Literature Review

Development and Validation of a Screening Tool for Surgery-Specific Neuropathic Pain: Neuropathic Pain Scale for Postsurgical Patients

Mathias Kvist Mejdahl, MD[1,2], Karl Bang Christensen, PhD[3], and Kenneth Geving Andersen, MD, PhD[4,5]

From: [1]Department of Breast Surgery, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark; [2]Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark; [3]Section of Biostatistics, Department of Public Health, University of Copenhagen, Copenhagen, Denmark; [4]Section for Surgical Pathophysiology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; [5]Department of Anaesthesiology, Zealand University Hospital, Køge, Denmark

Address Correspondence:
Kenneth G. Andersen, MD, PhD
Rigshospitalet
University of Copenhagen, Section for Surgical Pathophysiology
Blegdamsvej 9 Copenhagen, DK-2100 Denmark
E-mail: kenneth.geving.andersen@regionh.dk

Disclaimer: This study was funded by grants from the Danish Cancer Society.
Conflict of interest: Each author certifies that he or she, or a member of his or her immediate family, has no commercial association (i.e., consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted manuscript.

Manuscript received: 05-15-2018
Revised manuscript received: 07-25-2018
Accepted for publication: 09-18-2018

Free full manuscript: www.painphysicianjournal.com

Background: Despite the existence of several screening tools for neuropathic pain, none of these are specific to surgery. We have developed a simple questionnaire tool, the Neuropathic Pain scale for Postsurgical patients (NeuPPS), to measure neuropathic pain in postsurgical patients.

Objectives: The aim was to validate this tool in a breast cancer population using an item response theory model, resulting in an easy-to-use scale.

Study Design: Development was based on literature reviews and interviews with patients and experts and consisted of 6 items. It was tested among 2,217 long-term breast cancer survivors, and cross-validated in another data set of breast cancer survivors with 18 months follow-up.

Setting: We tested the questionnaire tool among breast cancer survivors in 2 cohorts, one nationwide and one of consecutively treated patients at Rigshospitalet, Copenhagen, Denmark.

Methods: Items were assessed for criterion-related construct validity using the Rasch model, and for convergent validity by comparison to another neuropathic pain screening tool, the self-report version of the Leeds Assessment of Neuropathic Symptoms and Signs (S-LANSS) pain scale.

Results: The selected 5-item model showed good fit, unidimensionality, monotonicity, and homogeneity. This result was reproducible in the cross-validation population. In a combined dataset with data from both studies, the model displayed a slightly lower fit, suggesting that items to some degree may vary over time. The Spearman rank correlation between the NeuPPS and S-LANSS was $P = 0.57$.

Limitations: We observed differential item functioning between the primary study population and the cross-validation population, meaning that some items behave differently at different follow-up times or study populations.

Conclusions: With the NeuPPS, we have validated a simple and easy-to-fill-out questionnaire tool for the measurement of neuropathic pain among postsurgical patients. The items are additive, giving a total score that measures neuropathic pain symptoms.

Key words: Scale validation, Rasch analysis, item response model, persistent postoperative pain, intercostobrachial nerve, neuropathy, neuropathic pain, quantitative sensory testing, breast cancer

Pain Physician 2019; 22:E81-E90

Persistent pain is a common consequence after surgery with a prevalence ranging from 5%-50% depending on type of surgery (1), and is often accompanied by sensory dysfunction (2,3). Nerve damage and neuropathy have been suggested to be one of the main mechanisms leading to persistent postsurgical pain (1,4). This is supported by studies using different quantitative sensory testing (QST) protocols, in which cutaneous hypoesthesia, hypoalgesia, hyperalgesia, and painful response to repetitive innocuous stimuli have been found among postsurgical patients (3,5-9). Neuropathic pain as a term consists of a variety of different causes, pain distributions, and pathophysiological mechanisms. Persistent postsurgical pain is special in the sense that we know nerves are injured during surgery and pain arises somewhere close to the injury. Thus, according to the classification, these patients have probable or definitive neuropathic pain depending on direct observation of nerve damage or not (10). Both questionnaire studies and physical testing often reveal pain and sensory dysfunction in the area of surgery (2,4,8,9,11). However, when considering the surgical population as a whole, one observes the following categories: 1) patients with persistent pain and sensory disturbances, 2) patients with persistent pain without sensory disturbances, 3) patients with sensory disturbances but without persistent pain, and 4) patients with neither. By focusing only on patients with pain, information is lost concerning patients with the same trauma, but a different symptomatology. Many questionnaire tools have been developed to measure neuropathic pain (12-15), but there has been less focus on tools without pain as a defining component among postsurgical patients. Persistent sensory disturbances due to iatrogenic peripheral nerve damage could be a continuum, in which increasing severity of symptoms greatly increases the likelihood of the presence of a neuropathic pain component. Based on this assumption, we have developed a measurement tool to assess neuropathy specifically in the area of surgery in patients both with and without pain: the Neuropathic Pain scale for Postsurgical patients (NeuPPS). The items have previously been content validated (16). The primary aim of this study was to test construct validity of the items in a breast cancer population with long-term follow-up (5-7 years since surgery) using a statistical model founded in item response theory (IRT), the Rasch model (17-19). Furthermore, the aim was to cross-validate the NeuPPS in another breast cancer population with a shorter follow-up time (8-20 months since surgery). Convergent validity of the instrument was assessed by comparing it with the self-report version of the Leeds Assessment of Neuropathic Symptoms and Signs (S-LANSS) pain scale, and the NeuPPS was evaluated against pain reporting in the long-term follow-up study.

