

Nichols, D. P. (1999). My coefficient α is negative! Retrieved from http://www.ats.ucla.edu/stat/spss/library/negalpha.htm

Niederhauser, A., Lukas, C. V., Parker, V., Ayello, E. A., Zulkowski, K., & Berlowitz, D.

(2012). Comprehensive programs for preventing pressure ulcers: A review of the literature. *Advances in Skin & Wound Care, 25*, 167-188.

Nightingale, F. (1860). *Notes on nursing: What it is, and what it is not (1st American ed.)* Retrieved from http://digital.library.upenn.edu/women/nightingale/nursing/nursing.html

Nixon, J. (2006). Randomised, controlled trial of alternating pressure mattresses compared with alternating pressure overlays for the prevention of pressure ulcers: PRESSURE (pressure relieving support surfaces) trial. *British Medical Journal (Clinical research ed.), 332*, 1-5. doi: 10.1136/bmj.38849.478299.7C

Nixon, J., Cranny, G., & Bond, S. (2007). Skin alterations of intact skin and risk factors associated with pressure ulcer development in surgical patients: A cohort study. *International Journal of Nursing Studies, 44*, 655-663. doi: http://dx.doi.org/10.1016/j.ijnurstu.2006.02.010

O'Reilly, K. B. (2008, January 7). No pay for 'never event' errors becoming standard. *American Medical News, 7*. Retrieved from http://www.amednews.com/article/20080107/profession/301079966/7/

Padula, W. V., Mishra, M. K., Makic, M. B., & Sullivan, P. W. (2011). Improving the quality of pressure ulcer care with prevention: A cost-effectiveness analysis. *Medical Care, 49*, 385-392. doi: 10.1097/MLR.0b013e31820292b3

Peirce, S. M., Skalak, T. C., & Rodeheaver, G. T. (2000). Ischemia-reperfusion injury in chronic pressure ulcer formation: A skin model in the rat. *Wound Repair & Regeneration, 8*, 68-76. doi: 10.1046/j.1524-475x.2000.00068.x

Pham, B., Teague, L., Mahoney, J., Goodman, L., Paulden, M., Poss, J., . . . Krahn, M. (2011). Early prevention of pressure ulcers among elderly patients admitted through emergency

departments: A cost-effectiveness analysis. *Annals of Emergency Medicine, 58*, 468-478.e463. doi: 10.1016/j.annemergmed.2011.04.033

Pieper, B., & Mattern, J. C. (1997). Critical care nurses' knowledge of pressure ulcer prevention, staging and description. *Ostomy Wound Management, 43*(2), 22-31.

Polit, D.E., & Beck, C.T. (2012). *Nursing research: Generating and assessing evidence for nursing practice* (9th ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice* (3rd ed.). Upper Saddle River, JNJ: Prentice-Hall.

Qualidigm. (2013a). Medicare *Patient Safety Monitoring System (MPSMS) overview*. Retrieved from http://www.qualidigm.org/wp-content/uploads-lrg/MPSMS%20Overview%20Final%20Report%206.2013.pdf

Qualidigm. (2013b). *MPSMS*. Retrieved from http://www.qualidigm.org/index.php/current-initiatives/mpsms/

QualityNet. (n. d.). *Overview hospital-acquired conditions (HACs)*. Retrieved from https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1228759479767

Rastinehad, D. (2006). Pressure ulcer pain. *Journal of Wound Ostomy Continence Nursing, 33*, 252-257.

Ratliff, C., Tomaselli, N., & The Guideline Task Force. (2010). *Guideline for prevention and management of pressure ulcers: 2 WOCN clinical practice guideline series*. Mount Laurel, NJ: Wound, Ostomy and Continence Nurses Society.

Reddy, M., Gill, S. S., & Rochon, P. A. (2006). Preventing pressure ulcers: A systematic review. *Journal of the American Medical Association, 296*, 974-984

Reed, K., & May, R. (2011). *The eighth annual HealthGrades patient safety in American hospitals*. Retrieved July 19, 2013, from http://www.healthgrades.com/business/img/HealthGradesPatientSafetyInAmericanHospitalsStudy2011.pdf

Rich, S. E., Margolis, D., Shardell, M., Hawkes, W. G., Miller, R. R., Amr, S., & Baumgarten, M. (2011a). Frequent manual repositioning and incidence of pressure ulcers among bed-bound elderly hip fracture patients. *Wound Repair & Regeneration, 19*(1), 10-18. doi: 10.1111/j.1524-475X.2010.00644.x

Rich, S. E., Shardell, M., Hawkes, W. G., Margolis, D. J., Amr, S., Miller, R., & Baumgarten, M. (2011b). Pressure-redistributing support surface use and pressure ulcer incidence in elderly hip fracture patients. *Journal of the American Geriatrics Society, 59*, 1052-1059. doi: 10.1111/j.1532-5415.2011.03446.x

Rintala, D. H., Garber, S. L., Friedman, J. D., & Holmes, S. A. (2008). Preventing recurrent pressure ulcers in veterans with spinal cord injury: Impact of a structured education and follow-up intervention. *Archives of Physical Medicine & Rehabilitation, 89*, 1429-1441. doi: 10.1016/j.apmr.2008.01.015

Roaf, R. (2006). The causation and prevention of bed sores. *Journal of Tissue Viability, 16*(2), 6-8.

Rosenthal, R. & Rosnow, R. L. (2008). *Essentials of behavioral research* (3rd ed., pp. 87-122). Boston: McGraw Hill.

Russo, A., Steiner, C., & Spector, W. (2008). *Statistical Brief 64: Hospitalizations related to pressure ulcers among adults 18 years and older.* Retrieved from http://www.hcup-us.ahrq.gov/reports/statbriefs/sb64.jsp

Saleh, M., Anthony, D., & Parboteeah, S. (2009). The impact of pressure ulcer risk assessment on patient outcomes among hospitalised patients. *Journal of Clinical Nursing, 18*, 1923-1929. doi: 10.1111/j.1365-2702.2008.02717.x

Salkind, N. J. (2006). *Tests & measurement for people who (think they) hate tests & measurement.* Thousand Oaks, CA: Sage Publications.

Scanlon, M. C., Harris II, M., Levy, F., & Sedman, A. (2008). Evaluation of the Agency for Healthcare Research and Quality Pediatric Quality Indicators. *Pediatrics, 121*, e1723-e1731. doi: 10.1542/peds.2007-3247

Schessel, E. S., Ger, R., & Oddsen, R. (2012). The costs and outcomes of treating a deep pressure ulcer in a patient with quadriplegia. *Ostomy Wound Management, 58*(2), 41-46.

Schindler, C. A., Mikhailov, T. A., Kuhn, E. M., Christopher, J., Conway, P., Ridling, D., . . . Simpson, V. S. (2011). Protecting fragile skin: Nursing interventions to decrease development of pressure ulcers in pediatric intensive care. *American Journal of Critical Care, 20*, 26-34; quiz 35. doi: 10.4037/ajcc2011754

Schone, E., Hubbard, M., & Jones, D. (2011, November 18). *Reporting period and reliability of AHRQ, CMS 30-day and HAC quality measures - revised*. Retrieved from http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based purchasing/Downloads/HVBP_Measure_Reliability-.pdf

Schubart, J. R., & Hilgart, M. (n. d.). The Pressure Ulcer Prevention and Management E-Learning Program. Retrieved from http://www.healthlearnpa.com/pva/PU_basics/player.html

Serpa, L. F., Santos, V. L. G., Peres, G. R. P., Cavicchioli, M. G. S., & Hermida, M. M. (2011). Validity of the Braden and Waterlow subscales in predicting pressure ulcer risk in hospitalized patients. *Applied Nursing Research, 24*(4), e23-28. doi: 10.1016/j.apnr.2010.05.002

Shekelle, P. G., Ortiz, E., Rhodes, S., Morton, S. C., Eccles, M. P., Grimshaw, J. M., & Woolf, S. H. (2001). Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: How quickly do guidelines become outdated? *Journal of the American Medical Association, 286*, 1461-1467.

