Craniofacial Pain and Disability Inventory (CF-PDI): Development and Psychometric Validation of a New Questionnaire

Roy La Touche, PT, MSc1,3, Joaquín Pardo-Montero PhD1,3, Alfonso Gil-Martinez PT, MSc1,4, Alba Paris-Alemany PT, MSc2,3, Santiago Angulo-Díaz-Parreño MSc2,3, Juan Carlos Suárez-Falcón PhD6, Manuel Lara-Lara, MD5, and Josué Fernández-Carnero, PT, PhD2,8

Background: Orofacial pain, headaches, and neck pain are very common pain conditions in the general population and might be associated in their pathophysiology, although this is not yet clarified. The development and validation of a prediction inventory is important to minimize risks. Most recent questionnaires have not focused on pain, but pain is the common symptom in temporomandibular disorders, headaches, and neck pain. It is necessary to provide tools for these conditions.

Objectives: The purpose of this study is to present the development and analysis of the factorial structure and psychometric properties of a new self-administered questionnaire (Craniofacial Pain and Disability Inventory [CF-PDI]) designed to measure pain, disability, and functional status of the mandibular and craniofacial regions.

Study Design: Multicenter, prospective, cross-sectional, descriptive survey design. A secondary analysis of the reliability of the measures was a longitudinal, observational study.

Setting: A convenience sample was recruited from a hospital and 2 specialty clinics in Madrid, Spain.

Methods: The study sample consisted of 192 heterogeneous chronic craniofacial pain patients. A sub-sample of 106 patients was asked to answer the questionnaire a second time, to assess the test-retest reliability. The development and validation of the CF-PDI were conducted using the standard methodology, which included item development, cognitive debriefing, and psychometric validation. The questionnaire was assessed for the following psychometric properties: internal consistency (Cronbach’s α); floor and ceiling effects; test-retest reliability (Intraclass Correlation Coefficient [ICC]; Bland and Altman method); construct validity (exploratory factor analysis); responsiveness (standard error of measurement [SEM] and minimal detectable change [MDC]); and convergent validity (Pearson correlation coefficient), by comparing visual analog scale (VAS), the Tampa Scale for Kinesiophobia (TSK-11), the Pain Catastrophizing Scale (PCS), the Neck Disability Index (NDI), and the Headache Impact Test-6 (HIT-6). Multiple linear regression analysis was used to estimate the strength of the associations with theoretically similar constructs.

Results: The final version of the CF-PDI consists of 21 items. Exploratory factor analysis revealed 2 factors (“pain and disability” and “jaw functional status”), both with an eigenvalue greater than one, explaining 44.77% of the variance. Floor or ceiling effects were not observed. High internal consistency of the CF-PDI (Cronbach’s α: 0.88) and also of the 2 subscales (Cronbach’s α: 0.80 – 0.86) was confirmed. ICC was found to be 0.90 (95% confidence interval [CI] 0.86 – 0.93), which was considered to be excellent test-retest reliability. The SEM and MDC were 2.4 and 7 points, respectively. The total CF-PDI score showed a moderate correlation with most of the assessed questionnaires (r = 0.36 – 0.52) and a strong correlation with the NDI (r = 0.65; P < 0.001). The NDI, VAS, and TSK-11 were predictors of CF-PDI.

Limitations: Only self-reported measures were considered for convergent validity. Future research should use physical tests to explore the clinical signs relating to pain and disability.
Conclusion: The CF-PDI showed good psychometric properties. Based on the findings of this study, the CF-PDI can be used in research and clinical practice for the assessment of patients with craniofacial pain.

Key words: Craniofacial pain, temporomandibular disorders, headache, neck pain, disability, development, questionnaire, reliability, psychometric validation, minimal detectable change

Pain Physician 2014; 17:95-108

Chronic orofacial pain and temporomandibular disorders (TMD) are commonly associated but may also arise from other sources (1). Orofacial pain is a common pain condition associated with the hard and soft tissues of the face and mouth. Its prevalence in the general population is approximately 13% (2). Headache and neck pain are also 2 of the most common symptoms seen in the general population (3,4).

TMD, headaches, and neck pain are related diseases and share signs and symptoms (5-7). Some clinical evidence of the interconnection between the cervical spine and TMD has been demonstrated (8). Plesh et al (9) showed that 53% of patients with TMD had severe headache and 54% had neck pain. Besides, 59% with TMD reported at least 2 comorbid pains, and women reported more comorbid pain than men (9). This relationship between headache and a causative disorder is a criterion for secondary headache diagnoses (10).

Although it has been suggested that TMD and headaches may be related in their pathophysiology (7,11) and that headache could be a possible risk factor for the development of neck pain (12), the pathophysiological mechanisms underlying these pain conditions are still not fully clarified. However, a biopsychosocial approach to the etiology, assessment, and treatment of chronic pain is widely advocated (13).

Nearly 60% of both men and women reported recent pain of moderate-to-severe intensity, with a quarter of them indicating interference or termination of work-related activities (14). Therefore, the correct diagnosis of these diseases is very important to reduce their huge economic impact (15,16).

A useful scale is the Jaw Functional Limitation Scale (JFLS), which consists of 3 constructs comprising a total of 20 items identified along a global scale (17). At present, there are no questionnaires in Spanish to assess these characteristics. This fact is especially relevant considering that Spain is one of the European Union countries with a high cost for these disabilities (18).

