A Proposal to Use Confidence Intervals for Visual Analog Scale Data for Pain Measurement to Determine Clinical Significance

Srinivas Mantha, MD, Ronald Thisted, PhD, Joseph Foss, MD, John E. Ellis, MD, and Michael F. Roizen, MD

Departments of Anesthesia and Critical Care and Statistics, University of Chicago, Chicago, Illinois

Visual analog scales (VAS) ranging from 0 cm (no pain) to 10 cm (worst imaginable pain) are used widely for pain measurement, but various investigators have not treated these data consistently. Conventional statistical tests of such data, although evaluating the "statistical significance" may obscure the clinical value of a treatment. On the other hand, confidence intervals (CIs) can illuminate both statistical and clinical importance. CIs give a range of values based on the observed data which contain, with a specified probability, a true but unknown variable typifying a population. We reviewed 112 articles published recently in anesthesia journals for statistical reporting of VAS data. Of the 112 articles, only two used CIs to report mean pain scores and one used CIs to report differences in median pain scores between the study groups. Only two articles presented 95% CI for the mean pain scores graphically. Analgesic techniques that produce VAS values in the range of 0–3 have been reported to represent adequate analgesia. A graphical method using CIs is proposed that allows ready interpretation of VAS data. With this approach, one evaluates whether the 95% CI for the mean pain score in a group during a particular period lies entirely within the zone defined as "analgesic success" (0–3). Such an analysis allows a visual assessment of whether a particular technique would produce clinically important effects in the population at large. This approach seems to provide more information than the use of conventional hypothesis testing in the interpretation of VAS data for pain measurement.

(Anesth Analg 1993;77:1041–7)

Visual analog scales (VAS) ranging from 0 cm (no pain) to 10 cm (worst imaginable pain) are used widely for pain measurement and are recommended by guideline committees for gauging therapy for each patient (1). But there is no standard approach for summary or inferential statistics for VAS data. Recent recommendations suggest that emphasis be placed on the use of confidence intervals (CIs) rather than on hypothesis testing by *P* values in the analysis of medical data regardless of whether the hypotheses are tested by parametric or nonparametric means (2–5). This policy has been endorsed by the International Committee of Medical Journal editors (6). Methods to analyze various types of data using CIs have been described (7), and computer software is also available for this application (8). We first examined statistical reporting of VAS data for pain measurement in six anesthesia journals and then developed a graphical method using CIs for ready interpretation of VAS data.

Analgesic techniques that produce VAS values in the range 0–3 cm have been reported to represent adequate analgesia (9–13). For this reason, the range 0–3 cm may be thought of as a "zone of analgesic success" although individual physicians may wish to use a slightly more or less stringent definition. CIs give a range of values based on the observed data within which, with a specified probability, the population value lies. In the context of VAS, the variable of interest is typically either the mean or median VAS response that would be observed in a large population of potential patients. The most common choice in the medical literature is the 95% confidence interval. In practical terms, if the CI for a particular study group lies entirely within the "zone of analgesic success," one can infer the analgesic technique used in that study group would produce clinically useful effects in the population at large.

Accepted for publication June 8, 1993.

Address correspondence and reprint requests to Joseph Foss, MD, Department of Anesthesia and Critical Care, University of Chicago, 5841 South Maryland Ave, Box MC 4028, Chicago, IL 60637.

0003-2999/93/$5.00

Methods

Literature Review

Articles from anesthesia journals (*Anesthesia & Analgesia, Anesthesiology, Canadian Journal of Anaesthesia, Regional Anesthesia, Anaesthesia,* and *British Journal of Anaesthesia*) published from January 1991 to September 1992 containing VAS data for pain measurement were examined. We determined the statistical methods employed to analyze the data in each article: Group descriptive statistics, hypothesis testing (group differences), or variable estimation (CIs for differences in means or medians). We also counted the number of articles with graphical representation of VAS data.