Shekelle, P. G., Pronovost, P. J., Wachter, R. M., McDonald, K. M., Schoelles, K., Dy, S. M., . . . Walshe, K. (2013). The top patient safety strategies that can be encouraged for adoption now. *Annals of Internal Medicine, 158*, 365-368. doi: 10.7326/0003-4819-158-5-201303051-00001

Shrout, P. E. (1998). Measurement reliability and agreement in psychiatry. *Statistical Methods in Medical Research, 7*, 301-307.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428. doi: 10.1037/0033-2909.86.2.420

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy, 85*, 257-268.

Simon, M., Klaus, S., Gajewski, B. J., & Dunton, N. (2013). Agreement of fall classifications among staff in U.S. hospitals. *Nursing Research, 62*, 74-81. doi: 10.1097/NNR.0b013e31827bf8c9

Simon, M., Yankovskyy, M. S., & Dunton, N. (2010). Regulatory readiness: Solving the mystery of patient days and the midnight census. *Nursing Management, 41*(2), 12-14.

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

Soban, L. M., Hempel, S., Munjas, B. A., Miles, J., & Rubenstein, L. V. (2011). Preventing pressure ulcers in hospitals: A systematic review of nurse-focused quality improvement interventions. *Joint Commission Journal on Quality and Patient Safety, 37*, 245-252

Spetz, J., Brown, D., Aydin, C., & Donaldson, N. (2013). The value of reducing hospital-acquired pressure ulcer prevalence. *The Journal of Nursing Administration, 43*, 235-241.

Stechmiller, J. K., Cowan, L., Whitney, J. D., Phillips, L., Aslam, F., Barbul, A., . . . Stotts, N. (2008). Guidelines for the prevention of pressure ulcers. *Wound Repair and Regeneration, 16*(2), 151-168. doi: 10.1111/j.1524-475X.2008.00356.x

Stotts, N. A., Brown, D. S., Donaldson, N. E., Aydin, C., & Fridman, M. (2013). Eliminating hospital-acquired pressure ulcers: Within our reach. *Advances in Skin & Wound Care, 26*, 13-18

Sullivan, N., & Schoelles, K. M. (2013). Preventing in-facility pressure ulcers as a patient safety strategy: A systematic review. *Annals of Internal Medicine, 158*, 410-416. doi: 10.7326/0003-4819-158-5-201303051-00008

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics (5th ed.)*. Boston: Pearson Education, Inc.

Tannen, A., Balzer, K., Kottner, J., Dassen, T., Halfens, R., & Mertens, E. (2010). Diagnostic accuracy of two pressure ulcer risk scales and a generic nursing assessment tool. A psychometric comparison. *Journal of Clinical Nursing, 19*, 1510-1518. doi: 10.1111/j.1365-2702.2009.03005.x

Taunton, R. L., Bott, M. J., Koehn, M. L., Miller, P., Rindner, E., Pace, K., . . . Dunton, N. (2004). The NDNQI-Adapted Index of Work Satisfaction. *Journal of Nursing Measurement, 12*(2), 101-122.

Thompson, W. D., Ravdin, I. S., & Frank, I. L. (1938). Effect of hypoproteinemia on wound disruption. *Archives of Surgery (1938), 36*, 509-518.

Tompkins, C. P., Higgins, A. R., & Ritter, G. A. (2009). Measuring outcomes and efficiency in Medicare value-based purchasing. *Health Affairs, 28*, w251-w261. Retrieved from http://content.healthaffairs.org/content/28/2/w251.full.html

Uebersax, J. (2009). *Raw agreement indices*. Retrieved from http://john-uebersax.com/stat/raw.htm

Uebersax, J. (2010a). *Kappa coefficients: A critical appraisal*. Retrieved from http://john-uebersax.com/stat/kappa.htm

Uebersax, J. (2010b). *Statistical methods for rater and diagnostic agreement*. Retrieved from http://www.john-uebersax.com/stat/agree.htm

United States Census Bureau. (2013). Metropolitan and micropolitan. Retrieved from http://www.census.gov/population/metro/

Vanderwee, K., Grypdonck, M. H. F., De Bacquer, D., & Defloor, T. (2007). Effectiveness of turning with unequal time intervals on the incidence of pressure ulcer lesions. *Journal of Advanced Nursing, 57*, 59-68. doi: 10.1111/j.1365-2648.2006.04060.x

VanGilder, C., Amlung, S., Harrison, P., & Meyer, S. (2009). Results of the 2008-2009 International Pressure Ulcer Prevalence Survey and a 3-year, acute care, unit-specific analysis. *Ostomy Wound Management, 55*(11), 39-45.

VanGilder, C., Lachenbruch, C., Harrison, P., & Meyer, S. (2013). Prevalence of suspected deep tissue injuries: Analysis of the 2012 International Pressure Ulcer Prevalence Survey. Retrieved from http://www.npuap.org/wp-content/uploads/2012/01/VanGilder-sDTI-talk-NPUAP-2013-VanGilder-31.pdf

Waltz, C. F., Strickland, O. L., & Lenz, E. R. (2010). *Measurement in Nursing and Health Research.* New York: Springer Publishing Company.

Wambach, K. (2012a, March). *Assessment of reliability*. PowerPoint presentation for NRSG 946, Kansas University Medical Center, Kansas City, KS.

Wambach, K. (2012b, March). *Measurement error and classical measurement theory*. PowerPoint presentation for NRSG 946, Kansas University Medical Center, Kansas City, KS.

Weaver, B. (2007, May 17). Re: SPSS doesn't calculate kappa when one variable is constant [Web log comment]. Retrieved from http://www.mofeel.net/1170-comp-soft-sys-stat-spss/2553.aspx

Weng, M.-H. (2008). The effect of protective treatment in reducing pressure ulcers for non-invasive ventilation patients. *Intensive & Critical Care Nursing, 24*, 295-299. doi: 10.1016/j.iccn.2007.11.005

Woolf, S. H., Grol, R., Hutchinson, A., Eccles, M., & Grimshaw, J. (1999). Clinical guidelines: Potential benefits, limitations, and harms of clinical guidelines. *British Medical Journal (Clinical research ed.), 318*, 527-530.

Young, J., Ernsting, M., Kehoe, A., & Holmes, K. (2010). Results of a clinician-led evidence-based task force initiative relating to pressure ulcer risk assessment and prevention.

Journal of Wound Ostomy Continence Nursing, 37, 495-503. doi: 10.1097/WON.0b013e3181edadcf

Young, T. (2004). The 30 degree tilt position vs the 90 degree lateral and supine positions in reducing the incidence of non-blanching erythema in a hospital inpatient population: A randomised controlled trial. *Journal of Tissue Viability, 14*(3), 88, 90, 92-86.

