Comparison of Multiple Against Single Pain Intensity Measurements in Complex Regional Pain Syndrome Type I: Analysis of 54 Patients

Tymour Forouzanfar, M.D., Marius Kemler, M.D., Ph.D., Alfons G. H. Kessels, M.D., Albere J. A. Köke, M.Sc., Maarten van Kleef, M.D., Ph.D., and Wilhelm E. J. Weber, M.D., Ph.D.

Department of Anesthesiology, Pain Management and Research Center, University Hospital Maastricht, Maastricht, The Netherlands

Abstract:

Objective: To describe the comparison of multiple and single pain ratings in patients with complex regional pain syndrome type I (CRPS I).

Design: Correlation, agreement, and reliability analyses were performed between the average pain intensity measured 3 times a day over a course of 4 days and one single pain rating (designated the “recalled average” pain, as assessed by the patient) before treatment and at 1-, 3-, and 6-month periods after treatment.

Patients: The patient population consisted of 54 patients with CRPS I in a randomized trial.

Results: The results show that both measurements correlate and have excellent agreement. Furthermore, both ratings measure significant pain reduction after treatment; “recalled average” pain, however, reflects greater change in pain intensity.

Conclusions: In patients with CRPS I a single pain rating is an accurate predictor of the average pain measured by a multiple pain-rating test. Moreover, both assessments are accurate enough to determine changes in pain over time with an effective treatment.

Key Words: Agreement analyses—Complex regional pain syndrome type I—Multiple pain rating—Single pain rating.

Complex regional pain syndrome type I (CRPS I), formerly known as reflex sympathetic dystrophy, is defined as a post-traumatic syndrome with pain, not related to the territory of a single nerve and disproportionate to the inciting event.1 No specific test is available to diagnose this syndrome. The diagnosis of CRPS I therefore has to be based solely on subjective measurements of the clinical symptoms, which include sensory abnormalities, diminished strength, hyperhidrosis, skin coloring changes, and pain.2 In clinical studies on CRPS I, pain assessment is a crucial measurement. In most of these studies a single pain rating is used as the primary outcome measure, under the assumption that this is equivalent to multiple ratings.3–9

In the assessment of pain, pain reports tend to change over time, even during the course of a day.10–12 Jensen et al.13,14 have shown that in patients with chronic pain a single pain intensity rating was the least reliable, whereas 3 measures of pain intensity a day over the course of 4 days showed excellent internal consistency and validity. Because multiple assessments are impractical both in the clinical and research setting, Dworkin et al.15 suggested that patients may be able to assess their own average pain levels over a period of time simply by asking them to rate their pain “on average” at a single point in time. For patients with back pain a single rating of pain “on average”
is an accurate estimate of “actual average” pain intensity measured over a course of 7 days. There are different methods to compare clinical measurements. According to Bland and Altman, such a comparison should be based on the differences between the two measurements performed on the same subject. The mean differences would be the estimated bias—the systematic difference between measurements. Further, the standard deviation (SD) of the differences would measure random fluctuations around this mean. In addition, the 95% limits of agreement (mean differences plus or minus 1.96 SDs) can be used to determine how far apart measurements by the two assessments are likely to be for most persons.

The present study was designed to investigate whether a single pain rating in patients with CRPS I can replace a multiple pain-rating test. For this purpose, we performed correlation and agreement analyses according to Bland and Altman between the average pain intensity measured 3 times a day over a course of 4 days and one single pain measurement of pain “on average” in 54 patients with CRPS I. Then we analyzed the accuracy of both measurements in assessing changes in pain ratings to determine the efficacy of pain treatment.

MATERIALS AND METHODS

Patients
The data for this study consisted of the pain intensity assessments measured in 54 patients with CRPS I (17 men, 37 women; mean age, 39 years; mean CRPS I duration, 38 months; Table 1).

