Title: The Leeds Assessment of Neuropathic Symptoms and Signs Scale (LANSS) is not an adequate outcome measure of pressure ulcer-related neuropathic pain

Running head: Validation of LANSS for PUs

Authors: *C. Rutherford[1,2], J. Nixon[2], J. M. Brown[2], M. Briggs[3], M. Horton[4]

[1]School of Psychology, University of Sydney, NSW, Australia,
[2]Leeds Institute of Clinical Trials Research, University of Leeds, UK,
[3]Institute of Health and Well Being, Leeds Metropolitan University, Leeds, UK,
[4]Psychometric Laboratory for Health Sciences, University of Leeds, Leeds, UK

*Corresponding author: QOL Office, University of Sydney, Sydney NSW 2006, Australia; Tel: +61(0)286271583; Fax +61 (0)290365292; claudia.rutherford@sydney.edu.au

Original Article

Funding sources: The authors would like to thank Molnlycke Health Care for funding this work.

Conflict of interest: The authors declare that they have no conflict of interest.

What does this study add?

- The Leeds Assessment of Neuropathic Symptoms and Signs scale (LANSS) is not suitable as an outcome measure of pressure ulcer-related neuropathic pain as it did not meet requirements for reliable and valid measurement in this population

Abstract

Background: Few pain assessment scales have been used in Pressure Ulcer (PU) research and none developed or validated for people with PUs. We examined the Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) scale to determine its utility as an outcome measure for people with pressure-area related pain.

Methods: LANSS data from 728 participants underwent psychometric analyses: traditional tests for data quality, scaling assumptions, reliability and validity and a Rasch analysis including tests of fit, spread and targeting of item locations, response dependency, person separation index (reliability) and differential item functioning.

Results: Our findings offer support for a unidimensional scale; confirmatory factor analysis indicated a non-significant Chi-Square test of model fit ((df =14) 23.48, p= 0.053). However, some misfit was identified at the overall scale and individual item levels, and internal construct validity of the LANSS as an outcome measure for neuropathic pain in people with pressure-area related pain was not supported; low to moderate item-total correlations (Chi Square (df = 28) 55.546, p = 0.002) and inter-item correlations (mean 0.117 and range from 0.063 - 0.415); and low Cronbach's alpha (0.549) and Person Separation Index (0.334).

Conclusions: Requirements for reliable and valid measurement do not support the use of the LANSS as an outcome measure in people with PUs at the individual level or as a generalised measurement scale of neuropathic pain across ulcer severity groups. Expanding the number of items to aid differentiation between neuropathic pain levels and improving scale reliability is recommended.

Introduction

Pressure ulcers (PUs) are painful chronic wounds of the skin and deeper soft tissue, occurring primarily in areas of bony prominence and in immobile individuals (National Pressure Ulcer Advisory & the European Pressure Ulcer Advisory, 2009; Coleman *et al.*, 2013). Patients report pain and discomfort as a major burden affecting quality of life (Szor & Bourguignon, 1999; Reddy *et al.*, 2003; Rastinehad, 2006; Gorecki *et al.*, 2009; Gorecki *et al.*, 2010) and that their pain has been underestimated by healthcare professionals, under-treated, and under-assessed (Gorecki *et al.*, 2012). For people with PUs, pain can be cyclic; intermittent discomfort occurring with repetitious treatments or movements, noncyclic; pain occurring during a particular event (e.g. debridement), and chronic; background pain, varying in severity, intensity and duration (i.e. persistent or intermittent), that occurs without manipulation (Krasner, 1995). Differentiating between different types of pain has important implications for diagnosis and treatments, and ultimately effects on patients, by identifying pain, targeting therapies and expediting relief.

As part of a NIHR funded programme of PU research (PURPOSE RP-PG-0407-10056) four related studies were undertaken. The systematic review proposed that to achieve the best possible outcomes important to patients, improved communication of pain experienced between the individual and their healthcare provider and pain assessment, interventions to help control or reduce PU pain, patient-centred concerns, and systemic barriers need to be considered when managing PUs to ensure effective PU pain management (Gorecki *et al.*, 2011). The hospital prevalence reported pressure area-related pain prevalence of 16.3% (327/2010) including 233 patients with no observable PU (Briggs *et al.*, 2013). The community prevalence study included only patients with PUs and reported pressure area-related pain of 75.6% (133/176) (McGinnis *et al.*, 2014). A prospective cohort study found that he distribution of pain intensity was similar for all PU grades, and both inflammatory and neuropathic pain was observed (Nixon *et al.*, 2015. The dominant type of pain in hospital patients was inflammatory pain (70.3% torso and 60.3% limb), whilst in the community patients neuropathic pain was dominant (54.5% torso and 61.1% limb). The combined work identifies the extent of the problem in the PU and at-risk PU populations, the impact upon patients, and indicates the need for improved assessment, targeted treatment, and tools to assess treatment outcomes (Nixon *et al.*, 2015).