Moreover, most questionnaires have focused on disability or dysfunction but not on pain, even though pain is the common symptom in TMD, headaches, and neck pain. Additionally, pain is been addressed by other validated scales (19).

For clinical practice and research, it is necessary to have tools to measure neck pain and the associated disability (20). In addition, the development and validation of a prediction inventory allows the minimization of risks and helps prevent the development of the disease.

The purpose of this study is to present the development and analysis of the factorial structure and psychometric properties of a new self-administered questionnaire (Craniofacial Pain and Disability Inventory [CF-PDI]) designed to measure pain, disability, and functional status of the mandibular and craniofacial regions.

Methods

The development and validation of the CF-PDI was conducted in a standardized manner, using an accepted measure development methodology that included 3 phases (21):

a) item development and identification of domains;

b) pilot testing on a small number of patients with cognitive debriefing; and

c) psychometric validation.

Item Development

Items were generated through a multi-step process (21):

1) literature review;
2) patient interviews and focus group;
3) examination by the research group;
4) item writing and selection; and
5) examination of the inventory by independent experts.

The relevant scientific literature search was conducted using electronic databases (Medline, Embase, CINAHL). The extracted information was related to the diagnosis, pathophysiology, comorbidities, and psycho-
Psychometric Validation of the Craniofacial Pain and Disability Inventory

social and disability factors associated with craniofacial pain. We found 5 published questionnaires that assess orofacial pain and jaw function (22-26). All of them were only validated in the English language. On the basis of the existing literature, a semi-structured interview guide was developed focusing on the following 3 main areas:

1) perception of physical and psychosocial health in relation to craniofacial pain;
2) patient-perceived physical impacts of the condition (including impacts on general physical functioning and specific jaw function); and
3) perception of disability and pain associated with their condition.

A total of 13 patients with chronic craniofacial pain underwent the semi-structured interview and 5 patients with the same condition participated in the focus group. Both processes ended with the question: “Do you think there are any other aspects of craniofacial pain we have not discussed?” The research group proceeded to analyze and compare the information extracted from semi-structured interviews, focus group, and review of the relevant literature to generate the construct concept of the CF-PDI and subsequently to write the items. A list of 30 draft items was generated. The research group then selected 22 items based on a finely structured consensus process (27) to not omit any necessary concepts.

The 22 items of the inventory were subjected to an external assessment by a group of experts in craniofacial pain (3 physiotherapists, one dentist, and one medical doctor). The 5 experts assessed whether each of the items had a relationship with the conditions of craniofacial pain and TMD, through a 3-level Likert scale (complete disagreement, neither agreement nor disagreement, and complete agreement).

Cognitive Debriefing
Cognitive debriefing of the preliminary CF-PDI was conducted with a small number of patients to assess their interpretations of the questions (24 patients with craniofacial pain in the pilot test). Patients were selected from 3 different educational levels (primary school, secondary school, and university) and the total response time for all items of the CF-PDI was calculated. These patients were asked to complete the preliminary CF-PDI, and were then interviewed about its comprehensiveness, relevance, and clarity of expression. This led to some minor alterations to the questionnaire.

Psychometric Validation

Sample/Patients
This study employed a prospective, cross-sectional, descriptive design. A consecutive convenience sample was recruited from outpatients of the Hospital Universitario La Paz (Madrid, Spain) and 2 private clinics specializing in craniofacial pain and TMD (Madrid, Spain). Patients were selected if they met all of the following criteria: 1) headache and facial pain, the diagnosis of which was made according to the guidelines of the International Classification of Headache Disorders (10); 2) headache or facial pain attributed to TMD (10), the diagnosis of which was based on the Research Diagnostic Criteria for TMD (28,29) to classify patients with painful TMD (myofascial pain, temporomandibular joint [TMJ] arthralgia, or TMJ osteoarthritis); 3) pain history of at least 6 months prior to the study; 4) at least 18 years of age; and 5) good understanding of the Spanish language. The exclusion criteria were as follows: cognitive impairment; the presence of psychiatric limitations that impede participation in the study assessments; and poor knowledge of the Spanish language. To assess the test-retest reliability of the CF-PDI, a sub-sample of 106 patients whose clinical conditions were stable were asked to answer the inventory a second time, after an interval of 12 days.

The study was conducted in accordance with the Declaration of Helsinki and was approved by the local ethics committee of the Hospital Universitario La Paz (PI-1241). Prior to their participation, subjects gave written informed consent.

After consenting to the study, recruited patients were given a battery of questionnaires to complete on the day of the visit. These included various self-reports for demographic and pain-related variables, including the CF-PDI to be validated, a visual analog scale (VAS) for pain intensity, and the validated Spanish versions of the Tampa Scale for Kinesiophobia (TSK-11), the Pain Catastrophizing Scale (PCS), the Neck Disability Index (NDI), and the impact associated with headache was assessed using the Headache Impact Test-6 (HIT-6). The sociodemographic questionnaire collected information about gender, date of birth, marital status, living arrangements, education level, and work status.

Pain intensity was measured with the VAS. The VAS consists of a 100 mm line, on the left side of which represents “no pain” and the right side represents “the worst pain imaginable.” The patients placed a mark on
the line where they felt best represented their pain intensity (30).

The Spanish version of the TSK-11 is a self-reported questionnaire that assesses fear of re-injury due to movement (31). The TSK-11 is an 11-item questionnaire that eliminates psychometrically poor items from the original version of the TSK (32) to create a shorter questionnaire with comparable internal consistency. The TSK-11 has a 2-factor structure: activity avoidance and harm, and has demonstrated acceptable psychometric properties (31).