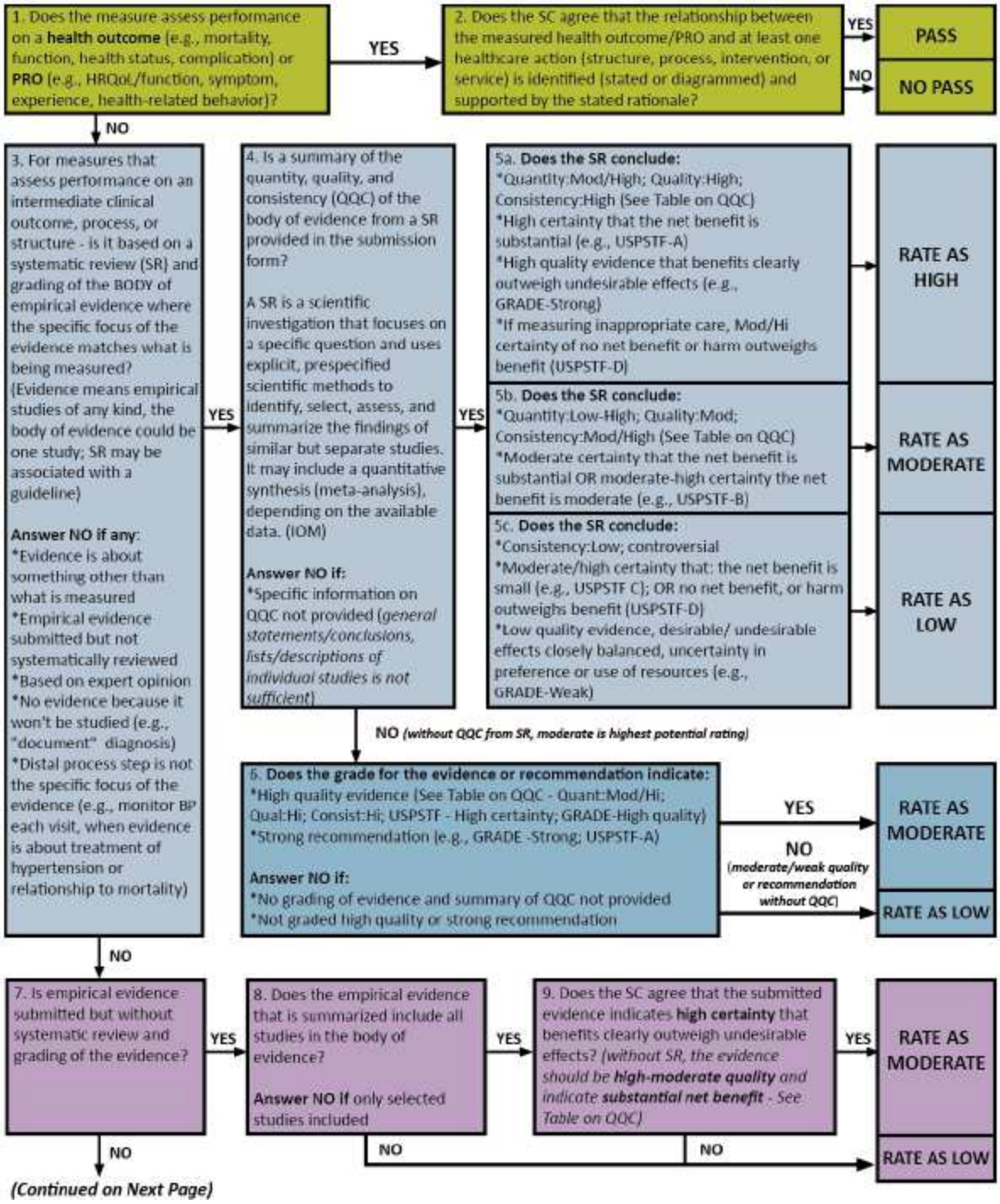
Zaiontz, C. (2013). *Real statistics using Excel: Fleiss' kappa*. Retrieved from http://www.real-statistics.com/reliability/fleiss-kappa/

Zhan, C., & Miller, M. R. (2003). Administrative data based patient safety research: A critical review. *British Medical Journal Quality & Safety*, ii58-ii63. Retrieved from http://qualitysafety.bmj.com/content/12/suppl_2/ii58.full.pdf+html

Zulkowski, K., Ayello, E. A., & Wexler, S. (2007). Certification and education: Do they affect pressure ulcer knowledge in nursing? *Advances in Skin & Wound Care, 20*, 34-38.

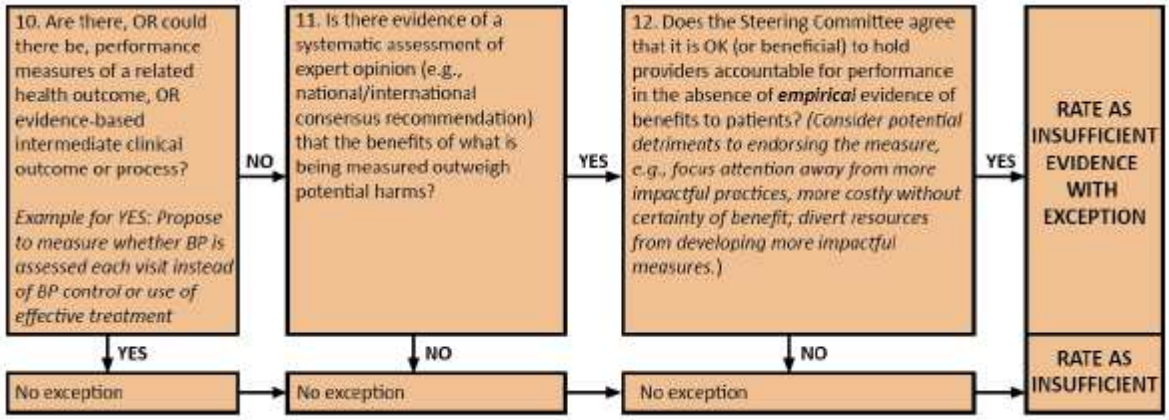
Appendix A

National Quality Forum Algorithm 1: Guidance for Evaluating the Clinical Evidence (NQF, 2013e, p. 8)



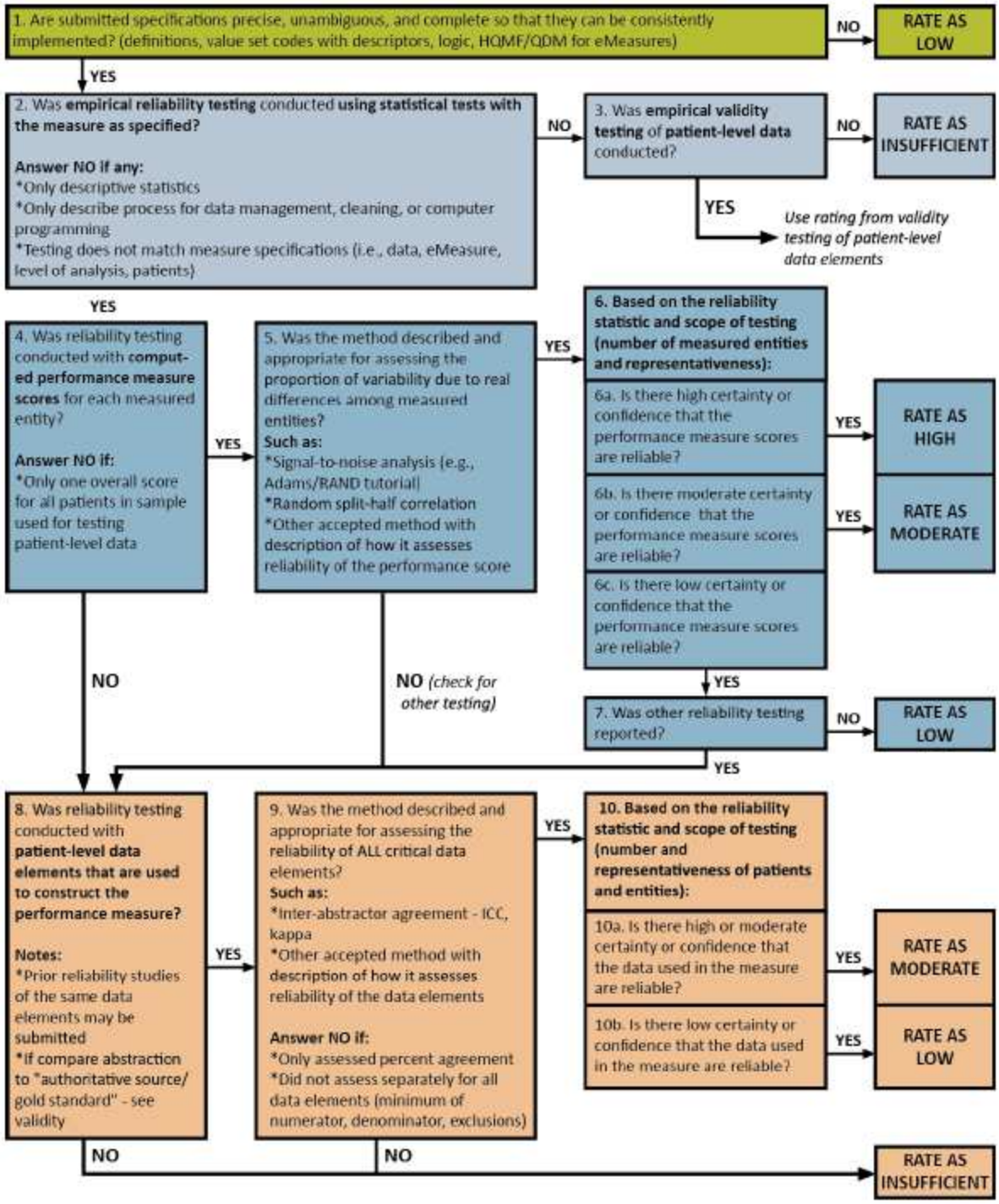
Appendix A (continued)