Selection of outcome measures to assess treatment response in research and clinical practice needs to consider instruments with evidence of reliability, validity and responsiveness (i.e. ability to detect change in patients) (Scientific Advisory Committee of the Medical Outcomes Trust, 2002; FDA., 2009). Few pain assessment scales have been used in PU research and none developed specifically for people with PUs (Gorecki *et al.*, 2011). In the prevalence and cohort work (Briggs *et al.*, 2013; Nixon *et al.*, 2015), screening of neuropathic pain was undertaken using the Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) (Bennett, 2001). The LANSS is a screening tool (Searle *et al.*, 2011) with evidence for reliability and validity of its ablity to discriminate between neuropathic and nociceptive pain (Briggs *et al.*, 2007). However, it has not been assessed for use as an outcome measure for research or clinical practice in chronic wounds. The aim of this study was to investigate whether the LANSS is suitable for use as an outcome measure in people with PUs by determining whether the LANSS satisfies traditional psychometric and Rasch model expectations as a summed score.

Methods

Design

To determine the LANSS's utility as a useful, scientifically robust and clinically meaningful outcome measure, the measurement properties of the LANSS were evaluated in people with reported pressure area related pain. A secondary analysis was undertaken on data collected from two cross-sectional PU prevalence studies (Briggs *et al.*, 2013; McGinnis *et al.*, 2014) and a large prospective cohort study (Nixon *et al.*, 2015). LANSS data was available from three patient populations: acute hospital prevalence (n=157), community prevalence (n=37), and acute and community cohort (n=534)(Nixon *et al.*, 2015). Data from these three samples was used to ensure representation of patients across PU severity groups. The studies were approved by the Leeds Central Research Ethics Committee (REC) prior to data collection; Ref No 09/H1313/14 for prevalence studies and Ref No 09/H1313/32 for cohort study. Additional ethical approvals were not required for this secondary analysis.

Eligibility and recruitment

Hospital prevalence study eligibility

All inpatients (aged 18-years or older) from 9 UK participating hospitals, including those with and without PUs and considered by the clinical team to be well and able to report the presence or absence of localised skin pain were asked about their pressure area related and PU pain. Patients in paediatric, obstetric and psychiatric care settings or considered ethically or clinically inappropriate to approach (e.g. severely ill or death was imminent) were excluded. Pain was assessed by asking two pain screening questions as follows:

1. At any time, do you get pain, soreness or discomfort at a pressure area (prompt: back, bottom, heels, elbows or other as appropriate to the patient)? (Yes or No)
2. Do you think this is related to either; your PU OR lying in bed for a long time OR sitting for a long time? (Yes or No)

Patients with reported pain on a pressure area skin site or PU where then consented to a full pain assessment including the LANSS (Briggs *et al.*, 2013).

Community prevalence study eligibility

All community nursing patients (aged 18 years and over), from 2 UK NHS community services including only those with PUs (Grade 1-4/Ungradable) and considered well and able to report the presence or absence of localised skin pain by the clinical team were asked about their PU pain. Patients were asked two screening questions (as above) and those with reported PU pain where then consented to a full pain assessment including the LANSS (McGinnis *et al.*, 2014).

Cohort study Eligibility

Acutely ill hospital and community patients 18 years and over and at-high risk of PU development were recruited to the cohort study. High risk of PU development was defined as the presence of one or more of the following: a) bedfast/chairfast and completely immobile/very limited mobility, b) localised skin pain on any pressure area skin site, or c) Category 1 PU (National Pressure Ulcer Advisory & the European Pressure Ulcer Advisory, 2009) on any pressure area skin site (see Nixon et al 2015) for full eligibility criteria). After written or witnessed, informed consent pain screening and full pain assessment including the LANSS was undertaken (Nixon *et al.*, 2015).