The Spanish version of the PCS assesses the degree of pain catastrophization (33,34). The PCS has 13 items and a 3-factor structure: rumination, magnification, and helplessness. The theoretical range is between 0 and 52, with lower scores indicating less catastrophizing. The PCS has demonstrated acceptable psychometric properties (33).

The Spanish version of the NDI measures perceived neck disability (20,35). This questionnaire consists of 10 items, with 6 possible answers that represent 6 levels of functional capacity, ranging from 0 (no disability) to 5 (complete disability) points. The NDI has demonstrated acceptable psychometric properties (20).

The Spanish version of the HIT-6 (36,37) is a 6-item questionnaire that measures the severity and impact of headache on the patient’s life. The HIT-6 has demonstrated acceptable psychometric properties (38).

**Statistical Analysis**

Socio-demographic and clinical variables of the patients were analyzed. Analysis of variance (ANOVA) was used to test for differences in socio-demographic and clinical characteristics between the groups of patients.

Weighted kappa statistics (39) were calculated to assess the percentage agreement between external expert evaluators. Kappa statistics were calculated for each item. The Kappa coefficient varies from -1 (complete disagreement) to +1 (complete agreement), with 0 representing neither agreement nor disagreement.

**Factor Analysis**

The factor structure was investigated using an explorative factor analysis (ie, principal component analysis [PCA]) with Oblimin rotation. The number of factors for extraction was based on Kaiser’s eigenvalue criterion (eigenvalue ≥1) and evaluation of the scree plot (40). The quality of the factor analysis models was assessed using Bartlett’s test for sphericity and the Kaiser-Meyer-Olkin (KMO) test. Bartlett’s test is a measure of the probability that the initial correlation matrix is an identity matrix and should be < 0.05 (41). The KMO test measures the degree of multicollinearity and varies between 0 and 1 (should be greater than 0.50 – 0.60) (42).

**Reliability**

For reliability, internal consistency and reproducibility were examined. Internal consistency was estimated using Cronbach’s α and item total correlation coefficients. For a questionnaire to be internally consistent, α levels should be above 0.7 (43).

The test-retest reliability (repeatability) was evaluated using the Intraclass Correlation Coefficient (ICC). An ICC value above 0.70 is considered acceptable (44). We also constructed a Bland Altman Plot by calculating the mean difference between 2 measurements and the standard deviation (SD) of the difference (45). In this plot, 95% of the differences are expected to be less than 2 SDs.

**Floor and Ceiling Effects**

Potential floor and ceiling effects were measured by calculating the percentage of patients indicating the minimum or maximum possible scores in the questionnaires. Floor and ceiling effects are considered to be present if more than 15% of respondents achieved the highest or lowest possible total score (44).

**Responsiveness Analyses**

Measurement error is the systematic and random error of a patient’s score that is not attributable to true changes in the construct to be measured (46). Measurement error is expressed as a standard error of measurement (SEM), which is calculated as:

\[
SEM = \frac{SD}{\sqrt{1 - ICC}}
\]

where SD is the SD of values from all participants and ICC is the reliability coefficient (47,48). Ostelo et al (49) suggested that the percentage of the SEM in relation to the total score of a questionnaire is an important indicator of agreement, and can be interpreted as follows: ≤ 5% very good; > 5% and ≤ 10% good; > 10% and ≤ 20% doubtful; and > 20% negative. Responsiveness was assessed with the Minimal Detectable Change (MDC). The MDC expresses the minimal magnitude of change required to be 95% confident that the observed change between the 2 measures reflects real change and not just measurement error (50). It is calculated as

\[
MDC = SEM \times \sqrt{2} \times 1.96
\]
Convergent Validity
The convergent validity was assessed by the Pearson correlation coefficient between the CF-PDI and the other questionnaires: VAS, TSK-11, PCS, NDI, and HIT-6. A strong correlation was considered to be over 0.60; a moderate correlation between 0.30 and 0.60; and a low (very low) correlation below 0.30 (44).

Linear Regression
Multiple linear regression analysis was used to estimate the strength of the associations with theoretically similar constructs, so multiple linear regression analyses were also performed including CF-PDI as a criterion variable to estimate the strength of the association between CF-PDI and NDI, PVAS, TSK, and PCS as predictor variables. As a measure of multicollinearity, the variance inflation factor (VIF) is presented (VIF < 10 indicates no problem with multicollinearity).

All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS), version 20 (IBM company, USA) except for the SEM and MDC values, which were calculated using Microsoft Excel. The critical value for significance was $P < 0.05$.

Results

Item Development and Cognitive Debriefing
A total of 18 patients with chronic craniofacial pain participated in the focus group and were also interviewed in May 2011 and July 2011, and 30 items were pooled as potential questions. After a review by the research group, some questions were added, and 22 items covering 4 aspects (quality of life, jaw functional status, avoidance behavior, and pain) were finally generated. There was agreement among the expert evaluators who reviewed the items, with a kappa coefficient of 0.83. The greatest disagreement occurred in items 8 and 20.

A pilot test for cognitive debriefing was performed in 24 patients in September 2011 to examine the content validity of the preliminary questionnaire in regards to relevance and clarity of the language. After a review by the research group, some questions were added, and 22 items covering 4 aspects (quality of life, jaw functional status, avoidance behavior, and pain) were finally generated. There was agreement among the expert evaluators who reviewed the items, with a kappa coefficient of 0.83. The greatest disagreement occurred in items 8 and 20.

