Development and Pilot Analysis of the Bruise Visibility Scale

Katherine N. Scafide, PhD, RN[1] (iD),
Ghareeb Bahari, PhD, RN[2] (iD),
Nesibe S. Kutahyalioglu, MSN, RN[1] (iD),
Mona Mohammadifirouzeh, MSN, RN[1] (iD), and
Susan M. Senko, DNP, APRN, AGPCNP-BC[1]

Abstract

Introduction: The accuracy of assessing and documenting injuries is crucial to facilitate ongoing clinical care and forensic referrals for victims of violence. The purpose of this cross-sectional, pilot study was to evaluate the inter-rater reliability and criterion validity of a newly developed Bruise Visibility Scale (BVS).

Methods: The instrument was administered to a diverse sample ($n = 30$) with existing bruises. Bruises were assessed under fluorescent lighting typical of an examination room by three raters who were randomly selected from a pool of eight experienced clinical nurses. Colorimetry values of the bruise and surrounding tissue were obtained using a spectrophotometer.

Results: The BVS demonstrated good single (ICC $= 0.71$, 95% CI $= 0.54 − 0.84$) and average agreement (ICC $= 0.88$, 95% CI $= 0.78 − 0.94$) between raters. A significant, positive moderate correlation was found between mean BVS scores and overall color difference between the bruise and surrounding skin (Pearson's $r = 0.614$, $p < 0.001$).

Conclusion: With further research, the BVS has the potential to be a reliable and valid tool for documenting the degree of clarity in bruise appearance.

Keywords

bruise, documentation, visibility, reliability, validity

Introduction

From intimate partner violence to child abuse, bruising is the most common manifestation of physical assault across the lifespan (Kemp et al., 2014; Savall et al., 2017). Clinically, the appropriate recognition and documentation of this type of injury is crucial to facilitate ongoing treatment. Strangulation, for example, often results in subtle injury with insidious effects (De Boos, 2019). The accurate assessment and documentation of bruises may also support referrals for further forensic investigations. As such, identifying multiple bruises in various stages of healing (e.g., from dark to faded) may be an indication of chronic abuse (Tsokos, 2015). Thus, nurses would benefit from a valid and reliable tool to support the documentation of bruises in the clinical setting.

Cutaneous bruising is characterized by discoloration resulting from blunt or crushing force trauma causing damage to underlying blood vessels (Lyons & Ousley, 2014). The injury is distinguishable from ecchymosis in which blood leakage originates from other mechanisms (e.g., punctures). Recommended standards for clinically assessing and documenting a bruise include recording its

[1]College of Health and Human Services, George Mason University, Fairfax, Virginia, United States
[2]College of Nursing, King Saud University, Riyadh, Kingdom of Saudi Arabia

Corresponding Author:
Katherine N. Scafide, College of Health and Human Services, George Mason University, 4400 University Drive, Mail Stop Number 3C4, Fairfax, VA 22030, United States.
Email: kscafide@gmu.edu

size (i.e., length and width), appearance (i.e., shape, pattern, location, color, margins), and presence of pain or induration (Lyons & Ousley, 2014). However, documenting the variable appearance of these injuries using qualitative descriptors is subjective. To date, documenting bruise visibility is typically limited to descriptive terms, such as "faint," "difficult to see," "noticeable," or "obvious." Although a body of research exists on instrumentation used to accurately measure bruise size and color (Cosman et al., 2015; Lawson et al., 2015; Scafide et al., 2016), quantitative scales for assessing bruise visibility in vivo are lacking. Only Limmen et al. (2013) has presented an ordinal, 5-point instrument (*not/barely/moderate/sufficient/good*) for measuring bruise visibility, but did not formally evaluate its reliability or validity.

Scale Development

To address the gap in available bruise appearance metrics, the Bruise Visibility Scale (BVS) was created by the first author (KNS), based on existing research on scales used to assess scars (Vercelli et al., 2015). The instrument was developed and revised as a result of several preliminary analyses. The original version was created as a Visual Analogue Scale (VAS) ranging in values from zero-to-ten with accompanying major (0–10) and minor (tenth) tick marks. The observer responded by placing a mark on the number line. The scale's three included labels (0 = Not visible, 5 = Moderately visible, 10 = Clearly visible) were adapted from Limmen et al. (2013). Since bruise visibility is defined in terms of clarity of perception, one of the anchors was changed from "good visibility" to "clearly visible."