Data collection

In all three studies, data was collected by a member of the tissue viability team (e.g. tissue viability nurse consultant/specialist or research nurse). Site initiations were performed prior to the study commencing at each centre. Clinical research and tissue viability nurses had training

in the completion of the CRFs by the Clinical Co-ordinator. Data recorded included: place of assessment; date of birth; gender and ethnicity. In the prevalence studies skin was assessed using the 1998 EPUAP(Panel, 1998) classification and recorded for a minimum of 13 skin sites (sacrum, left and right buttocks, hips, heels, ankles and elbows). In the cohort study, skin was assessed using the revised EPUAP/NPUAP 2009 classification (National Pressure Ulcer Advisory & the European Pressure Ulcer Advisory, 2009). In addition, the presence of an unstageable PU, other type of wound and normal skin were confirmed or skin status was recorded as not applicable (e.g. amputation) or unable to assess.

Full pain assessment

Pain was assessed by asking patients to report the pain intensity (for most severe pain over the past week) for all pressure areas using a visual analogue scale (VAS) of 0 – 10 (Royal College of Physicians, 2007; Dworkin *et al.*, 2008). Patients were also asked to identify their most painful torso and limb skin sites and these were assessed using the LANSS Pain Scale (Bennett, 2001). The LANSS consists of a brief assessment and is easy to score in the clinical setting. It contains five patient-reported symptom items and two clinical sensory testing items associated with neuropathic pain. The five symptom questions identify whether the patient experiences any phenomena associated with neuropathic pain using yes/no responses; a 'yes' response is associated with a weighted value, resulting in a maximum possible score of 16 depending on the number of phenomena experienced. The clinical assessment is designed to identify allodynia (pain to non-painful stimulus such as pain elicited by gentle skin stroking when normal sensation is experienced in a control site) and raised sensory threshold (e.g. sensory threshold assessed as present by a nurse when patients are unable to detect pressure from a 10 g monofilament (Bailey Instruments Ltd., Manchester, UK) at the ulcer site in contrast to a control site). A maximum score of 8 is available for the clinical assessment. Scores from the symptom questions and clinical assessment are combined; scores between 0 – 24 are generated. If the LANSS total score is less than 12, then neuropathic mechanisms are unlikely and the pain is classified as inflammatory pain. If the total score is greater than or equal to 12, then neuropathic mechanisms are likely to be contributing to the pain and it is classified as neuropathic (Bennett, 2001).

The LANSS was adapted for use in this population and to facilitate this secondary analysis. The responses to items were treated as dichotomous without any weighting. We added a third response option 'not applicable' (N/A) to the standard response question relating to skin colour. Patients with PUs cannot easily answer this question due to the location of their PU and/or presence of dressings. 'N/A' responses were treated as missing; no value was imputed for this item so the item calibrations were only based on the people for whom this item was relevant. This meant that the 'difficulty' of the item would not be over- or under-estimated by treating the symptom as 'not present' or 'present' respectively. Likewise, the person estimates were only based on the items that participants responded to, so the only impact of not responding to the item was that the standard error of their estimates would be larger as they had responded to fewer items. Methodologically, this is the most valid way to deal with missing data for these types of analyses, so the impact of this amendment serves to improve the accuracy of the LANSS and highlight the relevance of item 2 within this particular patient group.

The prevalence data included one assessment time-point for all patients. The cohort study assessed pain at baseline and multiple follow-ups, but only data from patients recruited to the study who had reported pressure area related pain at baseline was utilised in this analysis.

Analyses

Analyses were undertaken on the combined prevalence and cohort patient population data. Although these three patient samples may be considered different (i.e. hospital prevalence sample included any hospital patient with or without PUs; community prevalence sample included only patients with a PU; and the cohort sample included acutely ill high-risk patients), they all had a common feature; all patients reported to have pressure area related pain. As such, it was considered appropriate to combine these samples for the secondary analysis. For patients that had a LANSS completed on two skin sites (i.e. both a limb and a torso skin site), as the frequency of limb site assessments was fewer than torso, only data from the LANSS completed for the limb site was included. Statistical analysis was undertaken with SPSS (Francis, 2001) and RUMM2030 software (Andrich *et al.*, 2010).

Reliability

Internal consistency reliability was assessed by Cronbach's alpha coefficients (Cronbach, 1951). Alpha provides an indication of the degree of convergence between items hypothesised to represent the same variable. Adequate scale internal consistency estimates of 0.7 are considered acceptable for group comparisons (Nunnally & Bernstein, 1994). The relationship between an item and the total score was assessed using item-total correlations, expressed as a number between -1.0 to +1.0. Item-total correlations (ITC) in the range of around 0.4 - 0.6 indicate items are moderately correlated with scale scores; higher values indicate well correlated items with scale scores (Nunnally & Bernstein, 1994). Reliability was further tested by the Person Separation Index from the Rasch analysis (see below). The Person Separation Index should be interpreted in the same way as a Cronbach's alpha value, where a minimum value of 0.7 is required for group use and 0.85 for individual use as an outcome measure (Tennant & Conaghan, 2007).

