Brief Pain Inventory–Facial minimum clinically important difference

Sukhmeet K. Sandhu, BA,1 Casey H. Halpern, MD,1 Venus Vakhshori, BS,1 Keyvan Mirsaedi-Farahani, BS,1 John T. Farrar, MD, PhD,2 and John Y. K. Lee, MD1

Departments of 1Neurosurgery and 2Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania

OBJECT Neurosurgeons are frequently the primary physicians measuring pain relief in patients with trigeminal neuralgia (TN). Unfortunately, the measurement of pain can be complex. The Brief Pain Inventory–Facial (BPI-Facial) is a reliable and validated multidimensional tool that consists of 18 questions. It measures 3 domains of pain: 1) pain intensity (worst and average pain intensity), 2) interference with general activities of daily living (ADL), and 3) face-specific pain interference. The objective of this paper is to determine the patient-reported minimum clinically important difference (MCID) using the BPI-Facial.

METHODS The authors conducted a retrospective study of 234 patients with TN seen in a single neurosurgeon's office. Patients completed baseline and 1-month follow-up BPI-Facial questionnaires. The MCID was calculated using an anchor-based approach in which the defined anchor was the 7-point patient global impression of change (PGIC). Two statistical methods were employed: mean change score and optimal cutoff point.

RESULTS Using the mean change score method, the investigators calculated the MCID for the 3 domains of the BPI-Facial: 44% and 30% improvement in pain intensity at its worst and average, respectively, 54% improvement in interference with general ADL, and 63% improvement in interference with facial ADL. Using the optimal cutoff point method, they also calculated the MCID for the 3 domains of the BPI-Facial: 57% and 28% improvement in pain intensity at its worst and average, respectively, 75% improvement in interference with general ADL, and 62% improvement in interference with facial ADL.

CONCLUSIONS The BPI-Facial is a multidimensional pain scale that measures 3 domains of pain. Although 2 statistical methods were used to calculate the MCID, the optimal cutoff point method was the superior one because it used data from the majority of subjects included in this study. A 57% improvement in pain intensity at its worst and a 28% improvement in pain intensity at its average were the MCIDs for patients with facial pain. A greater improvement was needed to achieve the MCID for interference with general and facial ADL. A 75% improvement in interference with general ADL and a 62% improvement in interference with facial ADL were needed to achieve an MCID. While pain intensity is easier to measure, pain's interference with ADL may be more important for patient outcomes when designing or evaluating interventions in the field of TN. The BPI-Facial is a useful instrument to measure changes in multidimensional aspects of pain in patients with TN.

http://thejns.org/doi/abs/10.3171/2014.8.JNS132547

KEY WORDS trigeminal neuralgia; facial pain; Brief Pain Inventory; Brief Pain Inventory–Facial; minimum clinically important difference; functional neurosurgery

Neurosurgeons are often critically involved in the diagnosis, management, and surgical treatment of patients with facial pain. Neurosurgeons have decades of experience in the management of such patients, but the measurement of pain remains a challenge. Part of the challenge is that pain is a unique experience that no other person can feel or perceive on one’s behalf. Even if a group of individuals receives the same stimuli or undergoes the same intervention, there is a large range in reported pain ratings.11,32 The subjective nature of pain makes it difficult to study in a quantifiable manner, and it requires the use of patients’ verbal accounts and memo-
Thus, it is important to use patient-reported outcomes (PROs) in the measurement of pain relief following treatment for facial pain disorders.

The majority of neurosurgical studies on trigeminal neuralgia (TN) pain have not used PROs and make no attempt to quantify pain. For example, many published studies that have evaluated pain intensity outcomes following neurosurgical intervention used either a 3- or 5-point scale to measure pain, and these scales relied on the physician's assessment of pain relief and not on the patient's reporting. In addition, a literature review of surgical studies on TN revealed that 221 of 222 studies had not measured pain preoperatively, and therefore a comparison between baseline and postoperative pain could not be made. Despite decades of neurosurgical management of TN, quantifiable benefits are difficult to obtain. Recognizing the lack of quantifiable outcome measures in the study of TN and facial pain disorders, the senior author (J.Y.K.L.) adapted the Brief Pain Inventory (BPI) to evaluate patients with facial pain. The original BPI is an instrument that has been thoroughly vetted as a reliable and valid instrument. It has been translated into over 75 languages and has been used successfully in hundreds of clinical trials. The senior author's appended version of the BPI was named the Brief Pain Inventory–Facial (BPI-Facial). In 2010, the senior author published a study describing the instrument and carefully analyzing and defining its psychometric properties. In that study, the BPI-Facial was only administered at the initial visit; change score before and after intervention were not examined. In this current study, administration of the BPI-Facial before and after intervention provides insight into pain outcomes in a standardized format.

