Original Article

Clinically Important Changes in Acute Pain Outcome Measures: A Validation Study

John T. Farrar, MD, MSCE, Jesse A. Berlin, ScD, and Brian L. Strom, MD, MPH
University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA

Abstract

The purpose of this study was to validate the changes in acute pain measurement scales that are most strongly associated with a patient-determined indicator of clinical importance. Measures of pain intensity and pain relief are commonly used outcomes in therapeutic clinical trials. Recent studies of the properties of acute pain measures have provided data defining the cut-off points that are best associated with clinically important differences. Validation of these findings in another clinical trial data set is important. Data were obtained from the titration phase of a recently conducted randomized controlled clinical trial of oral transmucosal fentanyl citrate (OTFC), which compared OTFC to immediate release morphine sulfate (MSIR) for the treatment of cancer-related acute breakthrough pain. Changes in pain intensity and pain relief were recorded every 15 minutes for 60 minutes and global medication performance recorded at the end of each study pain episode. At any titration step, if the patient felt that the first dose of the study medication did not provide adequate relief within 30 minutes, an additional rescue medication could be taken. To find the level of each pain scale best associated with this measure of the adequacy of pain relief, the calculated sensitivity, specificity, and accuracy for different cut-off points of the measured pain scales were compared to whether or not the patient needed rescue medication. The overall ability of the pain measures to discriminate episodes for which a rescue was not needed was calculated using area under the receiver operating characteristics (ROC) curves. Data were analyzed from 134 OTFC-naive patients who collected data on 1307 episodes of breakthrough pain. Using the criteria of a balanced sensitivity and specificity, the best cut-off points were determined to be: 33% for the percent pain intensity difference; ≥2 for the raw pain intensity difference on a 0–10 numeric rating scale; ≥2 (i.e., moderate or better) for pain relief; ≥33% for the percent maximum total pain relief; and ≥2 (good or better) for global medication performance. ROC area under the curve ranged from 0.839 to 0.862 for each of the pain measures listed above, calculated at 60 minutes. These data indicate that the pain scale cut-off points that are best associated with a patient-derived measure of a clinically important difference closely approximate those found in an earlier study. ROC analysis provided evidence that the overall pain measures were strongly associated with not requiring an "additional dose of rescue medication." Thus, the cut-off points determined for these pain scales provide a good surrogate measure of a patient-determined clinically important response. This provides support for the usefulness of these values in future clinical trials of pain therapy. J Pain Symptom Manage

Address reprint requests to: John T. Farrar, MD, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Blockley Hall, Room 816, 423 Guardian Drive, Philadelphia, PA 19104, USA.

Accepted for publication: August 5, 2002.

0885-3924/03/$–see front matter
doi:10.1016/S0885-3924(03)00162-3

2003;25:406–411.

Key Words
Analgesics, pain, pain measurement, randomized controlled trials, treatment outcome, clinically important difference

Introduction

There is no objective measurement for the experience of pain. The measurements of this inherently subjective symptom rely primarily on the verbal reports of patients.[1] No matter how numeric the values provided by the measurement instrument appear to be, those values recorded for each patient represent a subjective interpretation of the pain experience and the patient's assignment of value to the measurement scale. Given the wide variation in the pain experience among individuals, there is a large variability in the reported pain ratings across individuals, despite seemingly similar stimuli or interventions.[2] In addition, specific values on pain measurement scales are often interpreted differently by researchers, clinicians, and patients, depending on the criteria they choose to apply. To better understand the way in which patients use these scales to report their pain, a previous report argued for data-driven cut-off points to help clarify the appropriate interpretation of clinical trial data.[3]

In considering the treatment of acute pain, a reason for a patient's request for additional rescue medication may have several components. These include consideration of whether or not the treatment provides meaningful pain relief in a pharmacokinetically appropriate time frame, whether the relief is adequate to permit increased levels of function, and how long the relief lasts. For this study, we were primarily interested in the first characteristic, namely, the amount of reduction in pain intensity achieved by a rapid onset pain medication at the time of its peak dose effect. In acute pain studies where the therapy being evaluated has a rapid onset of action, the patient can decide whether the pain relief is adequate or whether they want to take an additional dose of rescue medicine. Although the need for additional rescue medication is a measure of the clinical importance of the pain relief obtained by the patient, it is not regularly measured. A recently published study examined the relationship between levels of change on pain measurement scales and the need to additional rescue medication. Cut-off points were derived as surrogate measures for a clinically important improvement in pain.[3] A new database has become available which has permitted us to repeat the analysis to see if the cut-off point values previously determined could be confirmed.

