# Exploring factors associated with pressure ulcers: A data mining approach

Dheeraj Raju [a,*], Xiaogang Su [b], Patricia A. Patrician [a], Lori A. Loan [c], Mary S. McCarthy [c]

[a] School of Nursing, University of Alabama at Birmingham, United States
[b] University of Texas at El Paso, United States
[c] Center for Nursing Science & Clinical Inquiry, Madigan Army Medical Center, Tacoma, United States

## ARTICLE INFO

## ABSTRACT

Background: Pressure ulcers are associated with a nearly three-fold increase in in-hospital mortality. It is essential to investigate how other factors besides the Braden scale could enhance the prediction of pressure ulcers. Data mining modeling techniques can be beneficial to conduct this type of analysis. Data mining techniques have been applied extensively in health care, but are not widely used in nursing research.
Purpose: To remedy this methodological gap, this paper will review, explain, and compare several data mining models to examine patient level factors associated with pressure ulcers based on a four year study from military hospitals in the United States.
Methods: The variables included in the analysis are easily accessible demographic information and medical measurements. Logistic regression, decision trees, random forests, and multivariate adaptive regression splines were compared based on their performance and interpretability.
Results: The random forests model had the highest accuracy (C-statistic) with the following variables, in order of importance, ranked highest in predicting pressure ulcers: days in the hospital, serum albumin, age, blood urea nitrogen, and total Braden score.
Conclusion: Data mining, particularly, random forests are useful in predictive modeling. It is important for hospitals and health care systems to use their own data over time for pressure ulcer risk prediction, to develop risk models based upon more than the total Braden score, and specific to their patient population.

© 2014 Elsevier Ltd. All rights reserved.

### What is already known about the topic?

- The Braden scale is one of the widely used tool for assessing pressure ulcer risk.
- Data mining techniques have been applied extensively in health care, but are not widely used in nursing research.
- The nursing research studies use one or another technique but do not compare them.

### What the paper adds?

- This paper adds to our knowledge of how other factors enhance assessing the probability of developing pressure ulcers when combined with the Braden scale.
- The paper extends the knowledge of data mining to the nursing statistical toolbox.

* Corresponding author at: The Center for Nursing Research, 1720 2nd Avenue South, NB 1019F, Birmingham, AL 35294, United States.
Tel.: +1 205 975 5336.
E-mail address: seeth001@uab.edu (D. Raju).

## 1. Introduction

Pressure ulcers (PU) are a substantial burden for patients and for the health care system in general. The National Patient Care Safety Monitoring Study (Lyder et al., 2012) of over 51,000 patients found that 4.5% of Medicare beneficiaries developed a pressure ulcer during their hospital stay and 5.8% had a pressure ulcer on admission. Pressure ulcers regardless of whether they were present on admission were associated with a nearly three-fold increase in in-hospital mortality, 69% increase in 30-day mortality, and an increased length of stay of 6.4 days (Lyder et al., 2012). As of 2008, hospital acquired stage III and IV pressure ulcers are no longer reimbursed by the U.S. Centers for Medicare and Medicaid Services (CMS), leaving the hospitals themselves to absorb the cost of care for patients, which is estimated at $43,180 per patient (Armstrong et al., 2008). Thus it is imperative to discover factors associated with both community and hospital acquired pressure ulcers and institute additional care measures to prevent their occurrence.