The challenge when using PROs is interpreting the change in instrument score and determining what magnitude of change should be considered clinically meaningful. Changes in pain intensity or interference with daily activities may result in statistically significant changes that are not clinically relevant. The concept of the minimum clinically important difference (MCID) has been defined by several authors, such as Jaeschke and colleagues, who defined it as the “smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient’s management.” This study investigates the MCID of the BPI-Facial change score in a cohort of patients who were evaluated by the senior author for neurosurgical intervention.

Methods

Patient Sample

This is a single center, institutional review board–approved, retrospective review of 234 patients seen between January 2006 and December 2011. The analysis includes data from patients who were evaluated by a single surgeon for TN. Patients seen in the office were asked to complete a paper questionnaire that included the BPI-Facial. After review of history, physical examination, and imaging, the senior surgeon (J.Y.K.L.) categorized patients’ facial pain according to the Burchiel classification system and International Headache Society. Only patients with a diagnosis of TN (Type 1 or 2) were included in this study.

Outcome Measures

BPI-Facial

The BPI-Facial (Appendix 1) is composed of 18 questions that cover 3 domains of pain: 1) 4 questions that address pain intensity using the Numerical Rating Scale (NRS), 2) 7 questions that address pain interference with general activities of daily living (ADL) (general interference items), and 3) 7 questions that address pain interference with face-specific daily activities (facial interference items). The BPI-Facial uses an 11-point Likert scale, ranging from 0 to 10 for all 18 items.

Pain Intensity

Pain intensity was measured for the 4 items of the NRS using a Likert scale of 11 choices, ranging from 0 (“no pain”) to 10 (“pain as bad as you can imagine”). The 4 items are as follows: 1) pain at its worst, 2) pain at its least, 3) pain at its average, and 4) pain right now.

Although pain intensity was calculated at its worst, least, average, and at present, we only present data for NRS-worst and NRS-average pain, as this provides the greatest face validity. Hence, pain intensity is presented as 2 variables: 1) worst intensity of pain and 2) average intensity of pain.

General Interference Items

Interference with general ADL was measured for 7 different conditions using a Likert scale of 11 choices, ranging from 0 (“does not interfere”) to 10 (“completely interferes”). The 7 items are as follows: 5) general activity, 6) mood, 7) walking ability, 8) normal work (includes both work outside the home and housework), 9) relations with other people, 10) sleep, and 11) enjoyment of life.

Although interference in each aforementioned activity was calculated, a composite general interference score was determined by averaging the 7 components of the scale.

Facial Interference Items

Interference with face-specific activities was measured for 7 different conditions using a Likert scale of 11 choices, ranging from 0 (“does not interfere”) to 10 (“completely interferes”). The 7 items are as follows: 12) eating a meal, 13) touching your face (including grooming), 14) brushing or flossing your teeth, 15) smiling or laughing, 16) talking, 17) opening your mouth widely, and 18) eating hard foods like apples.

Although interference for each activity was calculated, a composite facial interference score was determined by averaging the 7 components of the scale.

Patient Global Impression of Change

In addition to completing the BPI-Facial at the time of their follow-up visit, patients completed the Patient Global Impression of Change (PGIC; Appendix 2). Follow-up visits were typically 30 days after the patient’s procedure. During that visit, patients were not shown their initial BPI-Facial questionnaires; therefore, they did not know the specific score they had chosen before treatment. However,
since the follow-up period was short, patients were more likely to recall their preoperative pain.

The PGIC is a 7-point Likert scale of overall change in which patients are asked to choose 1 of 7 descriptors that best describes how their symptoms have improved or worsened. The 7 categorical options are as follows: 1) very much improved, 2) much improved, 3) minimally improved, 4) no change, 5) minimally worse, 6) much worse, and 7) very much worse.

Data Analysis

There are many ways to calculate the minimal clinically important difference (MCID). One method is the anchor-based approach, where an external criterion, also known as the anchor, is used to compare the change observed in a PRO tool. In this study, the anchor was the PGIC, and the PRO tool was the BPI-Facial. Two statistical methods were used to compute the MCID using the anchor-based approach: mean change score and optimal cutoff point. Missing data were not imputed and all calculations were restricted to completed data elements. All statistical analyses were performed using STATA version 10 for Windows.