Internal construct validity

The internal construct validity of the LANSS was assessed using the Rasch unidimensional measurement model (Rasch, 1960; Andrich, 1988). The Rasch model defines how a set of items should perform to generate reliable and valid measurements (Andrich, 1989) and evaluates the legitimacy of summing items to generate measurements (Rasch, 1960; Andrich, 1988). A Rasch analysis examines the extent to which the observed data (patients' actual responses to scale items) are concurrent with ('fit') predictions of those responses from the Rasch model; whereby the difference between expected and observed scores indicates the degree to which rigorous measurement is achieved. The expected response structure for the Rasch model is a probabilistic Guttman pattern, which assumes that for the same person ability, the probability of endorsing an easy item is higher than the probability of endorsing a more difficult item, and vice versa(Andrich, 1985). When a rating scale is used to discriminate between persons with different abilities, someone with higher ability is expected to affirm all items endorsed by a person with lower ability in addition to items representative of higher ability. Within this context, the 'ability' of a person is a representation of their level of neuropathic pain, whereas the 'difficulty' of an item refers to the level of neuropathic pain as represented by an item. An 'easier' item will be more frequently endorsed, and therefore represents a lower level of neuropathic pain.

Data were fitted to the Rasch model to determine how well the LANSS scale items, both individually and collectively, conform to the Rasch measurement model. A range of fit statistics are available, which are described elsewhere (Pallant & Tennant, 2007; Tennant & Conaghan, 2007), and the main fit statistics used were chi-square fit statistics and individual item fit residuals. Generally, fit residuals should lie between +/-2.5 and chi-square values should be non-significant (Bonferroni adjusted). In addition, the traditional test theory statistics of item-

total correlations, alpha coefficient and homogeneity coefficient (inter-item correlation mean and range) provide evidence towards internal construct validity.

External construct validity

- Convergent validity

Convergent validity - the degree to which constructs (or scores on an outcome measure) expected to be related are, in fact, related (Campbell & Fiske, 1959) - was examined by computing a Spearman rank-order correlation coefficient between the LANSS and pain intensity VAS scores to determine how closely the LANSS was related to a similar construct (i.e. pain intensity). For exploratory purposes, we hypothesised that correlations may be high with the VAS based on the proximity of the constructs; criteria were used as a guide to the magnitude of correlation, as opposed to pass/fail benchmarks: high (>0.70); moderate (0.30 - 0.70); low (<0.30) correlation (Cohen, 1960; Burnand *et al.*, 1990).

- Discriminant validity

Discriminant validity (or divergent validity) – the degree to which constructs expected to not be related (have no relationship) are, in fact, not related (Campbell & Fiske, 1959) - was examined by computing a Spearman correlation coefficient between LANSS scores and age (<70 vs. ≥70-years of age) to determine the extent to which responses were related to these variables. A low (<0.3) correlation was predicted.

- *Known groups hypothesis testing*

Known-group comparisons are used to evaluate the clinical utility of outcome measures by assessing the extent to which scales are able to discriminate between subgroups of patients known to differ in terms of clinical presentations (Kerlinger, 1973). We explored the validity of group differences by examining LANSS scores for four groups of patients defined by 'healthy intact skin' or PU severity classified as category 1, category 2 and "severe" (category 3 and 4 combined due to small sample sizes) groups (see Table 1). A Kruskal-Wallis non-parametric ANOVA was used to test for statistically significant differences between the LANSS median scores of different ulcer severity groups. We predicted a gradient in LANSS scores across severity groups, with differences between adjacent categories to be moderate in size (i.e. clinical importance) and the difference between category 1 and the "severe" category to be large.

Differential Item Functioning

Differential item functioning (DIF) (Rasch, 1960; Andrich, 1988) provides a method for exploring conditional relationships between item response and group membership by examining the significance of differences observed between different groups within a Person Factor (Teresi *et al.*, 2008). DIF occurs when people from different groups (e.g. age groups), at the same level of the latent trait (e.g. neuropathic pain), have a different probability of giving a certain response to an item; respondents at any specific level of neuropathic pain should respond in similar ways to individual items irrespective of group membership. This was tested though DIF analysis across gender, age (18-69-years; 70-years and over), and healthcare setting (acute hospital prevalence; community prevalence; prospective cohort) groups.