Because the causative factors for pressure ulcers are "multifactorial and not well understood" (Benoit and Mion, 2012, p. 341), it is critical for hospitals, nursing homes, and home care agencies to systematically monitor patients for pressure ulcer rates, assess risk, and enhance prevention efforts. Although not all pressure ulcers are avoidable (Black et al., 2011), frequent monitoring may lead to better risk predictions and more thoughtful application of resources (i.e., evidence-based nursing preventive interventions such as turning and repositioning) to those who need it most. As more hospitals adopt electronic medical records, the large clinical data repositories could help improve clinical care through the study of their own best practices and lessons learned directly from their patients. Analyzing clinical data collected from discharge abstracts or directly from clinical records and comparing those who developed or did not develop a pressure ulcer can inform problem identification in quality improvement. Data mining modeling techniques can be beneficial to conduct this analysis. The purpose of this paper is to build and compare data mining models for pressure ulcer prevalence (both community and hospital acquired) and determine the variables that are associated with pressure ulcers based on a four year study database collected from 12 military hospitals. The variables included in the analysis are easily accessible demographical information and medical measurements. We carefully selected a group of data mining techniques that not only supply high predictive accuracy but also allow for meaningful interpretations. The Braden scale developed by Bergstrom et al. (1987) is one currently available tool for assessing pressure ulcer risk. Given the multifaceted nature of pressure ulcers, it is of keen interest to see whether and how other factors could enhance the performance of predicting pressure ulcers when combined with the Braden scale. The eventual wide scale use of electronic medical records will enable hospitals to apply these data mining techniques to their own patient level data to determine factors associated with pressure ulcers.

## 2. Background

Identifying patients who may have a pressure ulcer on admission to a health care facility or who may develop one during hospitalization is the starting point for primary, secondary and tertiary preventive activities aimed at reducing this costly and debilitating complication. A recent systematic review of 54 pressure ulcer studies (Coleman et al., 2013) identified three primary risk factors: mobility level, perfusion, and skin status. Other secondary risk factors that emerged from this literature were skin moisture, age, basic serum metabolic measures, nutrition and general health status.

Risk assessment scales have been used for many years to forecast the patients who are at high risk for developing pressure ulcers; however, their sensitivity and specificity are far from ideal (Pancorbo-Hidalgo et al., 2006; Schoonhoven et al., 2006). It is important that risk assessment scales predict the likelihood of getting a pressure ulcer, so that scarce resources can be applied in an evidence-based manner to the highest risk patients. The Braden Scale for Predicting Pressure Sore Risk (Bergstrom et al., 1987) is the most widely used and most widely researched risk prediction scale for pressure ulcers (Balzer et al., 2007; Pancorbo-Hidalgo et al., 2006). According to a meta-analysis of risk assessment scales, the Braden has the "best sensitivity and specificity balance" (weighted means were 57.1% and 67.5% respectively) and was the best at predicting risk (Odds Ratio of 4.08, 95% CI = 2.56–6.48) compared to Norton Scale, Waterlow Scale, and Nurses' clinical judgment (Pancorbo-Hidalgo et al., 2006).

The Braden scale consists of six subscales derived from two main components of a patients' skin, pressure and tissue tolerance. The subscales mobility, activity, and sensory perception are associated with pressure, whereas the subscales moisture, friction/shear, and nutritional status are associated with the level of tissue tolerance for pressure. These six subscales are rated on scales of 1–4, with the exception of the friction/shear scale which is rated 1–3. When added together, the scores range from 6 to 23, with lower numbers indicating higher risk for pressure ulcer development. Scores of 18–23 indicate no risk; 15–18 for low risk; 13–14 moderate risk; 10–12 high risk, and 6–9 very high risk (Cremasco et al., 2012).

Although the Braden scale is widely used for predicting patients at risk, it is not infallible. In the Coleman et al. (2013) review, mobility subscale scores were found more predictive than total risk assessment scores among all risk assessment scales. Furthermore, the Braden scale does not evaluate perfusion and skin status as defined in the Coleman study. There are other factors important to consider in predicting patients at risk. For example, in 7 of the 11 studies Coleman et al. (2013) reviewed that evaluated serum albumin, it was noted to be statistically associated with pressure ulcer development, such that lower albumin was associated with pressure ulcers (odds ratios of 0.4–0.8). Serum albumin is associated with nutritional status, which is of practical importance in wound healing. Because of the difficulty in capturing all aspects important to predict pressure ulcers, we used data mining techniques to provide a comprehensive assessment