EDUCATION AND PRACTICE

An Assessment of Clinical Wound Evaluation Scales

JAMES V. QUINN, MD, GEORGE A. WELLS, MD

Abstract. **Objective:** To compare 2 clinical wound scales and to determine a minimal clinically important difference (MCID) on the visual analog cosmesis scale. **Methods:** Using data from 2 previously published clinical trials, 91 lacerations and 43 surgical incisions were assessed on the 2 scales; a 100-mm visual analog scale (VAS) (0 = worst possible scar, 100 = best possible scar) and a wound evaluation scale (WES) assessing 6 clinical variables (a score of 6 is considered optimal, while a score of ≤5 suboptimal). All wound assessments on the VAS were done by 2 cosmetic surgeons who rated photographs on 2 occasions. A cohort of wounds on the WES were assessed by a second observer. The difference of the mean optimal and suboptimal VAS scores for each study was used to determine a MCID on the VAS scale. **Results:**

The VAS scale yielded intraobserver agreements of 0.93 and 0.87 (95% CI: 0.89–0.96 and 0.78–0.93) and interobserver agreements of 0.50 and 0.71 (95% CI: 0.32–0.65 and 0.52–0.84) for lacerations and incisions, respectively. Kappa coefficient measuring agreement on the WES was 0.79 (95% CI: 0.57–1.0). The mean (±SD) VAS scores of optimal wounds were 72 ± 12 mm and 65 ± 20 mm, while the mean scores of suboptimal wounds were 57 ± 17 mm and 50 ± 23 mm for lacerations and incisions, respectively. **Conclusions:** An MCID on the VAS cosmesis scale is 15 mm. Studies should be designed to have a sample size and power to detect this difference. **Key words:** wounds; lacerations; tissue adhesives; sutures; staples; infection; cosmetic appearance. ACADEMIC EMERGENCY MEDICINE 1998; 5:583–586

VALID, reliable outcome measures that can assess important clinical differences are essential tools for clinical trials.[1] Investigators have been interested in measuring important but subjective outcomes for some time. Outcome measures have been developed to evaluate subjective outcomes such as quality of life, and pain.[2,3] Recently there has been interest in clinical wound healing, and several scales have been developed to measure the cosmetic outcomes of healed lacerations and incisions.[4,5] A visual analog cosmesis scale, like the visual analog pain scale, has become a desirable method of clinical evaluation because of its continuous nature, ability to measure small differences, reproducible results, and ease of use.

The difficulty with visual analog scales (VASs) is that although they can measure small changes and statistical differences between groups, the clinical relevance of the numerical results and important clinical differences on these scales are not intuitive.[6,7] To improve the VAS scale as an outcome measure, the determination of minimal clinical important differences (MCIDs) is essential to their proper use. In addition, they should be combined with descriptive scales that can better define the numerical results of the VAS. To date, investigators have determined this MCID on VAS pain scales[7] and the VAS cosmesis scale has been combined and contrasted with descriptive scales.[5] This is the first study to derive an MCID for the VAS cosmesis scale.

From the Department of Surgery, Division of Emergency Medicine (JVQ), University of Michigan, Ann Arbor, MI; and the Loeb Research Institute, Clinical Epidemiology Unit (GAW), University of Ottawa, Ontario, Canada.
Address for correspondence and reprints: James V. Quinn, MD, Division of Emergency Medicine, The University of Michigan, TC B1354, Box 0303, 1500 East Medical Center Drive, Ann Arbor, MI 48109-0303. Fax: 313-936-9414; E-mail: jaquinn@umich.edu

METHODS

Study Design. We performed a retrospective analysis and comparison of 2 scales used to measure the cosmetic outcomes of healed lacerations and incisions of 2 previously published studies. Approval to analyze the data from published studies using these scales was granted from the institutional review board at the University of Michigan.

Population and Settings. The data from 2 prospective trials on laceration treatment[8] and surgi-

TABLE 1. Wound Evaluation Scale*

1. Stepoff borders (0 for yes, 1 for no)
2. Contour irregularity—puckering
3. Scar width—greater than 2 mm
4. Edge inversion—sinking, curling
5. Inflammation—redness, discharge
6. Overall cosmesis (0 = poor, 1 = acceptable)

*6/6 = optimal wound healing.

cal incision[9] treatment were used. The trials compared traditional suturing with the use of a new topical tissue adhesive. The laceration study included all facial lacerations as well as certain selected extremity and torso lacerations. The surgical incision study compared incisions from a wide variety of head and neck incisions.