Methods

The present study was based on 2 data sets, in this study referred to as the "long-term study" (20) and the "medium-term study" (16), from which we have previously reported outcomes on pain (16,20). Both studies consisted of breast cancer survivors who had been treated for primary invasive breast cancer, and both studies included questionnaires that contained an identical set of questions regarding sensory disturbances after breast cancer treatment.

The NeuPPS and S-LANSS

The sensory disturbances items were originally developed as part of a larger questionnaire for postsurgical breast cancer patients (16). They were developed by assessing current literature on neuropathic pain screening tools (21) and semi-structured interviews with patients and health care providers (16). The items were developed with the aim of describing symptoms of neuropathy and neuropathic pain in a postsurgical setting. These could have a neuropathic pain component, but pain itself was not a prerequisite for the questions. The questionnaire originally consisted of 7 dichotomous (yes/no) items: pins and needles sensations (item 1), electric shock/jumping sensations (item 2), burning sensations (item 3), numbness (item 4), pain triggered by light touch (item 5), pain triggered by cold (item 6), and simultaneous pain and itching sensations (item 7) (see Fig. 1 for the complete wording in the full questionnaire tool). Item 7, "pain and itching sensations" was removed before any analyses were performed, as it was considered not to fit the overall construct after in depth interviews with patients in relation to another study (unpublished data). The long-term study also included S-LANSS (13). The S-LANSS consists of 7 dichotomous items, a 0-10 numerical pain rating scale, and a body map on which the patient can specify the location of the pain. The S-LANSS uses a scoring algorithm different from the simple sum, making it possible to obtain a total score of 24 (13). The scores obtained from the different items were originally developed for the Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) pain scale (22). Here, selection and weighting

We would like to know if you are experiencing sensory disturbances such as pins and needles or prickling sensations, in the area where you had surgery or in the side of the chest and arm, you are operated.

Within the last month:

1. Have you had pins and needles, tingling or stabbing sensations in or around the area of your surgery?	☐ Yes	☐ No
2. Have you experienced an electric shock like sensation or jabbing feelings in the skin area in or around the area of your surgery?	☐ Yes	☐ No
3. Have you experienced heat or burning sensations in or around the area of your surgery?	☐ Yes	☐ No
4. Have the lightest of touches from e.g. clothes been cause of pain in or around the area your surgery?	☐ Yes	☐ No
5. Have cold temperatures been the cause of pain in or around the area of surgery?	☐ Yes	☐ No

Items removed:

- Have you had numbness or decreased sensitivity in or around the area of your surgery?

- Have you experienced a painful itch in or around the area of your surgery?

Fig. 1. *The Neuropathic Pain scale for Postsurgical patients (NeuPPS).*

of the descriptor items and examination items were based on a logistic regression model (22). A cut-off at 12 or more has been suggested to identify pain with a predominantly neuropathic origin (13). The S-LANSS was translated after obtaining permission from the corresponding author, by 2 native Danish speakers and 2 native English speakers in an iterative process. The S-LANSS was attached in the original lay-out to the main questionnaire as an appendix.

Studies and Design

Study 1

The "long-term study." In 2005 and 2006, 5,119 women aged 19 to 70 years were treated for unilateral primary breast cancer in Denmark, identified through the Danish Breast Cancer Group (DBCG) database (23,24). On February 1, 2012, a nationwide cross-sectional study was performed in which questionnaires were sent to these women, 2,828 of whom were eligible and with 2,411 returning the questionnaire (20,25). Exclusion criteria were death, emigration, recurrence, new primary disease, contralateral cancer, metastatic disease, other malignant disease, and reconstructive/corrective breast surgery. Data regarding treatment, cancer recurrence, and demographics were retrieved from the DBCG database (23,24), and mortality data from the Danish Civil Registration System. The patients were treated according to the 2004 DBCG protocol (24) and full detailed information can be found elsewhere (20).

Study 2

The "medium-term study." This study was a questionnaire validation study performed on breast cancer survivors who had been treated at Rigshospitalet, Copenhagen, Denmark between August 13, 2009 and July 30, 2010 (16). Inclusion criteria were patients treated for primary breast cancer according to the 2007 or 2010 DBCG protocol (26). Exclusion criteria were death, bilateral disease, recurrence, other malignancy, or disease in the nervous system (16). A total of 494 questionnaires were sent, 402 returned, and 389 were included in the final analysis. Treatment data and disease characteristics were retrieved from the DBCG database (23,24).

Ethical Approval

The studies were approved by the regional bioethics committee of the capital region in Denmark, H-D-2007-0099 and H-4-2013-FSP-045. The data collection of both studies were approved by the Danish Data Protection Agency.