To evaluate the new instrument's face validity, a purposive sample of eight experienced (>10 years) nurse clinicians individually reviewed the BVS and provided feedback using a mixed methods approach. The clinicians were asked to examine the scale's relevance and sufficiency in capturing the assessment of bruise visibility. The nurses unanimously agreed on the scale's ability to measure the construct and affirmed its adaptation from Limmen et al. as being appropriate. Reviewers were then instructed to apply the scale to two, published (Limmen et al., 2013), digital color photographs of the same bruise, one using standard examination lighting (fluorescent), the other using alternate lighting. As anticipated, the BVS values reported were higher for the alternate light image in which the injuries were more contrasted, with a mean difference of 5.8 (SD = 1). Collectively, the results supported the scale as being at a minimum superficially adequate in measuring the study construct.

Once face validity was established, initial psychometric testing was performed in order to assess test-retest reliability (Scafide & Sheridan, 2015). Since bruises change as they heal, test-retest reliability could only be evaluated using photographs. Thirty practicing emergency department and forensic nurses were asked to apply the BVS to 10 printed images of bruises that included a variety of skin colors. Approximately 30 days later, twenty of the nurses from the initial sample (67% response) repeated the assessment on identical images presented in a different order. Based on paired t-test results, the scale was stable without significant difference between the two measures across all 10 images ($p > 0.05$).

Qualitative feedback received from the test-retest reliability study noted the added subjectivity of a scale increased with multiple value options. Additionally, the construct of bruise visibility was refined based on the assumption the bruise had to be detectable in order to assess its visibility. Thus, the "zero" value was dropped. As a result, the BVS was modified to a 1–5 scale (see Figure 1). Three reference photographs were added to the VAS to demonstrate bruise visibility values on both light and dark skin. The images were selected from the test-retest reliability study based on the raters' mean BVS scores.

Study Purpose

The purpose of this cross-sectional, pilot study was to examine the inter-rater reliability and criterion validity of the newly developed BVS for use on existing, visible bruises. By evaluating the instrument's psychometric properties in the context of its earlier development, the goal was to create an instrument for future clinical and research applications to improve the assessment and documentation of bruises among victims of violence.

