Psychometric properties of the Neuropathic Pain Scale (NPS) in a knee osteoarthritis population

Emma C. Lape a,*, Faith Selzer a,b,c, Aileen M. Davis d,e,f,g, Jamie E. Collins a,b,c, Elena Losina a,b,c,i, Jeffrey N. Katz a,b,c,h

a Orthopaedic and Arthritis Center for Outcomes Research (OrACORe), Brigham and Women’s Hospital, Boston, MA, USA
b Department of Orthopedic Surgery, Brigham and Women’s Hospital, Boston, MA, USA
c Harvard Medical School, Boston, MA, USA
d Division of Health Care and Outcomes Research, Krembil Research Institute, University Health Network, Toronto, Canada
e Institute of Health, Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
f Institute of Rehabilitation Science, University of Toronto, Toronto, Canada
g Department of Physical Therapy, University of Toronto, Toronto, Canada
h Division of Rheumatology, Immunology, and Allergy, Brigham and Women’s Hospital, Boston, MA, USA
i Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

ARTICLE INFO

Keywords:
Psychometric properties
Validation
Neuropathic pain
Osteoarthritis
Total knee replacement

SUMMARY

Objective: Symptoms resembling neuropathic pain (neuropathic-like symptoms) are prevalent in osteoarthritis (OA) populations. Scales that measure neuropathic-like symptoms frequently were established in groups with true neuropathic pain conditions and have not been assessed in OA. We assessed the psychometric properties of the Neuropathic Pain Scale (NPS) in subjects with OA undergoing total knee replacement (TKR).

Design: In a prospective study of adults undergoing TKR for OA, we assessed baseline distributions, acceptability (completion rate), internal consistency (Cronbach’s alpha), responsiveness 12 months post-TKR, and construct validity of the NPS. We performed factor analysis and created subscales from the items loading onto each retained factor. We evaluated subscale properties and calculated the proportion of total scores attributable to each subscale and compared this with the proportion expected if each item contributed equally.

Results: Mean baseline NPS score among 263 participants was 42.7 (SD: 15.9). Cronbach’s alpha was 0.88. Factor analysis produced two factors: “bothersome” (items: intense/sharp/dull/unpleasant/deep; Cronbach’s alpha = 0.87), and “dysesthetic” (items: cold/sensitive/itchy/surface; alpha = 0.77). Bothersome items contributed more to total NPS scores (74%) than would be expected if each item contributed equally (50%). NPS scores correlated moderately with baseline pain and function, and decreased after TKR, with standardized response means (SRMs) of: total NPS: 1.77, Bothersome subscale: 2.03, Dysesthetic subscale: 0.70.

Conclusions: The NPS had acceptable completion, internal consistency, and construct validity, but was not optimal for use in OA; Bothersome subscale items disproportionately drive total NPS scores and may fail to discriminate between nociceptive and neuropathic-like symptoms.

1. Introduction

There is a growing understanding that pain in osteoarthritis (OA) often has contributors beyond the nociceptive pain due to tissue damage at the affected joint. In fact, the extent of tissue damage does not predict who will experience the most pain, likely because other personal factors and non-nociceptive mechanisms contribute to pain and dilute the pain-structure association [1]. This observation motivates the study of additional pain mechanisms. In OA, neuropathic-like pain symptoms are quite common [2–9]. A 2017 systematic review concluded that 29% of persons with hip and knee OA experience neuropathic-like pain [10]. The prevalence of objective clinical evidence of peripheral nerve abnormality in OA is around 4% according to a recent epidemiological study [11]; thus, for the larger group of OA patients with symptoms similar to pain of true
neuropathic origin, the term “neuropathic-like” is generally used. The prominence of neuropathic-like pain in a subset of persons with OA suggests that there may be a distinct OA pain phenotype that is defined by the quality of pain [7].

Subgroups of patients with qualitatively different pain experiences may also have distinct pain etiologies, or mechanisms, which in turn can be targeted by specific treatments [12,13]. Identifying such subgroups within OA based on descriptions of pain quality can help target treatments to pain subgroups. Several self-administered scales exist to measure the extent of neuropathic qualities of pain [14]. However, the reliability and validity of these neuropathic pain scales have generally been established in populations with true neuropathic pain conditions (e.g. diabetic neuropathy), and the measures used in OA frequently have not been assessed for reliability, validity, and responsiveness in OA patients [10]. As a result, we cannot be certain of the extent to which these instruments measure a unidimensional construct (internal consistency) and selectively measure neuropathic-like pain quality as opposed to overall pain intensity (validity). The aim of the present analysis was to assess the psychometric characteristics (internal consistency, validity, responsiveness) of a widely used scale, the Neuropathic Pain Scale (NPS), in a cohort of subjects with advanced knee OA undergoing TKR.