Confidence Intervals for a Single Group of Patients or Study Condition

To illustrate the CI approach to VAS analysis, we used examples drawn from the literature. Because most articles in the literature do not present CIs directly, we calculated 95% CIs from the published summary figures, using the expression

$$\bar{x} \pm t \times \text{SE}(\bar{x})$$

where $\bar{x}$ is the reported average VAS response, t is the critical value for a 5% two-sided test drawn from tables of the t distribution with $n - 1$ degrees of freedom (df), n is the number of subjects on which $\bar{x}$ is based, and $\text{SE}(\bar{x})$ is either the reported standard error or, when only the standard deviation (SD) is reported, the calculated value using the relationship

$$\text{SE}(\bar{x}) = \frac{\text{SD}}{\sqrt{n}}.$$

When the primary outcome of interest is the change in VAS, such as difference between pain levels before and after administering analgesia, then the same expressions can be used, except that individual changes in VAS, rather the VAS scores themselves, are used in calculating $\bar{x}$ and SE (or SD).

Confidence Intervals for Comparing Two Groups of Patients or Study Conditions

In comparing different treatments or findings based on separate groups of patients, CIs for difference between the mean responses in the two groups of patients can be calculated using the expression

$$\bar{x}_1 - \bar{x}_2 \pm t \times \text{SE}(\bar{x}_1 - \bar{x}_2)$$

where t is once again obtained from tables of the t distribution, and $\text{SE}(\bar{x}_1 - \bar{x}_2)$ is calculated from a formula that depends on SD_1, SD_2, n_1, and n_2, the standard deviations and sample sizes from the two separate groups, namely,

$$\text{SE}(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{\text{SD}_1^2}{n_1} + \frac{\text{SD}_2^2}{n_2}}.$$

This calculation is based on not using a pooled estimate of the standard deviation, because the variability in VAS scores under different treatments or in different populations is not likely to be the same. When using this unpooled version, an approximation for the number of df may be derived from the formula (14):

$$df = \frac{(V + W)^2}{\left[\frac{V^2}{n_1 - 1}\right] + \left[\frac{W^2}{n_2 - 1}\right]}, \quad V = \frac{\text{SD}_1^2}{n_1}, \quad W = \frac{\text{SD}_2^2}{n_2}$$

where df is rounded down to the nearest integer. Alternatively if the population variances are assumed to be equal in the two groups, the pooled version for estimating the SE of the difference between the two groups may be used (2). Users of Minitab statistical software may note that this program uses the unpooled version of the formula by default for estimating the 95% CIs for the difference in means between two groups (15).

Confidence Intervals for Multiple Comparisons

Confidence intervals for differences in means can be computed when there are multiple comparisons, using adjusted values q (statistic for Neuman-Keuls test or Tukey test) and q' (statistic for Dunnett's test-multiple comparisons against a single control group) in the place of t. For computational details the reader is referred to Zar (16).

Inference of Hypothesis Testing from Confidence Intervals

There is a close link between the use of a CI for the difference between two means and a two-sided hypothesis test. If the CI for difference in means is calculated, the result of the hypothesis test can be inferred at an associated level of statistical significance. For example, if the 95% CI for the difference in means in two groups does not include zero, then a statistically significant difference between the sample means at the 5% level results from applying the appropriate t-test (2). When two means based on independent samples with more than 10 observations in each group are compared, and the 95% CIs for their means do not overlap, then the means are significantly different at $P < 0.05$ (Figure 1A). If the CIs for the means overlap so much that the mean of one group is contained within the interval of the other group, one can conclude that there is no