A pilot test for cognitive debriefing was performed in 24 patients in September 2011 to examine the content validity of the preliminary questionnaire in regards to relevance and clarity of the language. After a review by the research group, some questions were added, and 22 items covering 4 aspects (quality of life, jaw functional status, avoidance behavior, and pain) were finally generated. There was agreement among the expert evaluators who reviewed the items, with a kappa coefficient of 0.83. The greatest disagreement occurred in items 8 and 20.

Characteristics of the Sample
The final study sample consisted of 192 heterogeneous chronic craniofacial pain patients (68.8% women, one patient was of unknown gender) aged 19 – 78 years (mean ± SD: 46.00 ± 13.06). The vast majority of patients (28.1%) had myofascial pain diagnoses; other patients suffered from TMJ arthralgia (15.1%), headache or facial pain attributed to TMD (myofacial pain/TMJ osteoarthritis or arthralgia) (24.5%), combined tension-type headache and myofascial pain (16.7%), and migraine (15.6%). The mean time of pain was 130.46 ± 151.44 months (range: 15 – 888), and 19 patients (9.9%) had received disability benefits. Educational levels in our sample were primary (23.4%), secondary (36.5%), and university graduates (25.0%); there was no information for 15.1% of our sample.

Distribution of Total CF-PDI Scores
The distribution of CF-PDI scores did not differ significantly from a normal symmetric distribution (skewness = 0.43; SE = 0.18; kurtosis = -0.36; SE = 0.35), Kolmogorov-Smirnov Z = 1.11 ($P = 0.172$). There were no significant differences in scoring between men (19.46 ± 9.04) and women (20.52 ± 9.22). There was no significant association between CF-PDI scores and age, marital status, average duration of pain, education level, or work status.

Only the type of diagnosis showed differences in the median score of CF-PDI, headache or facial pain attributed to the TMD (myofacial pain/TMJ osteoarthritic or arthralgia) group presented higher scores 28.62 ± 7.10; TMJ arthralgia, 14.2 ± 5.24; migraine, 17.93 ± 12.30; myofascial pain, 18.17 ± 6.44; combined tension-type headache and myofascial pain, 19.00 ± 7.05. The distribution of CF-PDI total scores and other principal scales are shown in Table 1.

Table 1. Descriptive statistics and estimates of internal consistency ($N = 192$).

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Mean (SD)</th>
<th>Range</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF-PDI</td>
<td>20.24 (9.15)</td>
<td>2–48</td>
<td>0.88</td>
</tr>
<tr>
<td>Pain and Disability</td>
<td>15.43 (6.77)</td>
<td>1–34</td>
<td>0.86</td>
</tr>
<tr>
<td>Jaw Functional Status</td>
<td>4.81 (3.57)</td>
<td>0–14</td>
<td>0.80</td>
</tr>
<tr>
<td>HIT-6</td>
<td>54.48 (7.67)</td>
<td>36–74</td>
<td>0.85</td>
</tr>
<tr>
<td>NDI</td>
<td>16.96 (6.00)</td>
<td>0–42</td>
<td>0.74</td>
</tr>
<tr>
<td>TSK-11</td>
<td>25.40 (7.08)</td>
<td>11–44</td>
<td>0.88</td>
</tr>
<tr>
<td>PCS</td>
<td>23.86 (8.90)</td>
<td>7–52</td>
<td>0.84</td>
</tr>
<tr>
<td>VAS</td>
<td>52.94 (13.83)</td>
<td>15–85</td>
<td>----</td>
</tr>
</tbody>
</table>

CF-PDI, craniofacial pain and disability inventory; VAS, visual analogue scale; TSK-11, Tampa Scale for Kinesiophobia; PCS, pain catastrophizing scale; NDI, Neck Disability Index; HIT-6, headache impact test-6
Internal Consistency

Cronbach’s α coefficient was 0.88 (95% CI = 0.86 – 0.91), indicating a high degree of internal consistency. The item-to-total correlation coefficients ranged from 0.32 to 0.73; no item dominated with an especially high correlation and no item appeared to be redundant. The previous item 20, “How long have you had pain?” was deleted in the final version and it increased slightly the Cronbach’s α coefficient. This item showed a strong positive skew, refers to the time of pain in our population, and shows limited information because all patients suffered from chronic pain. It was removed it; other results in Table 2.

Factor Analysis

In order to explore the factorial structure of the instrument, a PCA without rotation was conducted on the scores of our sample. We also attempted to construct one-, 2-, and 3-factor structures. A 2-factor solution emerged in our sample using a PCA that explained 40.8% of the variance. The KMO was found to be 0.85, which exceeds the recommended minimum value of 0.60. Bartlett’s Test of Sphericity was highly significant (Chi square = 1467.10 P < 0.001), supporting the suitability of the data for PCA.

When factor loading smaller than 0.30 was suppressed, but there were no cases. The first factor (30.43% of the explanatory variance) was composed of 14 items (1, 2, 3, 4, 5, 6, 7, 16, 17, 18, 19, 20 (previously 21), and 21 (previously 22); the second factor (10.39% of the explanatory variance) was composed of 7 items (9, 10, 11, 12, 13, 14, and 15). With these results and a visual inspection of the scree plot, a 2-factor solution was considered suitable (Fig. 1).