2. Patients and methods

We conducted an analysis of the acceptability, internal consistency (a form of scale reliability), and convergent validity of the NPS in a cohort with knee OA and the responsiveness of the NPS at 12 months following total knee replacement (TKR).

2.1. Participants

The Study of Total Knee Arthroplasty Responses (STARs) is a prospective cohort study of adults (ages ≥40–) with knee OA undergoing unilateral TKR between September 2012 and April 2014, designed to evaluate predictors of suboptimal pain outcomes [15,16]. We excluded persons with a primary arthritis diagnosis other than OA, such as inflammatory arthritis. Brigham and Women’s Hospital (BWH) served as data coordinating center, and the surgeries were performed at three clinical sites: one academic medical center in New York City (NYC) and two community orthopedic centers in Colorado and Maryland.

Each clinical center identified potentially eligible individuals. Research coordinators at BWH (or at New York University for subjects recruited in NYC) contacted each potential participant to explain the study, gauge interest, and confirm eligibility. Participants were compensated ($25) at each timepoint (baseline and 12 months) for completing and returning the questionnaire forms.

2.2. Questionnaires and measures

At baseline (≤6 weeks pre-TKR) and at 12 months post-TKR, participants were asked to complete paper questionnaires including measures of pain and function, among them the NPS [12]. The NPS is scored as the sum of 10 pain descriptor items, each rated on a scale from 0 to 10, with 10 representing the strongest attestation to that descriptor. Thus, the NPS total score ranges from 0 to 100 with 100 representing the highest degree of neuropathic-like symptoms [12]. The NPS was presented as a visual analogue scale, meaning that non-integer responses were possible.

The original NPS development paper did not specify how many items needed to be answered for a valid score. Thus, we required that at least 8 of the 10 items be completed, accepting 1–2 missing items and assigning them the mean value of completed items. In sensitivity analyses, we instead assigned missing items a score of 0. The NPS’ final two questions are ‘pain course’ questions, which ask participants to characterize the pattern of their pain over time. The pain course questions are not included in the NPS summary score.

To examine the relation of NPS score to other aspects of the pain experience, we included other measures relevant to pain severity and pain phenotype. We assessed the WOMAC (Western Ontario and McMaster Universities Arthritis Index) Pain (5-item) and Function (17-item) subscales, scored as the sum of the items, each rated from 0 to 10, where 0 indicates the least pain/best function and transformed to a 0 to 100 scale (100 worst) [17,18]. We included the five-item Mental Health Inventory (MHI-5; a validated measure of mental health status) [19]; the score was calculated as the sum of responses to 5 items (responses to two of which are reversed in scoring) and transformed to a 0 to 100 scale, with 0 worst. Other instruments included the Pain Catastrophizing Scale (PCS) [20] and a body pain diagram (as a measure of widespread pain), for which participants marked up to 19 body sites with pain during the past week [21,22]. We summed painful body sites (0–19) and excluded the index knee, since pain there is expected and does not reflect the construct of widespread pain. We then grouped body sites to derive the total painful non-index body regions [16] for a maximum of 7 regions.

2.3. Statistical analysis

2.3.1. Psychometric properties in baseline cohort

Our reporting on these analyses is guided by the COSMIN checklist for evaluating patient-reported outcome measures [23]. We evaluated acceptability of the NPS at baseline by assessing its completion rate among the 267 participants who returned questionnaires and comparing with the completion rate of the primary outcome (WOMAC Pain). We assessed the distributions of the NPS total score and of each descriptor item. We present the proportion of participants scoring near the bottom for the total NPS score (<10) and of each item (<1), since floor effects would reflect on whether the scale is applicable in an OA population. To evaluate internal consistency and the unidimensional nature of the measure, we calculated Cronbach’s alpha [24], considering an alpha value of 0.7–0.95 to be adequate [23]. We calculated inter-item correlations among the 10 descriptors and corrected item-to-total correlations. As described above, missing items were assigned the mean value of completed items.

2.3.2. Responsiveness of the NPS after TKR

In those completing the NPS at both baseline and 12 months after TKR, we assessed the standardized response mean (SRM; mean change in NPS score divided by the standard deviation of the change) and the effect size (mean change in NPS score divided by the standard deviation of scores at baseline) [25]. We also assessed the strength of linear association of these changes in NPS with change in WOMAC Pain using Pearson correlation coefficients.

Those included in the responsiveness analyses were compared against those lost to follow-up on several baseline measures using Wilcoxon rank sum tests.

