Discrepancies in Describing Pain: Is There Agreement Between Numeric Rating Scale Scores and Pain Reduction Percentage Reported by Patients with Musculoskeletal Pain After Corticosteroid Injection?

Daniel Cushman MD,* Zachary McCormick MD,* Ellen Casey MD,* and Christopher T Plastaras MD†

*The Rehabilitation Institute of Chicago/Northwestern McGaw Medical Center, Department of Physical Medicine and Rehabilitation, Chicago, Illinois; †Hospital of the University of Pennsylvania, Department of Physical Medicine and Rehabilitation, Philadelphia, Pennsylvania, USA

Reprint requests to: Daniel Cushman, MD, 345 East Superior St Suite 1600, Chicago, IL 60611, USA. Phone: 312-238-1250; Fax: 312-238-1219; E-mail: dcushman@ric.org.

Abstract

Objective. Pain intensity is commonly rated on an 11-point Numerical Pain Rating Scale which can be expressed as a calculated percentage pain reduction (CPPR), or by patient-reported percentage pain reduction (PRPPR). We aimed to determine the agreement between CPPR and PRPPR in quantifying musculoskeletal pain improvement at short-term follow-up after a corticosteroid injection.

Design. Retrospective cohort study.

Results. We examined 197 subjects with greater than 3/10 pain who underwent first-time fluoroscopically-guided corticosteroid injections. Ninety-three subjects reported higher PRPPR than CPPR values, and 41 subjects reported higher CPPR values. The CCC between CPPR and PRPPR was 0.44 (95% CI 0.35–0.54), with a precision of 0.54 and an accuracy of 0.81, and 95% limits of agreement ranging between −24% and +173%. Values for CCC, precision, and accuracy were higher for males compared to females and were highest in the youngest age group (18–40) and lowest in the middle age group (41–60).

Conclusions. PRPPR may not agree with CPPR at 3 week follow-up, as these individuals tend to report a higher estimated percentage improvement compared to the value calculated from their pain scores.

Key Words. Pain Measurement; Musculoskeletal Pain; Radiculopathy; Memory; Acute Pain; Chronic Pain

Introduction

Quantifying the subjective experience of pain is challenging but important in measuring treatment outcomes.
Discrepancies in Describing Pain

Table 1

**Primary analysis: Subjects with reported improvement in pain on follow-up (n = 150)**

<table>
<thead>
<tr>
<th>Age (years), mean ± SD</th>
<th>59 ± 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age range</td>
<td>21–89</td>
</tr>
<tr>
<td>Percentage of subjects 65 years or older</td>
<td>33%</td>
</tr>
<tr>
<td>Females (%)</td>
<td>62%</td>
</tr>
<tr>
<td>Follow-up time (days), mean ± SD</td>
<td>24 ± 31</td>
</tr>
<tr>
<td>Transforaminal epidural injections (% of total injections)</td>
<td>59%</td>
</tr>
<tr>
<td>Zygapophysial joint injections (% of total injections)</td>
<td>13%</td>
</tr>
<tr>
<td>Knee joint injections (% of total injections)</td>
<td>6.0%</td>
</tr>
<tr>
<td>Hip joint injections (% of total injections)</td>
<td>4.0%</td>
</tr>
<tr>
<td>Glenohumeral joint injections (% of total injections)</td>
<td>6.3%</td>
</tr>
<tr>
<td>All other injections combined (% of total injections, none greater than 3% individually)</td>
<td>13%</td>
</tr>
</tbody>
</table>

**Secondary analysis: subjects with reported unchanged pain on follow-up (n = 47)**

<table>
<thead>
<tr>
<th>Age (years), mean ± SD</th>
<th>53 ± 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age range</td>
<td>23–81</td>
</tr>
<tr>
<td>Percentage of subjects 65 years or older</td>
<td>24%</td>
</tr>
<tr>
<td>Females (%)</td>
<td>86%</td>
</tr>
<tr>
<td>Follow-up time (days), mean ± SD</td>
<td>22 ± 23</td>
</tr>
<tr>
<td>Transforaminal epidural injections (% of total injections)</td>
<td>66%</td>
</tr>
<tr>
<td>Knee joint injections (% of total injections)</td>
<td>11%</td>
</tr>
<tr>
<td>Zygapophysial joint injections (% of total injections)</td>
<td>6.3%</td>
</tr>
<tr>
<td>Glenohumeral joint injections (% of total injections)</td>
<td>6.3%</td>
</tr>
<tr>
<td>All other injections combined (% of total injections, none greater than 5% individually)</td>
<td>11%</td>
</tr>
</tbody>
</table>