Measurements. In the laceration study, patients were prospectively randomized to have the skin of their lacerations closed with either 5-0 monofilament suture or topical tissue adhesive. The average length of laceration in the study was 2.5 cm, and 12.5% of the lacerations required deep sutures. At 3 months the patients had a wound assessment done on a previously validated wound evaluation scale (WES).[4] The score was assigned by a research nurse aware of the method of closure and validated by a second nurse blinded to the method of closure and unaware of previous wound scores. Nurses and physicians have been shown to use the WES reliably.[4] The wound score addresses 6 clinical variables: absence of stepoff, contour irregularities, wound margin separation >2 mm, edge inversion, excessive distortion, and overall cosmetic appearance. Each of these categories is graded on a 0-or-1-point scale. A total cosmetic score is derived by the addition of the scores of the 6 categorical variables. A score of 6 is considered optimal, while a score of ≤5 suboptimal (Table 1).

At the same time the WES was determined, the patients had photographs of their healing wounds taken. The photographs were taken by the research assistant in a standard fashion using a 1:3 macro setting with a ring flash and using 100 Ektachrome slide film.[10] These photographs were rated for cosmesis on a previously validated cosmesis scale. This VAS for cosmesis has been demonstrated to be reliable and valid.[5] The VAS cosmesis scale is a 100-mm line with "worst scar" at the right end (0 mm) and "best scar" at the left end (100 mm) (Figs. 1 and 2). The photographs were rated on 2 occasions by 2 cosmetic surgeons blinded to which method was used to close the wounds.

The surgical incision study was a prospective, nonrandomized comparison study comparing the skin closures of various head and neck incisions with either 4-0 nylon subcuticular suture or octyl-

cyanoacrylate tissue adhesive. All wounds in this study had identical closures of the deep layer prior to skin closure. The average length of the incisions was 9.4 cm. The same 2 methods of wound evaluation (WES and VAS) were used, but were performed at 1 month rather than 3 months for the laceration study.

Data Analysis. The mean ± SD of the VAS cosmesis scores for optimal and suboptimal wounds was determined for each study. The MCID on the VAS was considered to be the distance between an average optimal score and an average suboptimal score. Observer agreement of the cosmetic surgeons on their VAS scores was determined using the intraclass correlation coefficient. Agreement on the WES was determined using the kappa (κ) coefficient. Confidence interval estimates for the agreement coefficients were calculated.

RESULTS

There were 91 patients in the laceration group, 67 with optimal scores and 24 with suboptimal scores. In the incision study, there were 43 patients, 23 with optimal and 20 with suboptimal wound scores. The mean difference in VAS cosmesis scores in both the laceration and incision studies was 15 mm (Figs. 1 and 2). The VAS scale yielded intraobserver agreements of 0.93 and 0.87 (95% CI: 0.89–0.96 and 0.78–0.93) and interobserver agreements of 0.50 and 0.71 (95% CI: 0.32–0.65 and 0.52–0.84) for lacerations and incisions, respectively. The κ coefficient measuring agreement on the WES was 0.79 (95% CI: 0.57–1.0).

DISCUSSION

Until recently most treatments for wounds were evaluated and reported as case series.[11] Objective comparison of treatments with a randomized trial was impossible because of the lack of valid, reliable, and clinically important outcome measures that could be assessed in a blinded fashion. Valid reliable scales such as the VAS cosmesis scale and the WES now make it possible to do this. Any good outcome measure should have 5 properties: it should measure an important outcome of the intervention, it should be reliable (i.e., the results should be reproducible), it should be valid (i.e., be a true measure of the outcome), one should be able to measure or determine clinically important differences on the scale, and it should be easy to use and understand.

Several years ago we chose cosmesis as our outcome measure for clinical wound evaluation. We determined that it was the most important outcome for patients by informally surveying 75 pa-

tients and parents of children with lacerations. While outcomes such as pain and time were very important, final cosmetic outcome (80% of the time) was the most important outcome to this group of surveyed parents and patients with facial lacerations. When considering other parts of the body it was still important (39% of the time), but time spent waiting was considered more important for these wounds (46%), while pain was third (15%) (unpublished data, 1991). Other important outcomes such as infection and dehiscence were considered, but using them as a primary outcome measure is difficult because they are relatively rare, and measuring them is problematic because many cannot agree on what truly is an infection.[12,13] We also believe that important infections and dehiscences will manifest themselves as poor cosmetic outcomes.

