The Visual Analog Scale in the Immediate Postoperative Period: Intrasubject Variability and Correlation with a Numeric Scale

Lauren J. DeLoach, MD*, Michael S. Higgins, MD†, Amy B. Caplan, MD*, and Judith L. Stiff, MD*

Departments of Anesthesiology, *Johns Hopkins Bayview Medical Center, Baltimore, Maryland; and †Vanderbilt Medical Center, Nashville, Tennessee

The visual analog scale (VAS) has been used to assess the efficacy of pain management regimens in patients with acute postoperative pain, but its usefulness has not been confirmed in postoperative pain studies. We studied 60 subjects in the immediate postoperative period. The specific data collected were: VAS scores versus an 11-point numeric pain scale; repeatability in VAS scores over a short time interval; and change in VAS scores from one assessment period to the next versus a verbal report of change in pain. The correlation coefficients for VAS scores with the 11-point pain scale were 0.94, 0.91, and 0.95. The repeatability coefficients were 17.6, 23.0, and 13.5 mm. Of the 56 patients who completed all three assessments, only 16 (29%) had repeatability within 5 mm on all three. Some of the changes in VAS scores between assessments were in the direction opposite the verbally reported changes in pain (31%); however, most (92%) were within 20 mm. There was no correlation between the level of sedation, previous pain experience, anxiety, or anticipated pain with consistency in VAS scores. We conclude that any single VAS score in the immediate postoperative period should be considered to have an imprecision of ±20 mm. **Implications:** The visual analog scale was developed for assessing chronic pain but is often used in studies of postoperative pain. This study finds that the visual analog scale correlates well with a verbal 11-point scale but that any individual determination has an imprecision of ±20 mm.

(Anesth Analg 1998;86:102–6)

The visual analog scale (VAS) is a simple and often used method for evaluating variations in pain intensity (1). Subjects are instructed to indicate the intensity of the pain by marking a 100-mm line anchored with terms describing the extremes of pain intensity. Its usefulness has been validated in the setting of chronic pain by several investigators (2–4). In this setting, the VAS is superior to fixed interval scales, relative pain scales, and verbal reports of pain (5–7).

More recently, the VAS has been used to measure pain in the immediate postoperative period to compare the effect of different analgesic regimens. Results are sometimes difficult to interpret. In a study comparing preoperative ibuprofen with intraoperative fentanyl in laparoscopic surgery, although the ibuprofen subjects had lower VAS scores at some of the times postoperatively, they had received more fentanyl during the recovery period (8). In other drug comparisons, differences in VAS scores are reported at various time intervals, but some patients have received rescue analgesia (9,10).

We hypothesize that the postoperative perceptual-cognitive impairment experienced by patients who have undergone anesthesia degrades the relationship of the VAS with the subjective pain experience, which leads to a range of imprecision of each individual measurement. Understanding this imprecision will help to interpret studies using VAS as the outcome measure.

Methods

After approval by our institutional review board, informed consent was obtained in the immediate preoperative period. Sixty subjects, ASA physical status I–III, aged 18–86 yr, undergoing various surgical procedures, were studied.

During the preoperative interview, subjects were familiarized with the recording of the VAS and digit

Accepted for publication September 9, 1997.

Address correspondence and reprint requests to Lauren J. DeLoach, MD, Department of Anesthesiology, Johns Hopkins Bayview Medical Center, 4940 Eastern Ave., Baltimore, MD 21224.

0003-2999/98/$5.00

symbol substitution test (DSST). Preoperatively, subjects marked three VAS: 1) worst pain in a previous experience; 2) anticipation of severity of postoperative pain in the postoperative anesthesia care unit (PACU); and 3) anxiety. Subjects also reported a postoperative pain anticipation rating using a 5-point verbal pain scale (VPS) (no pain, mild pain, moderate pain, severe pain, and horrible pain) and marked an unmarked VAS with the following categories: no pain, mild pain, moderate pain, severe pain, and horrible pain. The VAS pain scales were 100-mm vertical lines anchored with "no pain" at the bottom and "worst imaginable pain" at the top. The anesthetic plan was determined by the anesthesiologist caring for the subject. Subjects received general anesthesia, regional anesthesia with sedation, or local anesthesia with sedation.