Mean Change Score: Raw and Percentage Change

For each patient, the raw and percentage change score was calculated for the 3 domains of the BPI-Facial: pain intensity, composite general interference, and composite facial interference. To calculate the raw change score, each patient’s follow-up score was subtracted from his/her baseline score. Hence, positive change scores reflect a decrease in pain intensity and a decrease in pain interference. The percentage change was calculated by dividing the raw change score by the baseline score.

Optimal Cutoff Point

The MCID was calculated using the optimal cutoff point method by plotting receiver operating characteristic (ROC) curves. ROC curves have been used as a tool to measure the classification performance of diagnostic or screening tests, namely, the ability to distinguish diseased from healthy individuals. The diagnostic accuracy of the tool is determined by the area under the ROC curve (AUC). The AUC was also calculated. The AUC represented the probability that change scores correctly discriminated patients who were “better” from those who were “not better.” According to Hosmer and Lemeshow, an AUC of 0.7–0.79 was considered to be acceptable discrimination, 0.8–0.9 to be excellent discrimination, and greater than 0.9 to be outstanding discrimination.

Results

Patient Characteristics

Of the 234 subjects included in this analysis, 207 (88%) subjects completed majority of the 18-item BPI-Facial questionnaire at their baseline and follow-up visit, as well as the PGIC at the follow-up visit. The mean age of the subjects in the study was 62.1 years, and 62% of the subjects were female. The mean follow-up period between baseline and follow-up visit was 3.7 months, and the mean time between surgical procedure and follow-up visit was 1.4 months. Characteristics of the patients are summarized in Table 1.

The majority of patients were diagnosed with Burchiel Type 1 TN (79%); the remaining patients were diagnosed with Burchiel Type 2 TN. In addition, the V2 and V3 distributions were the most commonly affected distributions of pain, and just more than half of patients had right-sided pain (56%).

Two hundred thirty patients underwent at least one of the following neurosurgical interventions: microvascular decompression (46%), Gamma Knife radiosurgery (47%), and percutaneous glycerol rhizotomy (51%) (Table 2). Four patients who did not undergo a procedure attended follow-up visits at which they completed the BPI-Facial questionnaire. Those 4 patients were included in this analysis because their pain should theoretically not have changed and thus their scores should not have changed.

Patient Distribution of the PGIC

Of the 207 subjects, 77 characterized themselves as either “very much improved” or “much improved” (Fig. 1). Fourteen percent of all subjects had characterized themselves as having “no change,” being “minimally worse,” or being “much worse.” No patient rated himself or herself as “very much worse.”
Calculating the MCID
Mean Change Score: Raw and Percentage

The raw and percentage change scores were calculated for the 3 domains of the BPI-Facial (pain intensity, composite general interference, and composite facial interference) (Tables 3 and 4). Box plot analysis revealed that the average intensity of pain had the smallest distribution of change scores on the BPI-Facial, while the worst intensity of pain had the largest distribution of change score (Fig. 2).

The raw and percentage changes for the worst and average pain intensity were calculated for each of the 7 categories of the PGIC (Table 3). This allows the readers to choose the appropriate MCID for their own particular study or area of interest—since researchers with an interest in a medication with minimal side effects may choose a less stringent criterion for MCID, whereas researchers with an interest in a new invasive surgical procedure may choose a more stringent criterion for MCID. For the purposes of this study, the mean change score method of MCID calculation was determined by patients who reported “much improved” outcomes on the PGIC. The MCID was a 44% reduction in worst pain intensity and a 30% reduction in average pain intensity (Table 3).

Since pain was not only measured on an intensity scale but also with respect to how it interfered with ADL, the MCID for the composite general interference score and composite facial interference score of the BPI-Facial questionnaire was also calculated. Table 4 provides the average raw and percentage change score for each category of the PGIC. Patients who reported “much improved” outcomes on the PGIC determined the MCID for this method. The MCID was a 54% reduction in general interference and a 63% reduction in facial interference. In comparison with pain intensity, a greater percentage improvement was needed for both general and facial interference items for patients to rate their change as clinically significant.

Optimal Cutoff Point
We present the selected cutoff points (MCID) for the 3 domains of the BPI-Facial (pain intensity, composite general interference, and composite facial interference) in raw and percentage format in Tables 5 and 6. We present the results using the intermediate stringency model (patients who were “very much improved” and “much improved”). The MCID was a 57% reduction in worst pain intensity and a 28% reduction in average pain intensity. This result is similar to the mean change method of calculating the MCID (refer to Table 7). The optimal cutoff point method was also used to calculate the MCID for the composite general interference score and the composite facial interference score (Fig. 3). The same 3 models of varying stringency were used to generate ROC curves for the composite general and facial interference scores (Table 6). We present the results using the intermediate stringency model. The MCID was a 75% reduction in composite general interference and a 62% reduction in composite facial interference. Similar to the mean change score method, the optimal cutoff point method yielded MCID values that were higher for the interference items than for the pain intensity items (Table 7).

