Reliability of the Bates-Jensen Wound Assessment Tool (BWAT) for Pressure Injury Assessment: The Pressure Ulcer Detection Study

Barbara M. Bates-Jensen[1,*], **Heather McCreath**[2], **Anabel Patlan**[1], **Deniz Harputlu**[4]

[1]University of California, Los Angeles, School of Nursing, Los Angeles, California, United States.

[2]University of California, Los Angeles, David Geffen School of Medicine, Division of Geriatrics, Los Angeles, California, United States.

[4]Ankara Universitesi, Ankara, Turkey

Wounds are classified as chronic if fail to heal in a timely manner and they are associated with significant morbidity, cost, and suffering for patients. In the United States chronic wounds affect more than 6.5 million people and these numbers are expected to continue to grow as the population ages and diabetes, obesity and cardiovascular disease continue to increase[1]. Chronic wounds are burdensome for health care providers as many require weeks and months to heal, often necessitating complex treatment regimens and a multidisciplinary approach. One challenge for health care providers is evaluating treatment effects on chronic wounds. Historically, wound assessment has relied on provider observations, with inconsistent use of multiple assessment tools across providers[1,2]. Standardized, evidence-based tools allow consistent assessment of wound characteristics to monitor and measure wound healing progress. Use of wound assessment tools in clinical settings promotes communication among health care providers and is recommended in multiple clinical practice guidelines[2–4]. One such instrument, the Bates-Jensen Wound Assessment Tool (BWAT), is used in the U.S. and internationally. The BWAT consists of 13 wound characteristics: size, visible depth, wound edges, undermining and tunneling processes, necrotic tissue type and amount, exudate type and amount, surrounding skin discoloration, peripheral tissue edema, peripheral tissue induration, granulation tissue and epithelialization[5,6]. Nine characteristics are subjectively rated on a 1 to 5 scale, with a value

of 1 indicating the healthiest attribute and a value of 5 indicating the least healthy attribute of the characteristic. The remaining four characteristics (size, depth, edges, undermining) are rated from 0–5 with a value of 0 indicating "none present" and scored for wounds that have resolved. The 13 wound characteristic item scores can be summed (with no weighting) for a total score ranging from 9 to 65 (profound tissue degeneration)[5,6].

The BWAT is a 2001 revision of the Pressure Sore Status Tool (PSST), first published in 1990[5]. In the original study, Bates-Jensen and colleagues[5] demonstrated the BWAT reliability on ten adult patients with 20 PrIs in an acute care hospital. Two nurses with special training in wound care independently and concurrently assessed the PrI twice two hours apart. Inter-rater and intra-rater reliability was demonstrated with Spearman rho correlation coefficients greater than 0.90[5]. Inter-rater reliability with three licensed practical nurses and two physical therapists assessing NH residents with a variety of wound types demonstrated total score agreement of 75% in a second study[7]. In a third study, inter-rater reliability between two licensed treatment nurses on 40 NH residents with stage 2–4 PrI demonstrated a total score Spearman rho correlation coefficient of 0.78. Across the same PrIs, between a licensed treatment nurse and a wound care nurse, the total score was correlated (r= 0.89) with overall agreement of 80% for the treatment nurses and 95% for the wound care and treatment nurses[8]. An additional study examining reliability in the NH setting with a wound care nurse and licensed treatment nurses across 304 assessments of stage 1 through 4 PrIs found a higher agreement of 97% for total score and high agreement across individual items (all items above 92% agreement)[9]. However, there have been few reports of reliability since the original tool development. Studies have looked only at nurse or physician raters, and reliability analyses have been less rigorous than is now suggested in the literature[10]. The strengths of the BWAT are the use of the tool as a foundation for development of other instruments, use in clinical research trials, with a wide variety of chronic wounds, and with various populations across multiple settings. The BWAT has provided a basis for other wound assessment tools including the DESIGN-R in Japan[11], the Photographic Wound Assessment Tool[12,13], the Spinal Cord Injury Pressure Ulcer Monitoring Tool[14,15] and the diabetic foot ulcer assessment scale (DFUAS)[16]. In each of these cases, the BWAT has been used as the reference method to validate the new instrument.

Since its development, the BWAT has been used to assess wound healing in clinical research trials evaluating multiple wound treatments[17–24]. These clinical studies have examined diverse wound treatments ranging from topical dressings[17,21,23] and advanced wound therapies (negative pressure wound therapy[21,22], multiwave length light therapy and shock wave therapy[19]) to patient education interventions[25]. Takahashi et al used the total BWAT score to evaluate healing of stage 3 and 4 pressure injuries (PrI) treated with plastic wrap compared to moist gauze dressings and showed a significant difference in treatments with the BWAT total score as the outcome measure[17]. Similarly, Porso and colleagues used the BWAT to measure healing in lower leg ulcers treated with defocused shock wave therapy and showed improved BWAT total scores in their pre-experimental pilot study[19]. In their randomized controlled trial, Clark et al used the BWAT to evaluate PrI healing in persons with spinal cord injury receiving an occupational therapy intervention[26]. Some investigators have also used subsets of BWAT characteristics to evaluate certain effects of treatments. For

example, McCallon and Frilot used the BWAT total score and specific necrotic tissue characteristics to evaluate the effects of clostridial collagenase ointment and negative pressure wound therapy on healing of chronic PrIs[21]. The BWAT has also been used by investigators to specifically profile wounds. For example, Paul used the BWAT to assess characteristics of wounds that itch and others have used the BWAT to examine characteristics of recurrent PrI[27].

Though originally developed and tested to assess wound healing in adults with PrI, BWAT use has been extended to multiple wounds, both acute and chronic, including post-surgical wounds[22], infected wounds[18], neuropathic foot ulcers[24], and lower leg ulcers[19,23]. It has been used with a variety of populations and health care settings including older adults in nursing homes (NHs) and acute care hospitals[7–9,20,22,28], persons with spinal cord injury in acute care hospitals and in the community[26,29,30], children with spina bifida in an outpatient clinic setting[31], and adult patients in acute care hospitals[32]. In all the above examples, the BWAT was scored by a nurse or physician demonstrating setting and practice versatility.

In the majority of these studies, the reliability of BWAT score was rarely mentioned or examined. Similarly, data examining properties of individual BWAT characteristics is limited. Thus, there is a need for further reliability testing of the BWAT. Understanding the psychometric properties of the tool may suggest possible revisions and inform clinician guidelines for use. Further, there are no reports of reliability of any wound assessment tools across differing wound severity, on persons with different skin tones or from different ethnicity/racial groups, or specific to anatomic locations.