Local dependence and unidimensionality

A test of unidimensionality within the Rasch analysis framework, as suggested by Smith (Smith, 2000), is likely to be underpowered within the LANSS scale and was therefore not conducted. However, to assess the unidimensionality of the scale a confirmatory factor analysis based upon a tetrachoric correlation was carried out using MPlus version 6 (Muthén & Muthén, 2011).

Table 1 Analysis sample characteristics

	Acute prevalence (n=157)	Community prevalence (n=37)	Cohort baseline (n=534)	Combined (n=728)
Characteristics	**Range (Mean, SD)**	**Range (Mean, SD)**	**Range (Mean, SD)**	**Range (Mean, SD)**
Age	18 - 99 (66, 18.4)	23 - 98 (73, 15.3)	22 - 102 (79, 12.9)	18 - 102 (76, 15.3)
	Total n (%)	**Total n (%)**	**Total n (%)**	**Total n (100%)**
Gender				
Male	78 (49.7)	9 (24.3)	211 (39.5)	298 (40.9)
Female	79 (50.3)	28 (75.7)	323 (60.5)	430 (59.1)
Ethnicity				
White	154 (96.3)	37 (100.0)	533 (99.8)	707 (97.1)
Other	3 (3.7)	0 (0.0)	1 (0.2)	12 (1.7)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	9 (1.2)
Ulcer location				
Torso	90 (57.3)	19 (51.4)	258 (48.3)	367 (50.4)
Limb	67 (42.7)	18 (48.6)	276 (51.7)	361 (49.6)
N patients with pressure ulcer	75 (46.9)	37 (100)	475 (89.0)	587 (80.6)

N pressure ulcers	139	54	489	682
Ulcer severity[a]				
Healthy intact skin	-	0	70	70
Alterations to intact skin	-	0	150 (30.7)	150 (22.0)
Grade/Category 1	97 (69.8)	20 (37.0)	176 (36.0)	293 (43.0)
Grade/Category 2	32 (23.0)	17 (31.5)	140 (28.6)	189 (27.7)
Grade/Category 3/4	7 (5.0)	13 (24.1)	17 (3.5)	37 (5.4)
Unstageable	3 (2.2)	4 (7.4)	6 (1.2)	13 (1.9)
Missing	0	0	19	19

[a]Some patients had more than one pressure ulcer. The numbers and percentages represent the total number of ulcers in the sample.

Response dependency is where items are linked in some way, over and above what is explained by the underlying latent trait, such that the response to one item will overtly influence the response to another, different item. Response dependency was investigated by inspecting the residual correlations (Wright & Masters, 1982) for pairs of items with correlations exceeding 0.2. Correlated items imply that the response to one item is dependent on another, thus correlations among residuals should be lower than this benchmark value.

Results

The sample included 728 patients who reported to have pressure area related pain (157 patients from the acute hospital and 37 from the community prevalence survey, and 534 patients (excluding those with healthy intact skin) from the prospective cohort sample). A total of 367 LANSS were completed for a torso skin site and 361 for a limb skin site. Of these, 19 were excluded due to missing LANSS data, resulting in a total of 709 patients with complete data in the analysis sample. Table 1 presents participant characteristics.

Reliability

Internal consistency reliability was moderate as demonstrated by a Cronbach's alpha value of 0.549; below the standard criterion of 0.7. The Person Separation Index was 0.334 (Table 2).

Internal construct validity

Most items were found to fit the model; only one item had a significant chi-square value (Table 3). The item-trait interaction was significant; failing to support invariance of items (Chi Square (df = 28) 55.546, p = 0.002; Table 2). However, inter-item correlations were low to moderate (mean 0.117 and range from 0.063 - 0.415 indicates that LANSS items were not all correlated with total scale score) and item–total correlations were low to moderate (range from 0.156 - 0.421); not all items fulfilling the recommended criteria of >0.3 (Table 3).

Table 2 Reliability and scaling assumptions: validity within-scale analysis (n=709)

Scale (N items)	Rasch Analysis					Traditional Psychometric Analysis					
	Fit Residual		(df) Total-item Chi-square	Chi square p	PSI	Cronbach alpha	SEM	95% CI	Mean IIC	Range IIC	Scaling Assumptions Corrected ITC range
	Mean	SD									
LANSS (7)	1.270	1.467	(28) 55.546	0.002	0.334	0.549	0.227	7.79, 8.68	0.117	0.063 – 0.415	0.156 – 0.421

n number; SD standard deviation; PSI person separation index; SEM standard error mean; CI confidence interval; IIC inter-item correlation; ITC item-total correlation