Postoperatively, three sets of data and a set of recall data at PACU discharge were collected. Set 1 was collected within 20 min of admission to the PACU, and Sets 2 and 3 were collected at subsequent 20-min intervals. A set consisted of a pain VAS (VAS A), an 11-point verbal pain scale (0 = no pain, 10 = worst imaginable pain), a 5-point VPS, a DSST, and another pain VAS (VAS B). VAS B was obtained approximately 3 min after VAS A. Sets 2 and 3 also included a relative VPS (much better, better, same, worse, much worse) collected after VAS A. At the beginning of each set, the subject's level of consciousness was recorded by the investigator (fully awake, drowsy, arouses to voice, arouses to tactile stimulation, needs excessive stimulation). Analgesics were given as requested by the attending anesthesiologist. On discharge from the PACU, subjects were asked to mark a VAS recalling their worst pain and their first VAS score (VAS 1A).

Spearman correlation coefficients were calculated and regression lines were drawn using a least squares method to compare the VAS with the 11-point pain scale. Reproducibility between VAS A and B and consistency of recall were assessed according to the methods of Bland and Altman (11) for determining repeatability. Repeatability was assessed by plotting the difference between the two measurements at the beginning and end of each assessment period against the mean of the two. This reveals any systemic bias (e.g., measurements were less repeatable or showed wider differences at one end of the scale). The sum of all the differences reflects any tendency for a consistent difference (e.g., first VAS always being higher than the second). If it is close to 0, there is no consistent difference. The repeatability coefficient is 2 SD of the differences, and it is expected that 95% of the differences would lie within the repeatability interval.

To measure longer-term recall, subjects were asked to mark a VAS to correspond with what they remembered as the worst pain they had during the study. This was compared with the highest VAS score on any assessment. They were also asked to recall the first VAS they had marked, which had been done on pink paper to aid in the recall. These recall measures were analyzed both by using Spearman correlation coefficients and the methods of Bland and Altman (11). The change in VAS from one assessment to the next was compared with the VPS of change in pain.

We were unable to score the DSST, because many patients in the postoperative period started and then told us they could not do it or did not want to finish it.

Results

The correlation plots for the 11-point pain scale and VAS scores measured at the start of each assessment are presented in Figure 1. The Spearman correlation coefficients for each of the three assessment periods were 0.94, 0.91, and 0.95, respectively, showing good correlation. However, the regression lines derived from these data show that the slopes of the lines are