Methods

Study Design

To determine the clinically important cut-off points, data were obtained from a randomized, blinded, double–dummy-controlled clinical trial of breakthrough cancer pain comparing a drug delivery system for oral transmucosal fentanyl citrate (OTFC; ACTIQ™) to oral immediate-release morphine sulfate (MSIR).[4] All 134 outpatients enrolled had chronic cancer-related pain controlled with long-acting opioid medications and recurrent episodes of acute breakthrough pain that were adequately treated with MSIR rescue. The current study used the data from the open-label titration stage (Phase I). During this phase all patients started at the lowest available dosage strength of 200 μg per OTFC unit. The pharmacokinetic blood levels of OTFC are known to peak in less than 30 minutes.[5,6] If acceptable pain relief was not achieved 30 minutes after starting the initial OTFC dose, a second dose of the OTFC or a dose of the patient's original intermittent opioid drug could be taken as an "additional rescue." The dosage strength of the OTFC unit dose was increased (maximum: 1600 μg per unit) until a single dose was found that controlled more than 75% of the patient's target breakthrough pain episodes, namely, not requiring an additional dose of rescue medication for that episode. Thus, for most patients, some of the Phase I episodes were not satisfactorily treated (i.e., did require an addi-

tional dose of rescue medication) and others were satisfactorily treated (i.e., did not require an additional dose of rescue medication).

Data Collected

Data on pain intensity and pain relief were collected every 15 minutes for one hour during the titration phase. A global medication performance rating was obtained at the end of 60 minutes, or at the time a patient took an additional dose of rescue medication, whichever came first. Data from all time points were available, but for the primary sensitivity/specificity analysis we focused on the values at 30 minutes because this was the time point after which 95% of those patients who went on to take an additional dose of rescue medication had done so. The changes over time for the standard pain measures were evaluated, including raw difference in pain intensity (PID; 0–10 numeric rating scale [NRS]), percentage difference in pain intensity (%PID; 0–100% of the NRS), pain relief (PR; 0 [none], 1 [slight], 2 [moderate], 3 [lots], 4 [complete]), sum of the absolute difference pain intensity (SPID; sum of 4 PID measurements over 60 minutes divided by 4), percent of the maximum total pain relief (%Max TOTPAR over 60 minutes), and global medication performance (GMP; 0 [poor], 1 [fair], 2 [good], 3 [very good], 4 [excellent]). If patients did not complete the full 60 minutes (i.e., took an additional dose of rescue medication), the last pain observation was carried forward and summed for the hour to calculate the SPID and %Max TOTPAR. In addition, for the area under the curve analysis (see Results) the comparison of areas is made at both 30 minutes and 60 minutes using the same mechanism for censored data.

Analysis

The analyses described in our previously published report[3] were applied to the current data set of Phase I patients, so that we could compare the results. All patients with complete data from one or more episodes of breakthrough cancer pain were included. The primary analysis determined the degree to which each pain scale accounted for the measured outcome of the "need for additional dose of rescue medication." The pain measures recorded for the satisfactorily treated episodes (i.e., no extra rescue taken) were compared to the unsatisfactorily treated ones (i.e., extra rescue taken). Functionally, a 2×2 table was constructed using each possible cut-off point value for that scale to determine the sensitivity, specificity, and accuracy. Sensitivity was defined as the proportion of episodes for which the patient did not need additional rescue in which the change in pain score equaled or exceeded the cut-off point. Similarly, specificity was defined as the proportion of episodes for which the patients needed additional rescue in which the change in pain score failed to exceed the cut-off point. Accuracy was defined as the overall number of concordant observations over the total number of episodes.

Receiver operator characteristic (ROC) curves are derived by plotting the sensitivity against 1-specificity. In practice, the ROC curves were derived using logistic regression analyses.[7,8] Episodes of breakthrough pain treated with OTFC were the unit of analysis, which was adjusted using a method that takes multiple observations per patient into account when estimating variables.[9] For each analysis, the use of an additional dose of rescue medication served as the outcome variable. The covariates were various forms of the change in pain intensity (raw change, percent change, sum of the raw pain intensity difference), pain relief (raw pain relief, total of the pain relief by hour, and percent maximum total pain relief), or the global medication performance.