Statistical Analyses

Items were assessed with an IRT model (27), the Rasch model (17-19). IRT models are statistical models for ordinal categorical questionnaire items that indirectly measure a latent trait. Each study was initially assessed separately, starting with the long-term study. The IRT model was developed based on the long-term study. Subsequently, the fit of data from the medium-term study to the developed IRT model were assessed. Data from both study populations were then combined and the model assessed in the combined dataset. Patients with missing data were excluded. Item fit was assessed with the Rasch model in a systematic and iterative process, discarding items that did not fit the model. In the Rasch model the items are additive, and the sum of the items can be used as a measure of the underlying latent trait. Overall fit of the model was assessed by the Andersen conditional likelihood ratio (CLR) test, in which homogeneity of 2 score groups is tested (28). Individual item fit was evaluated by comparison of observed and expected correlations between the separate item scores and the sum of the remaining items (29), and by conditional infit and outfit statistics (30,31). Item fit was also assessed graphically by plotting empirical and theoretical item characteristic curves (32). Differential item functioning (DIF) (33) was evaluated using likelihood ratio tests in log linear Rasch models (34). We tested for DIF with respect to the exogenous variables: chemotherapy (yes/no), age (< 50, 50-59, 60-69, ≥ 70 years), radiotherapy (yes/no), axillary surgery (sentinel lymph node biopsy/axillary lymph node dissection), endocrine therapy (yes/no), and breast surgery (mastectomy/breast conserving surgery). Furthermore, we evaluated DIF with respect to study/follow-up period in the combined dataset.

The S-LANSS questionnaire was assessed using 2 IRT models: the Rasch model and the Birnbaum (2PL) model, in which item discrimination parameters are also estimated (35). With item discrimination parameters, we were able to estimate how well the items were able to capture the latent trait, and thereby compare discrimination parameters to the suggested scores for each item in the S-LANSS. Item fit was assessed graphically by comparing observed item mean scores to values simulated under the IRT model (36). Correlation between S-LANSS and the NeuPPS was assessed using both the original S-LANSS weights and the weights suggested by the Birnbaum model using the Spearman Rank test. We evaluated unidimensionality of a combined item set containing the S-LANSS and the NeuPPS using equally weighted S-LANSS items by comparing observed and expected sub-score correlation (37).

Evaluation of statistical significance for item fit statistics and for tests of DIF and local dependence used the Benjamini-Hochberg (38) procedure with a significance level at 0.05. Analyses were performed using DIGRAM software (Department of Biostatistics, University of Copenhagen, Copenhagen, Denmark) and SAS 9.4 for Windows (SAS Institute Inc., Cary, NC).

Results

The Long-Term Study: Development of the Model

All patients with missing data (n = 194) on either items or exogenous variables were excluded leaving 2,217 patients (92% out of 2,411) in the final data set. For patient characteristics, we refer to previous publications regarding this population (20,25). In the initial Rasch model, items 1 through 6 were included, but the model showed poor overall fit, large residuals, and DIF with respect to 2 exogenous variables. Subsequently item 4 ('numbness') was omitted because of poor fit, and a new Rasch analysis considered. This model showed good overall fit (Andersen CLR = 6.6 on 4 degrees of freedom, $P = 0.1584$) and no evidence of item misfit. Evidence of DIF for item 1 ('pins and needles') with respect to axillary surgery was disclosed. In a comparison of patients with the same total score, patients

Table 1. *Item Fit Statistics for Items Selected for the NeuPPS in the 2 studies.*

Item	Long term Study (n = 2217)			Medium term Study (n = 340)		
	Obs.	Exp.	*P*	Obs.	Exp.	*P*
1 - pins and needles sensations	0.757	0.762	0.802	0.761	0.735	0.632
2 - electric shock/jumping sensations	0.740	0.739	0.960	0.659	0.705	0.395
3 - burning sensations	0.798	0.741	0.048	0.784	0.707	0.226
4 - numbness*						
5 - pain triggered by light touch	0.699	0.738	0.147	0.642	0.713	0.295
6 - pain triggered by cold	0.747	0.741	0.830	0.737	0.751	0.870
DIF	Item 1 with respect to axillary procedure					

Abbreviations: DIF: Differential Item Functioning; Exp: Expected rank correlation; NeuPPS: Neuropathic Pain scale for Post-Surgical patients; Obs: Observed rank correlation. *: not used.

who had axillary lymph node dissection performed scored 0.05 (95% confidence interval [CI]: 0.03 to 0.07) higher on this item compared to patients with sentinel lymph node biopsy only. No evidence of local dependence was disclosed.

The Medium-term Study: Validation of the Model

After exclusion of patients with missing data, the dataset consisted of 340 patients (87% out of 389). Patient characteristics are presented elsewhere (16). Rasch analysis on the final model, developed in the long-term study (with items 1,2,3,5, and 6), showed good overall fit (Andersen CLR = 5.1 on 4 degrees of freedom, $P = 0.2785$) and excellent fit for the individual items. The model did not show any signs of DIF or local dependence.

Combined Data Set

The combined data set consisted of 2,557 cases. The final Rasch model (items 1-3, 5, and 6) showed overall fit (Andersens CLR = 9.35 on 4 degrees of freedom, $P = 0.0530$). There was evidence of DIF for item 1 ('pins and needles') with respect to axillary procedure and for item 6 with respect to study. Further, in the combined data set, local dependence between item 1 and 3 and between item 5 and 6 was present. Graphical evaluation of item fit for the 5 items in both data sets is shown in Fig. 2, which illustrates that the Rasch model fits well in both data sets, but the shape of the curves is seen to be different across data sets for some items. Item fit statistics based on observed and expected item-rest score correlations are shown in Table 1 for both study populations. The conditional versions of the infit and outfit item fit statistics showed a similar picture and can be seen in the supplementary material.