This study seeks to address some of these important gaps. Here, we present findings from the BWAT related to PrI characteristics and natural history over 16 weeks among 142 nursing home (NH) residents who began the study with a PrI or developed a PrI during the study. Our objectives were to:

1. Compare BWAT scores relative to PrI stage, anatomic location, skin tone group, ethnicity/racial group and natural history.

2. Identify reliability estimates for BWAT use across PrI stage, anatomic location, and skin tone and ethnicity/racial group.

3. Examine use of BWAT scores as indicators of PrI healing or deterioration across PrI stage and anatomic locations.

Providing updated reliability data for the BWAT as used in wounds of various severity and across anatomic locations as well as using the BWAT to describe PrIs among an ethnically diverse NH population may increase our understanding of PrI assessment among this vulnerable population.

METHODS

Subjects and Setting

This study was part of a larger study, the Pressure Ulcer Detection (PUD) study, and the methods have been previously reported[33]. The study protocol was approved by the

University of California, Los Angeles, Human Subject Protection Committee, and each nursing home administration approved the study based on facility guidelines. Research staff obtained written informed consent to participate in the study directly from those residents who were able to provide informed consent or from their designated representatives for residents unable to provide consent, with assent obtained from the resident. Only participants with PrI are reported in this paper. Figure 1 shows the flow of participants with PrI for this study.

Residents were recruited from 19 NHs in the greater Los Angeles area. The characteristics of the NHs have been previously reported[33]. Briefly, study NHs demonstrated relatively lower rates than national and state averages among short stay residents with PrIs and demonstrated relatively higher rates than national and state averages long stay residents with PrIs.

All data were collected by trained research staff that included non-health care staff and first year prelicensure nursing students. All research staff received training by one of the principal investigators (BBJ) for all data collection protocols, visual skin assessments, and subepidermal moisture measures as previously described[33].

Medical Record Data

Research staff abstracted medical and demographic information from participants' medical records and their most recent Resident Assessment Instrument Minimum Data Set (MDS), a multi-domain assessment tool that is completed on admission and at quarterly intervals for residents in Medicare and Medicaid accredited NHs in the U.S. During the study, the MDS was updated from version 2.0 to 3.0. This did not affect extracted data.

Braden Scale Risk Assessment Scores

The Braden Scale for Predicting Pressure Sore Risk (Braden Scale) was used for monthly assessment of PrI risk. The Braden Scale is composed of six subscales (five subscales rated from 1–4): (1) sensory perception, (2) moisture, (3) activity, (4) mobility, (5) nutrition, and (6) friction and shear (rated from 1–3) which are summed for a total score ranging from 6 to 23 (low scores indicate high risk for PrI development).[34] A score of less than 18 is considered "at risk"[35]. Reliability has been previously reported for the PUD study[33].

Skin Tone Assessment

Skin tone was assessed by research staff using the Munsell System of Color Notation (Munsell chart) to provide an objective assessment. The protocol has been presented elsewhere[36]. Skin tone categories were defined as light (Munsell values 7–8), medium (Munsell values 5–6), and dark (Munsell values 2.5–4) relative to the alphanumeric designations assigned on the Munsell chart based on three qualities: hue, value, and chroma of the best color match for the skin[36].

Visual Skin Assessment and Subepidermal Moisture Measures

Each week, two research staff members independently assessed skin health at nine anatomic locations: sacral as well as right and left buttocks, ischial tuberosities, trochanters, and heels.

No PrI were observed on trochanters so these locations are not reported here. Visual skin assessment training, definitions of blanchable and nonblanchable erythema and deep tissue injury (DTI), details of the protocol, and reliability have been previously reported[33]. Skin that demonstrated no erythema or PrI damage was considered normal skin.

Concurrent with visual skin and PrI assessments, subepidermal moisture was measured weekly using the Delfin MoistureMeter D (Delfin Technologies, LTD, Greenwich, CT) dermal phase meter on clean, dry skin at the lateral wound edge of the PrI. The details of the protocol and reliability of the instrument have been previously reported[33]. Briefly, subepidermal moisture reflects water *below* the epidermis to the depth of 2.5mm. Over 8 seconds electromagnetic waves of 300 MHz are transmitted via a coaxial line terminating in an open-ended coaxial wand on the skin. The induced electrical field interacts with water molecules closest to the wand while the portion of the electromagnetic energy that is not absorbed by tissue water is reflected, measured, and displayed in the measuring unit. Subepidermal moisture is displayed as the tissue dielectric constant (TDC) and is directly proportional to the hydration of the tissue. Thus, subepidermal moisture increases with increasing water content and edema (range 0–79 TDC). Normal sacral and ischial skin is approximately 37 TDC[33].

Pressure Ulcer Assessment

As previously reported, visual assessment agreement for PrI presence was 99% for PrI stage 2 and greater across all locations with weighted Kappas ranging from 0.86 (DTI) to 0.96 (stage 4)[33]. All PrIs were assessed each week with the BWAT. Research staff training was carried out by a certified wound care nurse (BBJ) using photographs that depicted each BWAT characteristic and choice. Initial training required an hour. Additional 30-minute training sessions occurred at the beginning of each new cohort of NHs (generally every five months) throughout the five-year study period. PrIs more severe than Stage 1 were classified using the National Pressure Ulcer Advisory Panel's (NPUAP's) 2009 staging system[37].

Independent PrI assessments were obtained by two research staff five to 10 minutes apart, at randomly selected observation visits. Paired research staff were blinded to each other's ratings.

Critically important to the study was identifying unique, separate PrIs on the same anatomic location as this was necessary for determining PrI history and incidence. Using a Q-sort method, the summaries of the visual assessments over 16 weeks of each anatomic location for each participant were evaluated independently by 3 blinded raters (chosen from research staff). Each rater categorized each summary into one of four history patterns observed during the study: 1) PrI all weeks, 2) PrI less than three weeks, 3) PrI three or more weeks but not all weeks, and 4) PrI resolved, no damage for two weeks. For those with PrI resolved, subsequent damage was considered a new PrI. Agreement between the three blinded raters' categorizations of all PrI history summaries was 95% (data not shown).

Statistical Analysis

The nested nature of the study design required data to be summarized or selected at the level appropriate for the analysis (e.g., participant, PrI, or weekly assessment). Descriptive

statistics illustrate the demographic characteristics of the study participants. To standardize analyses used to examine BWAT score and individual characteristics by PrI stage and location, the BWAT assessment of the first observed PrI at an anatomic location (either at study start or developed during the study) was used. In total, 36 subsequent PrIs were excluded from these analyses as they occurred on anatomic sites with prior PrIs. This avoided correlated observations of the BWAT for a subsequent PrI at the same anatomic location. Mixed models (to account for repeated observations by participant) were used to evaluate BWAT differences across stage, anatomic location, skin tone, and ethnicity/racial groups (SPSS 24, SPSS Inc., Chicago, IL).