Our criterion for the best cut-off point was to find the value with a good overall accuracy coupled with the best balance of sensitivity and specificity. This is identical to finding the tangent point to a 45° line on the receiver operator characteristic (ROC) curve, which assumes an equal importance to sensitivity and specificity. That is, given similar levels of accuracy, we would prefer a measure with a reasonable value for both sensitivity and specificity, so as not to over- or underestimate the efficacy. The area under the ROC curve, reported as the c-statistic from the logistic regression model, represented the total overall agreement between the various measures being evaluated and the use of rescue medication.

Results

In the titration phase of this trial, 134 patients recorded data on a total of 1307 episodes

of cancer-related breakthrough pain, of which 281 episodes were treated with an additional dose of rescue medication. The demographic data for this clinical trial have been previously described.[4]

The sensitivity, specificity, and accuracy for the pain measures being evaluated varied systematically over a wide range of cut-off points when compared to *not* needing an "additional rescue medication" (Table 1). The data point values that came closest to a point represented by the 45° tangent to the ROC curve are bolded, that is, where sensitivity and specificity are given equal weight and the accuracy remains high. For example, note that there is acceptable accuracy for the %PID cut-off points between 10% and 40%, but a cut-off point of ⩾33% provided the highest accuracy associated with an approximately balanced sensitivity and specificity. Using the same method, the best cut-off points for other scales are essentially identical to those found in the previous study,[3] namely: 33% or above for the %Max TOTPAR; 2 (moderate) or better for pain relief; 2 or better for absolute pain intensity difference; 2 or better for the SPID/hour; and 2 (good) or better for the GMP. Data for other values are also provided in Table 1, for readers who may want to consider different cut-off points with different characteristics.

The ROC area under the curve (AUC) analysis demonstrated a high level of discrimination between not needing an "additional dose of rescue medication" as an outcome, and %PID, raw PID, and pain relief at 30 minutes (range: 77–80%) and an even better discrimination (range: 83.9–86.2%) with all the measures at 60 minutes (see Table 2). The summary scores for the %Max TOTPAR and GMP were only available for the 60-minute time point. There is no statistically significant difference among the measured ROC-AUC within each time point.

Discussion

This analysis of a new data set has determined findings that are highly consistent with our prior study.[3] Specifically, our findings in

Table 1

Sensitivity, Specificity, and Accuracy of Various Pain Measures Over a Range of Cut-Off Points Compared with the Use of an Additional Dose of Rescue Medication at 30 Minutes[a]

	≥10%	≥15%	≥25%	≥33%	≥40%	≥50%	≥60%	≥75%
% PID								
Sensitivity (%)	95.3	89.3	80.4	**70.6**	61.6	52.6%	39.6	22.7
Specificity (%)	31.7	44.1	54.8	**67.3**	74.4	82.2%	94.0	97.5
Accuracy (%)	81.5	79.5	74.8	**69.8**	64.4	59.0%	51.4	38.8
% Max TOTPAR								
Sensitivity (%)	97.2	96.0	78.9	**75.3**	66.9	42.3	34.3	9.2
Specificity (%)	29.6	32.4	67.3	**77.8**	85.5	94.4	95.1	99.4
Accuracy (%)	78.6	78.4	75.7	**76.0**	72.0	56.7	54.6	34.0
	<1	<2	<3	<4	<5	<6		
PID (raw change)								
Sensitivity (%)	95.2	**77.4**	57.7	40.2	23.5	18.5		
Specificity (%)	31.0	**60.1**	79.7	93.6	97.9	99.8		
Accuracy (%)	81.3	**73.7**	62.5	51.8	39.5	14.6		
SPID (Sum of PID/hr)								
Sensitivity (%)	92.7	**71.8**	48.6	33.8	19.6	10.6		
Specificity (%)	36.1	**64.2**	83.1	91.7	96.6	98.9		
Accuracy (%)	76.4	**69.6**	58.5	50.5	41.8	36.0		
Pain Relief								
Sensitivity (%)	95.7	**71.7**	35.5	8.5				
Specificity (%)	28.7	**77.1**	97.5	100.00				
Accuracy (%)	80.9	**72.9**	47.1	28.6				
Global Medication								
Sensitivity (%)	89.4	**71.6**	42.2	14.3				
Specificity (%)	53.3	**87.2**	97.9	99.7				
Accuracy (%)	81.7	**74.9**	54.1	32.5				

[a]Using all available data on breakthrough pain episodes in titration Phase 1 (134 patients, 1307 episodes of which 281 required an additional dose of rescue medication).