For item 1 and 3, responses were consistently higher in the medium-term study, and for item 5 scores were consistently higher in the long-term study. Figure 3 shows the item profiles for the 2 studies.

S-LANSS Questionnaire in the Long-Term Study

In the long-term study, among 878 S-LANSS questionnaires, 677 S-LANSS questionnaires were answered without any missing items (77.1%). Estimated item threshold and discrimination parameters can be seen in Table 2. The graphical evaluation of item fit indicated good fit for all S-LANSS items except item 4. The Spearman rank correlation between the NeuPPS sum score and the original S-LANSS score was $P = 0.57$. In additional analyses, we recalculated the S-LANSS score using the estimated item discrimination parameters as weights. The rank correlation between the NeuPPS sum score and this re-weighted S-LANSS score was 0.54. Among patients with a score of zero in NeuPPS, 12% had a score of ≥ 12 on the original S-LANSS, and among patients with a score of 5, 89% reported an S-LANSS score of ≥ 12. The NeuPPS and the S-LANSS questionnaire, the latter using equal weighting of the 7 items, were also tested for unidimensionality (n = 621). Expected and observed gamma coefficients were 0.64 and 0.62 with 2-sided $P = 0.44$.

NeuPPS Scores and Persistent Postoperative Pain

Among patients reporting persistent postoperative pain, 20%-32% reported a score of zero in the included studies (Table 3). Approximately 43%-50% of patients with persistent pain reported a score of 2 or more, and > 90% of patients with a score of 4 or 5 reported persistent pain. Among patients not reporting

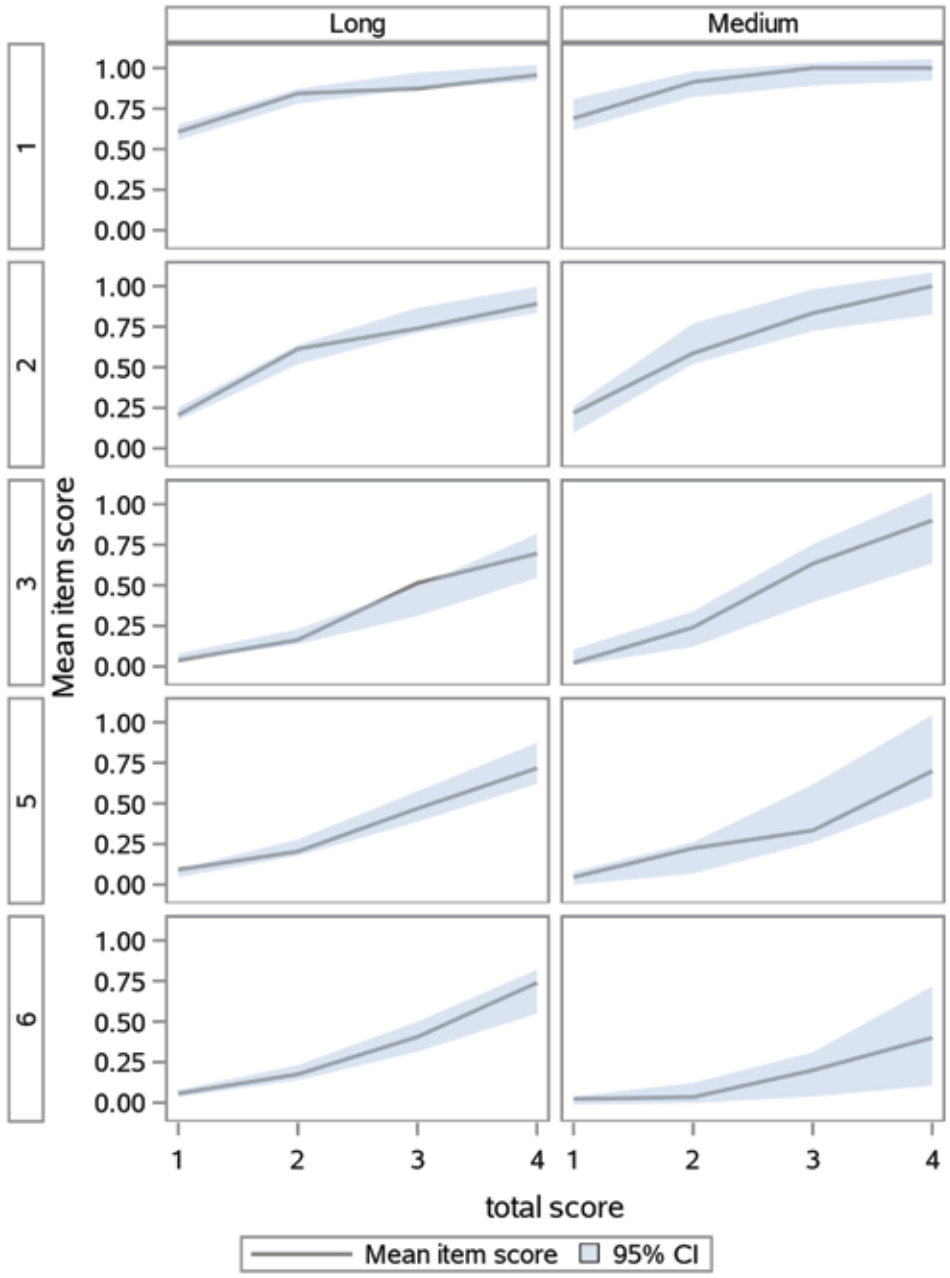

Fig. 2. *The observed item means across score groups (solid lines) and expected values (shaded areas) for the 5 NeuPPS items in each of the 2 studies.*

persistent postoperative pain, only 7% in the long-term study and 11% in the medium-term study reported a score of 2 or more, and 76% and 69% reported zero in score, respectively.

