Measurement properties of painDETECT by average pain severity

Joseph C Cappelleri1
E Jay Bienen2
Vijaya Koduru3
Alesia Sadosky4
1Pfizer, Groton, CT, 2Outcomes research consultant, New York, NY, 3Eliassen Group, New London, CT, USA; 4Pfizer, New York, NY, USA

Background: Since the burden of neuropathic pain (NeP) increases with pain severity, it is important to characterize and quantify pain severity when identifying NeP patients. This study evaluated whether painDETECT, a screening questionnaire to identify patients with NeP, can distinguish pain severity.

Materials and methods: Subjects (n=614, 55.4% male, 71.8% white, mean age 55.5 years) with confirmed NeP were identified during office visits to US community-based physicians. The Brief Pain Inventory – Short Form stratified subjects by mild (score 0–3, n=110), moderate (score 4–6, n=297), and severe (score 7–10, n=207) average pain. Scores on the nine-item painDETECT (seven pain-symptom items, one pain-course item, one pain-irradiation item) range from -1 to 38 (worst NeP); the seven-item painDETECT scores (only pain symptoms) range from 0 to 35. The ability of painDETECT to discriminate average pain-severity levels, based on the average pain item from the Brief Pain Inventory – Short Form (0–10 scale), was evaluated using analysis of variance or covariance models to obtain unadjusted and adjusted (age, sex, race, ethnicity, time since NeP diagnosis, number of comorbidities) mean painDETECT scores. Cumulative distribution functions on painDETECT scores by average pain severity were compared (Kolmogorov–Smirnov test). Cronbach’s alpha assessed internal consistency reliability.

Results: Unadjusted mean scores were 15.2 for mild, 19.8 for moderate, and 24.0 for severe pain for the nine items, and 14.3, 18.6, and 22.7, respectively, for the seven items. Adjusted nine-item mean scores for mild, moderate, and severe pain were 17.3, 21.3, and 25.3, respectively; adjusted seven-item mean scores were 16.4, 20.1, and 24.0, respectively. All pair-wise comparisons of scores between pain-severity groups showed sizable and statistically significant differences (P<0.0001). Cumulative distribution functions showed distinct separation between severity (P<0.0001). Cronbach’s alphas were 0.76 and 0.80 for the nine- and seven-item scales, respectively.

Conclusion: This study provides strong psychometric evidence on the validity and reliability of painDETECT for distinguishing average pain severity in patients with NeP.

Keywords: neuropathic pain, painDETECT, pain severity, psychometric properties

Introduction
Neuropathic pain (NeP), a common cause of chronic pain, results from a variety of diseases and medical conditions, and is defined as pain caused by a lesion or disease of the peripheral somatosensory nervous system.1 Although heterogeneity in epidemiologic studies has been a limitation in estimating its presence in the population, best estimates suggest that the overall prevalence of pain with neuropathic characteristics is between 7% and 10%.2 However, when NeP is present, it is associated with a substantial...
socioeconomic burden that has been characterized not only for conditions that are well recognized (eg, painful diabetic peripheral neuropathy and chronic low-back pain) but also for several that are less frequent (eg, small-fiber neuropathy). Furthermore, evidence suggests that the patient and economic burdens are higher at greater pain-severity levels.

Because there is only weak evidence for interventional management of NeP, treatment relies on a pharmacologic approach, and guidelines for pharmacologic management of NeP have been published. Despite these recommendations and the availability of medications from several drug classes, NeP management remains challenging, not least because of the need for appropriate identification of NeP when making treatment decisions.