Table 3 Rasch analysis summary statistics for LANSS items

Original scale item number and description[a]	Location (Difficulty)	Item FR	Chi Square	Chi Square p	Residual r
3. Pain makes affected skin abnormally sensitive to touch	-0.731	1.247	5.465	0.243	<0.2
5. Pain feel as if the skin temperature in the painful area has changed abnormally	-0.405	2.495	3.858	0.426	<0.2
7. Altered sensory threshold (clinical question 2)	-0.351	2.019	6.470	0.167	<0.2
2. Pain make skin in painful area look different from normal	-0.348	2.479	2.916	0.572	<0.2
1. Pain feels like strange, unpleasant sensations in your skin	0.271	1.580	5.067	0.281	<0.2
6. Allodynia (clinical question 1)	0.693	-1.746	27.362[b]	0.000[b]	<0.2
4. Pain come on suddenly and in bursts for no apparent reason when you're still	0.871	0.817	4.407	0.354	<0.2

[a]Items ordered by 'difficulty' (logit location) value, with the easiest to endorse listed first; [b]outside Bonferroni adjusted recommended criteria (indicates misfit)
FR fit residual; p probability; r correlation;

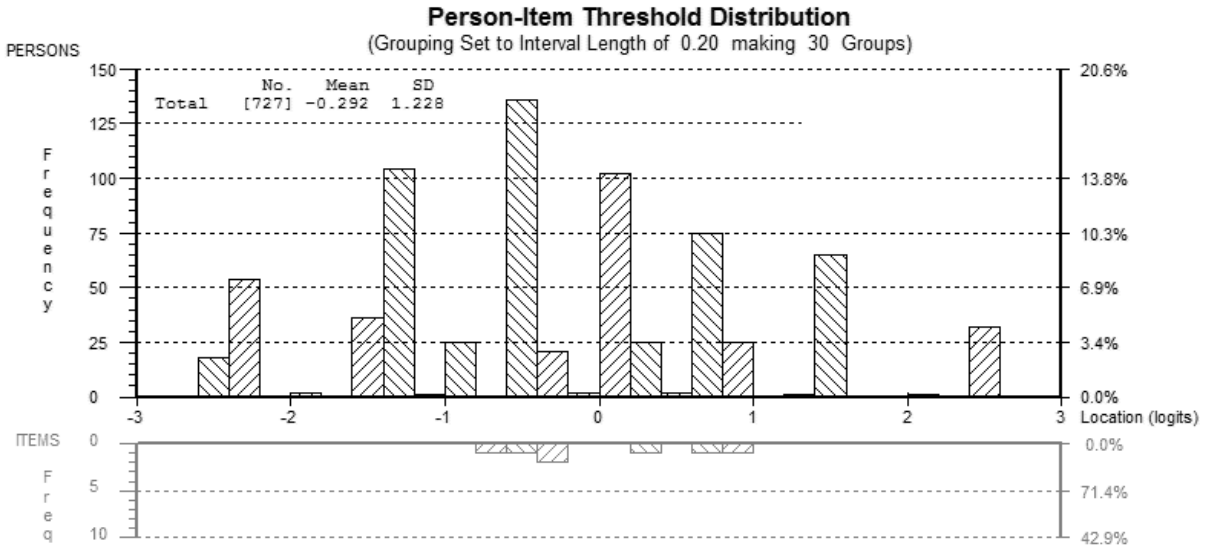

Figure 1 Targeting of scale to patient level of neuropathic pain: item-person threshold distribution for patients with pressure area related pain

Confirmatory factor analysis on the complete sample of 728 cases suggested evidence of a unidimensional construct. The confirmatory factor analysis indicated a non-significant Chi-Square test of model fit (value = 23.48, degrees of freedom = 14, p= 0.053), a Root Mean Square Error of Approximation of 0.04, a Comparative Fit Index of 0.96 and a Tucker-Lewis Index of 0.94.

Within the context of the LANSS proposed 'neuropathic pain' construct, the item locations spread out almost 2 logits (-0.731 – 0.871; Table 3) indicating that the LANSS defined a narrow measurement range. Figure 1 displays the distribution of person "ability" (in this case the level of neuropathic pain) and LANSS question difficulty on the same logit scale. The figure illustrates that the item difficulties are spread through the middle of the person ability range (Figure 1). This evidence of a floor and ceiling effect means that the scale is not measuring at lower and higher levels of neuropathic pain. A mean person score of -0.292 logits suggests that patients in this sample displayed a slightly lower average level of neuropathic pain than is being measured by the scale (Figure 1).

In this sample, all residual correlations were below the threshold. This implies that the responses to the items are independent of each other and that the items are locally independent (Table 3).