The patients were included in a randomized trial in which spinal cord stimulation plus physical therapy was compared with physical therapy alone. The results of this trial were described elsewhere. Diagnostic criteria for CRPS I developed by the International Association for the Study of Pain were used. In addition to these, patients had to show impaired function and extension of symptoms outside the area of trauma.

Procedure
All patients were asked to rate their pain, 3 times a day over a course of 4 days, on a 100-mm visual analog scale (VAS), which was anchored by 2 extremes of pain: “no pain” on the left and “the worst possible pain” on the right end. At the end of the fourth day they were asked to note their average pain of the last week using a VAS with the aforementioned extremes. The patients rated their pain before treatment and 1 month, 3 months, and 6 months after treatment.

Statistics
The pain intensity measured with the single VAS is referred to as “recalled average” pain, whereas the calculated average pain intensity noted over 4 days is the “actual average” pain. For comparison of these two entities, the equivalent statistic, Pearson’s R, was used. Correlation coefficients (R) were defined as follow: 0.5 to 0.6, moderate; 0.6 to 0.8, good; 0.8 to 0.10, excellent. To assess the agreement between the two measurements we compared the difference of these two pain measurements for each patient with the average of both measurements for each follow-up period, as described by Bland and Altman. The limit of agreement (95% confidence level) is defined as the mean differences of two measurements plus or minus 1.96 SD (M ± 1.96 SD). In assessing the accuracy of pain assessments in measuring treatment effects, the t values of the Student t test from before treatment to each follow-up were examined (e.g., a larger t value or lower p value as indicating greater accuracy).

RESULTS

“Recalled average” pain versus “actual average” pain
In Table 2 the correlation coefficients and the parameters of the agreement are listed for each follow-up period. The “recalled average” pain correlated excellently with the “actual average” pain, with correlation coefficients 0.81, 0.94, 0.95, and 0.95 for pretreatment period, 1-, 3-, and 6-month follow-up, respectively.

The plot presented in Figure 1 shows the correlation between “recalled average” and “actual average” pain.
Because the plots of the pretreatment, 1-, 3- and 6-month follow-up periods are similar, only the plot of 1-month follow-up is presented.

Because of the similarities of the results of the agreement analysis, only the analysis of the 1-month follow-up will be discussed. This analysis showed a mean difference, “recalled average” pain minus “actual average” pain, of −0.02 and a SD of 0.67. The lower 95% limit is \(-1.33 \text{ (} -0.02 - [1.96 \times 0.67]\text{)}\) and the upper limit is \(1.29 \text{ (} -0.02 + [1.96 \times 0.67]\text{)}\). In other words, for 95% of the subjects “recalled average” pain will be between 1.33 below and 1.29 above the “actual average” pain, as shown in Figure 2. Thus, there is a high degree of agreement between both measurements.

**FIG. 1.** The “recalled average” pain plotted against “actual average” pain for 1-month follow-up.

**FIG. 2.** Differences between “recalled average” pain and “actual average” pain plotted against their mean rating for 1-month follow-up; the horizontal lines are the mean differences and the limits of agreement.

### Accuracy in assessing treatment effects

In Table 3 the pain intensity for the pretreatment and the follow-up periods recording the measurement is listed. There were no significant differences between “actual average” and “recalled average” pain assessed on each measurement period. In addition, both measurements show pain intensity reduction during the follow-up periods.

Table 4 shows the t and p values from before treatment to each follow-up. The analysis showed a significant t value (all \(p < 0.001\)) at 3- and 6-month follow-up. Based on the size of the t value, the “recalled average” pain tended to reflect greater change than the “actual average” pain.

**TABLE 3.** The mean pain intensity for the pretreatment and the follow-up periods

<table>
<thead>
<tr>
<th>Actual average</th>
<th>1-month</th>
<th>3-month</th>
<th>6-month</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.9 ± 1.4</td>
<td>6.8 ± 1.9</td>
<td>5.5 ± 2.9</td>
<td>5.2 ± 2.9</td>
</tr>
<tr>
<td>Recalled average</td>
<td>7.0 ± 1.4</td>
<td>6.8 ± 1.9</td>
<td>5.5 ± 3.1</td>
</tr>
</tbody>
</table>

Values are mean ± SD.