Regarding factor analysis, item 9 was not clearly classified into the assumed factor (with loadings under 0.35 in each of them). The results showed similar weights for both factors. Despite the unexpected loading of this item, the CF-PDI still showed appropriate internal consistency; therefore, we incorporated it into the jaw functional status domain for theoretical reasons. Results of the PCA are shown in Table 3.

Floor and Ceiling Effects

No floor or ceiling effects were identified for the whole scale. Only 9.3% of the respondents scored the lowest possible score of 0 in the jaw functional status subscale, and none of the craniofacial pain patients scored the highest possible score of 63 points on the CF-PDI.

Table 2. Corrected item-total between CF-PDI items (N = 192)

<table>
<thead>
<tr>
<th>Item</th>
<th>Scale mean if item deleted</th>
<th>Corrected item-total correlation</th>
<th>Squared multiple correlation</th>
<th>Cronbach’s α if item deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.65</td>
<td>0.73</td>
<td>0.63</td>
<td>0.87</td>
</tr>
<tr>
<td>2</td>
<td>18.72</td>
<td>0.46</td>
<td>0.37</td>
<td>0.88</td>
</tr>
<tr>
<td>3</td>
<td>18.58</td>
<td>0.54</td>
<td>0.55</td>
<td>0.88</td>
</tr>
<tr>
<td>4</td>
<td>19.61</td>
<td>0.47</td>
<td>0.56</td>
<td>0.88</td>
</tr>
<tr>
<td>5</td>
<td>19.82</td>
<td>0.53</td>
<td>0.51</td>
<td>0.88</td>
</tr>
<tr>
<td>6</td>
<td>19.88</td>
<td>0.34</td>
<td>0.31</td>
<td>0.88</td>
</tr>
<tr>
<td>7</td>
<td>19.19</td>
<td>0.50</td>
<td>0.36</td>
<td>0.88</td>
</tr>
<tr>
<td>8</td>
<td>19.56</td>
<td>0.35</td>
<td>0.23</td>
<td>0.88</td>
</tr>
<tr>
<td>9</td>
<td>19.49</td>
<td>0.32</td>
<td>0.21</td>
<td>0.88</td>
</tr>
<tr>
<td>10</td>
<td>19.30</td>
<td>0.59</td>
<td>0.57</td>
<td>0.87</td>
</tr>
<tr>
<td>11</td>
<td>19.34</td>
<td>0.53</td>
<td>0.54</td>
<td>0.88</td>
</tr>
<tr>
<td>12</td>
<td>19.41</td>
<td>0.41</td>
<td>0.48</td>
<td>0.88</td>
</tr>
<tr>
<td>13</td>
<td>19.70</td>
<td>0.55</td>
<td>0.54</td>
<td>0.88</td>
</tr>
<tr>
<td>14</td>
<td>19.79</td>
<td>0.40</td>
<td>0.50</td>
<td>0.88</td>
</tr>
<tr>
<td>15</td>
<td>19.83</td>
<td>0.43</td>
<td>0.46</td>
<td>0.88</td>
</tr>
<tr>
<td>16</td>
<td>18.29</td>
<td>0.64</td>
<td>0.56</td>
<td>0.87</td>
</tr>
<tr>
<td>17</td>
<td>18.58</td>
<td>0.57</td>
<td>0.46</td>
<td>0.87</td>
</tr>
<tr>
<td>18</td>
<td>19.33</td>
<td>0.44</td>
<td>0.30</td>
<td>0.88</td>
</tr>
<tr>
<td>19</td>
<td>19.30</td>
<td>0.44</td>
<td>0.37</td>
<td>0.88</td>
</tr>
<tr>
<td>20</td>
<td>19.22</td>
<td>0.47</td>
<td>0.36</td>
<td>0.88</td>
</tr>
<tr>
<td>21</td>
<td>19.19</td>
<td>0.38</td>
<td>0.25</td>
<td>0.88</td>
</tr>
</tbody>
</table>
Test-Retest Reliability

The response to the CF-PDI provided by a random subsample of 106 patients (gender women: 70, 66.7%; age: 45.6 ± 12.9 years; duration of the disorders: 69.0 ± 46.2 months) showed satisfactory temporal stability of the scale after 12 days. ICC based on absolute agreement measures was 0.90 (95% CI: 0.86 – 0.93). The constructed Bland and Altman plot for test-retest agreement showed a good reliability for total CF-PDI score
and 2 subscales (Figs. 2-4). The results of reliability and responsiveness analyses are summarized in Table 4.

**Convergent Validity**

The total CF-PDI score was significantly associated with all the assessed questionnaires (Table 5), but the correlation with the NDI, was the most important in our sample.

**Linear Regression**

The resulting beta coefficients, ranging from 0.50 to 0.17, indicate independent contribution of each scale to the prediction of CF-PDI, the criterion variable. NDI, VAS, and TSK-11 were predictors of CF-PDI, significance < 0.05 (as illustrated by the higher standardized coefficients [beta] and P-values). NDI was the most important variable (Table 6). PCS and HIT-6 were excluded as predictor variables this time.

**Discussion**

The present study describes a methodical approach to the development and validation of a new self-administered questionnaire to measure disability, pain, and functional status of the mandibular and craniofacial region in patients with craniofacial pain. Our results demonstrate that the CF-PDI is psychometrically valid and reliable. In addition, the instrument has proven to be easy to complete, and only requires a relatively short time to administer. The CF-PDI was developed in Spain for Spanish patients with craniofacial pain. However, since the CF-PDI does not contain items that are specifically related to Spanish culture, it could be translated and used internationally.