Methods

A convenience sample of 30 healthy adults with at least one existing, accidental bruise were recruited from a college campus. To be eligible, bruises had to be visible, easily accessible on the arm or leg, and not in close proximity to other bruises (to avoid comparison). Only one bruise was examined per participant. Additionally, eight nurses were also recruited as raters from faculty and graduate students at a nursing school. To be eligible, the nurses had to have at least one year of clinical experience, have corrected vision of at least 20/30 determined by a Snellen Chart, and screen negative for color blindness using an Ishihara test. Glasses were worn at all times by three raters and for reading/distance by two. Raters' clinical nursing experience ranged from 1 to 43 years (mean = 17 yrs., SD = 17.7). Prior to study engagement, each nurse received standardized training on the BVS and its application which included a 6-minute video instruction followed by a practice session guided by the
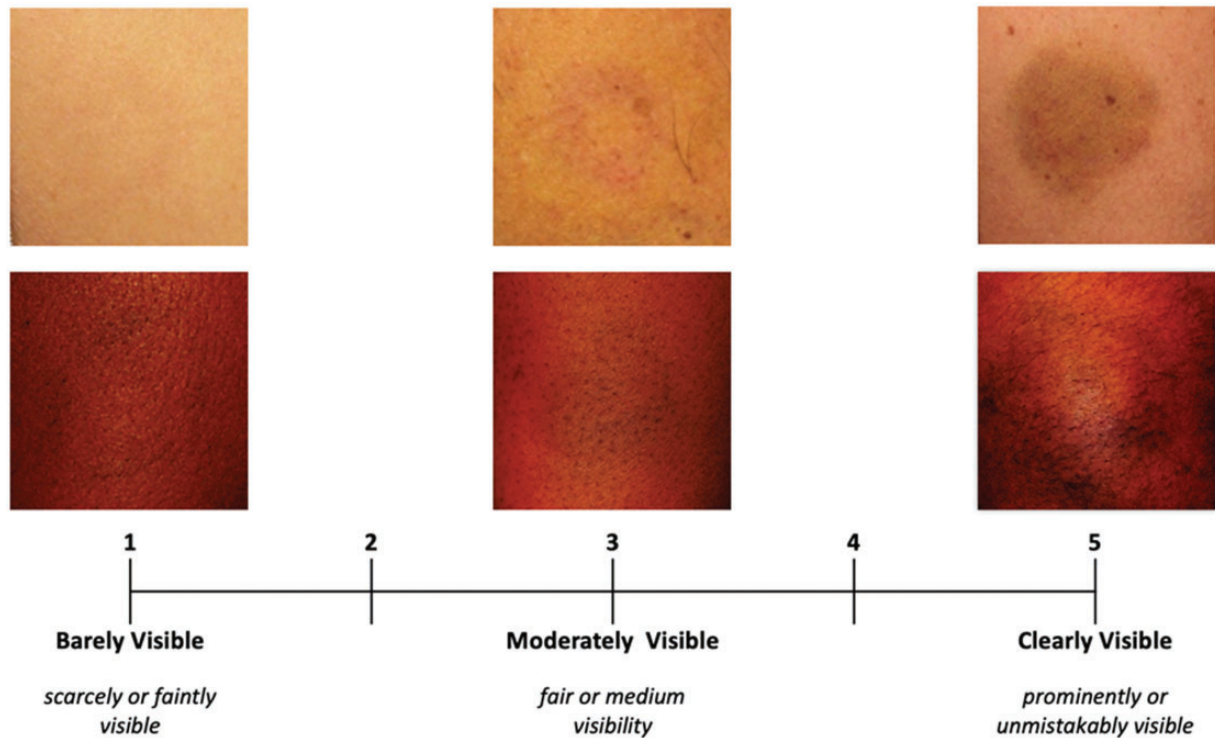
Figure 1. Bruise Visibility Scale (BVS).

primary researcher (KNS). Ethical approval for the study was obtained from the George Mason University Institutional Review Board. Both participants and raters completed an informed consent and were compensated for their time.

Data collection was performed in an office setting using fluorescent lighting typical of an examination room. The bruise was exposed and size measured. To assess the criterion validity of the BVS, the team compared the scale's values to the contrast in color difference between the bruise and surrounding skin. The color was measured using the Minolta® CM-600D spectrophotometer (Konica Minolta, Osaka, Japan). The handheld device is placed gently on the skin's surface and white light is projected through an 8mm aperture. The color of the reflected light is quantified using the Commission Internationale d'Eclairage (CIE) L*a*b* color space (Kruschwitz, 2018). In this system, color is defined using three values: L* indicating the concentration of white (100) vs. black (0); a* indicating the concentration of red ($+a*$) vs. green ($-a*$); and b* indicating the concentration of yellow ($+b*$) vs. green ($-b*$). The mean of three-color measurements was collected from the approximate bruise center and at three locations of non-discolored skin surrounding the injury. The overall color difference between the two mean measurements (bruise – skin) was calculated using the following Euclidian formula: $\Delta E*_{ab} = \sqrt{[(\Delta L*)^2 + (\Delta a*)^2 + (\Delta b*)^2]}$

(Kruschwitz, 2018). The threshold for perceivable color difference was a $\Delta E*_{ab} > 1$, with higher values being more visible (Kuehni & Marcus, 1979). Research has demonstrated excellent inter-rater and intra-rater reliability of bruise colorimetry measurements across all three L*a*b* parameters (ICC > 0.97) (Scafide et al., 2016). Calibration of the instrument was performed per manufacturer's recommendations.