To examine total and individual item score reliability, BWAT data collected during any weekly PrI assessment for which there was data from two independent observers was used. Multiple approaches were used to examine reliability: correlation coefficients, measures of agreement. and intraclass correlation coefficients. Spearman rho correlation coefficients were used to provide a measure of general association between observers across a variety of participant (e.g., skin tone, ethnicity/race) and PrI characteristics (e.g., stage, anatomic location).

Three measures of agreement were calculated. First, simple exact percent agreement was examined. Next, we calculated a moderate estimate of percent agreement for scores plus or minus 1 point. The moderate estimate of percent agreement may more accurately reflect clinically meaningful agreement as a 1-point difference in score may not be differentially actionable. Finally, Weighted Fleiss Kappa statistics were used to calculate agreement, correcting for chance agreements. This is a more stringent evaluation and may also be a better reflection of agreement in clinical use.

To formally assess reliability, intraclass correlation coefficients (ICC) were used. Shrout-Fleiss ICC (ICC 2, 4) estimates were calculated based on single rating and mean-rating absolute-agreement, using a 2-way random-effects model computed with SAS[38,39]. The ICC allows examination of reliability estimates from more than 2 observers. The Shrout-Fleiss agreement statistics were calculated from the four research staff who completed 90% of the BWAT assessments. Both the single rating and the mean rating agreement statistic to summarize the pattern of reliability across the various combinations of raters are presented. The 2-way random-effects model was used because not all raters assessed all PrIs, thus a sample of raters is reported.

Observation level data were analyzed with generalized multinomial logistic modeling[40–43] using Stata version 13. Effects for participant and measurement period were included in all models to account for the correlated nature of the data[43]. Multinomial logistic regression models were conducted to examine ability of BWAT initial assessment to predict PrI history pattern using covariates of size (surface area) and stage. Only PrI history pattern (four-level categorical outcome) on the trunk and heels were analyzed due to inadequate numbers of PrI on other locations.

RESULTS

Longitudinal skin health and PrI measures were obtained from 142 participants with 270 PrIs (Figure 1). Fifty-four participants had PrI on study entry (n=70) and 88 participants developed PrI during the study (n=200). Participants had a mean age of 78±14.5 years and 62% were female (Table 1). The majority of participants were ethnic minorities: 28% African American, 14% Asian American, and 18% Hispanic. About two-thirds of the participants had light skin tone (64%). Most participants were at high risk for PrIs, with a mean Braden Scale score of 14±2.7, normal weight (mean Body Mass Index = 26±6.1) and were functionally dependent. Nearly half had a PrI on more than one anatomic location and 23% had multiple PrIs at the same anatomic location.

The majority of PrIs occurred at the sacral area (n=171) and heels (n=61) (Table 2). PrIs ranged in size from 0.02 cm^2 to 174 cm^2 (mean surface area 6.4 ±14.9 cm^2). Most PrIs were stage 1 (n=44) and stage 2 (n=136) (Table 1). While not significant, subepidermal moisture values at the PrI edge were highest for stage 4 PrIs (41.7±8.4 TDC) and lowest for DTIs (28.3±7.5 TDC). Necrotic tissue was present in only 12% (n=33) of PrIs, predominantly as yellow slough. Fifteen percent (n=40) of the PrIs exhibited purulent exudate. Edema was present in 4% (n=10) of PrIs, induration was present in 19% (n=51), and undermining presented in 8% (n=22). Most PrIs were present more than 3 weeks but resolved during the study (69%, n=186) while a small number were present all weeks (10%, n=26) or less than 3 weeks (21%, n=57).

Comparison of BWAT scores relative to PrI stage, anatomic location, skin tone group, ethnicity/racial group, and natural history

Irrespective of the timing of PrI development, initial BWAT score was significantly higher for stage 3, 4 and unstageable PrIs (p<0.001) (Table 2). The mean BWAT score for stage 1 PrIs was 15.3±2.9 and 20.2±4.4 for DTIs±4.4 compared to mean BWAT scores of 31.5±5.2 and 36.2±6.7 for stage 3 and 4 PrIs, respectively. For PrIs that developed during the study, significant differences existed between initial BWAT score across anatomic locations with higher BWAT scores at the sacrum compared to heels (22.5±6.4 vs. 17.6±3.9). There was no difference in initial BWAT score across anatomic locations for PrIs present on admission to the study.

Based on the mixed models, there were significant differences between initial BWAT scores across different skin tones and ethnicity/racial groups. BWAT scores were significantly lower for persons with light skin tones compared to those with medium and dark skin tones (light skin tones 21.8±7.1, vs. medium skin tones 24.2±6.9, and vs. dark skin tones 24.9±8.2; both p=0.008). African Americans demonstrated higher BWAT scores for all PrIs compared to all other ethnicity/racial groups (Table 2).

Initial BWAT scores were significantly higher for PrIs that persisted and were present all weeks of the study compared to PrIs present less than 3 weeks (33.7±8.9 vs 19.6±5.0, F(3) = 34.44, p<.001). Persistent PrI initial BWAT scores were also significantly higher than for PrIs present more than 3 weeks but resolved during the study (33.7±8.9 vs 21.6±6.8, F (3) =34.44, p<.001).

Subepidermal moisture values were associated with BWAT characteristics of wound edges, skin color surrounding the wound, and exudate type. Subepidermal moisture increased as wound edges advanced in severity (p=.06) from 34.7±10.0 TDC (n=40) for "Indistinct, diffuse, none clearly visible" to 41.7±10.1 TDC (n=19) for "Well-defined, not attached to base, rolled under, thickened" and 45.5± 0.7 TDC (n=2) for "Well-defined, fibrotic, scarred or hyperkeratotic" (some categories not listed for clarity). In contrast, subepidermal moisture decreased as skin color surrounding the wound increased in severity (p=.003) as follows: 38.4±9.0 TDC (n=75) for "Pink or normal for ethnic group" and 40.8±9.7 TDC (n=58) for "Bright red &/or blanches to touch" to 36.6±9.9 TDC (n=72) for "Dark red or purple &/or non-blanchable" and 33.0±11.4 TDC (n=39) for "Black or hyperpigmented" (some categories not listed for clarity). Exudate types of bloody and serous demonstrated higher subepidermal moisture values (40.4±9.1 TDC, n=33 and 40.9±9.8 TDC, n=18, respectively) compared with exudate types of serosanguineous (31.3±7.1 TDC, n=7) and purulent (35.4±11.1 TDC, n=20). BWAT characteristics of peripheral tissue induration and edema, necrotic tissue type, and undermining did not show sufficient numbers across all characteristic choices making analyses not possible.

