The Braden Scale Cannot Be Used Alone for Assessing Pressure Ulcer Risk in Surgical Patients: A Meta-Analysis

Wei He, MSc; Peng Liu, MSc; and Hong-Lin Chen, MSc

Abstract
The validity and reliability of the Braden Scale for pressure ulcer development has been established in a variety of patient care settings, but studies suggest the scale does not capture risk factors in surgical patients. The purpose of this meta-analysis was to assess the predictive validity of the Braden Scale for pressure ulcer development in surgical patients. A literature search using PubMed and Web of Science databases (through July 2011) was conducted to identify all clinical studies on predicting pressure ulcers in surgical patients using the Braden Scale. To be eligible for inclusion, studies had to include sensitivity (true positive rate, TPR) and specificity (true negative rate, TNR) results or include sufficient data to calculate these factors. Study quality was assessed using the 14-item Quality Assessment of Diagnostic Accuracy Studies (QUADAS) instrument, and two-by-two tables of predictive validity were constructed from each article. Meta-analysis for predictive validity was performed, including calculation of pooled sensitivity, pooled specificity, diagnostic odds ratio (DOR), construction of summary receiver operating characteristic (SROC) curves, and overall diagnostic accuracy (Q*). Three studies (N = 609 patients) met the meta-analysis inclusion criteria. The pooled estimates for sensitivity and specificity were 0.42 (95% CI: 0.38 to 0.47) and 0.84 (95% CI: 0.83 to 0.85), respectively, yielding a combined DOR of 4.40 (95% CI: 2.98 to 6.50). The area under the ROC curve (AUC) was 0.6921 ± 0.0346, and the Q* was 0.6466 ± 0.0274. Significant heterogeneity was noted between the included studies with Q value 34.49 (P = 0.0321), and I² for pooled sensitivity, pooled specificity, and pooled DOR was 88.7%, 98.6%, and 39.1%, respectively. Although the observed heterogeneity between studies may have affected the results, the low values for overall diagnostic accuracy (Q*) and diagnostic capability (AUC) indicate the Braden Scale has low predictive validity for pressure ulcer risk in surgical patients. A new pressure ulcer risk assessment scale for surgical patients should be developed and tested.

Keywords: pressure ulcer, surgical patients, Braden Scale, predictive validity, meta-analysis


Potential Conflicts of Interest: none disclosed
factors for intraoperative pressure ulcer risk, including patient age, length of surgery, type of surgery, and comorbidities such as diabetes, preoperative hypertension, respiratory disease, and vascular disease are not included in the Braden Scale. The results of a descriptive study suggest the scale is not effective for determining pressure ulcer risk in surgical patients. However, the predictive validity of the Braden Scale for pressure ulcers in surgical patients has not been systematically reviewed.

The predictive validity of a test is determined by its sensitivity (also called true positive rate, TPR) and specificity (also called true negative rate, TNR). Sensitivity and specificity can be graphically represented by the receiver operating characteristic (ROC) curve that plots the TPR (sensitivity) against the false-positive rate (1-specificity). Based on data from a meta-analysis, the pooled sensitivities, pooled specificities, diagnostic odds ratios (DOR), summary receiver operating characteristic (SROC) curve, and overall diagnostic accuracy (Q*) have been recommended to represent the performance of predictive validity.

The purpose of this meta-analysis was to assess the predictive validity of the Braden Scale for pressure ulcer development in surgical patients.

Methods

Literature search strategy. A literature search was conducted to identify studies published in PubMed and Web of Science databases before July 2011 that evaluated the predictive validity of the Braden Scale for pressure ulcer development in surgical patients. Search terms, used alone or in combination, included sensitivity, specificity, accuracy, predictive value, ROC, likelihood ratio, pressure ulcer, surgical procedures, operative, and Braden. The “related articles” function in PubMed and the “citing articles” or “cited articles” function in Web of Science also were used to broaden the search.

Inclusion and exclusion criteria. All studies that met the following criteria were included: assesses the predictive validity of Braden Scale for pressure ulcers in surgical patients; provides sufficient information to construct the two-by-two contingency tables for individual study subjects; and includes actual sensitivity (TPR) and specificity (TNR) results or includes sufficient data to calculate these factors. Two reviewers independently judged study eligibility while screening the citations. Disagreements were resolved by a third reviewer.

Data extraction. To extract data, a standard form was created that included the following variables: author, year of publication, country in which the study was conducted, study design, population and number of patients, age, duration of follow-up (eg, postoperative days 1, 3, and 5), and pressure ulcer prevalence. The cutoff value used for the index test (Braden Scale score) and TPR, TNR, false negative, and false positive numbers also were collected to facilitate construction of the two-by-two tables. Two reviewers independently assessed the quality of studies using the 14-item Quality Assessment of Diagnostic Accuracy Studies (QUADAS) instrument. Disagreements were resolved by a third reviewer.