Differential item functioning

DIF was tested for age, gender, and healthcare setting. Across all seven items, no DIF was found for items between age and gender groups. Significant uniform DIF was observed in only one item between healthcare setting cohorts (item 4). This suggests that the LANSS scale items are largely measuring the same construct amongst gender, age, and healthcare setting groups.

External construct validity

- *Convergent validity*

The correlation between the LANSS and pain intensity VAS was low (r = -0.21; Table 4), indicating that neuropathic pain as measured by the LANSS was not related to pain intensity; moderate to high correlations (r >0.30) were predicted.

- *Discriminant validity*

Correlations between the LANSS and age groups were consistent with predictions (r <0.30; Table 4), thus suggesting that responses to the LANSS are not biased by age.

- *Known Groups hypothesis testing*

A Kruskal-Wallis test revealed a statistically significant difference in LANSS scores across the ulcer severity groups (Table 4). The category 2 ulcer group recorded a higher median score than category 1 or healthy skin groups (Table 4) suggesting that this group had the highest LANSS scores, with category 3/4 ulcer group reporting the lowest. This finding suggests that neuropathic pain is worst for people with superficial PUs (i.e. category 1 or 2 ulcers; Table 4). It is important to note that category 3 and 4 ulcers had small samples (n=24).

Table 4 External construct validity

Scale (pressure ulcer severity^)	Convergent Validity Pain intensity VAS r[a] (n)	Discriminant Validity Age r[a] (n)	Known Groups Median score (total n=497)	Known Groups Chi-Square (df), Asymp. Sig[bc] (mean rank)
LANSS	0.210[de] (699)	-0.138[df] (709)		10.985 (3), 0.012
(0-Healthy intact skin)			5.0 (175)	(227.48)
(Category 1)			7.0 (208)	(255.50)
(Category 2)			8.5 (90)	(284.24)
(Category 3/4)			4.5 (24)	(217.48)

CI confidence interval; VAS visual analogue scale; r correlation

[a] Spearman correlation; [b] Kruskal-Wallis test

[c] Pressure ulcer severity categorised into 4 groups: healthy skin, category 1, category 2, and category 3 and 4 combined

[d] Correlation significant at p=0.01 (2-tailed)

[e] Correlations falling outside of the predicted range; [f]Correlations consistent with predictions

Discussion

Our findings suggest that the LANSS items largely measure the same construct across gender, age, healthcare setting and skin status groups, and the confirmatory factor analysis offers support for a unidimensional scale. However, there is some misfit at the overall scale and individual item levels, and evidence towards internal construct validity of the LANSS as an outcome measure of neuropathic pain in people with pressure area-related pain was not supported: low to moderate item-total correlations and inter-item correlations, and low Cronbach's alpha and Person Separation Index values fail to support the LANSS as measuring a single construct (i.e. neuropathic pain) in people with pressure area related pain. This evidence combined suggests that the LANSS should not be used as an outcome measure in people with PUs.

A measurement scale is intended to measure one construct (unidimensional). Each item within a scale is a component of the construct being measured and each construct (scale) is operationalised by the items (content) within it. To enable measurement of clinical variables, each scale requires a set of items that map out the construct they purport to measure, and mark out a measurement continuum upon which patients can be located and monitored for changes (Hobart *et al.*, 2007). Importantly, measures employed for assessing construct validity found that some LANSS items were poorly correlated with the total score. It may be that the LANSS scale in fact measures two separate constructs and combining the five patient-reported and two clinical items to obtain a total score is not appropriate. The evidence for screening tools for neuropathic pain propose that it is more accurate to measure both constructs (patient reports of signs and clinical exam) and take the combined result (Haanpaa *et al.*, 2011). However, this may not apply to outcome measurement where scales should be unidimensional. This is a challenge when considering a scale that combines patient report and clinical exam. Important to highlight is that the LANSS was developed to be used as a screening tool. Screening tools provide a useful indicator of when to suspect neuropathic pain but may be limited as outcome measures to guide clinical decision making to improve the treatment of neuropathic pain and monitor changes. Further, the heterogeneous nature of the PU population may preclude the use of the LANSS in this way.

There is some evidence to suggest that the clinical items are loading together within the residual correlations and suggestion that item 6 (Allodynia) overdiscriminates. One of the five patient reported symptom questions (item 2) (relating to a change in skin colour) was assessed as 'N/A' by 75% of participants, calling into question this item's relevance in the context of PUs. Consequently, retaining an item that appears to not be relevant to PUs in the total score may reduce the severity of neuropathic pain in the sample.