To identify the presence of NeP in patients reporting chronic pain, differences in the symptoms and descriptions of NeP relative to nociceptive pain were investigated to develop various instruments to screen for the likelihood that pain is of neuropathic rather than nociceptive origin. While some instruments, such as the DN4 (Douleur Neuropathique 4) and the original Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) include a requirement for clinical examination, other instruments, such as a patient self-report version of the LANSS (S-LANSS), as well as painDETECT, do not require clinical examination. In particular, painDETECT is a simple, fully patient-reported questionnaire that has been validated for screening and identification of components of NeP. In contrast to S-LANSS, which evaluates sensory symptoms based on a dichotomous response (yes/no), painDETECT consists of 6-point Likert-type scales, which provide a broader range or continuum of assessment. Since pain severity is not only of immediate importance to patients but also increases the overall burden, characterizing pain-severity levels when identifying patients with NeP can help inform treatment decisions. Although an association between pain severity and painDETECT scores was suggested in several studies, including the original painDETECT-validation study, this relationship has not been psychometrically evaluated with the goal of differentiating NeP severity. Therefore, the purpose of this study was to determine whether painDETECT can differentiate among levels of average pain severity in patients with NeP.

Materials and methods

Data source

Data for this analysis were obtained during a cross-sectional, observational study on the burden of NeP across six NeP conditions: painful diabetic peripheral neuropathy, human immunodeficiency virus-related peripheral NeP, posttrauma/postsurgical NeP, spinal cord injury-related NeP, chronic low-back pain-associated NeP, and small-fiber neuropathy. The study was approved by a central institutional review board (Concordia Clinical Research, Cedar Knolls, NJ, USA), and collection and analysis of data were compliant with the Health Insurance Portability and Accountability Act.

Subjects

Subjects with a confirmed diagnosis of any of the six conditions were identified during routine office visits at 33 community-based physician practices across the US between September 2011 and June 2012. Study sites included general practitioners (n=9), neurologists (n=7), pain specialists (n=6), endocrinologists (n=3), and other specialists (eg, orthopedist, infectious disease specialist, podiatrist, rheumatologist; n=8).

For inclusion, subjects were required to be adults (≥18 years old) diagnosed with their NeP at least 6 months prior to enrollment, managed by the physician’s practice for at least 6 months, and had experienced NeP symptoms for at least the prior 3 months. Exclusion criteria were participation in an investigational drug study in the past 6 months; presence of a serious or unstable medical or psychological condition that in the opinion of the physician would compromise participation in the study; or having a concomitant illness unrelated to NeP that could have confounded the assessment of NeP.

Outcomes

Subjects were asked to complete a series of self-administered, one-time questionnaires during the office visit that included patient-reported measures of pain, function, quality of life, and productivity, as previously described. Of relevance to the current analysis, the questionnaire included the Brief Pain Inventory – Short Form (BPI-SF) and the painDETECT questionnaire.

Patients were stratified based on their level of average pain within the past 24 hours using the average pain item from the BPI-SF, which consists of an 11-point numeric rating scale (0= no pain to 10= pain as bad as you can imagine). Pain-severity levels based on previously established cut points were defined as mild (score 0–3), moderate (score 4–6), and severe (score 7–10). painDETECT is a nine-item questionnaire that consists of seven sensory symptom items for pain that are graded from 0= never to 5= strongly, one temporal item on pain-course pattern graded −1 to +1, and one spatial item on pain radiation...
graded 0 for no radiation or +2 for radiating pain. A total score that ranges from –1 to 38 can be calculated from the nine items, with higher scores indicating higher levels of NeP. Although the nine-item questionnaire demonstrated strong predictive ability for NeP, principal-component analysis identified the seven sensory items as driving the data structure of the questionnaire. Therefore, the seven-item version of painDETECT that only includes the sensory symptom items and a total score that ranges from 0 to 35 was also evaluated for its ability to differentiate pain severity.

Statistical analyses
Scores on painDETECT were derived for subjects who had no missing responses on the sensory items (seven- and nine-item scales), and response was also required on the pain-course item on the nine-item scale. Scale-level discrimination was evaluated using unadjusted and adjusted models. The ability of painDETECT to discriminate between average pain-severity levels of mild, moderate, and severe, based on the average pain item from the BPI-SF as described earlier, was determined using analysis of variance or covariance models (both models used Tukey’s adjustment for pair-wise comparisons) to obtain, respectively, unadjusted and adjusted mean painDETECT scores. Factors in the adjusted models were age, sex, race, ethnicity, time since NeP diagnosis, and number of comorbidities.