2.3.3. Exploratory factor analysis (EFA)

Given that the NPS has not been evaluated previously in those undergoing TKR for OA, we conducted an exploratory factor analysis. We used the principal factor method to extract factors, followed by a promax (oblique) rotation, which does not assume zero correlation among factors [26]. Two factors were retained, and factor retention was guided by a) a scree test, b) eliminating factors that explained less than 10% of total variance, c) parallel analysis [27], and d) criteria of interpretability. We considered a variable to load on a given factor if the factor loading was 0.4 or above for that factor, and less than 0.40 for the other factor [26].
The items loading onto each retained factor were summed to create two factor-based subscales. As a sensitivity analysis, subscales were scored as the sum of items weighted by their regression correlation coefficient on the factor.

2.3.4. Concurrent construct validity

We evaluated correlations between NPS score and other concurrently administered, validated measures that, based on the literature, were anticipated to be associated with neuropathic-like pain: pain (WOMAC Pain) [3,5,6,8,9,28] and function (WOMAC Function) [5,8], mental health (MHI-5 score) [28], and number of painful regions, a measure of widespread pain [3]. We used Spearman correlation coefficients for the comparison with another scale on the same questionnaire, the primary outcome (WOMAC Pain) was completed by 258 (96.6% of total) with no missing items and by 265 (99.3%) with enough items to receive a score.

2.3.5. Subscale properties

We evaluated the properties of each factor-based subscale, as described above for the full NPS: distribution, internal consistency, concurrent construct validity, and responsiveness. We calculated the proportion of total scores attributable to each subscale on average, and compared this with the proportion expected if each item contributed equally (in which case a subscale comprising 5 of 10 items would contribute 50% of the score).

2.3.6. Sensitivity analyses

There is an established 8-item version of the NPS (in addition to the 10-item version) that excludes the “intense” and “unpleasant” items, thus better isolating neuropathic-like components. To understand the performance of this scale version, we assessed NPS-8 scores (scaled to a 0 to 100 scale) at baseline and 12 months and also conducted an exploratory factor analysis.

3. Results

3.1. Participants

Of 267 participants who returned STARs questionnaires, 263 (98.5%) completed enough items to receive a NPS score, constituting the analytic cohort. This sample is considered adequate for the analyses conducted [23]. Mean age was 65.9 (standard deviation (SD): 8.7) and 62.7% were female (Table 1).

3.2. Completion and missing data

249 participants completed all 10 items of the NPS, representing 93.3% of the total cohort and 94.7% of the analytic cohort. For comparison with another scale on the same questionnaire, the primary outcome (WOMAC Pain) was completed by 258 (96.6% of total) with no missing items and by 265 (99.3%) with enough items to receive a score.

3.3. Psychometric properties in baseline analytic cohort

Fig. 1a depicts baseline NPS scores. The mean score was 42.7 (SD: 15.9) and median was 43.0. Individual item mean scores ranged from 1.32 (itchy) to 6.51 (deep) (Table 2). We also report the proportion of subjects scoring <1 for each item, which we considered minimal attestation to that item (Table 2).

### Table 1

| Participant characteristics at baseline. |
|-----------------------------------------|
| **Demographics**                        |
| Age (years)                             |
| 263                                     |
| Female                                  |
| 97                                      |
| Yes                                     |
| 163                                     |
| No                                      |
| 30                                      |
| Yes                                     |
| 233                                     |
| **Percent, mean (SD)**                  |
| 65.9 (8.7)                              |
| 37.3                                    |
| 62.7                                    |
| 11.4                                    |
| 88.6                                    |

### Table 2

| **Disease characteristics**            |
|----------------------------------------|
| WOMAC Function                         |
| 261                                    |
| Score 0-39                             |
| 100                                    |
| Score 40-69                            |
| 135                                    |
| Score >70                              |
| 20                                     |
| WOMAC Pain                             |
| 262                                    |
| Score 0-39                             |
| 107                                    |
| Score 40-69                            |
| 136                                    |
| Score >70                              |
| 19                                     |
| **Percent, mean (SD)**                 |
| 44.9 (17.9)                            |
| 39.2                                   |
| 52.9                                   |
| 7.8                                    |
| 41.6 (18.9)                            |
| 40.8                                   |
| 51.9                                   |
| 7.3                                    |

**Item-to-item Pearson correlation coefficients are presented in Table 3.** Most items had moderate-to-high correlations with the total NPS score; the lowest were “dull” (r = 0.49) and “cold” (r = 0.45). Cronbach’s alpha with raw variables was 0.88, classified as an adequate value for a unidimensional measure.