Discussion
This study provides quantitative data using a reliable and validated outcome tool in the measurement of pain in patients with TN before and after intervention. Patients self-rated their own improvement in this study, and the change scores are based on the patient’s perception of clinical relevance—the MCID. This is the first paper in

<table>
<thead>
<tr>
<th>TABLE 1. Patient characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Total no. of patients</td>
</tr>
<tr>
<td>No. of females</td>
</tr>
<tr>
<td>Mean age ± SD (yrs)</td>
</tr>
<tr>
<td>No. of op-treated patients</td>
</tr>
<tr>
<td>V1 distribution</td>
</tr>
<tr>
<td>V2 distribution</td>
</tr>
<tr>
<td>V3 distribution</td>
</tr>
<tr>
<td>No. w/ lt-sided pain (%)</td>
</tr>
<tr>
<td>No. w/ bilat pain</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2. Diagnosis stratified by treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Burchiel Type 1 TN</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Burchiel Type 2 TN</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
the neurosurgical literature that defines the MCID using a well-validated tool such as the BPI-Facial. It should help to set a standard by which future studies assess facial pain. Many working groups have suggested that outcome studies in pain should use multidimensional outcome tools. For example, the multimethod Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) defined 6 core domains to be considered in treatment trials of chronic pain: 1) pain intensity, 2) physical functioning, 3) emotional functioning, 4) participant ratings of improvement and satisfaction with treatment, 5) symptoms and adverse events, and 6) participant adherence to treatment regimen. This study investigated 3 domains of pain: pain intensity, interference with general ADL, and interference with face-specific ADL. The percentage threshold change required for patients to rate themselves as improved was higher for the interference items than for the 2 intensity items alone (Table 7). For example, a 30% improvement in pain intensity at its average was enough for patients to rate themselves as “much improved,” but a 30% improvement in pain interference with either general or facial ADL was insufficient for patients to rate themselves as “much improved.” Thus, although it is the severity of pain that leaves a lasting impression on practitioners who see TN patients in the office, it is the effect of that pain on an individual’s ADL that they want cured. Hence, the MCID was 57% for pain intensity at its worst and 28% for pain intensity at its average, but the MCIDs for the composite general and facial interference scores were 75% and 62%, respectively. This supports the concept that pain is multidimensional and cannot be simply measured on intensity alone.

There is one article in the available literature that presents data on the MCID in TN patients. Reddy and colleagues presented data on 60 patients who underwent microvascular decompression. These authors used the visual analog scale and the Barrow Neurological Institute Pain Scale and calculated the MCID using the anchor-based approach. Unfortunately, the visual analog scale is a 1-question measure of pain intensity and does not test additional dimensions of pain. In addition, the Barrow scale has not been validated for use as an outcome tool, although it is widely used. Perhaps because of a small sample size, Reddy et al. found widely varying results for the MCID and, in an effort to reconcile the large variance in the results, the authors simply averaged the high and low scores. This method is circumspect, considering that each method employs a different calculation technique and the methods are not statistically equivalent in weight or importance. Nevertheless, the authors do provide insight into some of the difficulties in measuring pain before and after intervention.

TABLE 3. Raw and percentage change of score for the pain intensity domain of the BPI-Facial: worst and average pain intensity*

<table>
<thead>
<tr>
<th>PGIC</th>
<th>NRS: Raw Change</th>
<th>NRS: Percentage Change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Worst</td>
<td>Average</td>
</tr>
<tr>
<td>Very much improved</td>
<td>7.0 ± 0.3</td>
<td>2.7 ± 0.2</td>
</tr>
<tr>
<td>Much improved</td>
<td>4.1 ± 0.6</td>
<td>1.8 ± 0.2</td>
</tr>
<tr>
<td>Minimally improved</td>
<td>1.9 ± 1.0</td>
<td>1.1 ± 0.3</td>
</tr>
<tr>
<td>No change</td>
<td>2.0 ± 1.0</td>
<td>0.5 ± 0.6</td>
</tr>
<tr>
<td>Minimally worse</td>
<td>0.3 ± 0.6</td>
<td>0.6 ± 0.4</td>
</tr>
<tr>
<td>Much worse</td>
<td>-1.0 ± 2.1</td>
<td>-0.3 ± 0.6</td>
</tr>
<tr>
<td>Very much worse</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

† The percentage change in this example is heavily influenced by 4 individuals. One individual rated pain at the worst as 1 but at next follow-up described their pain as 10. The percentage change for this individual is [(1 − 10) / 100] / 1, which is equal to a −900% change. Hence, the values are significantly skewed by this despite the raw change being on average only 6.9.