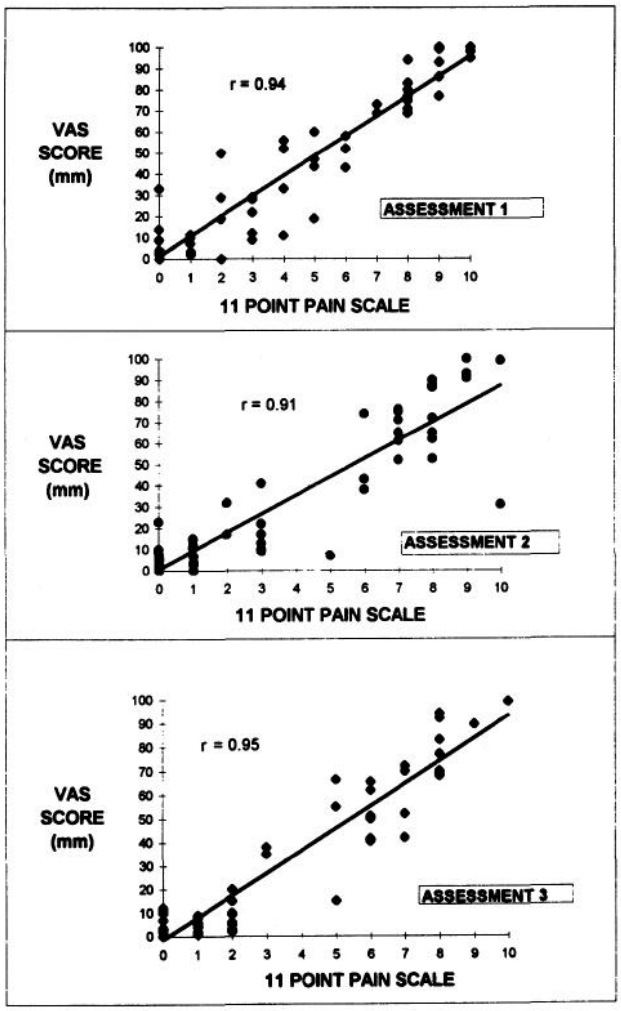

Figure 1. Visual analog scale (VAS) score versus 11-point pain scale score correlation plots for Assessments 1 (n = 60), 2 (n = 58), and 3 (n = 56).

not 1 (0.95, 0.86, and 0.95), which indicate that although the 11-point pain scale scores and the VAS scores correlate well, they are not equivalent.

The graphs for repeatability of VAS scores at the beginning (VAS A) and end (VAS B) of each assessment are shown in Figure 2. The repeatability coefficients for the three assessment periods were 17.6, 23.0, and 13.5 mm (dashed lines). From these graphs, it seems that there is no bias, because points are similarly dispersed at the high and low ends of the scale. The averages of each set are close to 0 (1.52, −0.83, −1.08), indicating no tendency of the first VAS score to be consistently higher or lower than the second within each assessment period.

The distribution of differences between the VAS scores at the beginning and end of each assessment period shows that, overall, 56% of subjects showed repeatability within 5 mm (Fig. 3). For this part of the analysis, patients who reported no pain and marked 0 both times were not included, because the zero line would be a cue and because the VAS is a pain score

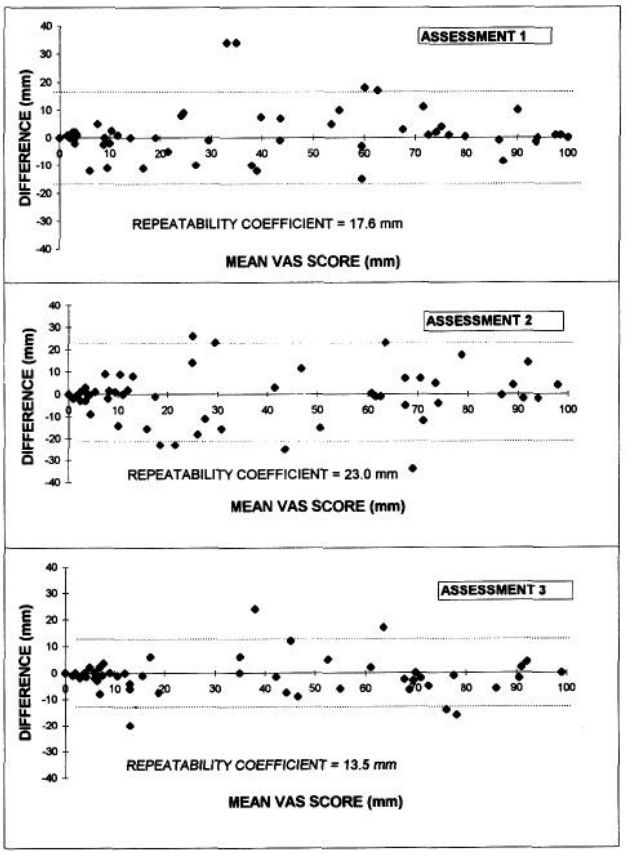

Figure 2. Repeatability plots for Assessments 1, 2, and 3. The repeatability coefficient is ± 1 SD from the zero line and is shown by the horizontal lines above and below the zero lines. VAS = visual analog scale.