3.4. Exploratory factor analysis (EFA)

The 249 participants who completed all 10 items at baseline were included in the factor analysis, such that there were no missing items in the EFA. This sample size is considered sufficient to perform an EFA (>7 times the number of items). Five items—intense, sharp, dull, unpleasant, and deep—loaded on the first factor, which we term “bothersome.” For the second factor, the constituent variables were cold, sensitive, itchy, and surface; we term this factor “dysesthetic.” “Hot” did not load on either factor. Factor loadings are presented in Table 4.

Additionally, we created subscales composed of the NPS items with significant loadings for each factor (5 items for Factor 1/“bothersome” and 4 items for Factor 2/“dysesthetic”). Thus, a participant’s score on the Bothersome subscale was the sum of item scores for intense, sharp, dull, unpleasant, and deep—loaded on the first factor, which we term “bothersome.” For the second factor, the constituent variables were cold, sensitive, itchy, and surface; we term this factor “dysesthetic.” “Hot” did not load on either factor. Factor loadings are presented in Table 4.

### Table 3

| **Comorbidities**                     |
|---------------------------------------|
| History of depression                 |
| No                                    |
| 198                                   |
| Yes                                   |
| 57                                    |
| **Percent, mean (SD)**                |
| 77.7                                  |
| 22.4                                  |
| **MHI-5 score category**              |
| 0-67                                  |
| 81                                     |
| 31.2                                  |
| 68-100                                |
| 179                                    |
| 68.9                                  |
| **Diabetes (Types I and II)**         |
| No                                    |
| 227                                   |
| 88.0                                  |
| Yes                                   |
| 31                                     |
| 12.0                                  |

3.5. Concurrent construct validity

Total NPS score showed [1] a moderate positive correlation with
3.6. Responsiveness after TKR

213 participants completed enough items to have NPS scores at both timepoints and were included in responsiveness analyses. The mean change in NPS score was 29.1 (SD: 16.4, SRM: 1.77) (Table 5). Fig. 1b shows 12-month NPS scores. WOMAC pain improved 31.4 (SD: 19.1, SRM: 1.65) points on average. Changes in NPS were linearly associated with changes in WOMAC Pain ($r = 0.58, P < 0.0001$).

Fig. 1. a–b. NPS scores (0–100 scale) at baseline and 12 months after TKR.
Scores for the Bothersome subscale changed more following TKR (mean change 42.9 (95% CI: 40.1–45.7, SD: 21.1), SRM 2.03) as compared with Dysesthetic subscale scores (mean change 11.3 (95% CI: 9.2–13.5, SD: 16.1), SRM 0.70) (Table 5; Appendix Fig. 1a-b). The sensitivity analysis using weighted subscales yielded correlation coefficients similar to the original unweighted versions, each differing by no more than 1%.

The 50 participants lacking 12-month scores, as compared with the 213 who had baseline and 12-month scores, had higher mean baseline NPS scores (49.1, SD: 18.0 versus 41.3, SD 15.1; P = 0.005) and mean baseline WOMAC Pain (48.7, SD 21.3 versus 39.9, SD 18.0; P = 0.02). Age and MHI-5 score did not significantly differ between these groups.

### 3.7. Sensitivity analyses using NPS-8

The NPS-8 baseline and 12-month scores (scaled 0 to 100) along with measures of responsiveness are presented in Table 5. Mean NPS-8 scores were slightly lower at baseline than NPS-10 (37.9 (SD: 16.1)) versus 42.7 (SD: 15.9) and, at 12 months, decreased slightly less (mean change of 25.2 versus 29.1; SRM: 1.55 versus 1.77). In the factor analysis using the NPS-8, “hot” and “surface” loaded onto Factor 1 (Bothersome), but the other 6 items grouped as they had in the original analysis (NPS-10) (Appendix Table 5).

### 4. Discussion

In a cohort of participants with knee OA scheduled to undergo TKR, we assessed the performance of the Neuropathic Pain Scale (NPS) in several key domains: acceptability (completion rate), internal consistency (scale reliability), concurrent construct validity, structural validity, and responsiveness following TKR. We conducted an exploratory factor analysis to determine whether the NPS may be capturing symptoms consistent with nociceptive pain, in addition to neuropathic-like symptoms. The NPS had high completion rates that were similar to those for other scales on the same questionnaire. The proportions of participants scoring at the very top or bottom of the scale were small. Item-to-total correlations were moderate to high. The Cronbach’s alpha of 0.88 suggests that the items do measure aspects of a shared underlying construct. Thus, the 10-item NPS appeared to function well as a unidimensional measure; however, exploratory factor analysis suggested that items sort into two distinct domains: “bothersome” (intense, sharp, dull, unpleasant, and deep) and “dysesthetic” (cold, sensitive, itchy, and surface). Two subscales constructed from the bothersome and dysesthetic items, respectively, each function well independently. Additionally, the bothersome domain, which describes symptoms that would be consistent with nociceptive pain, contributed disproportionately to total NPS scores.