Discussion

In the present study, the aim was to validate a tool measuring sensory dysfunction among postsurgical breast cancer patients. The scale was assessed for

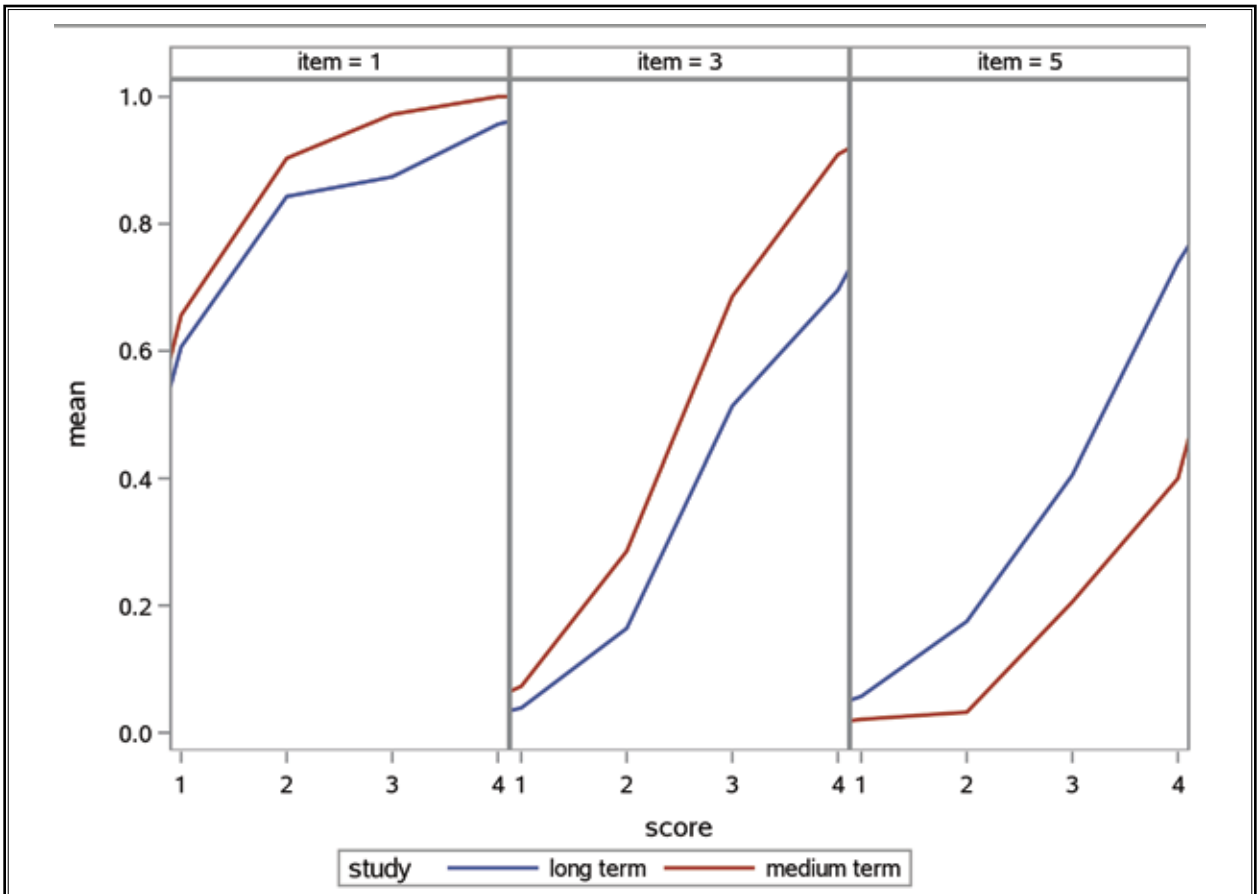

Fig. 3. *The observed item means across score groups for items 1, 3, and 5 in the 2 studies.*

content validity through interviews with patient and health care providers. Convergent validity by comparison with S-LANSS indicated that the present tool measures the same trait as this screening tool, and we observed a fair correlation. Testing of construct validity of the scale in a Rasch analysis demonstrated unidimensionality, monotonicity, good overall fit and homogeneity, additivity, and no local dependence in 2 studies with patients treated surgically for primary breast cancer. The fact that the scale measures the same latent trait as the S-LANSS, that the scale is strongly related to reporting of persistent pain, and that the item "numbness" did not fit with the other items, suggests that the scale is more specifically measuring neuropathic pain. We observed good fit for the individual items except item 4 (numbness),

Table 2. *Estimated item discrimination and item threshold parameters for S-LANSS items using the Birnbaum IRT model.*