National Quality Forum Algorithm 1: Guidance for Evaluating the Clinical Evidence (NQF, 2013e, p. 9)



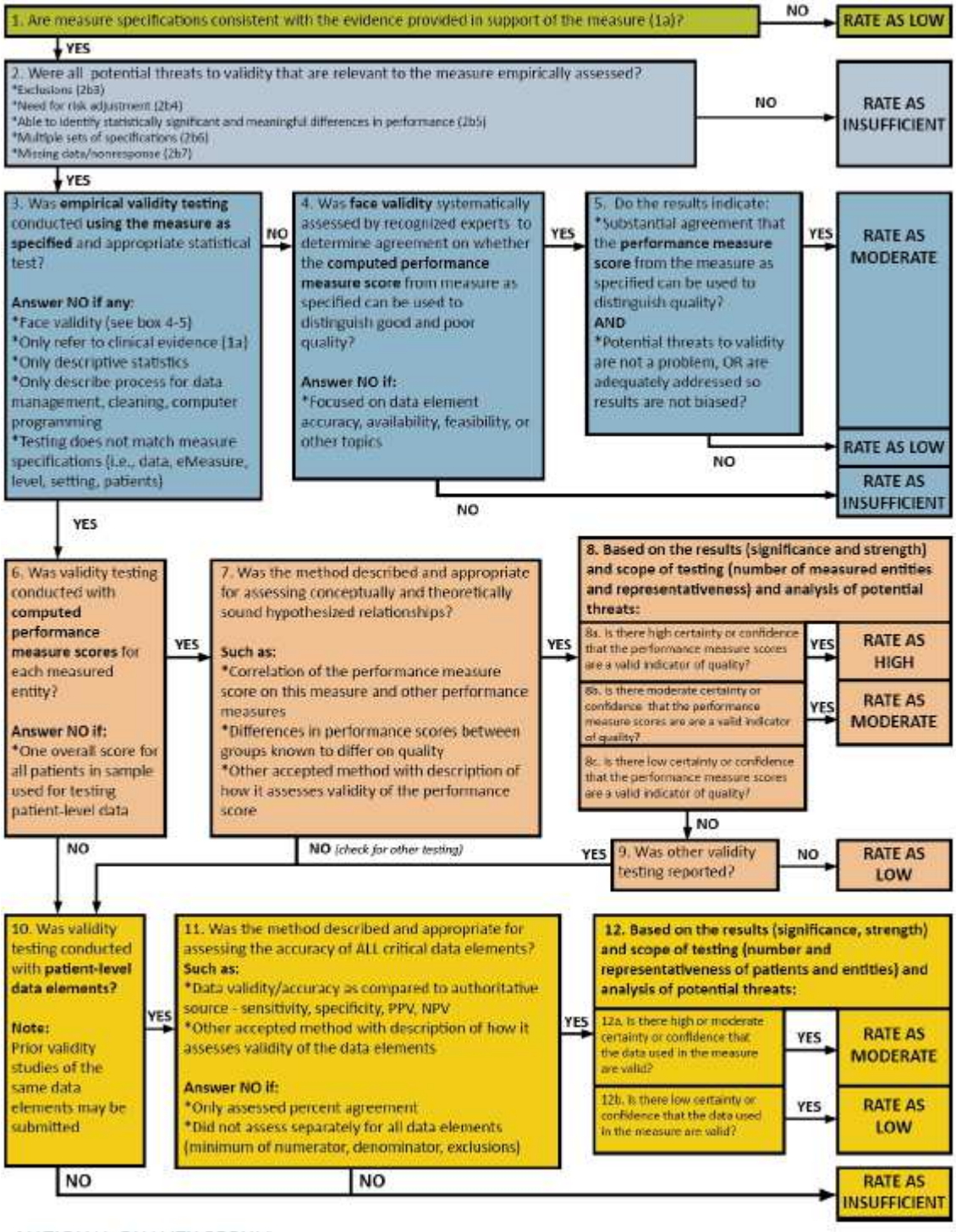
Appendix A (continued)

National Quality Forum Algorithm 2: Guidance for Evaluating Reliability (NQF, 2013e, p. 15)



Appendix A (continued)

National Quality Forum Algorithm 3: Guidance for Evaluating Validity (NQF, 2013e, p. 16)



Appendix B

Data Collection Form

The NDNQI® *Data Collection Form*© is copyrighted material, and could not be included in the published dissertation manuscript.

Appendix C

Pressure Ulcer Risk and Prevention Reliability Survey

Survey ID _____(This is the 7-digit alphanumeric code that you have been assigned. It is located next to your rater number on the print-copy of your *Data Collection Form*.)

1. What is your job title? (Select one.)

Staff Nurse

Clinical Nurse Specialist

Advanced Practice Nurse

Nurse Manager

Nursing Administrator

Quality Improvement

Wound/Skin Care Nurse

NDNQI Site Coordinator

Other (please describe):_____

2. Registered Nurse (RN) Nursing Education? (Select one.) [Appropriate list will open following Yes/No selection.]

Yes (Select one)	**No (Select one)**
Associate's Degree Nursing	High School Graduate/GED
Bachelor's Degree Nursing	Associate's Degree non-nursing
Master's Degree Nursing	Bachelor's Degree non-nursing
Doctorate Degree Nursing	Master's Degree non-nursing
	Doctorate Degree non-Nursing

Appendix C (continued)

Pressure Ulcer Risk and Prevention Reliability Survey

3. Please select what you have completed within the last 12 months: (Select all that apply.)

I have reviewed the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers.

I have completed all 4 modules of the NDNQI Pressure Ulcer Training program.

I have completed only some (not all) of the 4 NDNQI Pressure Ulcer Training program modules.

None of the above

4. Have you received education for data collection on pressure ulcers *other than* reviewing the NDNQI Guidelines for Data Collection and Submission of Pressure Ulcers *or* completing the NDNQI Pressure Ulcer Training program? (Select one.)

Yes No

5. How many years have you collected NDNQI pressure ulcer data? (Select one.)

< 1 year	5 years	11 years
1 year	6 years	12 years
2 years	7 years	13 years
3 years	8 years	14 years
4 years	9 years	15 years
	10 years	> 15 years

Appendix C (continued)

Pressure Ulcer Risk and Prevention Reliability Survey

6. In addition to chart abstractor, what other role(s) in the NDNQI Pressure Ulcer Survey have you had? (Select all that apply.)

Site Coordinator

Patient skin inspection—Rounding on all patients

Patient skin inspection—Rounding on selected patients to confirm pressure ulcer presence or stage

Training of pressure ulcer team

Data entry

No other roles

Other (please describe) _____

7. Who usually leads your NDNQI Pressure Ulcer Survey data collection team? (Select one.)

Someone certified in wound care

Someone who is not certified in wound care but is the wound/skin care nurse

Neither of the above

I don't know their certification status

Appendix C (continued)

Pressure Ulcer Risk and Prevention Reliability Survey

8. Which of the following active certifications do you hold? (Select all that apply.)

CWOCN—Certified Wound, Ostomy, Continence Nurse

CWCN—Certified Wound Care Nurse

COCN—Certified Ostomy Care Nurse

CCCN—Certified Continence Care Nurse

CWON—Certified Wound Ostomy Nurse

CWS—Certified Wound Specialist

WCC—Wound Care Certified

No certifications in wound care

Other (please describe):_____

Appendix C (continued)

Pressure Ulcer Risk and Prevention Reliability Survey

With regard to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements: (Select one)

	Never	Sometimes	Usually	Almost Always	Always	I Don't Know
1. I am comfortable with my skills to review patient records and collect NDNQI pressure ulcer risk and prevention data.						
2. Prior to each time that we collect NDNQI pressure ulcer data, our hospital provides training or reviews.						
3-9. The people at our hospital who review patient records to collect NDNQI pressure ulcer risk and prevention data include:						
• RNs						
• LPNs						
• Nurse aids						
• Physical Therapists						
• Respiratory Therapists						
• Clinical Secretaries						
• Others (please specify)						
10. At least once a year, we compare pressure ulcer team assessments of pressure ulcer risk and prevention to evaluate the reliability of this data.						
11. I collect data on the unit where I usually work.						
12. I retrieve documentation on pressure ulcer risk and prevention from the patient's paper health record.						
13. I retrieve documentation on pressure ulcer risk and prevention from the patient's electronic health record.						