* Values are presented as the mean ± SD. Boldface type represents the “suggested” threshold for the MCID.

TABLE 4. Raw and percentage change of scores for two of the domains of the BPI-Facial questionnaire: composite general interference and composite facial interference*

<table>
<thead>
<tr>
<th>PGIC</th>
<th>General Interference</th>
<th>Facial Interference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw Change</td>
<td>Percentage Change</td>
</tr>
<tr>
<td>Very much improved</td>
<td>5.5 ± 0.3</td>
<td>92.3 ± 2.1</td>
</tr>
<tr>
<td>Much improved</td>
<td>3.4 ± 0.6</td>
<td>53.5 ± 10.6</td>
</tr>
<tr>
<td>Minimally improved</td>
<td>2.9 ± 0.8</td>
<td>35.0 ± 12.1</td>
</tr>
<tr>
<td>No change</td>
<td>2.3 ± 0.7</td>
<td>41.6 ± 12.4</td>
</tr>
<tr>
<td>Minimally worse</td>
<td>0.5 ± 1.1</td>
<td>0.68 ± 17.9</td>
</tr>
<tr>
<td>Much worse</td>
<td>-1.6 ± 1.1</td>
<td>-20.1 ± 23.0</td>
</tr>
<tr>
<td>Very much worse</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

* Values are presented as the mean ± SD. Boldface type represents the “suggested” threshold for the MCID.
**Brief Pain Inventory–Facial minimal clinically important difference**

**Table 5. Raw and percentage change optimal cutoff points determined by the pain intensity domain of the BPI-Facial: worst and average pain intensity**

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Raw/Percentage Change</th>
<th>Model</th>
<th>Cutoff Point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>AUC</th>
<th>Percentage Correct (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Worst pain intensity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw</td>
<td>Very much improved</td>
<td>7</td>
<td>65.4</td>
<td>83.3</td>
<td>0.81</td>
<td>73.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Much &amp; very much improved</td>
<td>6</td>
<td>65.5</td>
<td>90.7</td>
<td>0.83</td>
<td>71.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimally, much, &amp; very much improved</td>
<td>4</td>
<td>71.3</td>
<td>82.1</td>
<td>0.82</td>
<td>73.0</td>
<td></td>
</tr>
<tr>
<td>Percentage</td>
<td>Very much improved</td>
<td>70</td>
<td>83.3</td>
<td>78.6</td>
<td>0.87</td>
<td>81.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Much &amp; very much improved</td>
<td>57</td>
<td>75.9</td>
<td>88.4</td>
<td>0.87</td>
<td>78.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimally, much, &amp; very much improved</td>
<td>40</td>
<td>79.0</td>
<td>82.1</td>
<td>0.85</td>
<td>79.4</td>
<td></td>
</tr>
<tr>
<td><strong>Average pain intensity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw</td>
<td>Very much improved</td>
<td>2.5</td>
<td>66.3</td>
<td>75.3</td>
<td>0.75</td>
<td>70.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Much &amp; very much improved</td>
<td>2</td>
<td>65.7</td>
<td>81.0</td>
<td>0.78</td>
<td>69.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimally, much, &amp; very much improved</td>
<td>1.5</td>
<td>73.0</td>
<td>70.4</td>
<td>0.79</td>
<td>72.6</td>
<td></td>
</tr>
<tr>
<td>Percentage</td>
<td>Very much improved</td>
<td>40</td>
<td>82.6</td>
<td>83.1</td>
<td>0.87</td>
<td>82.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Much &amp; very much improved</td>
<td>27.8</td>
<td>83.6</td>
<td>82.9</td>
<td>0.88</td>
<td>83.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimally, much, &amp; very much improved</td>
<td>27.8</td>
<td>77.6</td>
<td>88.5</td>
<td>0.86</td>
<td>79.3</td>
<td></td>
</tr>
</tbody>
</table>

* Cutoff points were calculated using the Youden index. Boldface type represents the "suggested" threshold for the MCID.