For each bruise participant, three nurses were randomly selected from the pool of 8 raters to perform an independent bruise assessment. To avoid influencing the rater by drawing attention to the bruise, the primary researcher asked the rater to point to the area being assessed. The nurse indicated his or her response to the BVS directly on the instrument using Qualtrics, a secured, online survey platform (Qualtrics, Provo, UT). The nurses were provided a laminated card containing images and definitions for the BVS values as a reference (Figure 1). During data collection, the nurses were blinded to each other's ratings and the spectrophotometer readings.

Statistical analysis was performed using SPSS version 25 (IBM Corp, Armonk, NY). Descriptive statistics were used to examine sample characteristics. Two BVS observations by single raters were missing. SPSS missing values analysis indicated those values were missing at random and were replaced with the mean (Polit & Yang, 2016). To evaluate criterion validity, Pearson's r

was used to determine the relationship between the mean bruise visibility scores and bruise color values and overall color difference (ΔE^*_{ab}). Scatterplots were also examined to visually assess for relationships between variables. Intra-class correlation coefficients (ICC) and Cronbach's alpha were used to analyze inter-rater reliability of the BVS on 29 subjects. One participant had only one BVS observation and was eliminated from this portion of the analysis. The Shrout and Fleiss (1979) ICC [2] equation was used which includes two-way random effects model analysis of variance for absolute agreement. This model is appropriate because the nurse raters were randomly selected from a larger population of similar raters. Results of both single [2,1] and average [2,k] ICC analyses were presented (Shrout & Fleiss, 1979). A recommended minimum value of ICC for reliability is 0.70 (Polit & Yang, 2016).

Results

Table 1 presents the characteristics of the participants and their bruises. Most bruises assessed were located on the leg (66.7%) and only 31% knew the age of their injury. The number of BVS measurements and distribution of scores for each nurse rater are presented in Table 2. The mean difference between highest and lowest

Table 1. Characteristics of Bruise Participants (n=30).

Characteristics	n (%)	Mean (SD)
Sex		
Female	28 (93.3)	
Male	2 (6.7)	
Age (years)		24.5 (8.46)
Race		
Asian/Pacific Islander	5 (16.7)	
Black	3 (10)	
Caucasian/White	16 (53.3)	
Middle Eastern	4 (13.3)	
Multiracial	2 (6.7)	
Bruise location		
Right arm	4 (13.3)	
Left arm	6 (20)	
Right leg	12 (40)	
Left leg	8 (26.7)	
Bruise size (cm^2)		8.17 (10.79)

BVS observation within bruises was 0.93 (SD = 0.62). Based on the ICC analysis, the BVS demonstrated moderate single (ICC = 0.71, 95% CI = 0.54 – 0.84) and good average agreement (ICC = 0.88, 95% CI = 0.78 – 0.94) between raters. The reliability statistics showed a Cronbach's alpha of 0.88.
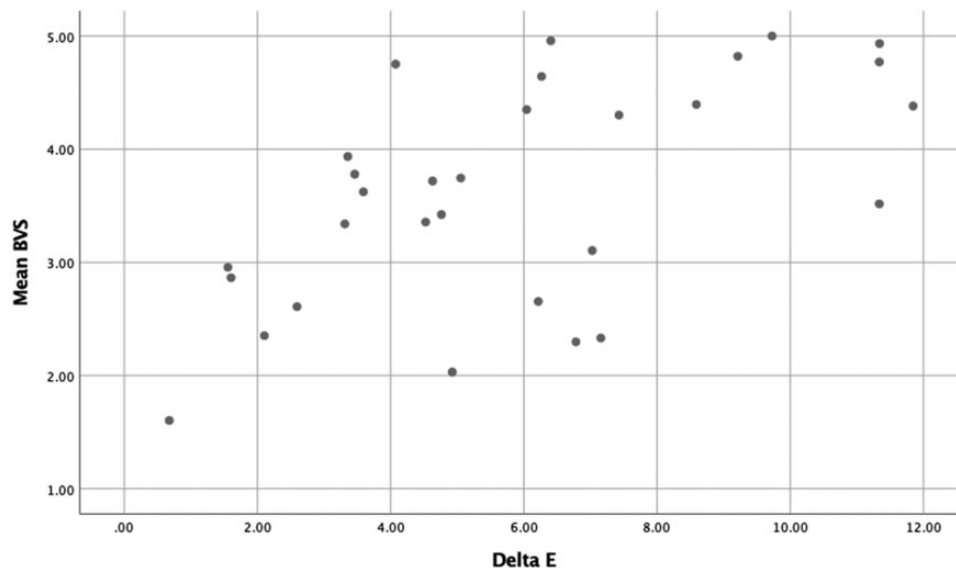
The difference in color between the bruise and surrounding skin (ΔE^*_{ab}) ranged from 0.67 to 11.84 with the average being 5.89 (SD = 3.14). There was a positive, moderate correlation between the mean BVS values and ΔE^*_{ab} (r = 0.57, p = < 0.001; see Figure 2). Further analysis found a greater contrast in lightness (black vs. white) as being negatively correlated with the BVS score (ΔL^*: r = −0.50, p = 0.005). However, specific color saturation was not associated with the scale (Δa^*: r = 0.20, p = 0.28; Δb^*: r = −0.26, p = 0.17). There was a low correlation between the size of the bruise and the mean BVS (r = 0.37, p = 0.045).