that are important to patients. Pain intensity is commonly rated [1] on an 11-point scale, from 0 (no pain) to 10 (worst pain imaginable), with different names including the Numerical Pain Rating Scale (NRS) [2,3] and the Pain Intensity Numerical Rating Scale [4,5]. This method of pain rating has been extensively studied and appears to have sufficient discriminative power to express pain intensity [6]; it has been shown to be one of the most responsive commonly used measures of pain intensity [7] and is commonly used in the literature [8]. Clinically meaningful improvement on this integer scale has been defined as 2–3 points [1,4,5,9], although this varies based on the affected area of the body and on the disease process [10]. Improvement in pain described on the NRS can be expressed as a calculated percentage pain reduction (CPPR) as well. The CPPR is a calculated score from comparisons of preintervention and postintervention NRS scores by percent change (e.g., patient reports scores of 6 and 3, respectively, and is viewed as a three-point or 50% improvement) [8]. Clinically meaningful improvement measured by the CPPR as also been defined [1]. Despite its ubiquity and significant assessment in the literature, the application of the NRS and CPPR varies amongst clinicians, payors, and researchers. Many quantify patient’s pain using patient-reported percentage pain reduction (PRPR). The PRPR is a self-reported percent improvement from a preintervention score (e.g., patient reports a 50% improvement from their preintervention score of 6) [8]. Although the CPPR and the PRPR may seem to provide identical quantification of improvements in pain, this assumption has only been verified in one study largely involving the short-term treatment of subjects with acute post-traumatic and postoperative pain (96% of subjects) using intravenous opioids (98.5% of subjects) [8]. To the best of our knowledge, no study has investigated agreement of these two means of quantifying pain over a longer time frame. Thus, the aim of our study was to determine whether the CPPR and the PRPR can be used interchangeably as a means of quantifying pain improvement for musculoskeletal or radicular pain over a 3-week period.

**Methods**

This study was approved by the Northwestern University Institutional Review Board. We performed a retrospective review of 1,116 consecutive fluoroscopically guided corticosteroid injections performed at an outpatient physical medicine and rehabilitation musculoskeletal and spine center between November 2004 and January 2009. The charts were then examined to look for follow-up information including PRPR and CPPR on follow-up. For subjects who received multiple injections, only the first injection was included in the analysis (i.e., all study participants received only first-time injections at our center). The following injections were performed on at least five patients: glenohumeral, hip, knee, transforaminal epidural, and zygapophysyal joint. The following injections were performed on fewer than five patients: greater trochanteric bursa, interlaminar epidural, and subacromial bursa. Patients were asked their NRS pain score prior to the injection on a medical intake form, and were again asked at follow-up (a CPPR was calculated from these two values) on a separate intake form before seeing a physician; they were not reminded of their initial score when filling out the second intake form. The intake form asked the following question, “Since your last physician visit, are your symptoms better, worse, the same? If better, by how much on a scale of 0–100%? (if 0 was the way you were and 100% was completely normal).” Patients were all asked to follow-up about 2 weeks after the date of the injection; the actual average follow-up time and standard deviation are shown in Table 1. They were also asked for their PRPR at the follow-up visit.

As our question for PRPR was phrased as percentage improvement, we excluded all patients who worsened.
Subjects who did not report any of the three pain variables (pain prior to the injection, pain at follow-up, and percentage improvement) were excluded from the study. Only participants with moderate to severe pain (NRS scores 4 to 10) prior to injection were included. We performed a secondary analysis on subjects who reported that their pain was unchanged (i.e., PRPPR = 0).

Statistical Analysis

To detect a 10% difference between CPPR and PRPPR with an alpha error of 0.05 and a beta error of 0.20, we estimated that we would require 73 subjects. As this retrospective review was part of a separate study, unpublished [11], we included more participants as we had additional available data.