The psychometric properties of the NPS have not previously been assessed in an OA population. While the presence of neuropathic-like symptoms in those with OA is well documented [2–10], previous studies have used other scales (PainDETECT, Douleur Neuropathique, Self-report Leeds Assessment of Neuropathic Symptoms and Signs) [10]. Our study provides key data on the performance in an OA cohort of a commonly used scale, the NPS [14,30]. Moreover, the factor analysis provides a closer investigation of the domains that are measured when a

### Table 2

Baseline distributions of NPS scores and item scores.

| Item       | N  | Mean (SD) | Range | Number (%) with score < 10 (total scale) or < 1 (individual items) |
|------------|----|-----------|-------|---------------------------------------------------------------|
| NPS total sum | 263 | 42.74 (15.92) | 5–93 | 5 (2%) |
| Deep       | 263 | 6.51 (2.20)   | 0–10 | 11 (4%) |
| Intense    | 258 | 6.27 (1.98)   | 0.5–10 | 8 (3%) |
| Unpleasant | 260 | 6.25 (2.08)   | 0.5–10 | 5 (2%) |
| Sharp      | 263 | 6.09 (2.56)   | 0–10 | 16 (6%) |
| Dull       | 261 | 5.40 (2.47)   | 0–10 | 19 (7%) |
| Surface    | 262 | 4.06 (2.75)   | 0–10 | 48 (18%) |
| Hot        | 260 | 3.75 (2.75)   | 0–10 | 62 (22%) |
| Sensitive  | 263 | 1.62 (2.15)   | 0–10 | 156 (58%) |
| Cold       | 262 | 1.53 (2.09)   | 0–10 | 160 (60%) |
| Itchy      | 262 | 1.32 (1.90)   | 0–10 | 175 (66%) |

Total possible scores range 0–100 with higher scores indicating more neuropathic-like symptoms; item possible scores range 0–10 with higher scores indicating greater attestation to the given pain descriptor.

### Table 3

Item-to-item and item-to-total correlations for NPS items at baseline (Pearson correlation coefficients).

| Correlation with other items and with total | Alpha with item deleted | Item | Intense | Sharp | Hot | Dull | Cold | Sensitive | Itchy | Unpleasant | Deep | Surface | Total NPS Score (corrected) | Alpha with item deleted |
|--------------------------------------------|------------------------|------|--------|------|-----|------|------|----------|-------|------------|------|---------|---------------------------|------------------------|
| Intense | – | 0.67 | 0.50 | 0.48 | 0.28 | 0.30 | 0.34 | 0.81 | 0.74 | 0.48 | 0.75 | 0.86 |
| Sharp | – | 0.35 | 0.32 | 0.26 | 0.27 | 0.27 | 0.58 | 0.63 | 0.51 | 0.65 | 0.86 |
| Hot | – | – | 0.22 | 0.42 | 0.45 | 0.40 | 0.42 | 0.50 | 0.50 | 0.63 | 0.86 |
| Dull | – | – | 0.27 | 0.27 | 0.27 | 0.47 | 0.46 | 0.40 | 0.49 | 0.87 |
| Cold | – | – | – | 0.44 | 0.47 | 0.22 | 0.22 | 0.32 | 0.45 | 0.88 |
| Sensitive | – | – | – | – | 0.63 | 0.29 | 0.23 | 0.47 | 0.53 | 0.87 |
| Itchy | – | – | – | – | – | 0.29 | 0.27 | 0.46 | 0.53 | 0.87 |
| Unpleasant | – | – | – | – | – | – | 0.73 | 0.50 | 0.70 | 0.86 |
| Deep | – | – | – | – | – | – | – | 0.41 | 0.67 | 0.86 |
| Surface | – | – | – | – | – | – | – | – | 0.65 | 0.86 |
formed our study using an already gathered set of data, we were unable 
in light of this, more information regarding comorbid neuropathic pain 
sections of the NPS does not instruct participants to think of pain only at the 
populations who also have nociceptive pain, as is the case in OA. The 
scores following a short-term lidocaine treatment [31,32]. Greater 
therapies, TKR. Two previous studies have assessed changes in NPS 
information about the responsiveness of NPS scores after a pain-relieving 
neuropathic pain conditions, is used in OA. We also contribute infor-

table like the NPS, which was developed in populations with true 
neuropathic pain conditions, is used in OA. We also contribute inform-