Discussion

Quantifying bruise appearance is a challenge due in large part to the complex psycho-physical process of color perception (Kruschwitz, 2018). By developing the BVS in consultation with nurses who frequently assess these injuries, researchers attempted to capture the clinical language and assessment features used to describe the degree of bruise visibility. VASs are already extensively used in the assessment of burn scars to document their severity over time and response to treatment (Vercelli et al., 2015). Some scales have included reference images in order to both improve inter-rater reliability and address diversity in skin color (e.g., Forbes-Duchart et al., 2007). Given the significant effect skin color has on bruise appearance, both light and dark skin examples were included when modifying the scale.

Discoloration is a defining quality of bruising, which results from the extravasation of hemoglobin from damaged blood vessels and its associated breakdown products (Lyons & Ousley, 2014). Not surprisingly, the BVS was associated with color difference; the significant factor being contrast. The degree of correlation may have been influenced, in part, by the procedure used for obtaining skin color readings. Prior research has demonstrated the reliability of repeated bruise colorimetry readings (Scafide

Table 2. Observers' Bruise Visibility Scale Ratings (n=30).

Observer #	1	2	3	4	5	6	7	8	Total
Number of assessments	20	22	4	20	4	7	8	5	90
Minimum score	1.5	1.3	3.8	1.7	4.0	2.0	3.0	1.0	1.0
Maximum score	4.9	5.0	5.0	5.0	4.7	5.0	5.0	5.0	5.0
Mean score	3.6	3.5	4.4	3.2	4.3	4.0	4.0	3.1	3.6
Standard deviation	1.0	1.2	0.5	1.1	0.4	1.4	0.8	1.7	1.1

Figure 2. Scatterplot of the Relationship Between Mean Bruise Visibility Scale (BVS) Values and the Overall Color Difference (ΔE^*_{ab}) Between the Bruise and Surrounding Skin ($r = 0.57$, $p = < 0.001$).

et al., 2016); however, the skin color measurements taken from locations around the bruise may not have been representative. This theory is supported by the one "visible" bruise noted that did not meet the threshold for perceived color difference ($\Delta E^*_{ab} > 1$).

Further limitations of the study should be noted. A minimum of 30 heterogeneous cases is often recommended for conducting inter-rater reliability testing (Koo & Li, 2016). This sample was heterogeneous in both BVS values and skin color; however, the small sample size may have contributed to lower correlation coefficients. Additionally, the BVS does not capture the important tactile properties of bruise assessment, including induration and pain (Lyons & Ousley, 2014). Overlying wounds and other skin lesions (e.g., lacerations and tattoos) and excessive hair were not observed during bruise assessments. Thus, the reliability of the scale when applied under these conditions may be affected.