Primary data analysis was performed similar to the work of Cepeda et al. [8]. Analysis of continuous variables was performed by means and standard deviations. A concordance correlation coefficient (CCC) [12] was calculated to discern the level of agreement between CPPR and PRPPR. A perfect correlation (CPPR and PRPPR were identical for each subject) would return a CCC of 1 and was categorized as excellent (0.81–1.00), good (0.61–0.80), moderate (0.41–0.60), fair (0.21–0.40), and poor (<0.20). Precision was categorized as excellent (>0.75), good (0.60–0.74), fair (0.40–0.59), and poor (<0.40). Visual representation of the data was performed with a Bland–Altman plot [13], with upper and lower 95% limits of agreement (LOA). The 95% LOA represents the likelihood of differences between the two measures; in other words, the LOAs give 95% certainty that measurement differences will lay within their bounds. Subgroup analysis for comparisons of precision was performed by Fisher transformation of the Pearson coefficients (precision portion of the CCC) with subsequent chi-square comparison. Finally, as pain is often viewed as an ordinal variable, each CPPR and PRPPR value was converted to ordinal variables and their correlation was examined with Goodman–Kruskal gamma, which has been shown to be a close estimate of the degree of association for these variables [14].

Secondary data analysis was performed on all the subjects who reported unchanged pain (PRPPR = 0). As the same statistical analysis cannot be calculated for this data as PRPPR will exclusively be zero, a paired t-test was performed on these subjects, which looks only for similar mean responses (not correlation between the measurement tests).

Statistical calculations were performed with PSPP 0.8.2 (Gnu Project, Boston, MA), except for the CCC calculations, which were performed with SAS version 8 (SAS Institute, Cary, NC).

Results

Of the 1,116 injections, 595 were first-time injections. Of the first-time injections, 345 subjects followed up, and

Figure 1 Comparison of subjects’ PRPPR and CPPR from their numerical rating scores. The dashed line represents an ideal correlation while the solid line represents the best-fit of the study data.

263 included both a PRPPR and CPPR. Of those 263 subjects: 26 had worsened pain, 47 were unchanged, and 190 were improved (40 subjects initially had mild pain, which was an exclusion criteria). Table 1 displays demographic and procedural data for the subjects who fit the inclusion criteria with improved pain (n = 150) and the subjects with unchanged pain (n = 47). Three subjects (2.0%) followed-up more than 90 days postinjection; a subsequent analysis was run with these three subjects excluded, which did not appreciably alter any of the reported statistics. The difference between CPPR and PRPPR for subjects who presented with pain of at least 3 months duration was the same as patients whose pain was of a duration less than 3 months (P = 0.58).

Subjects with Reported Improvement in Pain

The CCC between CPPR and PRPPR was 0.44 (95% CI 0.35–0.54), with a precision of 0.54 and an accuracy of 0.81. Figure 1 compares CPPR and PRPPR; the dashed line represents perfect correlation between the two and the solid line represents a best fit for the given data. The solid line leans toward the x-axis, displaying a trend for subjects to have higher PRPPR values compared to CPPR. In other words, subjects tended to endorse larger improvement in pain when self-reporting a percentage compared to when reporting integers on a 0 to 10 scale.

Figure 2 demonstrates a Bland–Altman plot of the two measurements. The solid line at 0% on the y-axis demonstrates the ideal perfect agreement between the two measurements. The dashed line represents the mean difference found between PRPPR and CPPR (the over-estimation of a self-reported percentage) at +16% (95% confidence intervals +11% to 21%). The upper and
lower lines demonstrate the 95% LOAs between the two measures (−41% and 73%, respectively). The overall range of differences between measurements was −35% to 128%. Three subjects reported a decrease in pain by PRPPR but their CPPR demonstrated an increase; 14 subjects reported an increase by PRPPR with a CPPR of zero. There were 93 subjects that had higher PRPPR values than CPPR (overestimated their improvement compared to their NRS scores) and 41 subjects with CPPR values greater than PRPPR. The largest measurement discrepancies occurred when changes in pain reduction were small. When stratified by injection location, axial injections had a higher CCC (0.47 vs 0.24) with higher accuracy (0.81 vs 0.78) and precision (0.58 vs 0.31, \( P = 0.09 \)) compared to peripheral injections.