As noted above, a good outcome measure must be reliable, i.e., reproducible. We have shown that if a cosmetic surgeon likes the scar the first time it is shown to him or her, and then the scar is shown to the same surgeon again some time later when he or she is unaware of the previous result, he or she will give the scar a very similar score (intraobserver reliability or agreement). Having 2 cosmetic surgeons agreeing on the score of the wound (interobserver agreement) not only gives it a measure of reliability, but also gives it a measure of validity. Not surprisingly, agreement of a surgeon with his or her own previous scores tends to be stronger than agreement between surgeons, and in this study one measurement was only 0.5, which is the lower acceptable boundary for such agreement.[1] However, when analyzing the incisions in this study, the interobserver agreement was 0.71, which was more consistent with previous studies where the interobserver agreement consistently ranges from 0.71 to 0.75.[5,14,15]

There is no criterion standard with which to measure and test the validity of the VAS cosmesis scale. For that reason we believe it is important to help define the scale with descriptive measures or outcomes that can help to define numerical values of the scale and improve its validity. In these 2 studies we chose to combine it with the WES, which assesses various clinical variables when determining an optimal vs a suboptimal scar. In the past we combined the VAS cosmesis scale with a descriptive scale that classified wounds into 3 categories: excellent, acceptable, and poor, based on clinical descriptors.[5] The importance of this is illustrated in the following example. If one were told that the average VAS cosmesis score of a group of traumatic lacerations was 68 mm, he or she would really have no idea what that means. It may be possible to compare them with another group, but was the score from this group actually good or bad?

By knowing that the average optimal score of all traumatic wounds was 72 mm, one can be fairly confident that the outcomes in that group are good. In another attempt to improve validity, we chose to use cosmetic surgeons to evaluate the scars, because they are the ones who deal with scars on a daily basis and make decisions regarding revision. Doing so gives the VAS cosmesis scale a measure of "face validity."

LIMITATIONS AND FUTURE QUESTIONS

This study is limited by its retrospective analysis of data from 2 studies, which, aside from using the same outcome measures, were very different in design. The study on lacerations was a prospective, randomized trial of traumatic lacerations, on a variety of sites of the body closed by 10 different physicians, with a 3-month blinded cosmetic evaluation. The study on incisions was a nonrandomized, prospective comparison of surgical incisions treated by 2 surgeons (1 sutured and 1 applied tissue adhesive) with a 1-month blinded cosmetic evaluation. To deal with this limitation we did not attempt to combine the data of these studies when doing our analysis. However, the fact that both scales are reliable and produce similar MCIDs despite vastly different clinical scenarios leads us to
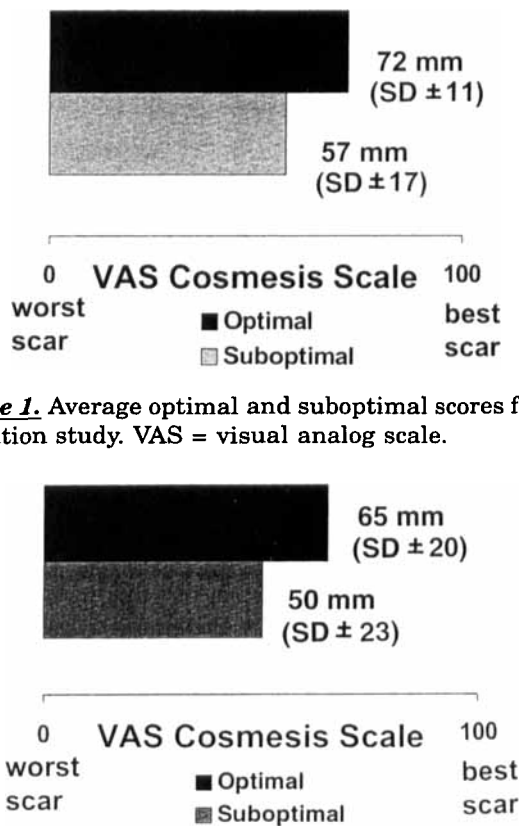

72 mm (SD ±11)

57 mm (SD ±17)

0 **VAS Cosmesis Scale** 100
worst scar ■ Optimal best scar
□ Suboptimal

Figure 1. Average optimal and suboptimal scores for the laceration study. VAS = visual analog scale.