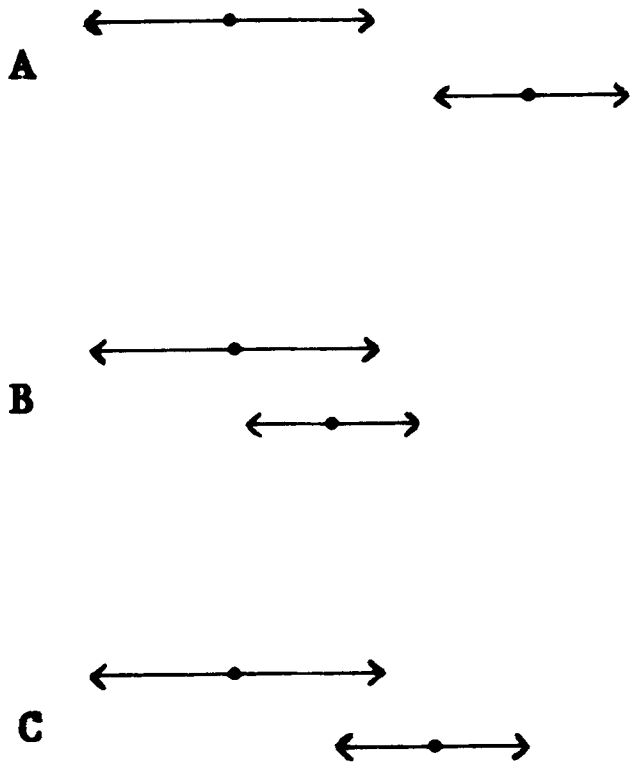

Figure 1. Visual assessment of the difference between two independent groups, using confidence intervals (CIs). In A, 95% CIs do not overlap; therefore, the means are significantly different ($P < 0.05$). In B, CIs overlap and the mean of one group is contained within the interval of the other group, showing no difference between the two means. In C, CIs overlap, but not to the extent that the mean of one group is contained in the interval of the other group; no inference about the statistical significance can be made. Hypothesis testing must be performed to determine whether the means are different.

difference between the means (Figure 1B). If the CIs for the means overlap but not so much that either mean is contained within the CI for the other group, the calculation for the hypothesis test must be performed before one can say whether the means are different (Figure 1C; 17,18).

Parametric versus Nonparametric Techniques and Confidence Intervals

CIs based on nonparametric tests also can be used for single samples and for differences between groups. These procedures give intervals for the median (or median difference) and can be calculated only from raw data, not from summary statistics. Investigators who prefer to use nonparametric hypothesis tests may use CI methods that correspond to each of these tests (19).

We compared inferences that could be drawn using CIs and those arising from standard statistical tests based on the data found in two articles to illustrate these points.

Table 1. Types of Statistical Analysis Used in the Treatment of Visual Analog Scale Data from 112 Articles in the Anesthesia Literature

	n	%
Group descriptive statistics		
Mean (SEM)	44	39.3
Mean (SD)	28	25
Nil	12	10.7
Median (Range)	11	9.8
Median	9	8
Median (IQR)	4	3.5
Mean (95% CI)	2	1.8
Mean	2	1.8
Hypothesis testing (group differences)		
Nonparametric *t*-test	53	47.3
Parametric ANOVA	31	27.7
Not clear	10	8.9
Nonparametric ANOVA	8	7.1
Parametric *t*-test	7	6.25
Contingency table	3	2.7
Parametric estimation (group differences)		
Nil	111	99.1
95% CI on difference in medians	1	0.9

CI = confidence interval, IQR = interquartile range; ANOVA, analysis of variance.

Results

Statistical Reporting of Visual Analog Scale Data in the Current Literature

There were 112 articles with VAS data for pain measurement in the journals we examined. For group descriptive data, only two articles (1.8%) used 95% CI for the mean or median. For hypothesis testing, 38 (33.9%) used parametric tests (Student's *t*-test, paired *t*-test, analysis of variance), and 64 (57%) used nonparametric tests (Mann-Whitney *U* test, Wilcoxon's rank sum test, Kruskal-Wallis test, or χ^2 test). In the remaining 10 (8.9%), it was not clear from the methods described which tests were used. For variable estimation (group differences), there was only one article that gave 95% CI for the difference in medians (Table 1). Only two articles presented the group descriptive statistics (mean and 95% CI) graphically, one as a line graph, the other as a bar graph (Table 2).

Analysis of Visual Analog Scale Data with Confidence Intervals

Example 1 A recent study (20) compared the efficacy of continuous low dose 3-in-1 nerve block technique for postoperative pain relief after total knee replacement to a control group of patients who received intermittent intramuscular papaveretum on request. Pain scores were measured at 4 h and 24 h. The 95% CIs for the mean VAS scores were computed using the summary data in Table 3 and are shown graphically as a to d in Figure 2. The 95% CIs for the difference in pain scores

Table 2. Number of Articles and Type of Graphical Representation of Visual Analog Scale (VAS) Data[a]

	Line graph (*n*)	Bar graph (*n*)
Mean (SD)	3	9
Men (SEM)	20	7
Only mean	3	1
Mean (95% confidence interval)	1	1
Median (range)	4	2
Median (IQR)	2	
Only median	6	

[a] Of 112 articles, 53 contained no graphical representation of VAS data.
IQR = Interquartile range.