Table 5 
NPS and NPS subscale change from baseline to 12 months. 

| Variable | Baseline N | Mean (95% CI) | SD | 12 months N | Mean (95% CI) | SD | Change | % change | SRM\(^a\) | ES\(^b\) |
|----------|------------|---------------|----|-------------|---------------|----|--------|----------|---------|---------|
| Total NPS-10 score | 263 | 42.7 (40.8-44.7) | 15.9 | 216 | 12.4 (10.9-13.8) | 10.9 | 213 | 29.1 (26.9-31.3) | 16.40 | 68.03 | 1.77 | 1.83 |
| NPS-8 score | 263 | 37.87 (35.9-39.8) | 16.13 | 216 | 11.4 (10.0-12.8) | 10.2 | 213 | 25.2 (23.0-27.4) | 16.28 | 66.68 | 1.55 | 1.57 |
| Factor 1 subscale\(^c\) | 249 | 61.0 (58.7-63.3) | 18.5 | 217 | 15.7 (13.8-17.7) | 14.7 | 216 | 42.9 (40.1-45.7) | 21.1 | 70.32 | 2.03 | 2.31 |
| Factor 2 subscale\(^d\) | 249 | 21.5 (19.3-23.6) | 17.4 | 217 | 8.66 (7.5-9.8) | 8.38 | 215 | 11.3 (9.2-13.5) | 16.09 | 57.20 | 0.70 | 0.71 |

\(^a\) SRM = Standardized response mean (mean change/SD of change). 
\(^b\) ES = Effect size (mean change/SD of Time 1 scores). 
\(^c\) The Factor 1 subscale consists of the items intense, sharp, dull, unpleasant, and deep. 
\(^d\) The Factor 2 subscale consists of the items cold, sensitive, itchy, and surface. 

scale like the NPS, which was developed in populations with true neuropathic pain conditions, is used in OA. We also contribute information about the responsiveness of NPS scores after a pain-relieving intervention, TKR. Two previous studies have assessed changes in NPS scores following a short-term lidocaine treatment [31,32]. Greater change occurred in intense, sharp, deep, and unpleasant than in cold, sensitive, or itchy [31]. Similarly, we observed substantial change in total scores, with greater responsiveness for the Bother subscale (SRM = 2.03) than for the Dysesthetic subscale (SRM = 0.70).

The factor analysis produced two subscales, one with descriptors typical of neuropathic pain specifically, such as abnormal temperature (“cold”) and skin sensitivity (“surface”) (Dysesthetic subscale), and one that would be consistent with nociceptive pain alone (Bothersome subscale). Similarly, a previous study in individuals with central neuropathic pain identified two domains typified by dysesthetic items (e.g., surface, itchy) and nociceptive-like items, respectively [14]. However, the interpretation of these nociceptive-like pain items is challenging in OA patients, for whom attesting to “intense” and “unpleasant” pain may reflect nociceptive pain experience. In the STARs OA cohort, nociceptive-like pain items contributed disproportionately to NPS total scores, which could mask or overwhelm the less prevalent neuropathic-like symptoms experienced by a subset of patients. The dominance of nociceptive-like pain was also evident in assessing responsiveness to TKR; postoperative change in nociceptive-like pain was much larger than in neuropathic-like pain.

This suggests a need for fundamental research either to ensure that scales used in OA have suitable content and measurement properties that permit nociceptive and neuropathic-like symptoms to be distinguished, or to build from scratch a scale tailored to OA. Existing scales such as the Douleur Neuropathique 4 (DN4), painDETECT, and Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) lack items that, like the NPS’ “intense” or “unpleasant”, seem to measure overall pain intensity [33–35]. Of note, the 4-item Dysesthetic subscale of the NPS also appears to be an internally consistent measure of typical neuropathic-like symptoms in OA (with an alpha of 0.77), and could be used in this population with proper caution in interpretation. The NPS-8, which aims to better isolate neuropathic-like symptoms, could be evaluated when administered as an 8-item scale. A rigorous process for developing a new neuropathic-like pain scale for OA would involve eliciting experiences from patients with OA (first to gather descriptors for possible inclusion and then to refine the instrument), establishing content validity among OA patients with neuropathic pain, and evaluating measurement properties [36].