Items	Original S-LANSS Scoring Weight*	Item Discrimination** (95% CI)	Item Threshold*** (95% CI)
1 – Dysesthesia	5	1.17 (0.89 to 1.44)	–0.26 (–0.43 to –0.08)
2 – Autonomic	5	0.91 (0.49 to 1.33)	3.36 (2.08 to 4.66)
3 – Evoked	3	1.79 (1.41 to 2.18)	0.41 (0.27 to 0.55)
4 – Paroxysmal	2	0.63 (0.41 to 0.84)	0.98 (0.58 to 1.37)
5 – Thermal	1	1.28 (0.89 to 1.66)	1.92 (1.49 to 2.34)
6 – Allodynia	5	4.96 (3.48 to 6.44)	0.15 (0.04 to 0.25)
7 – Tender/numb	3	2.81 (2.11 to 3.51)	–0.50 (–0.62 to –0.38)

*In the original S-LANSS, a maximum score of 24 is possible. "A score of ≥ 12 suggests pain of predominantly neuropathic origin.
**Item discrimination parameters indicates how well the item captures the latent trait it is measuring.
***Item threshold parameters indicates the location of the item. A score at 0 indicates the mean trait level, and the standard deviation is 1. A parameter of 0 then indicates that a person with a mean trait level will have a 50% chance of a positive answer in this item.

Table 3. *Pain reporting and NeuPPS score in the 2 studies.*

NeuPPS score	Persistent post-surgical pain			
	Long term study (n = 2214)		Medium term study (n = 332)	
	n	% with pain	n	% with pain
0	1326	19.5%	147	22.5%
1	436	46.3%	83	59.5%
2	274	71.2 %	41	73.2%
3	111	86.5%	30	86.7%
4	46	87.0%	9	100.0%
5	21	95.2%	6	100.0%

Abbreviations: NeuPPS: Neuropathic Pain scale for Post-Surgical patients.

suggesting that loss of sensation is somewhat different from the latent trait measured by the other items. A possible explanation could be that a large proportion of patients, also those without persistent postsurgical pain, experience loss of sensation after surgery. In a study using QST, as many as 85% reported hypoesthesia 1 year after breast cancer surgery (39). In our study, 34% of pain-free patients reported sensory loss, with similar findings in other surgical populations (8,40). Sensory loss following axonal deafferentation as a consequence of surgery would be expected, but neuropathic persistent postsurgical pain would probably require afferent input from injured nociceptors (41), arguing for sensory loss as a common consequence after surgery, but not distinctive for neuropathic pain.

The inherent properties of the Rasch model indicate a scale measuring a unidimensional latent trait likely to be neuropathic pain (21,42). This is also supported by our analysis of the S-LANSS and our questionnaire tool, indicating measurement of the same unidimensional latent trait. In the Birnbaum IRT model the S-LANSS items 6 (allodynia) and 7 (tenderness/numbness) weighted much higher than the other items in the S-LANSS. These items are the only 2 self-examination items in the S-LANSS (13,22). Therefore, this model suggests that the self-examination items tell more about the underlying latent trait than the patient descriptor items. Apart from the self-examination items, the scores suggested by the Birnbaum model are quite different from the originally suggested scores (13,22).

The NeuPPS is not limited to pain patients, and besides item 5 and 6 having the word "pain" included in their questions, the other items use the phrasing "sensation" as the specific sensory disturbance experienced by a patient (e.g., pins and needles) may not be interpreted as painful. The Douleur Neuropathique 4 questions (DN4), Neuropathic Pain Questionnaire (NPQ), painDETECT, and Self-complete Leeds Assessment of Neuropathic Symptoms and Signs (S-LANSS) ask specifically how the pain feels, restricting them to pain patients (12,13,15,43). As tools used in a postsurgical setting, assessing for instance the impact of certain procedures and treatments, these screening tools are at risk of losing information on all patients who experience these symptoms of abnormal sensation, but does not necessarily interpret them as painful. Nerve injuries can both be painful or painless (40,42). Provided discrimination between pain locations (surgical site vs. other pain), the S-LANSS, DN4, painDETECT, and NPQ are useful in a pain cohort. The NeuPPS however is location-specific to the surgical site as opposed to the other tools. Among breast cancer survivors, with the S-LANSS, DN4, and painDETECT it can be difficult to discriminate surgically induced neuropathy from other types of neuropathy (e.g., taxane induced neuropathy).

We only observed moderately good fit for the NeuPPS in the combined dataset, with evidence of DIF for item 6 (pain triggered by cold), with respect to study population in the combined dataset. We also observed, that although fit for the individual items was good in the 2 individual studies, there were discrepancies in item scores for item 1, 3, and 5 between the 2 study populations in the combined dataset. This suggests that the items may not be invariant over time, because time since initial treatment exerts a differential influence on these items irrespective of the latent trait we are examining. Another explanation could be different treatment protocols (e.g., changing chemotherapeutic regimens from cyclophosphamide, epirubicin and fluorouracil to cyclophosphamide, epirubicine and docetaxel) influencing items differently. DIF was also observed for item 1 (pins and needles sensations) with respect to axillary procedure in the long-term follow-up study. The DIF was not reproducible in the medium-term study, and the score difference on this item between axillary dissection and sentinel lymph node biopsy was only 0.05. The DIF is therefore of little significance and probably a chance finding.

The strengths of the present study are the use of an IRT model to validate our questionnaire tool, as supported by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) initiative (44). IRT models

have the advantage of assessing the questionnaire for unidimensionality (that the questionnaire is only measuring one trait), monotonicity (that expected item scores increase with an increase of the latent trait), local independence (conditional independence of items given the latent trait), and no DIF (conditional independence of items and exogenous variables given the latent trait). Furthermore, the NeuPPS was validated in a large sample size with high completeness of data, and cross-validated in another sample of breast cancer patients with different follow-up times. The NeuPPS was also validated against the S-LANSS in a subgroup of the long-term study.