at 4 h and 24 h between the two groups are shown in Figure 3. The authors found significantly lower pain scores at 4 h and 24 h postoperatively in the study group than in the control group ($P < 0.01$, Mann-Whitney *U*-test). Inspection of Graphs a and b in Figure 2 reveals that the control technique at both the periods can be expected to produce analgesic failure in the population at large. Inspection of Graphs c and d for the study group reveals that pain scores are lower than those in the control group. But these graphs occupy both the "analgesic success" (0–3) and the "analgesic failure" zones. The graphs indicate that it is uncertain whether the analgesic technique in the study group produces clinically important analgesia in the population at large, more so at 4 h. Statistical significance may be inferred by inspection of the graphs. The 95% CI graphs (Figure 2) at 4 h for control and study group (a and c) do not overlap, indicating that pain scores are significantly different at $P < 0.05$. The pain scores at 24 h are also statistically significant because 95% CI Graphs b and d do not overlap. The same inference about the statistical significance can be made from the inspection of 95% CI graphs for the difference in means (Figure 3). Neither Graphs a (control group versus study group at 4 h) nor b (control group versus study group at 24 h) enclose zero, indicating significant difference at $P < 0.05$.

Example 2 Swayze et al. (21) studied the efficacy of subarachnoid meperidine for labor pain relief. The data from this study (Table 3) were used to generate the 95% CI graphs (a, b, c, Figure 4) for mean VAS scores before the block, at maximum block, and 1 h later. The authors used analysis of variance to compare the VAS pain scores and found a significant reduction in VAS score from before the block to maximum block ($P < 0.0001$) and from before block to 1 h later ($P < 0.0001$). Inspection of 95% CI Graphs b and c (Figure 4) reveals that at both periods after the block, clinically important effects were produced. Pain relief seems to be better at 1 h after the maximum block because the 95% CI for the mean pain score at this period ranges from 0.1 to 0.6 compared to 0.2 to 1 at the period of immediate block.

Discussion

The model proposed takes advantage of CIs, of VAS for pain measurement, and of judgments about clinical success for pain treatments. An advantage of VAS is the ability to define analgesic failure, VAS > 3 cm (9–13). CIs have the advantage of expressing the results in the units in which the measurements are made and become important when an inference is made from the study results to the larger population (22). CIs provide extremely useful information for analyzing data such as VAS for pain measurement where clinically important effects can be demarcated on the scale of measurement (analgesic success versus failure). The model proposed here allows assessment of whether a particular technique at a given period will produce clinically important effects when the study results are extrapolated to the general population. Further, if a practitioner does not agree with 0–3 cm as an "analgesic success," another zone can be chosen.

CIs are an alternative to point estimates coupled with standard errors. They indicate both the best estimate of an effect and the degree of uncertainty about an effect, based on the variability in the sample. Because the effects of sampling variability decrease as sample increases, CIs become narrower as more subjects are studied. Thus a CI may not lie completely inside the zone of "analgesic success" either because the average effect in the population lies outside the interval, or

Table 3. Visual Analog Scale (VAS) and 95% Confidence Intervals (CI) for the Mean. Data from Examples 1 and 2

Example	Group	Mean VAS	SD	*n*	95% CI for the mean	Figure	Graphs in figure
1[a]	Control 4 h	6.43	2.20	18	5.3 to 7.5	2	a
	Control 24 h	5.56	1.82		4.7 to 6.5		b
	Study 4 h	3.29	2.52	19	2.1 to 4.5		c
	Study 24	2.51	1.87		1.6 to 3.4		d
2[b]	Preblock	8.57	1.43	14	7.9 to 9.3	4	a
	Postblock	0.62	0.89		0.2 to 1.0		b
	1 h later	0.33	0.57		0.1 to 0.6		c

[a] Data from reference 20.
[b] Data from reference 21.