Appendix C (continued)

Pressure Ulcer Risk and Prevention Reliability Survey

With regard to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements: (Select one)

	Never	Sometimes	Usually	Almost Always	Always	I Don't Know
14. Persons from information technology extract data on pressure ulcer risk and prevention from the electronic health record for use in the pressure ulcer survey.						
15. If a patient was admitted to the hospital less than 24 hours before the pressure ulcer survey, I exclude them from the survey.						
16. I decide if a patient is at risk for pressure ulcers by the score obtained on the risk assessment scale (i.e. Braden, Norton) at the time of the pressure ulcer survey.						
17. If a patient's risk assessment score (i.e. Braden or Norton Scale score) does not classify them as being "at risk for pressure ulcers", I look for documentation of other factors indicating the patient is at pressure ulcer risk.						
18-20. I directly observe patients in their rooms in order to rate: • *pressure-redistribution surface use* • *nutritional support* • *moisture management*						

Appendix C (continued)

Pressure Ulcer Risk and Prevention Reliability Survey

Questions 18-27 Apply to Patients "At Risk" for Pressure Ulcers. With regard to the NDNQI pressure ulcer prevention measures, please rate the following statements: (Select one)

	Never	Sometimes	Usually	Almost Always	Always	I Don't Know
21. If the nurse tells me a patient refused to be repositioned, but there is no documentation of this in the patient record, I select "Patient refused".						
22. I ask patient care staff for information in order to determine how to rate Routine repositioning as prescribed.						
23. If repositioning is appropriate for a patient; I allow a 30 minute leeway past the time repositioning was to occur before I rate Routine repositioning as prescribed as "No".						
24. To evaluate frequency of patient repositioning, I look in the patient record for physical therapy documentation of patient activity such as standing or ambulation.						
25. If a nutritional consult is ordered for a patient who has been hospitalized for 2 or more days with poor intake, but the consult has not yet been completed, I rate *Nutritional support* as "Yes".						

Appendix C (continued)

Pressure Ulcer Risk and Prevention Reliability Survey

With regard to the NDNQI pressure ulcer risk and prevention measures, please rate the following statements: (Select one)

	Never	Sometimes	Usually	Almost Always	Always	I Don't Know
26. Barrier skin cream at the bedside is sufficient for me to rate *Moisture management* as "Yes".						
27. If I rate that a patient did not receive a particular pressure ulcer prevention intervention, someone else reviews the patient record to make sure I selected the correct rating.						

Appendix D

Invitation to Participate

Dear NDNQI Site Coordinator,

We invite you to participate in the *Pressure Ulcer Risk and Prevention Reliability Study*. This study will examine the agreement of ratings on pressure ulcer (PrU) risk and prevention measures between persons who collect this data for the NDNQI Pressure Ulcer Survey.

No skin inspections (or data collection on pressure ulcer presence or staging) are involved. Study data collection is separate from and in addition to your quarterly NDNQI pressure ulcer data collection.

Study Protocol:

You will need to select 3 to 4 persons (participant raters) who usually collect data on pressure ulcer risk and prevention for the NDNQI Pressure Ulcer Survey. These participant raters will independently review 50 patient records, and independently rate these 50 patients on the study measures.

Pressure Ulcer Risk Measures
(All 50 patients)

1. Skin assessment on admission.
2. Pressure ulcer risk assessment on admission.
3. Time since last pressure ulcer risk assessment.
4. Last risk assessment scale and score.
5. Pressure ulcer risk status.

Pressure Ulcer Prevention Measures
(Those of the 50 patients who are at risk)

6. Pressure ulcer prevention.
7. Skin assessment documented.
8. Pressure-redistribution surface use.
9. Routine repositioning as prescribed.
10. Nutritional support.
11. Moisture management.

Participant raters will also complete an online survey about the methods they use to gather information on the risk and prevention measures.

Study Dates:

Participant raters will collect the data one morning during September 29 to October 13, 2014.

Eligible Units:

Only patients on adult medical-surgical, adult medical, and/or adult surgical units will be included in data collection.

Please see the attachment to this email for study details and the list of units from which study data may be collected in your hospital.

Please go to << INSERT REDCap URL >> to reply to this invitation. The deadline to reply is August 19, 2014.

Thank you for your consideration of participating in this study

Appendix D (continued)

Invitation to Participate (Email Attachment)

August 1, 2014

Dear NDNQI Site Coordinator,

We invite you to participate in the *Pressure Ulcer Risk and Prevention Reliability Study*. This study will examine the agreement of ratings on the NDNQI pressure ulcer risk and prevention measures between persons who collect this data for the NDNQI Pressure Ulcer Survey. The study is important because the data on pressure ulcer risk and prevention are used by hospitals for quality improvement purposes. The pressure ulcer risk and prevention study measures are:

Pressure Ulcer Risk:	Pressure Ulcer Prevention w/in the last 24 hours:
1. Skin assessment on admission.	6. Pressure ulcer prevention.
2. Pressure ulcer risk assessment on admission.	7. Skin assessment documented.
3. Time since last pressure ulcer risk assessment.	8. Pressure-redistribution surface use.
4. Last risk assessment scale and score.	9. Routine repositioning as prescribed.
5. Pressure ulcer risk status.	10. Nutritional support.
	11. Moisture management.

No skin inspections (or data collection on pressure ulcer presence or staging) are involved.

The study will be performed by Shirley Waugh for her PhD dissertation. Sandra Bergquist-Beringer, RN, PhD, CWCN, who helped develop these pressure ulcer measures, will supervise the study.

If you decide to participate in this study:

1. The site coordinator will direct the study and data collection. Data collection must be completed on a morning during the study period of September 29th through October 13th.

2. The site coordinator will identify 3 to 4 people at the hospital *who usually review the patient record during data collection on the pressure ulcer risk and prevention measures for the NDNQI Pressure Ulcer Survey* (participant raters).
    a. One of these people will be the most experienced and/or skilled in patient record review on your NDNQI pressure ulcer data collection team.
    b. The others will be persons who usually review patient records on pressure ulcer risk and prevention.
    c. The site coordinator may serve as a participant rater if they usually review patient records for the survey.

3. Participant raters will *independently review the records* of 50 patients (total = 50 patients) and *independently rate* these patients on the NDNQI pressure ulcer risk and prevention measures.
    a. Only patients on medical-surgical, medical, and surgical units will be included in this data collection.

Appendix D (continued)

Invitation to Participate (Email Attachment)

b. Units in your hospital from which study data may be collected are listed on the *Unit List*, which is attached as a separate document. These units are identified by name with corresponding unit ID and unit type.
c. Data collection for this study is separate from and in addition to your quarterly NDNQI pressure ulcer data collection.
d. Each participant rater will enter their data into an Excel file that will be sent to you. The Excel file is named the *Data Collection Form*.
e. It will take approximately 2.5 hours to collect the data.

4. Participant raters will complete an online survey after data collection is completed.
   a. This survey will take 20 – 25 minutes to complete:

5. Participant raters will submit (upload) their completed *Data Collection Form* Excel file to the online survey.

All information will remain confidential. Your hospital's identity will only be known by your NDNQI alphanumeric code. Participant rater identity will remain unknown to the research team. Only de-identified results will be included in the final report. Results will be reported to participating hospitals after the study is completed.

Participation in the study is important to us but it is voluntary. If you choose not to participate, your hospital's membership in the NDNQI will not be affected. You may contact Shirley Waugh with questions about this study at ndnqi@kumc.edu or 913-588-1691.

This study has been reviewed and received approval from the Human Subjects Committee at the University of Kansas Medical Center. If you have questions about the study, you can call NDNQI at 913-588-1691. If you have any questions about your rights as a research participant, call 913-588-1240 or write Human Subjects Committee, University of Kansas Medical Center, 3901 Rainbow Blvd, Mail Stop #1032, Kansas City, KS 66160-7700.