The design of the CF-PDI was based on a biopsychosocial approach. This conceptual model, recommended by the International Classification of Functioning Disability and Health (52,53), can assess the disease from a
Psychometric Validation of the Craniofacial Pain and Disability Inventory

broader perspective, and provides an understanding of health, functioning, and disability. In addition, research supports that clinical diagnosis is sometimes insufficient to explain patients’ pain and disability (54-56).

The scree plot and exploratory factor analysis revealed a 2-factor solution. Both factors had eigenvalues greater than 1. PCA indicated that a satisfactory percentage of total variance (40.8%) was explained by the 2 factors. The CF-PDI contains 21 items divided into 2 subscales according to their content and exploratory factor analysis: “pain and disability” and “jaw functional status.”

High internal consistency was shown for the CF-PDI (Cronbach’s α: 0.88) and also for the 2 subscales (Cronbach’s α: 0.80 – 0.86). These data are similar to the results from other research questionnaires used to assess facial pain and mandibular function (22-25,57,58).

Table 4. Descriptive statistics, test-retest reliability, and responsiveness results (N = 106)

<table>
<thead>
<tr>
<th>Domains</th>
<th>Test Mean</th>
<th>Test SD</th>
<th>Retest Mean</th>
<th>Retest SD</th>
<th>ICC</th>
<th>95% CI</th>
<th>SEM</th>
<th>MDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF-PDI</td>
<td>20.57</td>
<td>8.42</td>
<td>22.79</td>
<td>7.80</td>
<td>0.90</td>
<td>0.86-0.93</td>
<td>2.48</td>
<td>6.87</td>
</tr>
<tr>
<td>Pain and Disability</td>
<td>15.37</td>
<td>6.13</td>
<td>17.10</td>
<td>5.26</td>
<td>0.86</td>
<td>0.81-0.90</td>
<td>2.10</td>
<td>5.82</td>
</tr>
<tr>
<td>Jaw Functional Status</td>
<td>5.20</td>
<td>3.39</td>
<td>5.69</td>
<td>3.93</td>
<td>0.86</td>
<td>0.80-0.90</td>
<td>1.35</td>
<td>3.75</td>
</tr>
</tbody>
</table>

CF-PDI, craniofacial pain and disability inventory; SD, standard deviation; ICC, intraclass correlation coefficient; 95% CI, 95% confidence interval; SEM, standard error of measurement; MDC, minimal detectable change

Table 5. Pearson Correlation Coefficient of our principles scales (N = 192).

<table>
<thead>
<tr>
<th>VAS</th>
<th>CF-PDI 0.46**</th>
<th>Pain and Disability 0.50**</th>
<th>Jaw Functional Status 0.23**</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDI</td>
<td>0.65**</td>
<td>0.69**</td>
<td>0.37**</td>
</tr>
<tr>
<td>PCS</td>
<td>0.46**</td>
<td>0.50**</td>
<td>0.25**</td>
</tr>
<tr>
<td>PCS rumination</td>
<td>0.34**</td>
<td>0.35**</td>
<td>0.22**</td>
</tr>
<tr>
<td>PCS magnification</td>
<td>0.51**</td>
<td>0.52**</td>
<td>0.32**</td>
</tr>
<tr>
<td>PCS Helplessness</td>
<td>0.39**</td>
<td>0.45**</td>
<td>0.16*</td>
</tr>
<tr>
<td>TSK-11</td>
<td>0.40**</td>
<td>0.41**</td>
<td>0.26**</td>
</tr>
<tr>
<td>HIT6</td>
<td>0.38**</td>
<td>0.46**</td>
<td>0.09</td>
</tr>
</tbody>
</table>

** P < 0.01; * P < 0.05

Abbreviations: CF-PDI, craniofacial pain and disability inventory; VAS, visual analogue scale; TSK-11, Tampa Scale for Kinesiophobia; PCS, pain catastrophizing scale; NDI, Neck Disability Index; HIT-6, headache impact test-6

FIG. 4. Bland Altman plot illustrating the test-retest reliability of the Jaw Functional Status subscale. A total of 106 patients participated in the test-retest assessment. The central line representing the mean difference between test and retest scores, which was -0.49, and the 95% limits of agreement are presented as flanking lines.
In this study, we choose a retest interval of 12 days (approximately), in order to avoid variations in clinical status and patients remembering their previous answers. A longer interval for a test-retest study of health may be inappropriate as fluctuations in the patient’s health status can occur (59). In relation to this, Streiner and Norman suggested that a retest interval of 2 to 14 days is generally acceptable (60).

The test-retest reliability for the total CF-PDI score was considered to be excellent (ICC: 0.90; 95% CI: 0.86 – 0.93). Also, we were able to verify that the test-retest reliability was high for each subscale.

The measurement of SEM was 2.4 points, corresponding to 11.7% of the mean CF-PDI values and 3.8% of the maximum possible score. Based on the SEM, the MDC was 7 points (34.5% of mean values). Considering that the score of the questionnaire ranges from 0 to 63 points, 7 points represents 11.1% of the maximum possible score, which means that the CF-PDI is able to detect very small changes. Changes higher than the MDC can be interpreted as real and not due to measurement error, with an acceptable probability level. These results may help to calculate the sample size of future studies aiming to assess the effectiveness of craniofacial pain interventions.