The corresponding mean scores on painDETECT between pairs of severity groups were translated into standardized effect sizes, derived by taking the difference between pair-wise means and dividing them by the pooled standard deviation on painDETECT across the three severity groups. By convention, a standardized effect size of 0.20 is considered “small”, while 0.50 and 0.80 are “medium” and “large”, respectively. Additionally, cumulative distribution of painDETECT scores by (average) pain severity were plotted and compared using the Kolmogorov–Smirnov test.

Scale-level reliability was evaluated using Cronbach’s alpha (covariance-based formula) to determine the internal consistency reliability of the nine- and seven-item versions of painDETECT. Assessment of item-level discrimination was performed using corrected item-to-total correlations to determine how well each item discriminated across neuropathic levels of pain as measured by painDETECT. These were “corrected” to account for item overlap by deleting the item under consideration from the total score. All analyses and statistics were performed using SAS version 9.2 (SAS Institute, Cary, NC, USA).

Results
Subject characteristics
The study enrolled 624 subjects who were identified with the NeP conditions, with approximately the same number of subjects from each NeP condition. Data on average pain severity were available for 614 subjects, with an average pain-severity distribution of 17.9%, 48.4%, and 33.7% for mild, moderate, and severe pain, respectively (Table 1). These subjects were 55.4% male and predominantly white (71.8%), with a mean (standard deviation) age of 55.5 (13.7) years (Table 1). However, with increasing levels of pain severity, significant trends were observed across severity levels \( P<0.05 \) toward younger age, higher proportion of females, longer duration since NeP diagnosis, and more comorbid conditions (Table 1).

Among the 614 subjects with pain-severity data, the response rate for the painDETECT questionnaire was 96.9% and 98.7% for the nine- and seven-item versions, respectively; only two subjects (one mild and one moderate pain) had nine-item scores despite a missing pain-radiation item (which was taken to indicate no radiating pain and hence given a value of 0). Similar proportions of subjects across the pain-severity categories completed both versions (Table 1).

Scale-level discrimination
On the nine-item version of painDETECT, unadjusted mean scores were 15.2, 19.8, and 24.0, for mild, moderate, and severe pain, respectively (Figure 1A). These scores differentiated between levels of average reported pain, with all pair-wise comparisons statistically significant \( P<0.0001 \). The significant differentiation between all pair-wise comparisons was maintained for the adjusted mean scores of 17.3 (mild), 21.3 (moderate), and 25.3 (severe) \( P<0.0001 \) (Figure 1B). Similarly, for the seven-item version, both the unadjusted (Figure 1C) and adjusted painDETECT (Figure 1D) scores differentiated between average pain-severity levels, as indicated by the statistical significance for all pair-wise comparisons \( P<0.0001 \).

Standardized effect sizes were slightly lower (with no practical importance) for adjusted mean scores relative to unadjusted mean scores, and if anything slightly lower for the seven-item version than the nine-item version (Table 2). Both the nine-item scores and the seven-item scores showed separation with respect to the three levels of average pain severity, with strength of differentiation that was at least medium (effect sizes \( \geq 0.50 \)) across all comparisons, even after adjustment (Table 2). While the magnitude of separation was in the medium range for mild versus moderate pain severity
Table 1  Demographic and clinical characteristics by neuropathic pain severity