Please go to << INSERT REDCap URL>> to reply to this invitation. The deadline to reply is August 19th.

Thank you for this consideration.

Sincerely,

Shirley Waugh, PhD Candidate Researcher
University of Kansas, School of Nursing

Your *Unit List* is attached as a separate document.

Appendix E

Reply to Invitation to Participate

Note: This was available on REDCap.

Reply to Invitation to Participate:

The Pressure Ulcer Risk and Prevention Reliability Study:

1. NDNQI Hospital Code _____
(The alphanumeric code you use to sign-in to the NDNQI website.)

2. Would you like to participate in the Pressure Ulcer Risk and Prevention Reliability Study?

Yes No I'm not sure, I need more information

3. Site coordinator's first and last name:_____

4. Site coordinator's email address:_____

For more information, contact Shirley Waugh at 913-588-1691 or ndnqi@kumc.edu

Appendix F

Overview of the Pressure Ulcer Risk and Prevention Reliability Study

Overview of the Pressure Ulcer Risk and Prevention Reliability Study

The purpose of this study was to examine the agreement of ratings on pressure ulcer risk and prevention measures between persons who collect this data for the NDNQI Pressure Ulcer Survey. It is important that the reliability of the data on pressure ulcer risk and prevention be established as these data are used by hospitals for quality improvement purposes.

Thank you for accepting our invitation to participate and help us in this process.

The pressure ulcer risk and prevention study measures are:

Pressure Ulcer Risk:

1. Skin assessment on admission.
2. Pressure ulcer risk assessment on admission.
3. Time since last pressure ulcer risk assessment.
4. Last risk assessment scale and score.
5. Pressure ulcer risk status.

Pressure Ulcer Prevention w/in the last 24 hours:

6. Pressure ulcer prevention.
7. Skin assessment documented.
8. Pressure-redistribution surface use.
9. Routine repositioning as prescribed.
10. Nutritional support.
11. Moisture management.

No skin inspections (or data collection on pressure ulcer presence or staging) are involved.

The study will be conducted by Shirley Waugh for her PhD dissertation under the guidance of Sandra Bergquist-Beringer, RN, PhD, CWCN; who helped develop these NDNQI pressure ulcer measures.

PRIOR TO THE STUDY:

The site coordinator will direct the study and data collection at your hospital.

The site coordinator will identify 3 to 4 people at the hospital to participate in the study (participant raters). These people must be among those *who usually **review patient records** during data collection on the pressure ulcer risk and prevention measures for the NDNQI Pressure Ulcer Survey*

- One of these people will be the most experienced and/or skilled in patient record review on your NDNQI pressure ulcer data collection team.
- The others will be persons who usually review patient records during pressure ulcer data collection.
- The site coordinator may serve as a participant rater if they usually review patient records for NDNQI Pressure Ulcer Survey.

DATA COLLECTION:

The NDNQI Pressure Ulcer Risk and Prevention Measures:

Participant raters will *independently review the records* of 50 patients, and *independently rate* these patients on the NDNQI pressure ulcer risk and prevention measures. This will be like what they do for the NDNQI pressure ulcer survey except that this study will be separate from and in addition to your usual pressure ulcer survey. It is important that participant raters collect the data in their usual manner.

Data collection (for this study) is to begin on medical-surgical, then medical, and lastly surgical units until data have been collected on 50 patients. The goal is for participant raters to collect data on the same 50 patients. Therefore, the units in your hospital from where data can be collected, and the order that the participant raters will proceed through these units, have been provided to you in a separate document, the Unit List.

Participant raters will collect patient data on pressure ulcer risk and prevention using the *Data Collection Form* provided.

- This data collection form was created as an Excel file.
- Participant rater numbers and corresponding *Survey ID* are located on the form. The site coordinator will assign each participant rater a number (with corresponding *Survey ID*) following procedures identified in the *Site Coordinator Instructions.*

Participant raters will collect study data without discussing or sharing the information with the other participant raters.

Patients will be identified by room/bed number, but no link between patients and their room/bed number will be maintained. The identity of the participant raters and patients will be unknown to the research team.

The Methods Used to Collect NDNQI Pressure Ulcer Data:

After data collection is completed, participant raters will complete an online survey (on *REDCap*). The purpose of the online survey is to gather select information about the participant raters, and examine the methods and processes used to collect data on the NDNQI pressure ulcer risk and prevention measures. Also, each participant rater will upload their completed *Data Collection Form* Excel file onto the online survey site.

ADDITIONAL INFORMATION:

A teleconference for the site coordinators and participant raters has been scheduled for September 9th at 10:00AM (Central Daylight Time). Teleconference details are included in the *Teleconference Instructions*. The purpose of this teleconference is to explain the study and data collection procedures, review study materials, and answer any questions the site coordinator might have.

Thank you for taking the time to participate in the
Pressure Ulcer Risk and Prevention Reliability Study!

Appendix G

Site Coordinator Instructions

Site Coordinator Instructions

PRIOR TO THE STUDY:

1. Recruit 3 to 4 people at the hospital to participate in the study (participant raters). Recruit these people from those *who usually **review patient records** during data collection on the pressure ulcer risk and prevention measures for the NDNQI Pressure Ulcer Survey.*
    a. Recruit *the person most experienced and/or skilled in reviewing patient records.*

    b. Recruit 2 to 3 other team members.

    c. You may serve as a participant rater if you usually review patient records for the NDNQI Pressure Ulcer Survey.

2. Assign each participant rater a rater number.
    a. ***Assign the person most experienced and/or skilled in reviewing patient records to be "Rater 1".***

    b. Assign each of the other participant raters to be "Rater 2" or "Rater 3" if you have 2 other participant raters, and "Rater 4" if you have 3 other participant raters.

3. Establish the date and time of the study.
    a. Talk to the participant raters to coordinate the study date and time.

    b. The data collection period is September 29nd to October 13th.

    c. Select a single day during September 29th to October 13th to perform the study.

    d. The study will be performed between 0700 and 1200 the day of the study.

    e. Data collection (for this study) is separate from your quarterly NDNQI pressure ulcer data collection.

4. Note the units on which the data will be collected. Only adult medical-surgical, adult medical, and adult surgical units are eligible.
    a. The names of eligible units in your hospital and their NDNQI unit ID, and NDNQI unit type are listed here in your Unit List, which is attached as a separate document.

5. Distribute the *Overview of the Pressure Ulcer Risk and Prevention Reliability Study* and *Participant Rater Instructions* to each participant rater.

6. Distribute the *Data Collection Form* to Participant Raters.
    a. Print a *Data Collection Form* for each participant rater.

    b. On this form, mark the rater number that has been assigned to each participant rater.
        - The participant's corresponding *Survey ID* is located next to the rater number.

        - The *Survey ID* is the hospital alphanumeric code you use to sign-in to the NDNQI, followed by the number "1" if the rater is the most experienced and/or skilled in chart review; or "2", "3", or "4".

    c. The *Data Collection Form* will be used by participant raters to collect the data.

    d. Distribute the appropriate print form and electronic form to each participant rater. Be sure the most experienced/skilled chart reviewer is assigned to be Rater 1.

    e. You and the participant raters may choose to record data as it is collected (a) onto the print-copy of the *Data Collection Form*, or (b) directly into their electronic *Data Collection Form*.
        - If a print-copy is used for data collection, then *participant raters* will need to enter the data into their electronic *Data Collection Form* file after data collection has been completed.

        - Participant raters MUST ENTER THEIR OWN DATA into the e-file.

DATA COLLECTION – THE DAY OF THE STUDY:

7. Generate a list of patients (patient name and room/bed #) who are on each eligible unit at 0700. This list will be called the Patient List.
    a. Patients on these Patient Lists should be ordered by room/bed # (NOT alphabetically by patient name).

    b. ONLY PATIENTS ON THE PATIENT LISTS WILL BE INCLUDED in the study.