### DISCUSSION

In clinical studies with patients with CRPS I, pain is usually assessed by a single pain rating. This rating is consequently used as the primary outcome measure, under the assumption that it is equivalent to multiple pain ratings. However, to our knowledge this assumption has never been investigated in patients with CRPS I. The present study was undertaken to compare the validity of the single pain ratings in patients with CRPS I with multiple pain-rating tests. To this end we first analyzed whether a single pain rating (“recalled average” pain) can replace multiple pain ratings (“actual average” pain) in a group of patients with CRPS I. Next, we investigated the accuracy of both ratings measuring a treatment outcome. Further, the reliability (defined as stability over time) of both assessments was investigated.

The results show that both measurements correlate excellently. There is a high degree of agreement between

### TABLE 4. t values for change in pain intensity from the pretreatment to each follow-up period

<table>
<thead>
<tr>
<th>Actual average</th>
<th>Pretreatment to 1-month follow-up</th>
<th>Pretreatment to 3-month follow-up</th>
<th>Pretreatment to 6-month follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.10 ((p = 0.26))</td>
<td>4.0 ((p &lt; 0.001))</td>
<td>4.8 ((p &lt; 0.001))</td>
<td></td>
</tr>
<tr>
<td>Recalled average</td>
<td>1.8 ((p = 0.08))</td>
<td>4.4 ((p &lt; 0.001))</td>
<td>5.4 ((p &lt; 0.001))</td>
</tr>
</tbody>
</table>

\(p\) values are in parentheses.
both methods. Furthermore, both ratings measure significant pain reduction after treatment; however, “recalled average” pain reflects greater change in pain intensity.

The present study has several shortcomings. First, our patient population comprised only 54 patients. Second, the two pain ratings are not equivalent in their time frames. The patients were asked to provide a rating of their average pain over the preceding week. This measure was compared with the actual average pain of the preceding 4 days. However, Jensen et al.\(^\text{13}\) showed that increasing the number of assessments beyond 3 times daily for 4 consecutive days provides only minimal improvement in the psychometric properties of the measurement.

Having in mind that multiple measurements in outcome studies are impractical, Dworkin and Siegfried\(^\text{15}\) suggested that patients may be able to measure their average pain intensity over a period of time simply by asking them to rate their pain on average at a single point in time. When asking patients to rate their pain on average over the previous week, one assumes that patients can in fact recall their pain levels over this period of time with a high degree of accuracy. Several investigators have studied the accuracy of recall of pain intensity in chronic pain patients. This has yielded conflicting results, ranging from high levels of accuracy\(^\text{10}\) and no effect of current pain intensity on memory for past pain,\(^\text{22}\) to overestimated or underestimated pain levels depending on pain status at the time of recall.\(^\text{23}\) All studies have relied on correlation and prediction statistics in analyzing the relation between actual pain levels and recall of pain. According to Bolton et al.,\(^\text{16}\) correlation or regression statistics do not provide a real assessment of accuracy. These authors used both agreement statistic (intraclass correlation coefficient) and correlation statistic to investigate the accuracy of recalled pain in patients with back pain.\(^\text{16}\) The results of their study show that a single rating of pain “on average” is an accurate estimate of “actual average” pain intensity over a recording interval of 7 days.\(^\text{16}\)

Our results are in line with the results of Bolton et al.\(^\text{16}\) Despite the shortcomings of the present study, we conclude that, as in back pain patients, a single rating of pain “on average” (“recalled average” pain) is an accurate predictor of the actual “average pain” in patients with CRPS I. Moreover, both pain ratings proved to be accurate enough to determine reliable changes in pain over time.

REFERENCES