Construct validity was evidenced by significant correlations between the CF-PDI with all the questionnaires and scales used in the validation process. A moderate correlation between CF-PDI with the HIT-6 and the VAS (r = 0.38 – 0.46) was observed. In addition, the PCS and TSK-11 showed moderate correlation with the CF-PDI and the pain and disability subscale (r = 0.36 – 0.52). This is consistent with recent evidence demonstrating that patients with craniofacial pain or craniomandibular disorders report higher levels of catastrophizing (61-63). Furthermore, pain-related catastrophizing has been associated with the progression of pain intensity and signs of disability in chronic craniofacial pain (64-68).

---

Table 6. Multiple linear regression models with CF-PDI (A), pain and disability (B), and jaw functional status (C) as criterion variable, and NDI, VAS, TSK-11, PCS as predictor variables (N = 192).

<table>
<thead>
<tr>
<th>Criterion variable</th>
<th>Predictor variables</th>
<th>Regression coefficient (B)</th>
<th>Standardized coefficient (β)</th>
<th>Significance (P)</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. CF-PDI</td>
<td>NDI</td>
<td>0.77</td>
<td>0.50</td>
<td>0.000</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td>VAS</td>
<td>0.13</td>
<td>0.19</td>
<td>0.001</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>TSK-11</td>
<td>0.22</td>
<td>0.17</td>
<td>0.004</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>Excluded variables</td>
<td>PCS-Total</td>
<td>--</td>
<td>0.08</td>
<td>0.253</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HIT-6</td>
<td>--</td>
<td>0.01</td>
<td>0.905</td>
</tr>
<tr>
<td></td>
<td>B. Pain and Disability</td>
<td>NDI</td>
<td>0.55</td>
<td>0.49</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCS-Magnification</td>
<td>0.68</td>
<td>0.25</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VAS</td>
<td>0.10</td>
<td>0.21</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Excluded variables</td>
<td>PCS-Total</td>
<td>--</td>
<td>-0.50</td>
<td>0.480</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TSK-11</td>
<td>--</td>
<td>0.09</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HIT-6</td>
<td>--</td>
<td>0.08</td>
<td>0.159</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCS-Rumination</td>
<td>--</td>
<td>-0.06</td>
<td>0.314</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCS-Helplessness</td>
<td>--</td>
<td>-0.00</td>
<td>0.968</td>
</tr>
<tr>
<td></td>
<td>C. Jaw Functional Status</td>
<td>NDI</td>
<td>0.17</td>
<td>0.29</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCS-magnification</td>
<td>0.28</td>
<td>0.20</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>Excluded variables</td>
<td>PCS-Total</td>
<td>--</td>
<td>-0.13</td>
<td>0.207</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TSK-11</td>
<td>--</td>
<td>0.09</td>
<td>0.258</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HIT-6</td>
<td>--</td>
<td>-0.13</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCS-Rumination</td>
<td>--</td>
<td>-0.16</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PCS-Helplessness</td>
<td>--</td>
<td>-0.06</td>
<td>0.436</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VAS</td>
<td>--</td>
<td>0.06</td>
<td>1.22</td>
</tr>
</tbody>
</table>

CF-PDI, craniofacial pain and disability inventory; VAS, visual analogue scale; TSK-11, Tampa Scale for Kinesiophobia; PCS, pain catastrophizing scale; NDI, Neck Disability Index; HIT-6, headache impact test-6, VIF, variance inflation factor.
Previous research demonstrated the relationship between fear of jaw movements and craniofacial pain (69,70), but only limited evidence supports it. However, there is higher evidence showing that pain-related fear is associated with reduced activities in daily life and is also a strong predictor of disability in other chronic musculoskeletal disorders (71-75).

Pain catastrophizing and pain-related fear are 2 constructs that have been linked to the chronicity of musculoskeletal pain through the “fear avoidance model” (76). Based on the results of multiple linear regression analysis, pain intensity (VAS: $\beta = 0.19$, $P = 0.001$) and fear of pain and movement (TSK-11: $\beta = 0.17$, $P = 0.004$) were predictors of CF-PDI. For jaw functional status, and pain and disability, the variable predictor was pain catastrophizing (PCS-Magnification: $\beta = 0.25$, $P < 0.001$; $\beta = 0.20$, $P = 0.007$).

The principal predictor for CF-PDI and the 2 subscales was the variable of neck disability (NDI: $\beta = 0.29 – 0.50$, $P < 0.001$). In addition, a strong correlation was observed between CF-PDI and pain and disability factor with NDI ($r = 0.65 – 0.69$). This is in line with the results of Olivo et al (77) who described a strong relationship between neck disability and jaw disability ($r = 0.82$). Several studies have reported the high prevalence and comorbidity between orofacial pain, TMD, headache, and neck pain (65,78-81). Our findings suggest the importance of taking into account the neck disability questionnaires when assessing patients with craniofacial pain.

Limitations

Our study has several limitations. First, there is a gender disproportion as the sample had a smaller proportion of men. However, our findings showed no significant differences in scoring between genders. The evidence suggests that the prevalence of craniofacial pain is higher in women (82).

The second limitation of this study is that we did not assess the CF-PDI in healthy subjects; the sample consisted of patients with chronic pain. Further studies will need to be performed to assess the discriminant power of the CF-PDI for specific diagnostic entities.