8. Review study procedures with the participant raters just before data collection is to begin.
    a. Verify with each participant rater that they have the correct rater number and corresponding *Survey ID – the most experienced and/or skilled in patient record review must be "Rater 1" with the Survey ID ending in "1"*.

    b. Instruct participant raters to retain their print-copy of the *Data Collection Form* with their *Survey ID*.

    c. Instruct participant raters they will use their *Survey ID* later to login to the survey site to upload their completed *Data Collection Form* and complete an online survey.

    d. Organize data collection so that participant raters collect data INDEPENDENTLY of each other.

- While participant raters are to collect data on the same morning, they SHOULD NOT collect data together!

- Consider staggering data collection start time by 15 to 30 minutes to prevent collaboration among participant raters.

9. Instruct participant raters to INDEPENDENTLY REVIEW 50 patient records, and INDPENDENTLY RATE these patients on the NDNQI pressure ulcer risk and prevention measures.
    a. Begin with 3 documents; the Unit List, the Patient Lists, and the *Data Collection Form*.

    b. Start with the first unit on the Unit List. On this unit, begin with the lowest room bed #. Is there a patient in this room/bed #?

        No - move to the next bed. (If the room has 2 beds, this would be Bed 2. If the room has one bed you would move to the next highest room #.)
        Yes - Is THIS patient on the Patient List?
            No - move to the next bed. (If the room has 2 beds, this would be Bed 2. If the room has one bed you would move to the next room.)
            Yes - review this patient's record and enter the findings on the first line of the Data Collection Form ("Pt. 1"). Move to the next room/bed # where there is a patient.

        - Continue data collection until data collection is completed on this unit.

        - Be sure to identify patients by their room/bed # on *the Data Collection Form.*

        - If the patient in the room/bed # is different than the patient on the Patient List, DO NOT COLLECT DATA ON THAT PATIENT. Participant raters should NOT SEEK TO FIND patients who are no longer in their room.

        - It is important participant raters collect data in their *usual manner*.

    c. When data collection on this unit is completed, data collection should proceed to the next unit on the Unit List.

    d. Collect data on each unit in the order that the units are listed on the Unit List. Always start with the lowest room/bed # within that unit, and continue in order to the highest room/bed # on that unit—until 50 patients have been rated.
        - Collecting data starting with the 1st unit on the Unit List and continuing by unit in the order the units are listed on the Unit List is ESSENTIAL so that participant raters rate the same 50 patients!!

e. During data collection, CONSIDER ONLY PATIENT CARE PROVIDED THE 24-HOUR PERIOD BEGINNING 0700 THE DAY BEFORE DATA COLLECTION TO 0700 THE DAY OF DATA COLLECTION!!
   - Using only this 24-hour period is necessary to ensure that each participant rater is considering the same patient data as the other participant raters.

f. It will take approximately 2.5 to 3 hours to collect data. However, data collection may take a little longer than expected if data collectors are unfamiliar with the study's *Data Collection Form* (30 – 45 minutes) may also be needed.

10. Remind participant raters:
    a. Do not discuss their choices with others.

    b. Do not make any changes to their data.

    c. We will not know who they are.

    d. We expect variation between participant raters' ratings.

    e. This variation does not reflect negatively on the participant rater or their institution.

AFTER DATA COLLECTION IS COMPLETED:

11. Refer participant raters to their *Participant Instructions* for information regarding the online survey.
    a. Remind participant raters that they have until OCTOBER 27th to login to *REDCap* to:
       - Upload their completed *Data Collection Form* Excel file.

       - Complete the online survey.

    b. The *REDCap* survey is available at http://redcap.kumc.edu/surveys/?s=VFuoTGEmIZ

    c. Once participant raters access the online survey, they must enter their *Survey ID* located on their print-copy of the *Data Collection Form*.

THANK YOU for participating in the study!!

Appendix H

Participant Rater Instructions

Participant Rater Instructions

This study requires you ("participant rater") to *independently* review 50 patient records and rate these 50 patients on the NDNQI pressure ulcer risk and prevention measures. You will then access an online site to submit your data and complete a survey.

PRIOR TO DATA COLLECTION:

1. Establish the date and time of the study in conjunction with the site coordinator and other participant raters.
    a. The data collection period is September 29th to October 13th.

    b. Select a single day during September 29th to October 13th to perform the study.

    c. The study will be performed between 0700 and 1200 on the day of the study.

    d. Data collection (for this study) is separate from your quarterly NDNQI pressure ulcer data collection.

2. Receive a print-copy and an electronic copy of the *Data Collection Form* from the site coordinator – you will use this form to collect data.
    a. On this form, your site coordinator has identified your rater number and corresponding *Survey ID*. You will use this unique and anonymous *Survey ID* later when you access the online survey.

    b. The *Survey ID* does not contain information that would make it possible to identify you.

    c. You and the site coordinator may choose to record the data as it is collected (a) onto the print-copy of the *Data Collection Form*, or (b) directly into an electronic *Data Collection Form* file.

    d. If the print-copy is used for data collection, then YOU will need to enter your data into your own electronic *Data Collection Form* file after data collection has been completed.

    e. BE SURE TO DOUBLE CHECK YOUR DATA ENTRY—YOU MUST BE CAREFUL TO ENTER THE DATA WITHOUT ERROR!!

    f. No other person should enter the data you collected into your electronic file for you.

DATA COLLECTION:

3. Only adult medical-surgical, adult medical, and adult surgical units are eligible for inclusion in the study. Eligible units in your hospital and the unit order from which data are to be collected are listed in your Unit List, which is a separate document. This Unit List includes unit name, NDNQI unit ID, and NDNQI unit type.

4. On the morning of data collection, the site coordinator will generate a list of patients (patient name and room/bed #) who are on each eligible unit at 0700. ONLY PATIENTS ON THESE PATIENT LISTS WILL BE INCLUDED IN THE STUDY.

5. Prepare to collect data INDEPENDENTLY of the other participant raters.
    a. Please do NOT COLLECT data alongside other raters. Consider staggering data collection start time by 15 to 30 minutes to prevent collaboration among participant raters.

    b. Please DO NOT ask others for help.

6. INDEPENDENTLY REVIEW 50 patient records and INDEPENDENTLY RATE these patients on the NDNQI pressure ulcer risk and prevention measures on the day of the study between 0700 and 1200.
    a. Begin with 3 documents; the Unit List, the Patient Lists, and the *Data Collection Form*.

    b. Start with the first unit on the Unit List. On this unit, begin with the lowest room bed #. Is there a patient in this room/bed #?

        No - move to the next bed. (If the room has 2 beds, this would be Bed 2. If the room has one bed you would move to the next highest room #.)
        Yes - Is THIS patient on the Patient List?
            No - move to the next bed. (If the room has 2 beds, this would be Bed 2. If the room has one bed you would move to the next highest room #.)
            Yes - review this patient's record and enter the findings on the first line of the Data Collection Form ("Pt. 1"). Move to the next highest room/bed # on this unit.

    - Continue data collection until data collection is completed on the first unit.

    - Be sure to identify patients by their room/bed # on *the Data Collection Form.*

    - If the patient in the room/bed # is different than the patient on the Patient List, DO NOT COLLECT DATA ON THAT PATIENT. You should NOT SEEK TO FIND patients who are no longer in their room.

c. When data collection on this unit is completed, data collection should proceed to the next unit on the Unit List. Continue collecting data according to the order listed on the Unit List.
    i. Collecting data starting with the 1st unit on the Unit List and continuing by unit in the order the units are listed on the Unit List is ESSENTIAL so that participant raters rate the same patients!!

d. On each unit, always start collecting data from patients in the lowest room/bed # and continue in order by room/bed# to patients in the highest room/bed #.
    - Collecting data starting with the lowest room/bed# on the unit to the highest room/bed # on the unit is ESSENTIAL so that participant raters rate the same patients!!

e. During data collection, CONSIDER ONLY PATIENT CARE PROVIDED THE 24-HOUR PERIOD BEGINNING 0700 THE DAY BEFORE DATA COLLECTION TO 0700 THE DAY OF DATA COLLECTION!!
    - Using only this 24-hour period is necessary to ensure that each participant rater is considering the same patient data as the other participant raters.

    - It is not necessary for the patient to have been in that room for the entire 24-hour period to be included in the study

f. Continue collecting data until 50 patients have been rated on the pressure ulcer risk and prevention measures.