There are some limitations to be addressed. We did find DIF, which could introduce bias when comparing studies with large differences in follow-up times. It could also reflect a real change in symptoms of postsurgical neuropathy observed over time. Also, missing data may have introduced bias. Finally, we did not look at test-retest reliability, but given the highly fluctuating nature of pain and sensory disturbances (20), we would not expect a high reliability in a test-retest setup.

Conclusions

With the present study, we have examined and validated a simple and easy-to-fill-out scale, the NeuPPS, for the measurement of neuropathic pain in a setting of postsurgical patients. Displaying good psychometric properties, the scale simply assigns 0 (no) or 1 (yes) point per item, and the items are additive, giving a total score that measures neuropathy and neuropathic pain symptoms with increasing scores. In summary, we can conclude that the NeuPPS has excellent measurement properties but should not be used uncritically if there are patients with large differences in follow-up time.

References

1. Kehlet H, Jensen TS, Woolf CJ. Persistent postsurgical pain: Risk factors and prevention. *Lancet* 2006; 367:1618-1625.
2. Gartner R, Jensen MB, Nielsen J, Ewertz M, Kroman N, Kehlet H. Prevalence of and factors associated with persistent pain following breast cancer surgery. *JAMA* 2009; 302:1985-1992.
3. Aasvang EK, Gmaehle E, Hansen JB, Gmaehle B, Forman JL, Schwarz J, Bittner R, Kehlet H. Predictive risk factors for persistent postherniotomy pain. *Anesthesiology* 2010; 112:957-969.
4. Duale C, Ouchchane L, Schoeffler P, Dubray C. Neuropathic aspects of persistent postsurgical pain: A French multicenter survey with a 6-month prospective follow-up. *J Pain* 2014; 15:24.
5. Andersen KG, Duriaud HM, Kehlet H, Aasvang EK. The relationship between sensory loss and persistent pain 1 year after breast cancer surgery. *J Pain* 2017; 18:1129-1138.
6. Vilholm OJ, Cold S, Rasmussen L, Sindrup SH. Sensory function and pain in a population of patients treated for breast cancer. *Acta Anaesthesiol Scand* 2009; 53:800-806.
7. Gottrup H, Andersen J, Arendt-Nielsen L, Jensen TS. Psychophysical examination in patients with post-mastectomy pain. *Pain* 2000; 87:275-284.
8. Wildgaard K, Ringsted TK, Hansen HJ, Petersen RH, Werner MU, Kehlet H. Quantitative sensory testing of persistent pain after video-assisted thoracic surgery lobectomy. *Br J Anaesth* 2012; 108:126-133.
9. Duale C, Guastella V, Morand D, Cardot JM, Aublet-Cuvelier B, Mulliez A, Schoeffler P, Escande G, Dubray C. Characteristics of the neuropathy induced by thoracotomy: A 4-month follow-up study with psychophysical examination. *Clin J Pain* 2011; 27:471-480.
10. Finnerup NB, Haroutounian S, Kamerman P, Baron R, Bennett DL, Bouhassira D, Cruccu G, Freeman R, Hansson P, Nurmikko T, Raja SN, Rice AS, Serra J, Smith BH, Treede RD, Jensen TS. Neuropathic pain: An updated grading system for research and clinical practice. *Pain* 2016; 157:1599-1606.
11. Aasvang EK, Brandsborg B, Jensen TS, Kehlet H. Heterogeneous sensory processing in persistent postherniotomy pain. *Pain* 2010; 150:237-242.
12. Freynhagen R, Baron R, Gockel U, Tolle TR. painDETECT: A new screening questionnaire to identify neuropathic components in patients with back pain. *Curr Med Res Opin* 2006; 22:1911-1920.
13. Bennett MI, Smith BH, Torrance N, Potter J. The S-LANSS score for identifying pain of predominantly neuropathic origin: Validation for use in clinical and postal research. *J Pain* 2005; 6:149-158.
14. Bouhassira D, Lanteri-Minet M, Attal N, Laurent B, Touboul C. Prevalence of chronic pain with neuropathic characteristics in the general population. *Pain* 2008; 136:380-387.
15. Bouhassira D, Attal N, Alchaar H, Boureau F, Brochet B, Bruxelle J, Cunin G, Fermanian J, Ginies P, Grun-Overdyking A, Jafari-Schluep H, Lanteri-Minet M, Laurent B, Mick G, Serrie A, Valade D, Vicaut E. Comparison of pain syndromes associated with nervous or somatic lesions and development of a new neuropathic pain diagnostic questionnaire (DN4). *Pain* 2005; 114:29-36.
16. Andersen KG, Christensen KB, Kehlet H, Bidstup PE. The effect of pain on physical functioning after breast cancer treatment: Development and validation of an assessment tool. *Clin J Pain* 2015; 31:794-802.
17. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen, Denmark: Nielsen & Lydiche; 1960.
18. Christensen KB, Kreiner S, Mesbah M. *Rasch Models in Health.* Hoboken, NJ: John Wiley and Sons; 2013.
19. Fisher GH, Molenaar IW. *Rasch Models: Foundations, Recent Developments and Applications.* New York, NY: Springer Verlag; 1995.
20. Mejdahl MK, Andersen KG, Gartner R, Kroman N, Kehlet H. Persistent pain and sensory disturbances after treatment for

breast cancer: Six year nationwide follow-up study. *BMJ* 2013; 346:f1865.