**Figure 2.** Box plots for the raw change score for the 3 domains of the BPI-Facial. Each domain is stratified by the 7 PGIC categories. Average pain intensity has the smallest change in score in almost every category of the PGIC. The "much improved" group has a wide distribution of scores for the 3 domains. Dots represent extremes and hatches represent quartiles.
The results that we present appear to have internal consistency. Raw change scores are presented alongside percentage change scores, since studies have shown that a 1-point decrease in pain from 3 to 2 (33%) is more significant than a 1-point decrease in pain from 10 to 9 (10%). Additional research is needed to determine the underlying importance of each specific question of the BPI-Facial questionnaire and its correlation to TN. Nevertheless, we are encouraged because the AUC for the percentage values were larger than the raw value AUC (Tables 5 and 6).

We omitted 2 measures of pain intensity: “pain at its least” and “pain right now.” “Pain at its least” was excluded from the analysis because the values were similar to “pain at its average,” and it did not contribute any significant data to the study. “Pain right now” was excluded from the analysis because scores can vary greatly and may be confounded by many factors, including what the patient was doing before they filled out the survey.

The results that we present appear to have internal consistency. Raw change scores are presented alongside percentage change scores, since studies have shown that a 1-point decrease in pain from 3 to 2 (33%) is more significant than a 1-point decrease in pain from 10 to 9 (10%).

One caveat is that we have presented the average of 7 general interference items and the average of 7 facial interference items. Although the raw averaged numbers can be compared, we are less confident about the validity of presenting the percentage scores of the averaged 7 items because their statistical equivalence cannot be assured.

### Methodology

Of the 2 methods used in this paper for calculating the MCID, the optimal cutoff point method with ROC curve analysis was the superior method because it used all available data and maximized classification; therefore, this method was used to determine the threshold MCID values for this study. The mean change score method of calculating the MCID was simple to perform, but it ignored a significant amount of data. For example, if we had defined the “much improved” group as the MCID, only 47 patients would have been used to calculate the mean change score. The mean change score method would have discarded the change scores for all the other patients who had rated themselves in another category of the PGIC. In contrast, ROC curve analysis was performed utilizing the change scores for every patient to correctly classify patients who were “better” and those who were “not better.” Hence, the preferred method for calculation of the MCID was the optimal cutoff point method. Despite this, we recognize that sensitivity and specificity in this study ranged from 65% to 93% and that it was not uniformly greater than 90%. However, the AUC ranged from 0.85 to 0.9 for the percentage change scores for the 3 components of the BPI-Facial questionnaire. The AUC values indicated the BPI-Facial questionnaire’s ability to discriminate patients who were “better” from those who were not. According to Hosmer-Lemeshow criteria, our values are indicative of excellent discrimination between “better” and “not better” patients.

The results that we present appear to have internal consistency. Raw change scores are presented alongside percentage change scores, since studies have shown that a 1-point decrease in pain from 3 to 2 (33%) is more significant than a 1-point decrease in pain from 10 to 9 (10%).

One caveat is that we have presented the average of 7 general interference items and the average of 7 facial interference items. Although the raw averaged numbers can be compared, we are less confident about the validity of presenting the percentage scores of the averaged 7 items because their statistical equivalence cannot be assured.

### Future Directions

This study has practical value in an era of comparative effectiveness analysis. Future interventions in patients with TN and facial pain may need to specify criterion for success after medical or surgical intervention. For example, a prospective study may have to calculate a power analysis and sample size with an a priori plan to detect a 30% change in average pain intensity and a 70% change in composite general interference score and composite facial interference score. In this way, a new procedure or new treatment for patients with refractory facial pain can be validated and hopefully approved. Fortunately, patients

### Table 6: Raw and percentage change optimal cutoff points determined by two domains of the BPI-Facial: composite general interference and composite facial interference