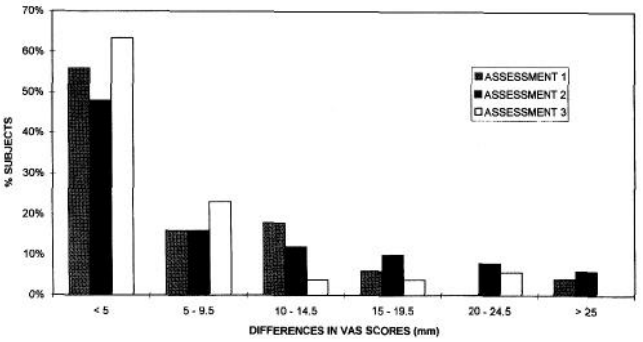

Figure 3. Distribution of the differences in visual analog scale (VAS) scores from the beginning to the end of each assessment period.

and these patients had no pain. There was no correlation between repeatability and observed level of consciousness. Also, subjects who had a repeatability coefficient of ±5 mm on one assessment did not necessarily have the same repeatability coefficient on another assessment. Of the 56 patients who completed all three assessments, only 16 (29%) had agreement within 5 mm on all three. These 16 had sedation scores ranging from "alert" to "arouses by voice." The anesthetic was local with sedation in 9 patients, regional with sedation in 3 patients, and general anesthesia in 4 patients.

In comparing the changes in VAS scores with the verbal report of change in pain from one assessment to the next, we found 15 reports of pain that were "worse," but only 10 showed an increase in VAS scores (more pain). Of those 5 with a decrease in VAS score, only 2 had a decrease of more than 20 mm. Of 30 reports of pain that was "better" or "much better," 21 had a decrease in the VAS score. Of the 9 with a VAS score in the opposite direction, 2 had an increase of more than 20 mm. If the repeatability of the VAS score is approximately 20 mm, as we found, then most of the inconsistencies in the change in VAS scores versus verbal change in pain can be explained as being a result of the degree of repeatability. There were 57 reports of the pain being the same. Of those, 4 showed more than a 20 mm difference from the previous VAS. Thus, in a total of 102 VAS ratings of change in pain, 8 (7.8%) were outside the 20-mm repeatability range. As described by Bland and Altman (11), 95% of measurements should fall within the repeatability range.

Two aspects of long-term recall of VAS scores were examined. At the time of discharge from the PACU, subjects were asked to recall their first VAS (VAS 1A) and to recall their worst pain. We found that 39% (24 of 62) of subjects had no recall of VAS 1A, and of those who did, the repeatability plot showed bias. Recall was better for those who had low VAS scores. Subjects tended to remember their first VAS as higher than they had marked (sum of differences was 5.42). Recall of the worst-pain VAS score tended to be higher than

what was actually marked during the assessments (sum of differences was 7.97).

We found no effect of type of anesthesia or level of sedation. However, our sample size was too small to conclude that there is no effect.

Discussion

The goal of this study was to assess the usefulness of the VAS in the immediate postoperative period. Confusing results from clinical studies assessing the efficacy of postoperative analgesic regimens may be explained, in part, by problems intrinsic to using the VAS in postoperative subjects.

For a pain measurement instrument such as the VAS to be useful and valid, it must be easily understood and used by the subject, and it should compare well with other established methods of assessing pain. When we presented a VAS line to the subjects preoperatively, all marked mild, moderate, and severe pain in an ascending fashion, which would indicate understanding. The VAS was more difficult to use in the postoperative period because of residual anesthesia, blurred vision, or nausea, and several subjects required additional instructions to complete the VAS. An 11-point verbal scale does overcome most of these difficulties. Four subjects did not complete all the data sheets.

The validity of the VAS or any pain measurement scale cannot be determined directly. One aspect of validity is a scale's agreement with another recognized measurement scale. There was good correlation between the VAS scores and the 11-point pain scale in each of the three assessment periods. However, it is possible that the VAS cued the subjects as to where to rate themselves on the 11-point pain scale, thus leading to good correlation.

Another suggested method of assessing validity is by the response of the scale to pharmacological pain interventions. Our study did not assess responses to analgesic regimens; pain medications were given on demand. However, we compared the change in VAS scores to the verbally described change in pain regardless of whether pain medication was administered.