Statistical analysis. The overall pooled sensitivity, specificity, and DOR, with 95% confidence intervals (CI), were estimated by DerSimonian and Laird’s random-effects model. In addition, SROC analysis was performed to examine the interaction between sensitivity and specificity and to quantify test performance using the area under the curve (AUC) and Q* value. Heterogeneity was analyzed by Cochran’s Q test and I² statistic. A P value of <0.05 by Cochran’s Q test indicated significant heterogeneity; an I² >50% indicated substantial heterogeneity. Threshold analyses by Spearman correlation between Logit (TPR) and Logit (1-TNP) also were performed. All analyses were performed using MetaDisc 1.4 (version 0.6, a program developed by a Unit of the Clinical Biostatistics Team, Ramon y Cajal Hospital, Madrid, Spain).

Results

Literature search. During the first literature search, two English-language publications were retrieved from PubMed and three from Web of Science; of those, only one was used. One was excluded because the study population included patients admitted to surgical and internal medicine, as well as to neurology and geriatric wards; and the other did not meet the meta-analysis inclusion criteria because the sample included patients in medical and infectious disease wards. No publications were identified, added, or excluded after broadening the search criteria, leaving a sample of three studies (N = 609 patients) for inclusion in this meta-analysis. (see Table 1). The quality of all three studies was good; 90% of the 14 QUADAS items received a yes score (see Table 2).
<table>
<thead>
<tr>
<th>First Author</th>
<th>Country</th>
<th>Study design</th>
<th>Population</th>
<th>Number of patients (male/female)</th>
<th>Assessment time</th>
<th>Prevalence Braden Scale Cut-Off Score</th>
<th>Age (years)</th>
<th>Study quality assessment results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kim 2007</td>
<td>South Korea</td>
<td>Prospective</td>
<td>Surgical ICU</td>
<td>219 (145/74)</td>
<td>Not mentioned</td>
<td>18.3% 16–98 (58.1 ± 1.2)</td>
<td>20, 16, 11, 10, 9</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>Feuchtinger 2007</td>
<td>Germany</td>
<td>Prospective</td>
<td>Cardiac surgery ICU</td>
<td>53 (31/22)</td>
<td>POD 1</td>
<td>49.0% 25–83 (62.0 ± 12.1)</td>
<td>20, 16, 11, 10, 9</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>Lewicki 2000</td>
<td>USA</td>
<td>Not mentioned</td>
<td>Not mentioned</td>
<td>337 (254/83)</td>
<td>Preoperative POD 1, 3, 5</td>
<td>4.7% 22–86 (62.0 ± 11.6)</td>
<td>17, 18, 19, 20, 21, 22, 23</td>
<td>Y Y Y</td>
</tr>
</tbody>
</table>

**Table 2. Study quality assessment results**

<table>
<thead>
<tr>
<th>QUADAS Instrument items</th>
<th>Study ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was the spectrum of patients representative of the patients who will receive the test in practice?</td>
<td>Y N N</td>
</tr>
<tr>
<td>2. Were selection criteria clearly described?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>3. Is the reference standard likely to correctly classify the target condition?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?</td>
<td>U U U</td>
</tr>
<tr>
<td>6. Did patients receive the same reference standard regardless of the index test result?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>7. Was the reference standard independent of the index test (ie, the index test did not form part of the reference standard)?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>8. Was the execution of the index test described in sufficient detail to permit replication of the test?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>9. Was the execution of the reference standard described in sufficient detail to permit its replication?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>10. Were the index test results interpreted without knowledge of the results of the reference standard?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>11. Were the reference standard results interpreted without knowledge of the results of the index test?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>13. Were uninterpretable/intermediate test results reported?</td>
<td>Y Y Y</td>
</tr>
<tr>
<td>14. Were withdrawals from the study explained?</td>
<td>U U U</td>
</tr>
</tbody>
</table>

_Y = Yes, N = No, U = Unclear per references 9, 16, and 17_
Statistical data synthesis. The pooled sensitivity for the three studies was 0.42 (95% CI: 0.38 to 0.47; $\chi^2_{(2)} = 185.57, P = 0.000$), indicating the probability that the test result will be positive when applied to a person who actually developed a pressure ulcer was 42% (see Figure 1a). The pooled specificity was 0.84 (95% CI: 0.83 to 0.85; $\chi^2_{(2)} = 1501.04, P = 0.000$) (see Figure 1b). The pooled DOR was 4.40 (95% CI: 2.98 to 6.50) (see Figure 1c). The SROC curve for all three studies is shown in Figure 1d. The overall weighted AUC was 0.6921 ± 0.0346, and the overall diagnostic accuracy ($Q^*$) was 0.6466 ± 0.0274. The low AUC and a $Q^*$ of <0.7 suggest the Braden Scale has low predictive validity for pressure ulcer development in surgical patients.18

Study heterogeneity. The Cochran’s $Q$ test value was 34.49 ($P = 0.0321$).17 for pooled sensitivity, pooled specificity, and pooled DOR was 88.7%, 98.6%, and 39.1%, respectively. The results showed significant heterogeneity between the studies. Threshold analyses showed heterogeneity came from the Braden cut-off score (ie, the diagnostic threshold (Spearman correlation coefficient: 0.875, $P = 0.000$).