Table 2 demonstrates subgroup analyses performed on our study population. Males had higher values for the CCC, precision, and accuracy compared to females. The youngest age group, 18–40, had the highest values for CCC, with the greatest precision and accuracy (\( P = 0.13 \) for precision). This was followed by the oldest age group, 61+, and the lowest values were seen in the Age 41–60 group (\( P = 0.69 \) between groups for precision).

When converted to ordinal variables, comparison of CPPR and PRPPR can be seen in Figure 3. It demonstrates that the highest direct group-to-group agreement was the 1–50% groups, which agreed 59.5% of the time, and the lowest direct agreement was the 100% relief group, which agreed only 9.62% of the time. The Goodman–Kruskal gamma value for all groups was 0.70.

### Subjects with Unchanged Pain (PRPPR = 0%)

A total of 47 subjects fit the inclusion criteria and noted no change in pain. All subjects are shown in the first line for comparison. A total of 12 subjects had a negative CPPR (pain worsened on NRS), 18 had positive CPPR values (pain improved on NRS), and 17 had zero CPPR values (unchanged pain on NRS). The mean difference between PRPPR and CPPR was 3.2%, \( P = 0.50 \), CI (−13%, +6.3%), with a range from −100% to +83%. The 95% LOAs for the measurements were −62% to 68%.

### Discussion

Our data show only a fair-to-moderate association between PRPPR and CPPR on the NRS for an approximately 3-week follow-up after corticosteroid injection. There was moderate accuracy between the two methods with fair precision; in other words, patients as a whole had a reasonable ability to use the two methods interchangeably, but there was large individual variation. We found that patients tend to report a significantly higher PRPPR compared to the CPPR, and that this difference has a large range (95% of the time, it will be within a −41% and +73% difference). Patients reported a larger PRPPR than CPPR value more than twice as often (i.e., patients overestimated their improvement by percentage compared to the percentage calculated by NRS scores). This suggests that patients are not able to accurately assess their percent improvement compared to their preintervention pain score.

Our findings contrast with the work of Cepeda et al. [8], who demonstrated a good overall agreement between the two methods of measurement. The clinical importance of this discrepancy between patient populations, to the authors, is meaningful—based on our findings, clinicians should expect a large amount of variability if using the rating systems interchangeably. Although Cepeda et al. found a 95% LOA of −12% to +17%, ours was much larger (−41% to +73%). The likely cause of this discrepancy between studies is the difference in time for follow-up between the two groups. In the study by Cepeda et al., subjects rated their
perceived percent improvement compared to pain that they experienced minutes prior, whereas in our study, subjects had to recall what their pain level was approximately 3 weeks prior to perform the mental calculation of percent improvement postintervention. Given this increase in the time between initial and follow-up pain rating, patients likely did not remember their initial pain level as accurately. Studies that have examined the accuracy of subjects’ memory of pain have demonstrated inaccurate recall after as little as 1–2 days [15]. The inaccuracy of these memories may be related to variable expectations of pain, bias toward recall of unrepresentative brief painful moments or by associated emotions, and current pain intensity [16–21]. Studies in which patients were asked to recall their pain 1–2 weeks prior demonstrated overestimation of prior pain [22,23]. Additionally, our study population included subjects with a history of chronic pain, which has been associated with overestimation of prior pain [24]. Other possible causes for the different outcomes could include the more subsacute patient population, increased anxiety prior to an injection, and the possible improvement in mood over the longer term. Subjects may have also noted a significant improvement in their pain, greater than 20% [1], which could have been reported as a higher PRPPR—an example of this would be a subject saying they are “99% better” even though they still have pain.

It appeared that men had better agreement between their PRPPR and CPPR compared to women. This also contrasts with the previous work of Cepeda et al. [8]. Even though men’s reports of pain seemed to be better-correlated, they still showed relatively poor precision, with a relatively large variation. Middle-aged subjects demonstrated the lowest agreement between percentages, primarily due to poor precision (a large variation amongst subjects), although there was no statistical difference between groups.

Importantly, patients may view pain more commonly in an ordinal manner (e.g., unchanged, mild, moderate, severe, etc.), rather than on a continuous scale. Furthermore, there are often “responders” and “non-responders” after therapeutic injections [25]. When the PRPPR and CPPR are compared as ordinal variables, there are relatively few exact matches between subjects’ CPPR and PRPPR, as can be seen in Figure 3, although the level of association overall is good. Interestingly, this likely suggests that patients’ PRPPR and CPPR trend in a similar direction, although notably, this ordinal analysis does not provide information related to the variability of their PRPPR. One important implication of this finding is that subjects with less pain relief appear to be more accurate than those with greater relief of pain, although this could be a byproduct of smaller percentage relief groups for high pain relief.