Conclusion

Clinical nurses encounter patients across the lifespan who experience bruising as a result of violence or other forms of non-fatal blunt force trauma. Communicating with the healthcare team about alterations in the integumentary system can be challenging given the subjectivity of qualitative descriptors. In the absence of photographs, using more reliable, evidence-based measures could enhance both the accuracy and consistency of bruise documentation. This study evaluated a newly developed scale specifically for the clinical assessment of bruise appearance. The sample size was relatively small, but still demonstrated the BVS as having the potential for being a reliable and valid instrument for application on visible bruises. More research is needed to further evaluate the psychometric properties of the BVS on a larger sample of more diverse bruise conditions and clinical environments, including alternate light assessments.

ORCID iDs

Katherine N. Scafide https://orcid.org/0000-0002-9443-0006
Ghareeb M. Bahari https://orcid.org/0000-0003-3949-3650
Nesibe S. Kutahyalioglu https://orcid.org/0000-0001-8220-4290
Mona Mohammadifirouzeh https://orcid.org/0000-0003-2257-1421

References

Cosman, T. L., Arthur, H. M., Bryant-Lukosius, D., & Strachan, P. H. (2015). Reliability of vascular access site bruise measurement. *Journal of Nursing Measurement*, *23*, 179–200.

De Boos, J. (2019). Non-fatal strangulation: Hidden injuries, hidden risks. *Emergency Medicine Australasia*, *31*, 302–308.

Forbes-Duchart, L., Marshall, S., Strock, A., & Cooper, J. E. (2007). Determination of inter-rater reliability in pediatric burn scar assessment using a modified version of the

Vancouver Scar Scale. *Journal of Burn Care & Research*, *28*, 460–467.

Kemp, A. M., Maguire, S. A., Nuttall, D., Collins, P., & Dunstan, F. (2014). Bruising in children who are assessed for suspected physical abuse. *Archives of Disease in Childhood*, *99*, 108–113.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*, 155–163.

Kruschwitz, J. D. T. (2018). *Field guide to colorimetry and fundamental color modeling*. Society of Photo-Optical Instrumentation Engineers (SPIE).

Kuehni, R. G., & Marcus, R. T. (1979). An experiment in visual scaling of small color differences. *Color Research & Application*, *4*, 83–91.

Lawson, Z., Dunstan, F., Nuttall, D., Maguire, S., Kemp, A., Young, S., Barker, M., & David, L. (2015). How consistently do we measure bruises? A comparison of manual and electronic methods. *Child Abuse Review*, *24*(1), 28–36.

Limmen, R. M., Ceelen, M., Reijnders, U. J. L., Joris Stomp, S., de Keijzer, K. C., & Das, K. (2013). Enhancing the visibility of injuries with narrow-banded beams of light within the visible light spectrum. *Journal of Forensic Sciences*, *58*, 518–522.

Lyons, F., & Ousley, L. (2014). Common dermatological conditions. In *Dermatology for the advanced practice nurse* (pp. 59–345). Springer Publishing Company.

Polit, D. F., & Yang, F. (2016). *Measurement and the measurement of change*. Wolters Kluwer Health.

Savall, F., Lechevalier, A., Hérin, F., Vergnault, M., Telmon, N., & Bartoli, C. (2017). A ten-year experience of physical intimate partner violence (IPV) in a French forensic unit. *Journal of Forensic and Legal Medicine*, *46*, 12–15.

Scafide, K. N., & Sheridan, D. J. (2015). [*Unpublished raw data on the test-retest and intrarater reliability of the Bruise Visibility Scale using photographs of bruises*]. George Mason University.

Scafide, K. N., Sheridan, D. J., Taylor, L. A., & Hayat, M. J. (2016). Reliability of tristimulus colourimetry in the assessment of cutaneous bruise colour. *Injury*, *47*, 1258–1263.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420.

Tsokos, M. (2015). Diagnostic criteria for cutaneous injuries in child abuse: Classification, findings, and interpretation. *Forensic Science, Medicine, and Pathology*, *11*, 235–242.

Vercelli, S., Ferriero, G., Sartorio, F., Cisari, C., & Bravini, E. (2015). Clinimetric properties and clinical utility in rehabilitation of postsurgical scar rating scales: A systematic review. *International Journal of Rehabilitation Research*, *38*, 279–286.