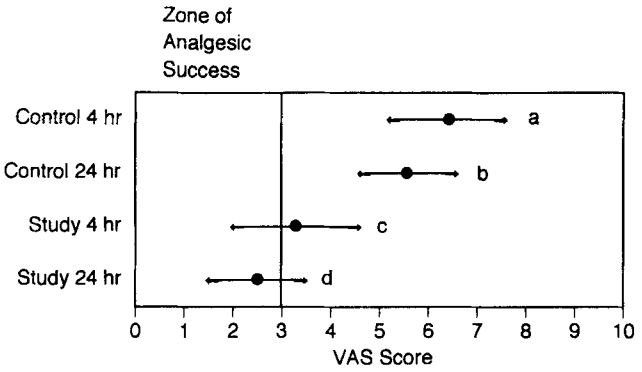

Figure 2. Assessment of the clinical significance of an analgesic technique using confidence intervals (CIs) for the mean. The 95% CI Graphs a and b for the mean visual analog scale (VAS) pain score in the control group for a particular technique at 4 h and 24 h, respectively, occupy the "analgesic failure zone." The control group technique at both these periods is expected to produce analgesic failure in the population at large. The 95% CI Graphs c and d for the study group at 4 h and 24 h, respectively, occupy both the "analgesic success" (0–3 cm) as well as the "analgesic failure" zones. This indicates that it is uncertain whether the analgesic technique in the study group produces clinically important analgesia in the population at large, more so at 4 h. Pain scores of the control group and the study group at 4 h are significantly different ($P < 0.05$) because 95% CI Graphs a and c do not overlap. Similarly the difference at 24 h is significant ($P < 0.05$) because the Graphs b and d do not overlap.

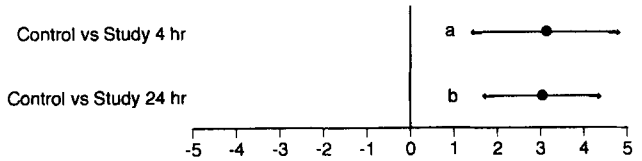

Figure 3. Inference of statistical significance from confidence intervals (CIs) for the difference in the two means. If the 95% CI for the difference in means in two groups does not include zero, then a statistically significant difference between the samples at 5% level results from applying the appropriate *t*-test. Both a (95% CI for the difference in pain scores between study group and control group at 4 h) and b (95% CI for the difference in pain scores at 24 h) do not enclose zero; therefore, a statistically significant difference ($P < 0.05$) in pain scores exists between the two groups at both the periods.

because the uncertainty associated with a given sample size is too great to permit more precise inference, as in Example 1. Often the quantity of greatest clinical significance is the average change in pain levels from a baseline measured by VAS. Unfortunately, most authors do not report the mean and standard deviation of the change in scores so that more appropriate CIs can be constructed and more appropriate inferences drawn. In the study of subarachnoid meperidine (21) referred to in Example 2, the duration and magnitude of levels after maximum block would be of interest to clinicians. A CI for the change would have answered this question. Thus, the method proposed can also suggest measurements important to clinical care for study design.

The choice of parametric versus nonparametric tests and CIs in the context of VAS is controversial (23–25).

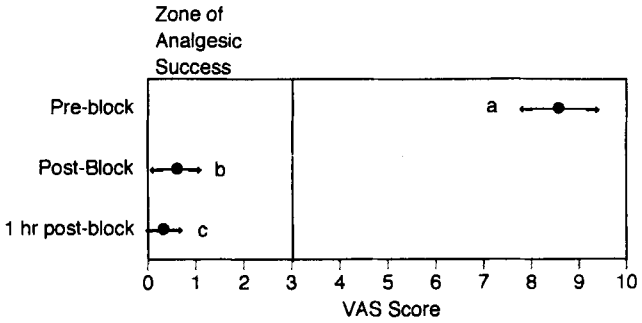

Figure 4. Comparison of visual analog scale (VAS) pain scores in a single group at different periods. The 95% confidence interval (CI) Graph a for the mean pain score before the block, occupies the "analgesic failure zone." The 95% CI Graphs b (postblock) and c (1 h postblock) occupy the "analgesic success zone," indicating that the technique produces clinically important analgesia in the population at large at both these periods.

The advantages of procedures based on the normal distribution (parametric procedures) are that they are widely understood, they are robust against modest departures from normality due to the central limit theorem, and they produce CIs that are generally narrower than nonparametric intervals. Parametric CIs can be calculated readily from summary statistics, all too often the only results reported in the literature. The main advantages of procedures based on the rank order of data (nonparametric procedures) are that they are less sensitive to departures from the normality assumption and they produce CIs for the median that always lie within the range of the observed data.