For comparison of subjects with unchanged pain on follow-up, where they had reported a PRPPR of 0%, subjects appeared to show good accuracy. We had a smaller sample size than our improved subjects, but still had a relatively small standard error resulting in tight 95% confidence intervals. This suggests that patients who report unchanged pain appear to be accurate in their

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Concordance Correlation Coefficient</th>
<th>Precision</th>
<th>Accuracy</th>
<th>Lower 95% Limit of Association</th>
<th>Upper 95% Limit of Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>(All subjects)</td>
<td>0.44</td>
<td>0.54</td>
<td>0.81</td>
<td>−41%</td>
<td>+73%</td>
</tr>
<tr>
<td>Males</td>
<td>0.66</td>
<td>0.70</td>
<td>0.95</td>
<td>−36%</td>
<td>+54%</td>
</tr>
<tr>
<td>Females</td>
<td>0.29</td>
<td>0.38</td>
<td>0.75</td>
<td>−45%</td>
<td>+82%</td>
</tr>
<tr>
<td>Age 18–40</td>
<td>0.62</td>
<td>0.62</td>
<td>0.99</td>
<td>−47%</td>
<td>+46%</td>
</tr>
<tr>
<td>Age 41–60</td>
<td>0.20</td>
<td>0.27</td>
<td>0.74</td>
<td>−39%</td>
<td>+77%</td>
</tr>
<tr>
<td>Age 61+</td>
<td>0.42</td>
<td>0.54</td>
<td>0.78</td>
<td>−42%</td>
<td>+80%</td>
</tr>
</tbody>
</table>

Figure 3  Comparison of subjects’ PRPPR and CPPR, defined by groups of pain relief. The vertical axis refers to the percentage of all subjects in a PRPPR group who matched the corresponding CPPR group (for example, the first left/bottom column indicates that 33.3% of patients who reported a PRPPR as unchanged also had a CPPR that was unchanged). All matched columns show white border and have data labels above.

874
assessments. Precision, conversely, again appeared to be lower, as the limits of association (LOAs) were wide.

Our study is primarily limited by its retrospective nature with a 57% follow-up rate. We studied only subjects who returned for follow-up and included both a postinjection NRS score and PRPPR on their intake form, which introduces selection bias. The follow-up timing was quite variable as well, with a standard deviation of 31 days, although 98% of subjects followed-up within 3 months. A prospective study examining subjects at a specific time could reduce many of these biases. Additionally, the minimal clinically important level of change for reported improvement in pain differs based on disease process and area of the body affected [10], but we combined several processes/areas into our single analysis. The NRS is a global assessment of pain, and thus new types of pain present on follow-up could introduce further discrepancy between the two measurements (e.g., a patient with hip and knee pain may have a flare of their knee pain after a hip has been properly injected). Both CPPR and PRPPR have limitations in their own regards; CPPR may be more difficult for those who are not accustomed to discussing percentages or have poor math skills while PRPPR and CPPR both are widely multifactorial in origin and prone to numerous biases [26]. Moreover, treatment success is often viewed by patients in a discrete, ordinal (e.g., unchanged, mild, moderate, severe, etc.), or categorical (e.g., “poor response,” “good response,” etc.) manner, which can be inaccurately represented by a numerical measure such as the NRS. This study purely examined numerical pain rating and did not take into account other important variables in the overall measure of treatment outcome, such as return to work, decrease in medication use, improvement in function, improvement in quality of life, or decreased health care expenditures.

Conclusion

PRPPR may not agree with CPPR at 3 week follow-up after an injection for musculoskeletal or radicular pain, as these individuals tend to overestimate PRPPR. We do not recommend using these two methods interchangeably in this population.

References


8 Cepeda MS, Africano JM, Polo R, Alcala R, Carr DB. Agreement between percentage pain reductions calculated from numeric rating scores of pain intensity and those reported by patients with acute or cancer pain. Pain 2003;106:439–42.


16 Redelmeier DA, Kahneman D. Patients’ memories of painful medical treatments: Real-time and
Cushman et al.