21. Bennett MI, Attal N, Backonja MM, Baron R, Bouhassira D, Freynhagen R, Scholz J, Tolle TR, Wittchen HU, Jensen TS. Using screening tools to identify neuropathic pain. *Pain* 2007; 127:199-203.
22. Bennett M. The LANSS Pain Scale: The Leeds assessment of neuropathic symptoms and signs. *Pain* 2001; 92:147-157.
23. Blichert-Toft M, Christiansen P, Mouridsen HT; Danish Breast Cancer Cooperative Group–DBCG. History, organization, and status of scientific achievements at 30-year anniversary. *Acta Oncol* 2008; 47:497-505.
24. Moller S, Jensen MB, Ejlertsen B, Bjerre KD, Larsen M, Hansen HB, Christiansen P, Mouridsen HT. The clinical database and the treatment guidelines of the Danish Breast Cancer Cooperative Group (DBCG): Its 30-years experience and future promise. *Acta Oncol* 2008; 47:506-524.
25. Gartner R, Mejdahl MK, Andersen KG, Ewertz M, Kroman N. Development in self-reported arm-lymphedema in Danish women treated for early-stage breast cancer in 2005 and 2006: A nationwide follow-up study. *Breast* 2014; 23:445-452.
26. Danish Breast Cancer Group (DBCG). Treatment guidelines (2017). *http://dbcg.dk/PDF%20Filer/DBCG_2014_protokoller_oversigt_01.08.2014.pdf*.
27. van der Linden WJ, Hambleton RK. *Handbook of Modern Item Response Theory*. New York, NY: Springer Verlag; 1997.
28. Andersen EB. A goodness of fit test for the Rasch model. *Psychometrika* 1973; 38:123-140.
29. Kreiner S. A note on item-restscore association in Rasch models. *Appl Psychol Meas* 2011; 35:557-561.
30. Kreiner S, Christensen KB. Exact evaluation of bias in rasch model residuals. In: Baswell AR (ed). *Advances in Mathematics Research. 1 ed. Vol. 12*. New York, Nova Science Publishers, 2011: pp. 19-40.
31. Christensen KB, Kreiner S. Item Fit Statistics. In Christensen KB, Kreiner S, Mesbah M (eds). *Rasch Models in Health*. Hoboken, NJ, John Wiley and Sons, 2013: pp. 83-104.
32. Christensen KB. Conditional maximum likelihood estimation in polytomous rasch models using SAS. In: *ISRN Computational Mathematics*. 2013:
33. Holland PW, Wainer H. *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum; 1993.
34. Kelderman H. Loglinear Rasch model tests. *Psychometrika* 1984; 49:223-245.
35. Birnbaum A. Some latent trait models and their use in inferring an examinee's ability. In: Lord FM, Novick MR (eds). *Statistical Theories of Mental Test Scores*. Reading, MA, Addison-Wesley, 1968: pp. 395-479.
36. Olsbjerg M, Christensen KB. *LIRT: SAS Macros for Longitudinal IRT Models*. Copenhagen, Denmark: University of Copenhagen; 2014.
37. Horton M, Marais I, Christensen KB. Dimensionality. In: Christensen KB, Kreiner S, Mesbah M (eds). *Rasch Models in Health*. Hoboken, NJ, John Wiley and Sons, 2013: pp. 137-158.
38. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995; 57:289-300.
39. Andersen KG, Duriaud HM, Jensen HE, Kroman N, Kehlet H. Predictive factors for the development of persistent pain after breast cancer surgery. *Pain* 2015; 156:2413-2422.
40. Aasvang EK, Kehlet H. Persistent sensory dysfunction in pain-free herniotomy. *Acta Anaesthesiol Scand* 2010; 54:291-298.
41. Campbell JN, Meyer RA. Mechanisms of neuropathic pain. *Neuron* 2006; 52:77-92.
42. Hansson P. Neuropathic pain: Clinical characteristics and diagnostic workup. *Eur J Pain* 2002; 6 Suppl A:47-50.
43. Krause SJ, Backonja M. Development of a neuropathic pain questionnaire. *Clin J Pain* 2003; 19:306-314.
44. Turk DC, Dworkin RH, Burke LB, Gershon R, Rothman M, Scott J, Allen RR, Atkinson JH, Chandler J, Cleeland C, Cowan P, Dimitrova R, Dionne R, Farrar JT, Haythornthwaite JA, Hertz S, Jadad AR, Jensen MP, Kellstein D, Kerns RD, Manning DC, Martin S, Max MB, McDermott MP, McGrath P, Moulin DE, Nurmikko T, Quessy S, Raja S, Rappaport BA, Rauschkolb C, Robinson JP, Royal MA, Simon L, Stauffer JW, Stucki G, Tollett J, von Stein T, Wallace MS, Wernicke J, White RE, Williams AC, Witter J, Wyrwich KW; Initiative on Methods, Measurement and Pain Assessment in Clinical Trials. Developing patient-reported outcome measures for pain clinical trials: IMMPACT recommendations. *Pain* 2006; 125:208-215.