Reliability estimates for BWAT use across PrI stage, anatomic location, skin tone, and ethnicity/racial group

The average time to complete a BWAT assessment was 1.6±1.5 minutes with a high value of 28 minutes for a large sacral stage 3 PrI. Most (98%) BWAT assessments were completed in less than 5 minutes. Inter-rater reliability estimates of the BWAT scores are shown in Table 3. BWAT reliability estimates were highest for full thickness PrIs (stage 3 and 4 PrIs, r=0.64 and r=0.69, respectively) and lowest for stage 2 PrIs (r=0.38). Reliability coefficients of BWAT assessments for PrIs involving skin discoloration were r= 0.48 for DTI and r=0.59 for stage 1 PrI. Reliability estimates were high across all anatomic locations, with the sacrum showing the highest reliability coefficient (r=0.92) and heels demonstrating lower reliability (r=0.83 and r=0.80 for left and right heel, respectively). Importantly, BWAT reliability estimates were similar for all ethnicity/racial groups and skin tones.

The BWAT score percent agreement was 89% using 1,161 paired observations with two observers across all PrIs on all locations (Table 4). Weighted Kappas of individual BWAT characteristics ranged from 0.46 (color surrounding the wound) to 0.79 (undermining). Two characteristics generated weighted Kappas above 0.70 (undermining and epithelialization) and two characteristics were below 0.50 (color surrounding the wound and induration). Agreement was higher for all characteristics when examined with moderate agreement (plus or minus one point) (Table 4).

Mean observer agreement ICCs were in the good to excellent range (0.60—0.74 for good and 0.75—1.00 for excellent)[44] indicating that the mean rating by observers had a good degree of agreement across wound characteristics (Table 4). Single observer agreement ICCs were mostly 0.40—0.59[44], indicating observers had a fair degree of agreement, suggesting most wound characteristics were rated reasonably similarly across observers. Three characteristics had single observer agreement ICCs below 0.40 suggesting poor agreement: periwound edema, necrotic tissue type and necrotic tissue amount. These values must be

interpreted with caution as minimal heterogeneity in presentation of these wound characteristics existed across the PrIs in this study.[44]

Use of BWAT scores as indicators of PrI healing or deterioration across PrI stage and anatomic locations.

The multinomial logistic regression models included covariates of surface area & stage and showed, initial BWAT score was a significant predictor of PrI duration (less than three weeks versus three weeks or greater) at the trunk (Wald chi^2 47.1, p <0.001; relative risk ratio (rrr) = 0.86, p = 0.005, 95% confidence interval (CI) = 0.081—0.707 for less than three weeks and rrr = 0.86, p =0.003, 95% CI = 0.339—0.943 for three weeks or greater). Initial BWAT score was significant for predicting resolved versus persistent heel PrI damage (Wald chi^2 83.75, p <0.001; rrr = 1.15, p = 0.01, 95% CI 1.03—1.30; rrr = 1.19, p < 0.001, 95% CI 1.09 —1.31, resolved and persistent, respectively).

DISCUSSION

In this report, two goals are accomplished. First, updated reliability estimates for the BWAT that demonstrate acceptability when used by non-health care research staff to monitor PrI status are provided. Second, wound characteristics across different PrI stage, anatomic location, race/ethnicity group, and varying skin tone are described. This enables examination of the ability of BWAT scores to predict natural history of PrIs.

The BWAT showed moderate to high reliability estimates (coefficients at or above r = 0.80) when used for assessment of PrIs of differing stages, at various anatomic locations, among different ethnicity/racial groups, and across skin tones. The reporting of reliability estimates across these categories is the first of such reports. The reliability estimates have clinical implications related to BWAT score interpretation and training in use of the BWAT. Specifically, lower reliability estimates for persons with medium and dark skin tones suggest the need for education targeted at identification of wound characteristics among persons of color[45]. The slightly lower reliability estimates for heel PrI may indicate a need for refining descriptors of macroscopic wound characteristics for PrI on heels[46–48]. The less than optimal reliability estimates for stage 2 PrIs and DTIs may be related to difficulty in assessing skin discoloration reflective of DTIs and problems differentiating moisture associated skin damage from stage 2 PrI[49,50].

The BWAT ICCs indicate lower levels of agreement, with single observer ICCs in the fair agreement range. The ICC values must be interpreted cautiously for some wound characteristics as there was minimal heterogeneity across the PrIs (notably peripheral tissue edema, necrotic tissue type and amount). When ICCs are applied to homogenous populations (as in these cases) the ICC will be low as it is strongly influenced by the variance of the trait in the sample in which it is assessed[44]. Chan and Lai[51] used the BWAT to examine wound healing in clean open wounds excluding wounds with edema and induration and their sample of 126 wounds also showed minimal heterogeneity across necrotic tissue type and undermining. Jesada and colleagues[32] also found lower reliability estimates for the wound characteristic of necrotic tissue type (kappa range 0.37—0.60) in

their study with certified wound care nurses assessing wounds at the bedside compared with experts assessing the same wounds based on photographs.

Reliability of wound assessment tools, and the BWAT in particular, has not been reported when used by non-health care workers and those with limited wound care experience. We demonstrate that adequate to good reliability can be obtained when persons with limited to no experience with wounds (such as non-health care research staff or beginning nursing students) use the BWAT with adequate training. The average time to complete the BWAT was only 1.6 minutes with 95% of assessments completed within 5 minutes. The use by persons with limited to no experience in wound care and the quick assessment times suggest the tool is practical for clinical use. The reliability estimates reported in this study may be higher when the BWAT is used by experienced wound care nurses and providers.

Not surprisingly, initial BWAT scores differ significantly by PrI stage. Higher BWAT scores were observed for more severe stage PrI. Higher scores would be expected because as skin and tissue damage increases, visual observation of more severe wound characteristics is available. These results are very similar to an early smaller study with the precursor to the BWAT where more severe Stage 3 and 4 PrI BWAT scores were significantly higher than less severe stage 1 and 2 PrI scores (31.8 versus 23.3, p < 0.001)[9]. Compared to the earlier study results, in this study, similar BWAT scores for stage 2 PrI and stage 3 PrI but higher BWAT scores for stage 4 PrI (mean BWAT score 36.2) are reported.