Our findings should be viewed in light of several limitations. As discussed above, NPS scores must be interpreted with caution in clinical populations who also have nociceptive pain, as is the case in OA. The stem of the NPS does not instruct participants to think of pain only at the index knee, limiting our ability to attribute responses to knee pain alone. In light of this, more information regarding comorbid neuropathic pain conditions would be useful, but was not collected. Given that we performed our study using an already gathered set of data, we were unable to assess test-retest reliability, an important aspect of the scale’s measurement properties. Finally, 19% of the analytic cohort lacked a score at 12 months, and this group presented to surgery with a higher baseline WOMAC Pain (reaching the minimal clinically important difference of 8–11 points) [37] and neuropathic-like symptoms. While a clinically meaningful difference in NPS scores has not been defined, the differences of about 1/2 of a standard deviation between those included and those excluded may be clinically important as well. The exclusion of more severely affected subjects may limit generalizability.

In summary, the NPS had good measurement properties in a population with end-stage knee OA: high internal consistency and convergent validity, as evidenced by the correlations of NPS score with measures of pain, function, and mental health status [38]. The NPS contains a 4-item subscale that functions as a reasonable measure of neuropathic-like symptoms in OA, consisting of the items cold, sensitive, surface, and itchy. Yet, the NPS is not optimal for measuring neuropathic-like pain when nociceptive pain is also present, due to the inclusion of several items that assess pain intensity and may span both domains.

Author contributions

Authors contributed to the 3 aspects of manuscript preparation as described below.

ECL: Conception and design of the study, analysis and interpretation of data (1); drafting and critical revision of the article (2); final approval of submitted version (3).

FS: Analysis and interpretation of data (1); critical revision of the article (2); final approval of submitted version (3).

AMD: Conception and design of the study (1); critical revision of the article (2); final approval of submitted version (3).

JEC: Analysis and interpretation of data (1); critical revision of the article (2); final approval of submitted version (3).

EL: Conception and design of the study, analysis and interpretation of data (1); critical revision of the article (2); final approval of submitted version (3).

JNK: Conception and design of the study, analysis and interpretation of data (1); critical revision of the article (2); final approval of submitted version (3).

Role of the funding sources

The project was funded by the Department of Orthopedic Surgery, Brigham and Women’s Hospital; and National Institutes of Health (NIH-NIAMS) grants P30AR072577 and K24AR 057827. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Funding sources had no role in the analysis or interpretation of data, the writing of the manuscript, or the decision to submit the manuscript for publication.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest.
Appendix Fig. 1. a-b. Factor 1/Bothersome subscale scores (0–100 scale) at baseline and 12 months after TKR.
Appendix Fig. 2. a–b. Factor 2/Dysesthetic subscale scores (0–100 scale) at baseline and 12 months after TKR.
Appendix Table 1
Subscale item-to-item and corrected item-to-total correlations (Pearson)

| Subscale item-to-item correlations (Pearson) | Factor 1 subscale | Factor 2 subscale | Alpha with item deleted* |
|-------------------------------------------|-------------------|-------------------|--------------------------|
| Intense | 1.00 | 0.68 | 0.49 | 0.81 | 0.74 | 0.83 | 0.82 |
| Sharp | 0.68 | 1.00 | 0.32 | 0.61 | 0.64 | 0.66 | 0.86 |
| Dull | 0.49 | 0.32 | 1.00 | 0.47 | 0.46 | 0.49 | 0.90 |
| Unpleasant | 0.81 | 0.61 | 0.47 | 1.00 | 0.74 | 0.79 | 0.82 |
| Deep | 0.74 | 0.66 | 0.46 | 0.74 | 1.00 | 0.79 | 0.82 |

*Subscale Cronbach's alpha with all items = 0.87.
†Subscale Cronbach's alpha with all items = 0.77.

Appendix Table 2
Correlations of NPS and NPS subscales with baseline pain, function, and psychosocial measures (Pearson unless otherwise noted; Estimate (95% CI))

| | NPS score | Factor 1 subscale | Factor 2 subscale | MHI | WOMAC pain | WOMAC function | PCS score |
|---|---|---|---|---|---|---|---|
| NPS score | – | 0.89 (0.87–0.92) | – |
| Factor 1 subscale | 0.83 (0.79–0.87) | 0.52 (0.42–0.61) | – |
| Factor 2 subscale | – | –0.28 | –0.26 |
| MHI | –0.31 | –0.42–0.19 | –0.39–0.16 | –0.37–0.14 | 0.66 (0.58–0.72) | 0.51 (0.41–0.60) | 0.35 |
| WOMAC pain | –0.28 | –0.40–0.16 | –0.37–0.14 |
| WOMAC function | 0.72 (0.65–0.77) | 0.68 (0.61–0.74) | 0.54 (0.45–0.63) | –0.31 | 0.60 (0.51–0.67) | 0.51 (0.41–0.60) | 0.77 |
| PCS score | 0.40 (0.28–0.51) | 0.42 (0.30–0.53) | 0.27 (0.14–0.39) | –0.53 | 0.40 (0.31–0.54) | 0.43 (0.31–0.54) | – |
| Pain regions* | 0.29 (0.17–0.40) | 0.22 (0.10–0.34) | 0.25 (0.13–0.36) | –0.27 | 0.18 (0.06–0.30) | 0.24 (0.12–0.36) | 0.24 |

*Correlations with pain regions are Spearman correlation coefficients.
†The Factor 1 subscale consists of the items intense, sharp, dull, unpleasant, and deep.
‡The Factor 2 subscale consists of the items cold, sensitive, itchy, and surface.