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Raw/Percentage Change</th>
<th>Model</th>
<th>Cutoff Point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>AUC</th>
<th>Percentage Correct (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Composite general interference</strong></td>
<td></td>
<td>Very much improved</td>
<td>4</td>
<td>75.9</td>
<td>70.6</td>
<td>0.75</td>
<td>73.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Much &amp; very much improved</td>
<td>3.7</td>
<td>68.3</td>
<td>80</td>
<td>0.77</td>
<td>71.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimally, much, &amp; very much improved</td>
<td>2.7</td>
<td>75.2</td>
<td>77.3</td>
<td>0.81</td>
<td>75.5</td>
</tr>
<tr>
<td></td>
<td>Percentage</td>
<td>Very much improved</td>
<td>94.5</td>
<td>78.6</td>
<td>84.6</td>
<td>0.85</td>
<td>81.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Much &amp; very much improved</td>
<td>75</td>
<td>81.9</td>
<td>85.3</td>
<td>0.85</td>
<td>82.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimally, much, &amp; very much improved</td>
<td>70</td>
<td>78.3</td>
<td>90.5</td>
<td>0.84</td>
<td>80.0</td>
</tr>
<tr>
<td><strong>Composite facial interference</strong></td>
<td></td>
<td>Very much improved</td>
<td>4.7</td>
<td>79.3</td>
<td>73.9</td>
<td>0.81</td>
<td>77.0</td>
</tr>
<tr>
<td></td>
<td>Percentage</td>
<td>Much &amp; very much improved</td>
<td>2.28</td>
<td>80.0</td>
<td>71.9</td>
<td>0.84</td>
<td>78.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimally, much, &amp; very much improved</td>
<td>3.7</td>
<td>73.1</td>
<td>90.9</td>
<td>0.85</td>
<td>75.7</td>
</tr>
<tr>
<td></td>
<td>Raw</td>
<td>Very much improved</td>
<td>85.5</td>
<td>92.9</td>
<td>74.6</td>
<td>0.89</td>
<td>85.0</td>
</tr>
<tr>
<td></td>
<td>Percentage</td>
<td>Much &amp; very much improved</td>
<td>61.5</td>
<td>87.2</td>
<td>83.3</td>
<td>0.90</td>
<td>86.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimally, much, &amp; very much improved</td>
<td>45.8</td>
<td>88.9</td>
<td>85.7</td>
<td>0.88</td>
<td>88.4</td>
</tr>
</tbody>
</table>

* Cutoff points were calculated using the Youden index. Boldface type represents the “suggested” threshold for the MCID.

### Table 7: Summary of percentage MCID using “much improved” as anchor criterion

<table>
<thead>
<tr>
<th>BPI-Facial Domain</th>
<th>Mean Change (percentage)</th>
<th>Optimal Cutoff (percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worst pain intensity</td>
<td>44</td>
<td>57</td>
</tr>
<tr>
<td>Average pain intensity</td>
<td>30</td>
<td>27.8</td>
</tr>
<tr>
<td>Composite general interference</td>
<td>54</td>
<td>75</td>
</tr>
<tr>
<td>Composite facial interference</td>
<td>63</td>
<td>61.5</td>
</tr>
</tbody>
</table>
with TN have had great success with microvascular decompression, Gamma Knife treatment, and glycerol rhizotomy; however, patients with atypical facial pain continue to suffer from a lack of validated treatment options.

Limitations

One of the difficulties of studying subjective data, like pain, is that statistically significant differences between numeric pain scores may yield scientifically valid results, but those results may only provide negligible benefit for patients. Recognizing this problem, many investigators have described methods to calculate the MCID. The anchor-based method has a common flaw—that is, the reliance on a single question or anchor—which in our case is the PGIC. The use of a single anchor to validate study results has been challenged. For example, the PGIC may be confounded by variables such as complications, comorbidities, and cognitive status. However, the PGIC has been well validated, and many investigators have used this single scale as an anchor or measure of comparison for other outcome tools. The PGIC also takes other factors into account that patients consider when rating their overall improvement, such as the following: sleep disturbance, cognitive dysfunction, depressive symptoms, and physical functioning. Despite its limitations, the PGIC, we believe, is a valid anchor for calculation of the MCID.

Recall bias is a significant concern in a study like this. The present study was intentionally designed to study time points that were in close proximity to minimize recall bias. Authors of many clinical studies on TN pain have been interested in the maximum follow-up time so that patients’ experiences can be studied longitudinally. However, subjects have difficulty recalling their prior state of health; they tend to forget their pain intensity and interference over time. In addition, the scores they select on follow-up PRO measures are often derived from their current state of health, since they have difficulty recalling their baseline scores. Thus, patients may inaccurately report...

**FIG. 3.** ROC curves for the 3 domains of the BPI-Facial. The AUC for the 4 respective ROC curves are 0.83, 0.78, 0.77, and 0.83. The AUCs represent the probability of correctly discriminating patients who are “better” from those who are “not better.” The scores fall within acceptable to excellent ranges.
their perceived progress. Crosby et al. had found that over an extended period of time, global rating scores might be flawed because scores might inaccurately portray the patient’s true clinical change.\textsuperscript{8} We believe that by minimizing follow-up time, subjects would be able to provide more accurate estimates of change in pain intensity and interference. Since the goal of the study was to define properties of the BPI-Facial instrument, longitudinal changes over time and the treatment received are less important than the characteristics of the specific instrument under study. Future studies will address long-term change in the BPI-Facial results.