65 mm (SD ±20)

50 mm (SD ±23)

0 **VAS Cosmesis Scale** 100
worst scar ■ Optimal best scar
▨ Suboptimal

Figure 2. Average optimal and suboptimal scores for the incision study. VAS = visual analog scale.

believe that our conclusions can be generalized to most clinical settings.

We also assumed that an important clinical difference corresponded to the difference between an optimal and a suboptimal WES score. The assumption that a scar is clinically less desirable because it fails to score on all 6 clinical variables may be debated, although the majority of wounds score either a 4–5/6 (suboptimal) or 6/6 (optimal). Only the occasional wound scores less than 4/6.

Finally, it can be argued that the satisfaction of patients with regard to assessment of their own wounds is the most important outcome. Although we have considered the MCID that physicians are willing to accept, we have not addressed the MCID that patients are willing to accept. Singer et al. have shown good concordance between physician and patient assessments of wounds.[16] However, using patients to assess their wounds has methodologic problems. It is impossible to blind patients to the treatment of their wounds and when patients are not blinded, they are more likely to bias their assessments to please the investigators (Hawthorne effect).[17] The VAS used by patients, unlike the VAS cosmesis scale, has not yet been validated as an outcome measure. Prospective, randomized trials using blinded validated outcome measures should be the design of choice for clinical wound studies, and future rigorously designed studies should similarly evaluate patient assessment of cosmesis.

CONCLUSION

The MCID on the VAS cosmesis scale is 15 mm. Studies should be designed with a sample size to have the power to determine this difference. Studies that have a large sample size may have the power to determine statistical differences less than this, but the clinical relevance of such smaller differences is uncertain.

References

1. Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use, ed 1. New York: Oxford University Press, 1991, pp 1–3, 8.
2. Huskisson EC. Visual analogue scales. In: Melzack R. Pain Measurement and Assessment, New York: Raven Press, 1983, pp 33–7.
3. Guyatt GH, Berman LB, Townsend M, et al. A measure for quality of life in clinical trials for chronic lung disease. Thorax. 1987; 42:773–8.
4. Hollander JE, Singer AJ, Valentine S, Thode HC Jr, Henry MC. Wound registry: development and validation. Ann Emerg Med. 1995; 25:675–85.
5. Quinn JV, Drzewiecki AE, Stiell IG, Elmslie TJ. Appearance scales to measure the cosmetic outcomes of healed lacerations. Am J Emerg Med. 1995; 13:229–31.
6. Wells GA, Tugwell P, Kraag GR, et al. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. J Rheumatol. 1993; 20:557–60.
7. Todd KH, Funk KG, Funk JP, Bonacci R. The clinical significance of reported changes in pain severity. Acad Emerg Med. 1995; 2:369–70.
8. Quinn J, Wells G, Sutcliffe T, et al. A randomized trial comparing octylcyanoacrylate tissue adhesive and sutures in the management of lacerations. JAMA. 1997; 277:1527–30.
9. Maw JL, Quinn JV, Wells GA, et al. A prospective comparison of octylcyanoacrylate tissue adhesive and sutures for the closure of head and neck incisions. J Otolaryngol. 1997; 26: 26–30.
10. Storrow AB, Stack LB, Peterson P. An approach to emergency department photography. Acad Emerg Med. 1994; 1: 454–62.
11. Watson DP. Use of cyanoacrylate tissue adhesive for closing facial lacerations in children. Br Med J. 1989; 299:1014.
12. Thomson PD, Smith DJ. What is infection? Am J Surg. 1994; 167(1A suppl):7S–10S.
13. Quinn JV, Maw JL, Ramotar K, et al. Octylcyanoacrylate tissue adhesive wound repair versus suture wound repair in a contaminated wound model. Surgery. 1997; 122:69–72.
14. Bruns TB, Simon HK, McLario DJ, et al. Laceration repair using a tissue adhesive in a children's emergency department. Pediatrics. 1996; 98:673–5.
15. Quinn JV, Drzewiecki AE, Li MM, et al. A randomized, controlled trial comparing a tissue adhesive with suturing in the repair of pediatric facial lacerations. Ann Emerg Med. 1993; 22:1130–5.
16. Singer AJ, Church AL, Forrestal K, et al. Comparison of patient satisfaction and practitioner satisfaction with wound appearance after traumatic wound repair. Acad Emerg Med. 1997; 4:133–7.
17. Fletcher RH, Fletcher SW, Wagner EH. Clinical Epidemiology: The Essentials, ed 2. Baltimore: Williams and Wilkins, 1988, p 137.