BWAT scores differed by anatomic location, a somewhat surprising finding. BWAT scores were higher for PrIs located on the trunk compared to heel PrIs. This may be related to differences in anatomic structure of the heel[46,47]. The anatomy of the heel may not allow for development and visualization of all the macroscopic wound characteristics evaluated by the BWAT characteristics.

BWAT scores were higher for African Americans and for persons with medium and dark skin tones. The higher scores may be due to difficulty in observing some of the macroscopic wound characteristics on persons with dark skin tones. For example, the characteristic "skin color surrounding the wound" evaluates skin discoloration, which is problematic in dark skin tones[45].

Subepidermal moisture was higher for stage 2, 3 and 4 PrIs (approximately 41 TDC), compared to subepidermal moisture at normal sacral skin (37 TDC)[33] suggesting some level of edema at the wound edge. This was not true for all heel PrI subepidermal moisture values. Subepidermal moisture for stage 2 heel PrIs was 34 TDC which is higher than subepidermal moisture at normal heel skin (29 TDC)[48]. Subepidermal moisture values for Stage 3, DTI, and unstageable heel PrIs (21, 27, and 28 TDC, respectively) were all lower than subepidermal moisture at normal heel skin. Again, these differences may be related to the anatomic structure of the heel[46,47]. Subepidermal moisture varied as would be expected across several characteritics on the BWAT. Subepidermal moisture values increased as wound edge increased in severity and with exudate types of bloody and serous. As skin color surrounding the wound increased in severity (typically demonstrating more tissue destruction below the skin surface), subepidermal moisture values also decreased. This is

similar to findings using subepidermal moisture values to detect PrI where values have been reported as lower for deep tissue injury.[48]

Few PrIs healed during the study and this limited ability to examine BWAT scores as a predictor of wound healing. The BWAT score was an indicator of the natural history pattern for PrIs occurring on the trunk and heels. In addition, BWAT score predicted PrI duration of greater than three weeks compared to those less than three weeks at the trunk and predicted resolved versus persistent heel PrIs. Others have shown that wound assessment instruments can be useful as predictors of wound healing[52,53]. These findings with the BWAT should be evaluated in a larger population with sufficient time to allow for PrI healing.

The calculation and reporting of multiple measures of reliability provides a realistic and robust examination of the BWAT when used by non-health care providers and non-wound care experts. The findings confirm the utility of the BWAT for reliably assessing PrI progress and healing progress and suggest several areas for improved training and education in wound assessment.

Acknowledgements

References:

1. Sen CK, Gordillo GM, Roy S, Kirsner R, Lambert L, Hunt TK, Gottrup F, Gurtner GC, Longaker MT. Human skin wounds: A major and snowballing threat to public health and the economy. Wound Rep Reg 2009 17(6): 763–771.

2. European Pressure Ulcer Advisory Panel (EPUAP), National Pressure Ulcer Advisory Panel (NPUAP) and the Pan-Pacific Alliance (PPA) International Guidelines, 2014.

3. AMDA—The Society for Post Acute and Long Term Care Medicine. Pressure Ulcers and Other Wounds in the Post-Acute and Long-Term Care Setting Clinical Practice Guideline. Columbia, MD: AMDA 2017.

4. Gould L, Stuntz M, Giovannelli M, Ahmad A, Aslam R, Mullen-Fortino M Whitney J, Calhoun J, Kirsner RS, Gordino GM. Wound healing society 2015 update on guidelines for pressure ulcers. Wound Rep Reg 2016; 24: 145–162.

5. Bates-Jensen BM, Vredevoe D, Brecht ML. Validity and reliability of the Pressure Sore Status Tool. Decubitus, 1992;5(6):20–28. [PubMed: 1489512]

6. Harris CL, Bates-Jensen B, Parslow N, Raizman R, Singh M, Ketchen R The Bates-Jensen Wound Assessment Tool (BWAT)© Pictorial Guide validation project. J Wound Ostomy Continence Nurs. 2010; 37(3):253–259. [PubMed: 20386331]

7. Bates-Jensen B & McNees P (1995). Toward an intelligent wound assessment system. Ostomy/ Wound Management Supplement, (41)7A: 80–87.

8. Bates-Jensen B (1997). The Pressure Sore Status Tool a few thousand assessments later. Advances in Wound Care, 10(5), 65–73. [PubMed: 9362584]

9. Bates-Jensen B, McNees P. The Wound Intelligence System: early issues and findings from multi-site tests. Ostomy Wound Manage. 1996 Nov-Dec;42(10A Suppl):53S–61S. [PubMed: 9397883]

10. Streiner DL, Kottner J. Recommendations for reporting the results of studies of instrument and scale development and testing. J Adv Nurs. 2014 9;70(9):1970–9. [PubMed: 24684713]

11. Matsui Y, Furue M, Sanada H, Tachibana T, Nakayama T, Sugama J, Furuta K, Tachi M, Tokunaga K, Miyachi Y. Development of the DESIGN-R with an observational study: an absolute evaluation

tool for monitoring pressure ulcer wound healing. Wound Rep Regen. 2011 May-Jun;19(3):309–15.

12. Thompson N, Gordey L, Bowles H, Parslow N, Houghton P. Reliability and validity of the revised photographic wound assessment tool on digital images taken of various types of chronic wounds. Adv Skin Wound Care. 2013 8;26(8):360–73. doi: [PubMed: 23860221]

13. Houghton PE, Kincaid CB, Campbell KE, Woodbury MG, Keast DH. Photographic assessment of the appearance of chronic pressure and leg ulcers. Ostomy Wound Manage. 2000 4;46(4):20–6, 28–30. [PubMed: 10788924]

14. Thomason SS, Luther SL, Powell-Cope BM, Harrow JJ, Palacios P Validity and reliability of a pressure ulcer monitoring tool for persons with spinal cord impairment. J of Spinal Cord Med. 2014 37(3): 317–327. [PubMed: 24621044]

15. Thomason SS, Powell-Cope G, Peterson MJ, Guihan M, Wallen ES, Olney CM, Bates-Jensen B A multi-site quality improvement project to standardize the assessment of pressure ulcer healing in veterans with spinal cord injuries/disorders. Adv Skin Wound Care. 2016 6; 29(6):269–76. [PubMed: 27171255]

16. Arisandi D Oe M, Yotsu Roselyne, Matsumoto R, Ogai M, Nakagami K, Tamaki G, Suriadi T, Sanada H, Sugama J. Evaluation of validity of the new diabetic foot ulcer assessment scale in Indonesia. Wound Rep Reg. 2016 24(5): 876–884.