For comparing groups, we used parametric methods that permit the degree of variability (SD) to vary from one group to the next because that reflects clinical reality. Unfortunately, nonparametric methods for comparing groups require that the variability be the same for each group, and the tests can be rather sensitive to this assumption (26). Although it is not within the scope of this paper to settle the parametric-nonparametric controversy, we generally prefer the parametric methods for VAS scores for the reasons noted above. In any case, we strongly believe that CIs rather than point estimates alone are more useful and informative.

Rarely does one encounter CIs for the mean (computed by parametric procedures) that extend beyond the defined VAS range (0–10 cm). For example, in a recent study, the mean pain scores (SEM) for intraarticular morphine ($n = 11$) and intraarticular morphine plus bupivacaine ($n = 11$) at 4 h postoperatively after arthroscopic knee surgery were 1.39 (0.7) and 0.86 (0.4), respectively (27). These data produce 95% CIs for their means −0.2 to 2.9 and −0.03 to 1.8. Such CIs result either from asymmetry in the VAS distribution, small sample size, or both. Asymmetry can be remedied in several ways: by truncating the CI at 0 or 10 as the case may be, by calculating CIs based on a transformation of the

VAS scale, or by using rank-based nonparametric CIs. Each approach has advantages, and none is uniformly superior to the others.

There are several advantages of analyzing VAS data by the approach proposed here. It is graphical and allows easy interpretation of the data. Different techniques at different periods can be compared to discover clinically useful effects in the population at large, as shown in Example 1. Repeated measures in a single group also can be compared easily as shown in Example 2. Conventional statistical tests, which evaluate statistical significance, may obscure the clinical value of a treatment. Although the authors of the study quoted in Example 1 found a highly significant difference ($P < 0.01$) in the pain scores between the study group and the control group at 4 h, CI analysis shows that the study technique is not expected to produce clinically adequate analgesia in much of the population at large at 4 h, because most of the 95% CI graph including the point estimate (Graph c in Figure 2) lies in the "analgesic failure zone." Thus, the technique proposed makes it possible to differentiate statistical significance from clinical significance. A clear benefit can be seen when two treatments have nonoverlapping CIs and are statistically different but both lie entirely in the "analgesic failure zone." The proposed method would help to clarify that although pain scores in the two techniques were statistically different, neither was clinically acceptable. CIs may be especially useful for small samples to help to avoid misinterpretation of nonsignificant ($P > 0.05$) results by showing which data are compatible with clinically useful effects (28). For example, two treatments may produce nonsignificant difference in pain scores, but the CI for the mean of one treatment may lie completely in the "analgesic success zone" whereas the CI for the other may include both "analgesic success and failures zones." Finally, meta-analysis may be easier to perform if the results of VAS data are presented as 95% CIs for the means and their differences.

A few precautions must be exercised in following this approach. Obviously, side effects of any clinically important technique should be considered. For example, a technique may produce a 95% CI for the mean VAS score within the zone defined as analgesic success (0–3 cm), but may be associated with a high incidence of respiratory depression. Although the commonly used 95% CIs are presented in this model, Gardner and Altman (29) cautioned that any standardization of 95% would not be desirable. At the same time they recognized the potential problems if different CIs were used for comparisons (29). Therefore, it would seem ideal if authors provide 95% CIs for the mean VAS scores graphically as proposed and also give mean, SD, and sample size in a table or text to enable readers to make their own choice of CIs.

An alternative to CIs for graphical representation are figures based on the 10th, 25th, 50th, 75th, and 90th percentiles of the observed distribution. These figures, called box plots in the statistical literature (30), are better at describing the range of results that will occur in the population, inasmuch as they directly exhibit the degree of variability in response across individuals studied. These plots are not affected by skewed or otherwise nonnormal distributions, but rather, they show the extent of skewness. For primary data presentation, the box plot (for describing the population) could be presented in addition to the CI (which indicates the precision of statistical inference).

In summary, a new graphical approach using CIs to analyze VAS data for pain measurement is proposed. If the 95% CI for mean VAS score (calculated by parametric approach) for a particular technique lies within the "zone of analgesic success" (0–3 cm, or another selected by the practitioner), one can infer that the technique is likely to produce clinically important effects in the population at large. This approach seems to be more informative than that provided by the conventional tests and allows ready interpretation of VAS data by the reader. In addition, the 95% CI for the difference in pain scores may also be useful to define precision of the point estimate of the difference.

The authors acknowledge the editorial assistance of Sally Kozlik.