Appendix Table 3
Factor loadings for the 8 NPS-8 items, from the rotated factor pattern matrix and factor structure matrix

| Item | Factor Pattern | Factor Structure |
|---|---|---|
| | Factor 1 | Factor 2 | Factor 1 | Factor 2 |
| Sharp | 0.82* | –0.04 | 0.80* | 0.35 |
| Hot | 0.50* | 0.31 | 0.65* | 0.55* |
| Dull | 0.39 | 0.16 | 0.47* | 0.35 |
| Cold | 0.07 | 0.57* | 0.34 | 0.60* |
| Sensitive | –0.02 | 0.81* | 0.37 | 0.80* |
| Itchy | 0.00 | 0.78* | 0.37 | 0.78* |
| Deep | 0.84* | –0.08 | 0.80* | 0.32 |
| Surface | 0.42* | 0.38 | 0.61* | 0.59* |

* Asterisk indicates that the item loading on the given factor exceeds 0.4.
†In the factor pattern (left), the values are standardized regression coefficients of an item on the factor.
‡In the factor structure (right), the values are the product-moment correlations between an item and the factor.
Acknowledgements

None.

References

[1] T. Neogi, D. Felson, J. Niu, M. Nevitt, C.E. Lewis, P. Aliabadi, et al., Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies, Br. Med. J. 339 (2009 Aug 21) b2844.
[2] J.R. Hochman, M.R. French, S.L. Bingham, G.A. Hawker, The nerve of osteoarthritis pain, Arthritis Care Res. 62 (7) (2010 Jul) 1019–1023.
[3] J.R. Hochman, I. Gagliese, A.M. Davis, G.A. Hawker, Neuropathic pain symptoms in a community knee OA cohort, Osteoarthritis Cartilage 19 (6) (2011 Jun) 647–654.
[4] J.R. Hochman, A.M. Davis, J. Ellayam, L. Gagliese, G.A. Hawker, Neuropathic pain symptoms on the modified painDETECT correlate with signs of central sensitization in knee osteoarthritis, Osteoarthritis Cartilage 21 (9) (2013 Sep) 1236–1242.
[5] Y. Garaj, F. Eser, A. Kiliçkarlan, H. Bodur, Prevalence of neuropathic pain in Rheumatoid arthritis: disorder with disease activity, functional status and quality of life, Arch. Rheumatol. 30 (3) (2015) 231–237.
[6] A. Oteo-Alvaro, M.A. Ruiz-Iban, X. Miguens, A. Stern, J. Villoria, I. Sanchez-Magro, High prevalence of neuropathic pain features in patients with knee osteoarthritis: a cross-sectional study. Pain practice, Off J Word Inst Pain 15 (7) (2015 Sep) 618–626.
[7] B.J. Moreton, V. Tew, R. das Nair, M. Wheeler, D.A. Walsh, N.B. Lincoln, Pain phenotype in patients with knee osteoarthritis: classification and measurement properties of painDETECT and self-report Leeds assessment of neuropathic symptoms and signs scale in a cross-sectional study, Arthritis Care Res. 67 (4) (2015 Apr) 519–528.
[8] C. Roubille, J.P. Raynauld, F. Abram, P. Paiement, M. Dorais, P. Delorme, et al., The presence of meniscal lesions is a strong predictor of neuropathic pain in symptomatic knee osteoarthritis: a cross-sectional pilot study, Arthritis Res. Ther. 16 (6) (2014 Dec) 507.
[9] S. Ohtori, S. Orita, M. Yamashita, T. Ishikawa, T. Ito, T. Shigemura, et al., Existence of a neuropathic pain component in patients with osteoarthritis of the knee, Yonsei Med. J. 53 (4) (2012 Jul) 801–805.
[10] H.P. French, K.M. Smart, F. Doyle, Prevalence of neuropathic pain in the nerve of osteoarthritis and Cartilage Open 2 (2020) 100027.

Osteoarthritis and Cartilage Open 2 (2020) 100027

2020 E.C. Lape et al.