PID = pain intensity difference; %PID = percent pain intensity difference (i.e., 100 × PID/Baseline pain); SPID = sum of the pain intensity difference per hour; %MAX TOTPAR = percent maximum total pain relief.

Table 2
Agreement Between Various Pain Measures and Patients Not Needing to Take an Additional Rescue Medication: Receiver Operating Characteristic (ROC) Area Under the Curve Analysis

Scale	Area (30 min)	Area (60 min)
%PID (percent)	0.773	0.862
PID (raw)	0.770	0.855
Pain Relief	0.801	0.839
%Max TOTPAR	N/A	0.840
GMP	N/A	0.839

PID = pain intensity difference; %PID = percent pain intensity difference (i.e., 100 × PID/Baseline pain); %Max TOTPAR = percent maximum total pain relief; GMP = global medication performance.

this second data set confirm the cut-off points of 33% or greater change in %PID or %Max TOTPAR, or a moderate or better rating for pain relief (PR score of ⩾2), or global medication performance (GMP score of ⩾2) are best associated with not needing an "additional dose of rescue medication."

In considering clinical trials of pain therapy, especially phase III efficacy trials, benefit should be demonstrated not only to a level the researcher thinks is appropriate, but also to a level important to the patients who would use the treatment, as well as the clinicians who are treating those patients. In an acute pain setting, the action of a patient not taking an additional treatment can be considered an easily measured and clinically appropriate indicator of the patient's perception of an effective initial therapy. However, direct measurement of this variable is possible only in situations where the onset of action of the medication is rapid enough for the subject to notice an improvement in a short period of time. In other situations, different mechanisms of action and longer times to onset of action make the measurement of the need for additional rescue therapy more difficult. In such situations, our data show that cut-off points described above in the change in pain intensity, pain relief, and global medication performance can serve as surrogate measures.

This analysis was possible because of a combination of features of the disease process and pain treatment. The study design of the OTFC clinical trial has several features that are important to consider. First, the study medication, fentanyl, is known to be effective for the treatment of breakthrough cancer pain.[10] Second, the onset of action of the medications is well known to be rapid.[5,6] Third, the intermittent and recurrent nature of breakthrough pain allowed for multiple treatments in the same patient over a relatively short period of time. Fourth, the study included a run in period during which patients had their dose of OTFC titrated to find a single dose strength that worked 75% of the time. As such, we have episodes adequately treated and inadequately treated on all patients against which to compare the change in pain intensity, pain relief, and the global medication performance in a similar fashion to the previous paper.[3]

Therefore, in this study the patient's perception of efficacy (i.e., not requiring an additional dose of rescue medication for that episode of breakthrough pain) is a clinically appropriate and easily measured outcome. The need for additional rescue medication to treat a specific episode of breakthrough pain is action-oriented, can be objectively measured, and reflects each patient's determination of the overall efficacy of the study medication. However, this approach is not often an option outside of specific acute pain models. More commonly used are pain intensity, pain relief, and a global measure of change. The high degree of agreement between pain intensity change, pain relief, and GMP, with the need for an additional dose of rescue medication in this study indicates that these measures are valid surrogate outcome measures.

Our study also provides evidence for the level of change, within subject, that should be considered clinically important. Although it is necessary to recognize that there is disagreement among researchers as to the need to define a clinically important difference at the level of the individual, if we assume that at least some questions being addressed in clinical trials are best answered by an analysis of the proportion of responders, an appropriate cut-off point needs to be defined to separate responders and non-responders. In most cases, the change in pain score that defines a clinically important difference for the individual patient is the most appropriate cut-off value. Our current findings are consistent with a number of analgesic studies in which the percentage change used as the definition of a positive outcome ranging from 33% to 50%.[11,12]