17. Takahashi J, Nakae K, Myagawa M, Yokota O, Fujiki Y, Ide M, Nishida S, Aoki H, Aoki T. Plastic wrap as a dressing material to treat stage III-IV pressure ulcers in the inflammatory phase: a randomized controlled trial.

18. Tao Q, Ren J, Ji Z, Wang B, Zheng Y, Li J. Continuous topical irrigation for severely infected wound healing. J Surg Res. 2015 10;198(2):535–40. [PubMed: 25958167]

19. Porso M, Loreti S, Nusca SM, Luziatelli S, Caccia D, Taborri G, Trischitta D, Taurino M, Padua L, Saraceni VM, Vulpiani MC, Vetrano M. Defocused shock wave therapy for chronic soft tissue wounds in the lower limbs: a pilot study. Ultrasound Med Biol. 2017 1;43(1):362–369. [PubMed: 27745716]

20. Bellingeri A, Falciani F, Traspedini P, Moscatelli A, Russo A, Tino G, Chiari P, Peghetti A Effect of a wound cleansing solution on wound bed preparation and inflammation in chronic wounds: a single-blind RCT. J Wound Care. 2016 3;25(3):160, 162–6, 168. [PubMed: 26947697]

21. McCallon SK, Frilot C A retrospective study of the effects of clostridial collagenase ointment and negative pressure wound therapy for the treatment of chronic pressure ulcers. Wounds 2015 3;27(3):44–53. [PubMed: 25786076]

22. de Leon JM, Barnes S, Nagel M, Fudge M, Lucius A, Garcia B. Cost-effectiveness of negative pressure wound therapy for postsurgical patients in long-term acute care. Adv Skin Wound Care. 2009 3;22(3):122–7. [PubMed: 19247013]

23. Litwiniuk M, Bikowska B, Niderla-Bieli ska J, Jó wiak J, Kami ski A, Skopi ski P, Grzela T. Potential role of metalloproteinase inhibitors from radiation-sterilized amnion dressings in the healing of venous leg ulcers. Mol Med Rep. 2012 10;6(4):723–8. [PubMed: 22798012]

24. Sipahi S, Gungor O, Gunduz M, Cilci M, Demirci MC, Tamer A. The effect of oral supplementation with a combination of beta-hydroxy-beta-methylbutyrate, arginine and glutamine on wound healing: a retrospective analysis of diabetic haemodialysis patients. BMC Nephrol. 2013 1 12;14:8. [PubMed: 23311725]

25. Chan LN, Lai CK, The effect of patient education with telephone follow-up on wound healing in adult patients wth clean wounds: a randomized controlled trial. J Wound Ostomy Continence Nurs. 2014; 41(4): 345–55. [PubMed: 24988513]

26. Clark F, Pyatak EA, Carlson M, Blanche EI, Vigen C, Hay J, Mallinson T, Blanchard J, Unger JB, Garber SL, Diaz J, Florindez LI, Atkins M, Rubayi S, Azen SP; PUPS Study Group. Implementing trials of complex interventions in community settings: the USC-Rancho Los Amigos pressure ulcer prevention study (PUPS). Clin Trials. 2014 4;11(2):218–29. [PubMed: 24577972]

27. Paul J Characteristics of chronic wounds that itch. Adv Skin Wound Care. 2013 7;26(7):320–32; quiz 333–4. [PubMed: 23777882]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

28. Karahan A, Abbasoglu A, Isik SA Cevik B, Saltan C, Elbas NO, Yall A. Factors affecting wound healing in individuals with pressure ulcers: a retrospective study. Ostomy Wound Manage. 2018; 64(2): 32–39.

29. Guihan M, Bombardier CH, Ehde DM, Rapacki LM, Rogers T, Bates-Jensen B, Thomas FP, Parachuri R, Holmes SA. A Multi-Component Intervention to Improve Skin Care Behaviors And Prevent Recurrence In Veterans Hospitalized For Severe Pressure Ulcers. Arch Phys Med Rehabil. 2014; 1 29.

30. Carlson M, Vigen CL, Rubayi S, Blanche EI, Blanchard J, Atkins M, Bates-Jensen B, Garber SL, Pyatak EA, Diaz J, Florindez LI, Hay JW, Mallinson T, Unger JB, Azen SP, Scott M, Cogan A, Clark F. Lifestyle intervention for adults with spinal cord injury: Results of the USC-RLANRC Pressure Ulcer Prevention Study. J Spinal Cord Med. 2017 4 17:1–18.

31. Ebid AA, El-Kafy EM, Alayat MS. Effect of pulsed Nd:YAG laser in the treatment of neuropathic foot ulcers in children with spina bifida: a randomized controlled study. Photomed Laser Surg. 2013 12;31(12):565–70. [PubMed: 24160904]

32. Jesada EC, Warren JI, Goodman D, Iliuta RW, Thurkauf G, McLaughlin MK, Johnson JE, Strassner L. Staging and defining characteristics of pressure ulcers using photographs by staff nurses in acute care settings. Wound Ostomy Continence Nurs. 2013; 40(2): 150–6.

33. Bates-Jensen BM, McCreath HE, Patlan A. Subepidermal Moisture Detection of Pressure Induced Tissue Damage on the Trunk: The Pressure Ulcer Detection (PUD) Study Outcomes. Wound Repair Regen. 2017 25(3): 502–511. [PubMed: 28494507]

34. Bergstrom N, Braden B A prospective study of pressure sore risk among institutionalized elderly. J Am Geriatr Soc, 1992; 40(8): 747–58. [PubMed: 1634717]

35. Miller N1, Frankenfield D, Lehman E, Maguire M, Schirm V. Predicting pressure ulcer development in clinical practice: evaluation of Braden Scale scores and nutrition parameters. J Wound Ostomy Continence Nurs. 2016; 43(2):133–9. [PubMed: 26680628]

36. McCreath HE, Bates-Jensen BM, Nakagami G, Patlan A, Booth H, Connolly D, Truong C, Woldai A Use of Munsell Color Charts to objectively measure skin color in nursing home residents at risk for pressure ulcer development. J Adv Nurs. 2016; 72(9): 2077–85. [PubMed: 27062396]

37. European Pressure Ulcer Advisory Panel and National Pressure Ulcer Advisory Panel. Prevention and treatment of pressure ulcers: quick reference guide. Washington DC: National Pressure Ulcer Advisory Panel; 2009.