Additionally, the measuring device should yield repeatable and reliable results. In the case of the VAS, repeatability means that a patient would give the same rating to the same amount of pain. To test the reliability of the VAS, subjects marked a VAS at the beginning and end of an assessment period, which lasted less than three minutes. In between, they completed an 11-point pain scale and a verbal report of pain and attempted to complete a DSST. Subjects were not asked to remember what they marked but to rerate their pain. It was felt that the pain should be the same in that short period of time and that they really would not be marking the scale based on what they remembered marking a few minutes before, after having been distracted by other tasks. We found that the repeatability coefficients ranged from 13.5 to 23.0 mm. Thus, in the immediate postoperative period, one can expect that a single VAS score has an imprecision of ±20 mm. Therefore, in using the VAS to show a change in pain or an effect of medication, the change must be greater than 20 mm to confirm a change. Also, the repeatability coefficients are large enough that any single VAS score may not be a true measure of pain but is probably within 20 mm. Jensen and McFarland (4) contend that clinicians who use a single, or even a few, measures of pain intensity in chronic pain studies are at risk of having unreliable and invalid results. Clearly, performing multiple pain assessments in the immediate postoperative period would be problematic and could delay the administration of analgesics.

We did find a subset of patients (n = 16) who had consistent reproducibility within 5 mm for all three assessments. However, they had undergone a variety of anesthetic techniques and had various levels of sedation.

We confirmed the imprecision by looking at changes in pain. The direction of change in pain was determined by a simple verbal scale and compared with change in VAS score. Although many subjects' VAS scores moved in a direction opposite to their verbal report, the change in VAS score, for the most part, was within the amount of uncertainty (repeatability) that we had found. Thus, the change in pain measurements was consistent if it is considered within the context of the repeatability (±20 mm) of the VAS.

At the time of discharge from the PACU, many subjects were unable to recall the first VAS they had completed on admission. Lack of recall may eliminate subject bias; however, the subjective evaluation of changes in pain involves the recall of previous pain intensity.

In conclusion, the VAS seems to be a valid measure of pain in the immediate postoperative period. It is easily understood and correlates well with an 11-point verbal scale, but it has some limitations. Any single VAS measurement should be considered as accurate ±20 mm. In using the VAS for treatment decisions or for the measurement of the effect of pharmacological interventions, one needs to be aware of this imprecision.

References

1. Chapman CR, Case KL, Dubner R, et al. Pain measurement: an overview. Pain 1985;22:1–31.
2. Scott J, Huskisson EC. Graphic representation of pain. Pain 1976;2:175–84.
3. Gaston-Johansson F, Gustafsson M. Rheumatoid arthritis: determination of pain characteristics and comparison of RAI and VAS in its measurement. Pain 1990;41:35–40.

4. Jensen MP, McFarland CA. Increasing the reliability and validity of pain intensity measurement in chronic pain patients. Pain 1993;55:195–203.
5. Joyce CRB, Autshi DW, Mason RM. Comparison of fixed interval and visual analog scales for rating chronic pain. Eur J Clin Pharmacol 1975;8:415–20.
6. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analog scale. Pain 1983;16: 87–101.
7. Ohnhaus EE, Adler R. Methodological problems in the measurement of pain: a comparison between the verbal rating scale and the visual analog scale. Pain 1975;1:379–84.
8. Rosenblum M, Weller RS, Conard PL, et al. Ibuprofen provides longer lasting analgesia than fentanyl after laparoscopic surgery. Anesth Analg 1991;73:255–9.
9. Comfort VK, Code WE, Rooney ME, Yip RW. Naproxen premedication reduces postoperative tubal ligation pain. Can J Anaesth 1992;39:349–52.
10. Ding Y, White PF. Comparative effects of ketorolac, dezocine, and fentanyl as adjuvants during outpatient anesthesia. Anesth Analg 1992;75:566–71.
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two method of clinical measurement. Lancet 1986;1:307–10.