The potential limitations of this study must also be considered. Factors other than pain level

may play a role in a patient's decision to take an additional dose of rescue medication. Many clinically relevant factors such as mood, expectation, and learning affect a patient's report of pain. However, we would expect these factors to remain constant for each individual over the two-week study period. Because each patient contributes both rescue and non-rescue episodes, inter-individual differences should not affect the outcome. The results of the current analysis could also be questioned because of limitations inherent in the original study.[4] All randomized clinical trials raise concerns about generalizability. The response to pain and its meaning in cancer patients may be different than for patients with other painful conditions. In addition, cancer-related breakthrough pain is often self-limited. This is highlighted by the observation that no additional dose of rescue medication was required in 66% of the episodes treated with placebo units in the randomized phase.[10] Thus, there may be concerns as to the applicability of this model to clinical trials of other types of acute or chronic pain or pain that requires longer duration of treatment. Ideally, these values would be calculated for each population of patients before conducting the actual clinical trial. It is reassuring to note the similarities between our values and the ones that have been successfully used in clinical trials of pain and analgesics in several distinctly different conditions.[12–15]

In conclusion, we have used data from a clinical trial that uses a clinically relevant and easily measured indicator of the cancer patient's assessment of analgesic efficacy to evaluate the agreement between different pain scales and this clinically relevant outcome. Our findings validated, as optimal, the cut-off points of 33% for %PID and %Max TOTPAR and 2 for PR, SPID, and GMP to be optimal, consistent with previously published studies.[3] Although the choice of scales and analyses for future clinical trials will depend on the study question and design, this information should lead to a more standardized approach to the design of studies of pain treatments and better validity, comparability, and clinical utility of the results of future clinical trials.

Acknowledgments

This study was supported in part by grants from Cephalon, Inc., West Chester, PA, and NIH KO8 NS1865 and RO1 CA73797.

References

1. Bromm B. The measurement of pain in man. In: B. Bromm, ed. Pain Measurement in Man. Amsterdam: Elsevier Science Publishers, 1984: 13–33.

2. Turk DC, Rudy TE, Sorkin BA. Neglected topics in chronic pain treatment outcome studies: determination of success. Pain 1993;53:3–16.

3. Farrar JT, Portenoy RK, Berlin JA, et al. Defining the clinically important difference in pain outcome measures. Pain 2000;88:287–294.

4. Coluzzi PH, Schwartzberg L, Conroy JD, et al. Breakthrough cancer pain: a randomized trial comparing oral transmucosal fentanyl citrate (OTFC) and morphine sulfate immediate release (MSIR). Pain 2001;91:123–130.

5. Lichtor JL, Sevarino FB, Joshi GP, et al. The relative potency of oral transmucosal fentanyl citrate compared with intravenous morphine in the treatment of moderate to severe postoperative pain. Anesth Analg 1999;89:732–738.

6. Streisand JB, Varvel JR, Stanski DR, et al. Absorption and bioavailability of oral transmucosal fentanyl citrate. Anesthesiology 1991;75:223–229.

7. Hanley JA. Receiver operating characteristic (ROC) methodology: the state of the art. Crit Rev Diag Imag 1989;29:307–335.

8. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983; 148:839–843.

9. Zeger SL, Liang KY, Albert PA. Models for longitudinal data: a generalized estimating equation approach. Biometrics 1988;44:1049–1060.

10. Farrar JT, Cleary J, Rauck R, et al. Oral transmucosal fentanyl citrate: randomized, double-blinded, placebo-controlled trial for treatment of breakthrough pain in cancer patients. J Natl Cancer Inst 1998;90:611–616.

11. Goldsmith CH, Boers M, Bombardier C, Tugwell P. Criteria for clinically important changes in outcomes: development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. OMERACT Committee. Rheumatology 1993;20:561–565.

12. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomized controlled trials of analgesics. Pain1996;66:229–237.

13. Sheiner LB. A new approach to the analysis of analgesic drug trials, illustrated with bromfenac data. Clin Pharmacol Ther 1994;56:309–322.

14. Max MB, Lynch SA, Muir J, et al. Effects of desipramine, amitriptyline, and fluoxetine on pain in diabetic neuropathy. N Engl J Med 1992;326:1250–1256.

15. Leijon G, Boivie J. Central post-stroke pain—a controlled trial of amitriptyline and carbamazepine. Pain 1989;36:27–36.