References

1. Acute Pain Management Guideline Panel. Acute pain management in adults: operative procedures. Quick reference guide for clinicians. AHCPR pub. no. 92-0019. Rockville, MD: Agency for Health Care Policy and Research. US Department of Health and Human Services, 1992.
2. Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. BMJ 1986;292: 746–50.
3. Rothman KJ. Significance questing. Ann Intern Med 1986;105: 445–7.
4. Braitman LE. Confidence intervals assess both clinical significance and statistical significance. Ann Intern Med 1991;114: 515–7.
5. Mantha S. Scientific approach to presenting and summarizing data. Anesth Analg 1992;75:469–70.
6. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. BMJ 1988;296:401–5.
7. Gardner MJ, Altman DG, eds. Statistics with confidence. London: British Medical Journal, 1989.
8. Gardner MJ, Altman DG. Confidence interval analysis (CIA). Microcomputer program manual and disk. London: British Medical Journal, 1989.
9. Salomaki TE, Laitinen JO, Nuutinen LS. A randomized double-blind comparison of epidural versus intravenous fentanyl infusion for analgesia after thoracotomy. Anesthesiology 1991; 75:790–5.

10. Huckaby T, Skerman JH, Hurley RJ, Lambert DH. Sensory analgesia for vaginal deliveries: a preliminary report of continuous spinal anesthesia with a 32-gauge catheter. Reg Anaesth 1991; 16:150–3.
11. George KA, Chisakuta AM, Gamble JAS, Browne GA. Thoracic epidural infusion for postoperative pain relief following abdominal aortic surgery: bupivacaine, fentanyl or a mixture of both? Anaesthesia 1992;47:388–94.
12. White MJ, Berghausen EJ, Dumont SW, et al. Side effects of continuous epidural infusion of morphine and fentanyl. Can J Anaesth 1992;39:576–82.
13. Tauzin-Fin P, Maurette P, Vincon G, et al. Clinical and pharmacokinetic aspects of the combination of meperidine and prilocaine for spinal anaesthesia. Can J Anaesth 1992;39:655–60.
14. Snedecor GW, Cocharan WG. Statistical methods. 8th ed. Iowa City: Iowa State University Press, 1989:83–106.
15. Ryan BF, Joiner BL, Ryan TA. MINITAB Handbook. 2nd ed. Boston: PWS-Kent, 1992:181–92.
16. Zar JH. Biostatistical analysis. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1984:185–205.
17. Browne RH. On visual assessment of the significance of a mean difference. Biometrics 1979;35:657–65.
18. Dawson-Saundars B, Trapp RG. Basic and clinical biostatistics. Los Altos, CA: Appleton & Lange, 1990:99–123.
19. Campbell MJ, Gardner MJ. Calculating confidence intervals for some non-parametric analyses. BMJ 1988;296:1454–6.
20. Edwards ND, Wright EM. Continuous low dose 3-in-1 nerve blockade for postoperative pain relief after total knee replacement. Anesth Analg 1992;75:265–7.
21. Swayze CR, Skerman JH, Walker EB, et al. Efficacy of subarachnoid meperidine for labor analgesia. Reg Anaesth 1991; 16:309–13.
22. Gardner MJ, Altman DG. Estimating with confidence. BMJ 1988; 296:1210–11.
23. Forrest M, Andersen B. Ordinal scale and statistics in medical research. BMJ 1986;292:537–8.
24. Meyer RM. Ordinal data are not interval data. Anesth Analg 1990;70:569–70.
25. Philip BK. Parametric statistics for evaluation of visual analog scale. Anesth Analg 1990;71:710.
26. Hollander M, Wolfe DA. Nonparametric statistical methods. New York: John Wiley and Sons, 1973:67–113.
27. Khoury GF, Chen ACN, Garland DE, et al. Intraarticular morphine, bupivacaine and morphine/bupivacaine for pain control after knee videoarthroscopy. Anesthesiology 1992;77:263–6.
28. Altman DG, Gardner MJ. Confidence intervals for research findings. Br J Obstet Gynaecol 1992;99:90–1.
29. Gardner MJ, Altman DG. Using confidence intervals. Lancet 1987;i:74.
30. Cleveland WS. The elements of graphing data. Monterey, CA: Wadsworth, 1985:129–34.