The sample size was sufficient to test the new instrument’s reliability, convergence validity, and exploratory factor analysis. However, it was too small to be able to carry out confirmatory factor analysis. Kline has suggested a sample size of 10 – 15 subjects per item to perform this statistical analysis (83). It should be noted that statisticians disagree on the issue of appropriate sample size for confirmatory factor analyses. In relation to this, DeVellis stated that as the sample size becomes larger, the relative number of respondents per item can diminish (84), and that a sample of 200 is adequate in most studies (85).

Another limitation is that only self-reported measures were considered for convergent validity. Future research should use physical tests to explore the clinical signs relating to pain and disability, and assess whether these are associated with the CF-PDI.

The last limitation of the study is the cross-sectional design, which prevented us from investigating the ability of the CF-PDI to detect responsiveness to change over time. Although in this study we investigated in a short period of time the reproducibility and the MDC, a longitudinal study or one with an experimental design with a follow-up period would be required to understand how CF-PDI scores change over time. Furthermore, such a study would allow us to obtain information such as the Minimum Clinically Important Difference.

Conclusion

Evidence has shown that the CF-PDI has a good structure, internal consistency, reproducibility, and construct validity, and provides an objective tool for assessing pain and disability in craniofacial pain patients. Neck disability showed a strong association with the CF-PDI, and is also a significant predictor of the construct. Based on the findings of this study, the CF-PDI could be used in research and clinical practice for the assessment of treatment outcomes.

References

4. Shan CL, Bin Adon MY, Rahman AB, Hassan ST, Ismail KB. Prevalence of neck pain and associated factors with personal characteristics, physical workloads
30. Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968; 70:213-220.
32. Bartlett MS. A note on the multiply-
Psychometric validation of the Craniofacial Pain and Disability Inventory

53. Stucki G, Cieza A, Melvin J. The Inter
52. Stucki G, Boonen A, Tugwell P, Cieza
51. Wyrwich KW. Minimal important differ-
49. Ostelo RWJG, De Vet HCW, Knol DL, Van
48. Harvill L. Standard error of measure-
47. Weir JP. Quantifying test-retest reliabil-
46. Mokkink LB, Terwee CB, Patrick DL,
45. Bland JM, Altman DG. Statistical meth-
44. Terwee CB, Bot SDM, De Boer MR, Van
42. Kaiser H. Analysis of factorial simplicity.
the rehabilitation strategy.
J Rehab Med
34:600-606.
are proposed for measurement prop-
Bland JM, Altman DG. Statistical meth-
Weir JP. Quantifying test-retest reliability using the intraclass correlation coeffi-
Harvill L. Standard error of measure-
Ostelo RWJG, De Vet HCW, Knol DL, Van den Brandt PA. 24-item Roland-Morris Disability Questionnaire was preferred out of six functional status question-
Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and mea-
sures used in physical therapy. Phys Ther 2006; 86:735-743.
Wywich KW. Minimal important differ-
Stucki G, Cieza A, Melvin J. The Inter-
Turner JA, Franklin G, Turk DC. Pre-
dictors of chronic disability in injured workers: A systematic literature synthe-
Braga S, Sandanger I, Nygård JF. Emo-
tional distress as a predictor for low back disability: A prospective 12-year popula-
Ohrbach R, Larsson P, List T. The Jaw functional limitation scale: Develop-
Holt AE, Shaw NJ, Shetty A, Greenough CG. The reliability of the Low Back Out-
Streiner DL, Norman GR. Health Mea-
surement Scales: A Guide to Their Develop-
Campbell CM, Kronfli T, Buenaver LF, Smith MT, Berna C, Haythornthwaite JA, Edwards RR. Situational versus dis-
positional measurement of catastrophizing: Associations with pain respons-
Fillingim RB, Ohrbach R, Greenspan JD, Knott C, Dubner R, Bair E, Bararian C, Slade GD, Maixner W. Potential psycho-
social risk factors for chronic TMD: Descriptive data and empirically identified domains from the OPPERA case-control study. J Pain 2011; 12:746-60.
Quartana PJ, Buenaver LF, Edwards RR, Klick B, Haythornthwaite JA, Smith MT. Pain catastrophizing and salivary cor-
tisol responses to laboratory pain test-
Rantala MA, Ahberg J, Suvinen Ti, Nis-
sinen M, Lindholm H, Savolainen A, Ko-
nonen M. Temporomandibular joint re-
lated painless symptoms, orofacial pain, neck pain, headache, and psychosocial factors among non-patients. Acta Odon-
Buenaver LF, Edwards RR, Smith MT, Gramling SE, Haythornthwaite JA. Cata-
Buenaver LF, Quartana PJ, Grace EG, Sarlani E, Simango M, Edwards RR, Hay-
Holroyd KA, Drew JB, Cottrell CK, Ro-
manek KM, Heh V. Impaired function-
ing and quality of life in severe micro-
Visscher CM, Ohrbach R, Van Wijk AJ, Wilkoss M, Naeje M. The Tampa Scale for Kinesiophobia for Temporomandib-
Buer N, Linton SJ. Fear-avoidance be-
Crombez G, Vlaeyen JW, Heuts PH, Ly-
sens R. Pain-related fear is more dis-
Walton D, Elliott JM. A higher-order analysis supports use of the 11-item version of the tampa scale for kinesiopho-
Vlaeyen JW, Seelen HA, Peters M, De Jong P, Artes E, Beisiegel E, Weber WE. Fear of movement/reinjury and mus-