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mild* (n=110)</th>
<th>Moderate* (n=297)</th>
<th>Severe* (n=207)</th>
<th>P-valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years, mean (SD)</td>
<td>58.3 (15.1)</td>
<td>55.7 (13.2)</td>
<td>53.6 (13.3)</td>
<td>0.0293</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.0222</td>
</tr>
<tr>
<td>Male</td>
<td>71 (64.5)</td>
<td>169 (56.9)</td>
<td>101 (48.8)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>39 (35.5)</td>
<td>128 (43.1)</td>
<td>106 (51.2)</td>
<td></td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.0015</td>
</tr>
<tr>
<td>American Indian or Alaska native</td>
<td>1 (0.9)</td>
<td>3 (1.0)</td>
<td>5 (2.4)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>1 (0.9)</td>
<td>2 (0.7)</td>
<td>2 (1.0)</td>
<td></td>
</tr>
<tr>
<td>Black or African American</td>
<td>13 (11.8)</td>
<td>37 (12.5)</td>
<td>47 (22.7)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>89 (80.9)</td>
<td>230 (77.4)</td>
<td>122 (58.9)</td>
<td></td>
</tr>
<tr>
<td>Multiracial</td>
<td>2 (1.8)</td>
<td>4 (1.3)</td>
<td>5 (2.4)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3 (2.7)</td>
<td>15 (5.1)</td>
<td>22 (10.6)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>1 (0.9)</td>
<td>6 (2.0)</td>
<td>4 (1.9)</td>
<td></td>
</tr>
<tr>
<td>Ethnicity, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.0151</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>100 (90.9)</td>
<td>266 (90.6)</td>
<td>168 (81.2)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>5 (4.5)</td>
<td>21 (7.1)</td>
<td>27 (13.0)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>5 (4.5)</td>
<td>10 (3.4)</td>
<td>12 (5.8)</td>
<td></td>
</tr>
<tr>
<td>BPI-SF pain-severity index, mean (SD)</td>
<td>2.0 (1.1)</td>
<td>5.2 (0.8)</td>
<td>7.7 (1.1)</td>
<td>NA</td>
</tr>
<tr>
<td>Time since NeP diagnosis, years, mean (SD)</td>
<td>6.3 (5.9)</td>
<td>7.9 (6.8)</td>
<td>8.5 (7.2)</td>
<td>0.0059</td>
</tr>
<tr>
<td>Number of comorbid conditions, mean (SD)</td>
<td>2.0 (1.8)</td>
<td>2.5 (2.2)</td>
<td>3.3 (2.4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Neuropathic pain condition, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Painful diabetic peripheral neuropathy (n=111)</td>
<td>22 (20.0)</td>
<td>57 (19.2)</td>
<td>32 (15.5)</td>
<td></td>
</tr>
<tr>
<td>Human immunodeficiency virus-related peripheral NeP (n=102)</td>
<td>24 (21.8)</td>
<td>41 (13.8)</td>
<td>37 (17.9)</td>
<td></td>
</tr>
<tr>
<td>Postrauama/postsurgical NeP (n=97)</td>
<td>14 (12.7)</td>
<td>48 (16.2)</td>
<td>35 (16.9)</td>
<td></td>
</tr>
<tr>
<td>Spinal cord injury-related NeP (n=101)</td>
<td>21 (19.1)</td>
<td>53 (17.8)</td>
<td>27 (13.0)</td>
<td></td>
</tr>
<tr>
<td>NeP in chronic low-back pain (n=104)</td>
<td>6 (5.5)</td>
<td>55 (18.5)</td>
<td>43 (20.8)</td>
<td></td>
</tr>
<tr>
<td>Small-fiber neuropathy (n=99)</td>
<td>23 (20.9)</td>
<td>43 (14.5)</td>
<td>33 (15.9)</td>
<td></td>
</tr>
<tr>
<td>painDETECT completion, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Nine-item version</td>
<td>106 (96.4)</td>
<td>291 (98.0)</td>
<td>198 (95.7)</td>
<td></td>
</tr>
<tr>
<td>Seven-item version</td>
<td>108 (98.2)</td>
<td>292 (98.3)</td>
<td>206 (99.5)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *Average pain severity was classified based on the Brief Pain Inventory – Short Form scores of 0–3 for mild, 4–6 for moderate, and 7–10 for severe pain; ^P-values across pain-severity levels are from the Kruskal–Wallis test for continuous variables; from the χ² test for number of comorbid conditions, and from Fisher’s exact test for the remaining categorical variables. 