38. Shrout PE and Fleiss JL. IntraclassCorrelations: Uses in Assessing Rater Reliability. Psychological Bulletin 1979; 86:2, 420–428 [PubMed: 18839484]

39. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability. J Chiropr Med. 2016; 15(2): 155–163. [PubMed: 27330520]

40. Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression, 3rd Edition Wiley, New York, 2013 (See Section 8.1).

41. Long JS, Freese J Regression Models for Categorical Dependent Variables Using STATA. STATA Press College Station, TX: 2014.

42. Bender R and Grouven U. Using binary logistic regression models for ordinal data with non-proportional odds. J Clin Epidemiol 1998; 51 (10): 809–816. [PubMed: 9762873]

43. StataCorp. 2013 Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.

44. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment. 1994;6(4):284–290 38–42

45. Rosen J, Mittal V, Degenholtz H, Castle H, Mulsant BH, Nace D, Rubin FH. Pressure ulcer prevention in black and white nursing home residents: a QI initiative of enhanced ability, incentives, and management feedback. Adv Skin Wound Care 2006; 19(5): 262–8. [PubMed: 16732072]

46. Arao H1, Shimada T, Hagisawa S, Ferguson-Pell M. Morphological characteristics of the human skin over posterior aspect of heel in the context of pressure ulcer development. J Tissue Viability. 2013; 22(2):42–51. [PubMed: 23561224]

47. Cichowitz A, Pan WR, Ashton M. The heel: anatomy, blood supply, and the pathophysiology of pressure ulcers. Ann Plast Surg. 2009; 62(4):423–429. [PubMed: 19325351]

48. Bates-Jensen BM, McCreath HE, Nakagami G Patlan, A. Subepidermal moisture detection of heel pressure injury: the Pressure Ulcer Detection study outcomes. International Wound Journal. 2017; 15(2): 297–309. [PubMed: 29250926]

49. Beele H, Smet S, Van Damme N, Beeckman D Incontinence-Associated Dermatitis: Pathogenesis, Contributing Factors, Prevention and Management Options. Drugs Aging, 2018; 35(1): 1–10. [PubMed: 29243033]

50. Bruce TA, Shever LL, Tschannen D, Gombert J Reliability of pressure ulcer staging: a review of literature and 1 institution's strategy. Crit Care Nurs Q. 2012; 35(1); 85–101. [PubMed: 22157495]

51. Chan LN, Lai CKY. The effect of patient education with telephone follow-up on wound healing in adult patients with clean wounds. J Wound Ostomy Continence Nurs. 2014; 41(4): 345–355. [PubMed: 24988513]

52. Sanada H, Iizaka S, Matsui Y, Furue M, Tachibana T, Nakayama T, Sugama J, Furuta K, Tachi M, Tokunaga K, Miyachi Y; Scientific Education Committee of the Japanese Society of Pressure Ulcers. Clinical wound assessment using DESIGN-R total score can predict pressure ulcer healing: pooled analysis from two multicenter cohort studies. Wound Repair Regen. 2011 Sep-Oct;19(5): 559–67. [PubMed: 22092794]

53. Gunes UY. A prospective study evaluating the Pressure Ulcer Scale for Healing (PUSH tool) to assess stage II, stage III, and stage IV pressure ulcers. Ostomy Wound Man. 2009; 55(5): 48–52.
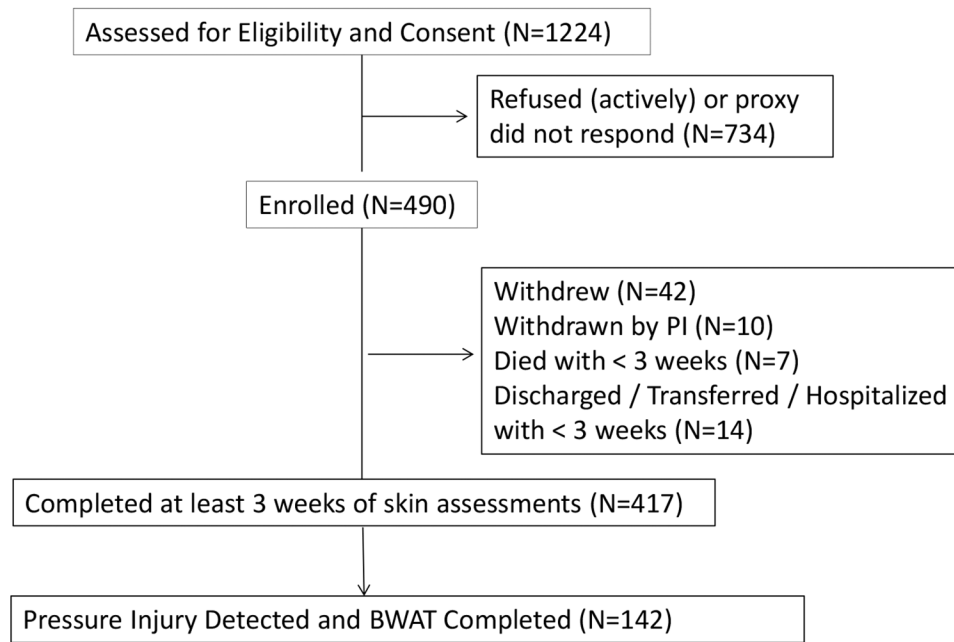
Figure 1:
Flow of Participants through Study

Table 1

Demographic characteristics of participants (n=142) and pressure injuries with BWAT data (n=270)

Characteristics	Mean (SD) *or* N (%)
Age	78.06 (14.52)
Female	88 (62.0)
Race/Ethnicity	
Asian American	20 (14.1)
African American	41 (28.1)
Hispanic	25 (17.6)
Caucasian	57 (40.1)
Munsell Skin Tone [*]	
Dark (2.5––4 values)	21 (14.8)
Medium (5—6 values)	27 (19.0)
Light (7–8 values)	91 (64.1)
Braden Score for Predicting Pressure Sore Risk	14.09 (2.72)
Body Mass Index (N=86)	25.71 (6.16)
MDS [**] Bed Mobility-Extensive assistance + Total dependence	112 (80.3)
MDS Transfer assessment-Extensive assistance + Total dependence	125 (89.4)
Diabetes Mellitus	63 (44.4)
Peripheral Vascular Disease	21 (14.8)
Pressure injury on more than one anatomic location	67 (47.2)
Multiple pressure injury at the same anatomic location	33 (23.2)
Pressure Injury size in cm^2	6.4 (14.9)
Initial Subepidermal Moisture value by stage [#] and anatomic location	Trunk Heels
Stage 1	38.73 (8.61) 28.66 (8.48)
Stage 2	41.04 (8.62) 33.80 (9.47)
Stage 3	41.40 (8.96) 20.75 (1.20)
Stage 4	41.59 (9.22) ----
Deep Tissue Injury	35.47 (6.22) 27.03 (6.92)
Unstageable	---- 27.82 (7.30)

[*] Munsell skin tone range 2.5—8.0 where higher values indicate lighter skin tones.