7. RECORD your ratings on pressure ulcer risk and prevention as you collect this data on your *Data Collection Form.*
    a. Please record:
        i. Date and time you began collecting data.

        ii. Patient room/bed #, patient age, and patient gender.

        iii. NDNQI unit ID number. [Unit name, NDNQI unit ID, and NDNQI unit type are included in the Unit List.]

8. Do not share your ratings with others. This is crucial to the integrity of the study!!

AFTER DATA COLLECTION IS COMPLETED:

9. Do not discuss your choices with others! Do not make any changes to your data.
    a. Remember, we will not know who you are and we expect variation between participant raters' ratings.

    b. Understand this variation does not reflect negatively on you or your institution.

    c. Know that individual results will not be shared or distributed.

10. Login to *REDCap* to upload your electronic Excel file containing the completed *Data Collection Form*. This must be done **on or before OCTOBER 27th**.
    a. Go to https://redcap.kumc.edu/surveys/?s=VFuoTGEmIZ . You can use any computer with internet access.

    b. Enter your *Survey ID* in the place requested on the online survey. (This is a 7-digit alphanumeric code.)
        - Your *Survey ID* corresponds to your rater number and has been recorded by the site coordinator on the top of the print-copy of your *Data Collection Form*.

    c. Instructions for uploading your completed *Data Collection Form* are on the survey site.

11. Complete the *Pressure Ulcer Risk and Prevention Reliability Survey*.
    a. You will be able to access the survey only once, therefore you must complete the survey and file upload in one sitting.

    b. It will take 20 - 25 minutes to complete the survey and *Data Collection Form* Excel file upload.

    d. WE RECOMMEND YOU COMPLETE THE SURVEY AND ELECTRONIC FILE UPLOAD AS QUICKLY AS POSSIBLE AFTER COLLECTING DATA!! The survey, however, will be open until OCTOBER 27th.

    e. Complete the survey independently of others and without consulting any other resources.

    f. Answers will be stored in a separate protected database.

    g. Please remember your responses are anonymous.

THANK YOU for your effort and time!

Appendix I

Email Sent with Study Materials

Dear Site Coordinator,

Thank you for volunteering to participate in the Pressure Ulcer Risk and Prevention Reliability Study. Your site has been selected to participate. This email contains all necessary documents (7 TOTAL) related to this project.

A 1-hour teleconference is scheduled for Tuesday, September 9th at 11:00AM (ET), 10:00 AM (CT), 9:00 AM (MT) and 8:00 AM (PT). Details for the teleconference are included in the Teleconference Instructions, which are attached.

1. Overview of the Pressure Ulcer Risk and Prevention Reliability Study Overview of the Pressure Ulcer Risk and Prevention Reliability Study.pdf

2. Site Coordinator Instructions
Site Coordinator Instructions.pdf

3. Participant Rater Instructions
Participant Rater Instructions.pdf

4. Data Collection Form
Data Collection Form <<Hospital Code>>.xls

5. List of Eligible Units in your hospital PrU Eligible Units <<Hospital Code>>.pdf

6. Your hospital's Data Submission Link
https://redcap.kumc.edu/surveys/?s=VFuoTGEmIZ
Use this link and only this link to submit data for the Pressure Ulcer Risk and Prevention Reliability Study; you cannot submit Pressure Ulcer Risk and Prevention Reliability Study Data on the NDNQI website.

8. Teleconference Instructions
Teleconference Instructions.pdf

Please select your participant raters by the date of the teleconference. See the Site Coordinator Instructions for details of how to select these individuals. We encourage you to review the study materials with the participant raters prior to the teleconference and ask the participant raters to attend.

Appendix I (continued)

Email Sent with Study Materials

Please do not hesitate to contact us with any questions or concerns. Thank you again for your time and effort in completing this research.

Thank You,
Shirley Waugh
ndnqi@kumc.edu
913-588-1691

Appendix J

Teleconference Instructions

PrU Risk and Prevention Reliability Study
1-hour TELECONFERENCE Instructions

Tuesday, September 9, 2014 at
11:00 am Eastern
10:00 am Central
9:00 am Mountain
8:00 am Pacific

Please note: These times reflect Daylight Savings Time. Please adjust if your area does not observe Daylight Savings Time.

You are receiving this email because your site has registered for the **Pressure Ulcer Risk and Prevention Reliability Study**. Due to the volume of participants, be sure to dial in to the teleconference **10 minutes** ahead of the scheduled time.

Select your participant raters and encourage them to attend the teleconference.

- It is imperative, however, that each site does not use more than its **one** allotted port into the conference. Therefore, all participants at your site must call from the same telephone/speaker phone.

To access the teleconference, please call **1-800-268-5851** 10 to 30 minutes prior to the start time. All calls are operator assisted. **There is NO PASSWORD for this conference.** You will be placed in the listen-only mode for the conference. However, you will have an opportunity to ask questions. Press '*1' and the operator will announce you for your question of the panel.

To maintain confidentiality during the question and answer, teleconference participants should **not** identify the hospital they represent. Only first names and State should be used. "Sally from Kansas," for example.

If at any point during the conference call you require operator assistance you may access the operator by pressing '*0'.

If you are unable to attend, an audio-recording will be available upon request.

We are looking forward to your participation.

Shirley Waugh
ndnqu@kumc.edu
913-588-1691

Appendix K

Pilot Study Feedback and Responses

1. Did you understand the study based on information presented in the *Overview of the Pressure Ulcer Risk and Prevention Reliability Study*?
   Yes **4(100%)**
   Somewhat
   No, not at all – please comment _____
   Not applicable, I am not the site coordinator and did not read this information.

2. Did you understand the *Site Coordinator Instructions?*
   Yes **3(75%)**
   Somewhat
   No, not at all – please comment _____
   Not applicable, I am not the site coordinator and did not read the information.
   1(25%)

3. Did you understand the *Participant Rater Instructions*?
   Yes **4(100%)**
   Somewhat
   No, not at all – please comment _____

4. Was *the Data Collection Form* readable?
   Yes **3(75%)**
   Somewhat **1(25%)**
   No, very difficult to read – please comment _____

5. What form of the *Data Collection Form* was used to record data as it was collected?
   Used a print-copy during data collection **4(100%)**
   Entered data directly into the electronic file as it was collected

Appendix K (continued)

Pilot Study Feedback

6. On average, how long did it take to complete the data collection for one patient?
   0 – 5 minutes

   > 5 – 10 minutes **2(50%)**

   > 10 – 15 minutes **1(25%)**

   > 15 minutes **1(25%)**

7. How long did it take to complete the data collection for all 50 patients?
   0 – 30 minutes

   > 30 – 60 minutes

   > 1 hour – 2 hours

   > 2 hours – 3 hours

   > 3 hours – 4 hours **1(25%)**

   > 4 hours **3(75%)**

8. Did you have difficulty linking to the online survey site?
   No **3(75%)**
   Yes, but I figured it out **1(25%)**
   Yes, I had a lot of difficulty – please comment _____

9. Did you have problems uploading the *Data Collection Form* Excel file?
   No **2(50%)**
   Yes, but I but I figured it out **2(50%)**
   Yes, I had a lot of difficulty – please comment _____

Appendix K (continued)

Pilot Study Feedback

10. Did you have any difficulty completing the online survey?
    No **4(100%)**
    Yes, but it I figured it out
    Yes, I had a lot of difficulty – please comment _____

11. How long did it take you to complete the entire online survey, and upload your *Data Collection Form* Excel file?

    less than or equal to 15 minutes **2(50%)**

    > 15 minutes – 30 minutes **2(50%)**

    > 30 minutes – 45 minutes

    > 45 minutes – 1 hour

    > 1 hour

12. Could you read the questions on the online survey?
    Yes **4(100%)**
    No, not really
    No, very difficult to read – please comment _____

13. Did you understand the questions on the online survey?
    Yes **4(100%)**
    No, not really
    No, not at all – please comment _____

14. Please describe the issues or problems in participating in this study. **3(75%)**

    "The electronic form had some issues with auto-fill, but was fixed."

    "The length of time and no access to bedside items."

    "There were a few bugs getting into the Excel file, but we figured it out."

Thank you for your time and valuable feedback!