Discussion
During surgery, patients are immobile and not able to feel pain caused by prolonged pressure on the operating table. Shear and friction injury also can occur as patients are repositioned and transferred. Some patients cannot be turned for extended periods of time due to postoperative use of balloon pumps or other devices.1 These and other factors may cause pressure ulcers in surgical patients. The Braden Scale is a good instrument for predicting pressure ulcer risk in acute care and long-term care settings, and often is used in surgical settings.19,20

A meta-analysis7 of the Braden Scale for assessing pressure ulcer risk among clinical settings showed that Braden and Norton scales are more accurate than nurses’ judgment in predicting pressure ulcer risk. In the current study, the authors used the meta-analysis method for diagnostic tests to assess predictive validity in order to ascertain the predictive validity assessed not only by pooled sensitivity and specificity, but also by SROC, AUC, and $Q^*$ value. This approach provides information for each cut-off point and is a comprehensive method to evaluate predictive validity.

The current meta-analysis included three publications with sample sizes ranging from 53 to 337 patients. The quality of these three studies, as evaluated using the QUADAS instrument for assessing the quality of diagnostic studies, was good, with 90% of items receiving a yes rating.18 A summary score estimating the overall quality of an article was not calculated, because the interpretation of such summary scores is problematic and potentially misleading.21

In this meta-analysis, the pooled sensitivity (TPR) was 0.42 (95% CI: 0.38 to 0.47), and the pooled specificity (TPN) was 0.84 (95% CI: 0.83 to 0.85). Such results imply that the probability that the test result will be positive when the test is applied to a person who developed a pressure ulcer was 42%, and the probability that the test result will be negative when the test is applied to a person who actually does not develop a pressure ulcer was 84%. The predictive validity was poor, especially for the TPR. The overall weighted AUC in this meta-analysis was 0.6921 ± 0.0346. An AUC value of 0.5 indicates that a test has no discriminatory ability, whereas an AUC value of 1.0 indicates perfect diagnostic capability.22 Others have suggested that, to demonstrate excellent accuracy, the AUC should be in the region of 0.97 or above; an AUC of 0.93 to 0.96 is very good; 0.75 to 0.92 is good; but an AUC less than 0.75 has obvious deficiencies in its diagnostic accuracy.23 The result showed that the Braden Scale has low predictive validity for pressure ulcer in surgical patients.

$Q^*$ has been suggested as a single number summarization of predictive validity, being the point closest to the ideal top-left corner of the SROC space for symmetric curves. $Q^*$ is the point on the SROC where TPR = 1−FPR.11 The higher the $Q^*$ value, the better the accuracy of predictive validity. When $Q^*$ is <0.7, predictive validity is poor.24 The $Q^*$ in this meta-analysis was 0.6466 ± 0.0274, suggesting that the Braden Scale cannot be used alone for predicting pressure ulcer development in surgical patients or that a new pressure ulcer risk assessment should be developed for surgical patients.

Limitations
The current meta-analyses have some limitations. First, it included only three studies with a total of 609 patients. This small sample size may have affected predictive validity. Second, the results showed a significant heterogeneity between the studies. Threshold effect in the meta-analysis was noted. This heterogeneity could be the result of variations in patient assessment time differences, actual pressure ulcer prevalence rates, or the Braden Scale Score cut-off value. Third, the authors did not conduct a publication bias analysis with so few studies available for inclusion. Although this could be a limitation of the meta-analysis, publication bias will not change the predictive validity very much, because the pooled diagnosis effect was too low.

Conclusion
A meta-analysis was conducted to assess the predictive validity of the Braden Scale score for pressure ulcer risk assessment in surgical patients. The study found that the pooled sensitivity was low (0.42 [95% CI: 0.38 to 0.47]), and the pooled specificity was high (0.84 [95% CI: 0.83 to 0.85]].
Figure 1. Meta-analysis forest plots of the included studies. A: the forest plots of pooled predictive sensitivity; B: the forest plots of pooled predictive specificity; C: the forest plots of pooled diagnostic odds ratios; D: SROC plot for all cut-off values combined and for the included studies.
to 0.85), indicating that scale score is specific but not very sensitive. In addition, both the overall weighted AUC and Q² value, calculated to determine diagnostic accuracy and predictive validity, were low (0.6921 ± 0.0346 and 0.6466 ± 0.0274, respectively). The results suggest that the Braden Scale has low predictive validity for assessing pressure ulcer risk in surgical patients. Thus, it cannot be used alone to predict pressure ulcer risk in these patients, and a new pressure ulcer risk assessment scale for surgical patients should be developed and tested.

References