[**] MDS Minimum Data Set; BMI only available for those with MDS 3.0

[#] Staging according to NPUAP 2009 guidelines

Table 2

Initial Bates-Jensen Wound Assessment Tool scores for initial observation by stage, location, skin tone and race/ethnicity for those pressure injuries present on admission to the study and acquired during the study.

Wound Characteristics-All Initial assessments	BWAT Total Score Mean (SD) N (%)		
	All Pus N=270	PU POA to study N=70	Acquired PU during study N=200
By Stage [*]			
Stage 1	15.29 (2.88) 44 (16.3)	15.33 (0.58) 3 (4.3)	15.29 (2.98) 41 (19.2)
Stage 2	21.48 (3.20) 136 (50.4)	21.37 (2.52) 24 (34.3)	21.51 (3.34) 112 (56.0)
Stage 3	31.55 (5.17) 29 (10.7)	30.71 (4.94) 14 (20.0)	32.33 (5.43) 15 (7.5)
Stage 4	36.2 (6.72) 20 (7.4)	35.36 (7.57) 14 (20.0)	38.17 (3.97) 6 (3.0)
Unstageable	39.2 (3.42) 5 (1.9)	39.20 (3.42) 5 (7.1)	--
Deep Tissue Injury	20.17 (4.42) 36 (13.3)	20.60 (4.62) 10 (14.3)	20.00 (4.42) 26 (13.0)
By Location [**]			
Trunk	23.86 (7.23) 171 (63.3)	27.6 (8.17) 46 (65.7)	22.50 (6.36) 125 (62.5)
Ischials	22.39 (6.42) 38 (14.1)	24.14 (7.43) 7 (10.0)	22.00 (6.24) 31 (15.5)
Heels	20.00 (7.23) 61 (22.6)	26.41 (9.78) 17 (24.3)	17.64 (3.93) 44 (22.0)
By Munsell Skin Tone [***]			
Light (values 7–8)	21.8 (7.1) 169 (62.6)	20.9 (6.3) 129 (64.5)	24.5 (8.7) 40 (57.1)
Medium (values 5–6)	24.2 (6.9) 61 (22.6)	22.2 (5.4) 46 (23.0)	30.4 (7.4) 15 (21.4)
Dark (values 2.5–4)	24.9 (8.2) 32 (11.8)	21.7 (7.5) 19 (9.5)	29.5 (7.0) 13 (18.6)
By Race/Ethnicity [****]			
African American	25.1 (8.0) 69 (25.6)	22.3 (6.6) 45 (65.2)	30.4 (7.7) 24 (34.8)
Asian American	23.7 (7.7) 55 (20.4)	23.5 (7.6) 46 (23.0)	24.6 (8.6) 9 (16.4)
Caucasian	21.3 (6.1) 93 (34.4)	20.4 (4.8) 72 (36.0)	24.6 (8.7) 21 (30.0)
Hispanic	21.4 (7.0) 53 (19.6)	19.4 (5.3) 37 (18.5)	26.2 (8.2) 16 (22.9)

Note: Mixed models used to account for multiple PrIs by participant on different anatomic locations.

[*] Main effect for stage significant (F(5,122)=91.15, p <0.0001)

[**] Interaction of location and POA significant (F(2,127)=3.07, p=0.05)

[***] Main effect for skin tone significant (F(2,124)=5.09, p=0.008)

*****
Main effect for race/ethncity significant (F(3,128)=3.08, p=0.03)

Table 3

Bates-Jensen Wound Assessment Tool total score for all observations inter-rater agreement across pressure ulcer stage, anatomic site, race/ethnicity groups and Munsell skin tone

Item	Spearman Rho correlation coefficient (N observations)
PU SEVERITY (n=940 observations) [*]	
Stage 1	0.59 (88)
Stage 2	0.38 (229)
Stage 3	0.64 (175)
Stage 4	0.69 (208)
Unstageable	0.57 (39)
Deep Tissue Injury	0.48 (201)
ANATOMIC LOCATION (n=1,093 observations)	
Sacral & coccyx	0.92 (332)
R buttock/L buttock	0.87 (149) / 0.80 (132)
R ischial/L ischial	0.94 (82) / 0.84 (45)
R heel/L heel	0.83 (174) / 0.80 (179)
RACE/ETHNICITY (n=1,155 observations)	
Asian American	0.89 (231)
African American	0.88 (352)
Hispanic non-caucasian	0.87 (260)
Caucasian	0.91 (312)
MUNSELL SKIN TONE (n=1,128 observations)	
Dark (2.5––4 values)	0.84 (175)
Medium (5—6 values)	0.85 (311)
Light (7–8 values)	0.92 (642)

[*] Differing number of observations due to missing data.

Table 4

Bates-Jensen Wound Assessment Tool item inter-rater reliability estimates (n= 1,161 observations)

BWAT Item	Percent Agreement Exact[*]	Percent Agreement Moderate[**]	Weighted Fleiss Kappa	Intraclass Correlation Coefficient single observer (mean rating of observers)[#]
Depth	73.21	90.19	0.66	0.47 (0.78)
Edges	59.93	89.94	0.62	0.38 (0.71)
Undermining	86.40	96.52	0.79	0.62 (0.81)
Necrotic Tissue Type	83.35	93.56	0.61	0.38 (0.71)
Necrotic Tissue Amount	81.40	90.98	0.56	0.37 (0.70)
Exudate Type	79.78	87.03	0.68	0.44 (0.76)
Exudate Amount	71.20	90.97	0.64	0.42 (0.75)
Skin Color Surrounding Wound	54.36	73.47	0.46	0.54 (0.83)
Peripheral Tissue Edema	91.77	97.90	0.51	0.08 (0.26)
Peripheral Tissue Induration	72.40	92.26	0.49	0.42 (0.75)
Granulation Tissue	65.30	82.36	0.57	0.47 (0.78)
Epithelialization	76.82	85.82	0.74	0.50 (0.80)
TOTAL Score (range 9–65)				**0.58 (0.84)**

[*] exact agreement between scores

[**] agreement based on scores +/− 1 point

[#] Shrout-Fleiss ICC (2,4) 2-way random effects model with 4 raters; single rater agreement (mean rater agreement)