Abbreviations: NA, not applicable; SD, standard deviation; BPI-SF, Brief Pain Inventory – Short Form; NeP, Neuropathic Pain.

and moderate versus severe pain, it was consistently large for mild versus severe pain. The cumulative frequency-distribution plots for painDETECT scores are shown in Figure 2 for the nine-item version (Figure 2A) and the seven-item version (Figure 2B). Both versions demonstrated differentiation between pain-severity levels that was statistically significant for all pair-wise comparisons (P<0.0001); for each painDETECT score, the proportion of patients at that score or lower (more favorable score) was greater at lower levels of pain severity. For example, for the seven-item score, approximately 20% of subjects in the severe group had scores of 18 or lower (better), compared with approximately 45% and 70% in the moderate and mild pain groups, respectively.

Scale-level reliability

Across all subjects and pain-severity categories, corrected item-to-total correlations showed discernible and sizable amounts of item discrimination for all items on the seven-item version, which ranged from 0.49 for “slight pressure triggers pain” to 0.62 for “light touching painful” (Table 3). The same set of items on the nine-item version was equally discriminating, as indicated by similar values for the corrected item-to-total correlation (Table 3). However, the two additional items showed low or no discrimination (pain radiation, 0.21; pain course, −0.08). Despite having two fewer items, the seven-item version gave a slightly higher Cronbach’s alpha (0.80) than the nine-item version (0.76).

Discussion

This study compellingly demonstrated the psychometric validity and reliability of painDETECT for distinguishing mild, moderate, and severe levels of average pain in subjects with NeP, with pain severity rated using the average pain item from the BPI-SF. Specifically developed screening instruments, such as painDETECT, have been shown to be...
useful in identifying the likelihood that pain is of neuropathic origin. Although there is a need for characterizing and quantifying NeP concurrently with its identification in patients with various NeP conditions, to our knowledge, no other NeP instrument has yet been specifically validated for distinguishing between levels of average pain severity. These results therefore complement and extend studies that have suggested that NeP-screening instruments may help distinguish among sensory symptom profiles and pain-severity levels.

Both the seven-item and nine-item scores of painDETECT provided good discrimination (known-group validity) of average pain severity, with significant differentiation ($P<0.0001$) for all pair-wise comparisons between mild, moderate, and severe pain. The extent of the separation was substantial based on estimation of standardized effect sizes: medium magnitude for mild versus moderate pain and for moderate versus severe pain, and large magnitude for mild versus severe pain. Although it can be expected that a similar set of results would apply to the similarly phrased BPI-SF questions on worst pain, least pain, and current pain, further research would be needed for their confirmation.

These findings, while new and clinically relevant, are not entirely unexpected, given the previously observed relationship between painDETECT and pain severity. The initial painDETECT-validation study reported a significant association ($P<0.001$) between pain severity (average over

---

**Table 2** Standardized effect sizes for pair-wise comparisons of painDETECT scores

<table>
<thead>
<tr>
<th>Comparison between average pain-severity levels</th>
<th>Standardized effect size$^a$</th>
<th>Nine-item version</th>
<th>Seven-item version</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Unadjusted</td>
<td>Adjusted</td>
</tr>
<tr>
<td>Mild versus moderate</td>
<td>$-0.65$</td>
<td>$-0.56$</td>
<td>$-0.63$</td>
</tr>
<tr>
<td>Mild versus severe</td>
<td>$-1.25$</td>
<td>$-1.13$</td>
<td>$-1.23$</td>
</tr>
<tr>
<td>Moderate versus severe</td>
<td>$-0.60$</td>
<td>$-0.57$</td>
<td>$-0.60$</td>
</tr>
</tbody>
</table>

*Note:* Negative values represent less neuropathic pain in the group with the lower level of average pain severity.