Lastly, another limitation of the study is its generalizability. Our prioritized patients with TN undergo surgery, since these patients have been referred for neurosurgical consult. Only 4 patients treated with medical management were included in this study. Future research could include a larger cohort of medically managed TN patients, which will improve generalizability of the BPI-Facial for this population.

Conclusions

In conclusion, we hope that this paper will set a standard for pain outcome research in neurosurgery for patients with TN. It is important to use PRO measures, like the BPI-Facial questionnaire, to evaluate pain intensity and how pain interferes with ADL. The BPI-Facial instrument is a multidimensional tool that appears to accomplish these tasks. We calculated the MCID for the 3 domains of the BPI-Facial: 57% and 28% improvement in pain intensity at its worst and average, respectively, 75% improvement in interference with general ADL, and 62% improvement in interference with facial ADL. The BPI-Facial is sensitive to change and measures 3 domains of facial pain. We encourage practitioners to employ the BPI-Facial in their outcomes studies of these challenging but often grateful patients.

Appendix 1

The Brief Pain Inventory–Facial

Circle the ONE number that describes how, during the past week, pain has interfered with your:

1. General activity
   \begin{tabular}{cccccccccc}
   0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
   \\
   Does not interfere & Completely interferes \\
   \end{tabular}

2. Mood
   \begin{tabular}{cccccccccc}
   0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
   \\
   Does not interfere & Completely interferes \\
   \end{tabular}

3. Walking ability
   \begin{tabular}{cccccccccc}
   0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
   \\
   Does not interfere & Completely interferes \\
   \end{tabular}

4. Normal work (includes both work outside the home and housework)
   \begin{tabular}{cccccccccc}
   0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
   \\
   Does not interfere & Completely interferes \\
   \end{tabular}

5. Relations with other people
   \begin{tabular}{cccccccccc}
   0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
   \\
   Does not interfere & Completely interferes \\
   \end{tabular}

6. Sleep
   \begin{tabular}{cccccccccc}
   0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
   \\
   Does not interfere & Completely interferes \\
   \end{tabular}

7. Enjoyment of life
   \begin{tabular}{cccccccccc}
   0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
   \\
   Does not interfere & Completely interferes \\
   \end{tabular}

8. Eating a meal
   \begin{tabular}{cccccccccc}
   0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
   \\
   Does not interfere & Completely interferes \\
   \end{tabular}

9. Touching your face (including grooming)
   \begin{tabular}{cccccccccc}
   0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
   \\
   Does not interfere & Completely interferes \\
   \end{tabular}

10. Brushing or flossing your teeth
    \begin{tabular}{cccccccccc}
    0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
    \\
    Does not interfere & Completely interferes \\
    \end{tabular}

11. Smiling or laughing
    \begin{tabular}{cccccccccc}
    0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
    \\
    Does not interfere & Completely interferes \\
    \end{tabular}

12. Talking
    \begin{tabular}{cccccccccc}
    0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
    \\
    Does not interfere & Completely interferes \\
    \end{tabular}

13. Opening your mouth widely
    \begin{tabular}{cccccccccc}
    0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
    \\
    Does not interfere & Completely interferes \\
    \end{tabular}

14. Eating hard foods like apples
    \begin{tabular}{cccccccccc}
    0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
    \\
    Does not interfere & Completely interferes \\
    \end{tabular}
Circle the ONE number that describes your pain at its WORST in the last week.

No pain

Circle the ONE number that describes your pain at its LEAST in the last week.

No pain

Circle the ONE number that describes your pain at its AVERAGE in the last week.

No pain

Circle the ONE number that describes your pain RIGHT NOW.

No pain

Appendix 2
Patient Global Impression of Change (PGIC)

Choose one of these 7 descriptions that best describes how your symptoms have improved or worsened.

1 – Very much improved
2 – Much improved
3 – Minimally improved
4 – No change
5 – Minimally worse
6 – Much worse
7 – Very much worse

References


**Author Contributions**

Conception and design: Lee, Sandhu, Halpern. Acquisition of data: Sandhu, Vakhshori, Mirsaedi-Farahani. Analysis and interpretation of data: Lee, Sandhu, Halpern, Vakhshori, Farrar. Drafting the article: Lee, Sandhu, Halpern, Farrar. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Lee. Statistical analysis: Lee, Sandhu, Halpern. Administrative/technical/material support: Sandhu. Study supervision: Lee, Sandhu, Halpern.

**Correspondence**

John Y. K. Lee, Department of Neurosurgery, University of Pennsylvania, 235 South Eighth St., Philadelphia, PA 19106. email: leejohn@uphs.upenn.edu.