---

**Figure 1** (A–D) Mean painDETECT scores across pain-severity categories for average pain. (A) Nine-item unadjusted scores; (B) nine-item adjusted scores; (C) seven-item unadjusted scores; (D) seven-item adjusted scores.
Two previous studies have also suggested that pain-symptom descriptors may be significant predictors of NeP severity. In those studies, the presence of a higher number of descriptors, as well as specific descriptor qualities, had greater predictive value for more severe pain. The range of descriptors constituting painDETECT overlaps with those descriptors that have shown predictive ability for pain severity. Taken in context, the current results confirm and supplement the relevance of using descriptors for characterizing NeP, including for quantifying severity, and provide reassuring evidence that the observed results enhance the clinical usefulness of painDETECT in ways not previously validated.

Both the seven-item version and the nine-item version demonstrated internal consistency reliability that was acceptable (above 0.70). However, the seven-item version resulted in a slightly higher reliability (Cronbach’s α=0.80) than the nine-item version (Cronbach’s α=0.76) and more consistent item-level discrimination. The reliability of the seven-item version is not only similar to that reported in the validation study (Cronbach’s α=0.83), but its higher value than the nine-item version is also in concordance with previous observations, suggesting that the two nonsensory items are not psychometrically compatible with the sensory items. While further confirmation for the greater reliability of the seven-item version is needed, it can nevertheless be recommended that the seven sensory items be used together; the two nonsensory items, on temporal and spatial characterization of the pain pattern, that are included in the nine-item version can be analyzed and interpreted separately.

It should also be noted that since the population in the current study consisted of patients with confirmed NeP, the cumulative distribution curves are in marked contrast to those in studies evaluating discrimination between main pain subtypes (nociceptive, pure neuropathic, mixed) where the nociceptive group clearly resulted in lower (more favorable) scores on painDETECT. In addition, in the current study, the cumulative distributions convincingly showed that painDETECT scores were dependent on the degree of average pain severity for patients with confirmed NeP; fitting only one curve for all patients with NeP (ie, without regard to average pain severity), as has been done previously, would mask this finding.

At least two possible limitations of this research should be noted. One is that the study was cross-sectional rather than longitudinal. The opportunity to link changes in painDETECT scores to changes in average pain severity was not available. However, painDETECT has generally been used...
as a screening tool at a single time point, since its responsiveness to change over time or with treatment has not been adequately evaluated.

The second limitation is that this study could be potentially criticized for evaluation of a population that included different NeP conditions, since it can be argued that there may be differences among the conditions resulting from their various etiologies. However, it has been reported that there appear to be greater similarities than differences overall with regard to NeP regardless of etiology.25,26

Furthermore, a recent study suggested that there may be up to four distinct pain-characteristic profiles that are independent of etiology.26 Therefore, since the scoring of painDETECT is based on an amalgam of frequency and severity for relevant individual sensory symptoms, severity discrimination will likely be maintained even across the different sensory symptom profiles, supporting a robust ability of painDETECT to both identify NeP and profile its sensitivity with respect to average pain regardless of the etiology or sensory profile. It could be argued that the wide range of average pain severity and NeP across etiologies is intended to foster a diverse and expansive base of patients that resounds in drawing stronger psychometric results and inferences to patients with NeP regardless of etiology.

Conclusion
This study provides strong psychometric evidence on the validity of painDETECT for distinguishing scores on NeP based on scores of average pain severity for patients with NeP. Therefore, the ability of painDETECT to discriminate among mild, moderate, and severe levels of average pain further supports the measurement properties and clinical relevance of this instrument for identifying and characterizing NeP. The nine-item and seven-item versions of painDETECT also show satisfactory reliability, but the higher reliability of the seven-item version suggests it may be more prudent to use this version, with the two nonsensory items used separately to further characterize NeP. Additional investigation of the psychometric properties of painDETECT and other NeP-screening instruments are warranted, since more comprehensive characterization of NeP at screening can inform management strategies that may help improve outcomes and reduce the disease burden.

Disclosure
This study was funded by Pfizer. Joseph C Cappelleri and Alesia Sadosky are employees and shareholders of Pfizer, the sponsor of this study; Vijaya Koduru is an employee of the Eliassen Group, who were paid consultants to Pfizer in connection with the analysis described in this study; E Jay Bienen is an independent scientific consultant who was funded by Pfizer in connection with manuscript development.

References