The Rasch analysis detected another important limitation; problems with targeting. Inspection of threshold distributions demonstrated adequate scale-to-sample targeting but potentially poor item-to-sample targeting; the range mapped out by the LANSS items was poorly matched to the range of variable in this sample (items did not span the full range of the patient sample). This finding indicates that the measurement range should be increased if it was to be used as an outcome measure. For example, the items spread 2-logits compared to a person spread of 5-logits, indicating a narrow measurement range.

Ideally, the range of outcome measured by the LANSS should be well-matched to the range of outcome (i.e. neuropathic pain) present in the study sample; we expect to see a good match between the scale and sample ranges, with people falling within the range of the items, so that the scale has the ability to detect variability among and within individuals. This is important because the limited item level targeting will impact on the overall ability of the LANSS to detect differences between people and groups and potentially be less sensitive to the effects of treatments. Poorly targeted scales are likely to underestimate changes over time and differences

between groups (Hobart *et al.*, 2007), which is relevant for future PU clinical trials that tend to recruit people with superficial ulcers (Nixon *et al.*, 2006). Our findings demonstrated greater variability in the sample than measured by the LANSS items. However, it is expected that in a sample of people with mixed severity ulcers, people would have different types of pain as well as some neuropathic pain. Worthy of note is that the sample was not evenly distributed, with many people reporting pain scores at the lower range of the scale; which may be a reason for this finding. The results suggest therefore, that the LANSS is not suitable for measuring change at the individual person level in this patient population. This finding is consistent with others who did not find support for the measurement properties of the LANSS and concluded that the LANSS is not suitable for clinical use (Mathieson *et al.*, 2015).

Some might argue that combining the three patient samples may have introduced too much clinical heterogeneity and in part, contributed to mistargeting. The study sample consists of prevalence (all hospital patients on day of survey; some have PUs but most do not but they do experience pain on pressure areas, and community sample; all have a PU) and cohort study patients (high-risk patients with and without PU but experience pain on pressure areas). Despite some differences between the samples, these patient groups represent the 'real' clinical world and they all have the same thing in common, they all have pain attributed to pressure. Further, DIF by healthcare setting mostly supported combining the samples, with only one item demonstrated significant DIF suggesting a potential problem. Of note is that the LANSS is only likely to be a candidate for outcome measurement in very narrow patient populations (e.g. diabetic neuropathy or post-surgical pain) (Searle *et al.*, 2011).

The targeting issues and findings from the known group analysis suggest that the LANSS may be limited in its ability to detect clinical change when it occurs and some items may underestimate differences in pain experienced for people with a severe ulcer. This has important implications for the inferences of future research using the LANSS as an outcome measure. Another consideration is that during the study period, as is standard practice, many patients would have received some form of treatment for their PU. This information was not accounted for in the analysis (e.g. amount of analgesia). As such, the true amount of pressure area-related pain may not have been captured (lower severity represented in the sample due to treatment effect) and be the reason for, at least in part, misrepresentation of known groups testing. However, due to small samples, the known groups validity results are preliminary and require further empirical evidence to determine whether the LANSS can differentiate known groups (e.g. ability to detect small differences in neuropathic pain by PU severity). The LANSS may benefit from further examination of the measurement properties in severe PU samples.

Our findings display similar results to those of Searle et al (Searle *et al.*, 2011), albeit in a different clinical context; the reliability of the LANSS does not allow statistical interpretation of LANSS results in people with pressure area-related pain, and cannot reliably be used to measure change in these patients. In addition, the item-person distribution map illustrates that the LANSS is unable to discriminate between people at either end of the scale. This is likely to reflect the relatively low number of items in the LANSS scale; seven items does not present a high degree of precision, nor are they suitable for outcome measurement as they have extremely poor precision at the margins of the scale. For scales where a measurement hierarchy is required to capture less of or more of a construct (e.g. different types of pain), multiple descriptors (items) are needed. The LANSS remains primarily a screening tool; properties that were not investigated in this study but that have been established in other settings (Hardy *et al.*, 2013).

The issue of targeting could be improved with the addition of items that span a wider measurement range. This can be achieved without affecting the scale as it stands, because the item locations are calibrated relative to each other. Future scale developments can be empirically driven as the distribution of item locations highlight where 'gaps' in the measurement continuum are (notable distances in item locations could be filled with items),

and the distribution of person measurements indicate that it may be valuable to extend the measurement range at the extreme ends of the continuum. Qualitative work may be beneficial to explore the addition of items to extend the LANSS's measurement range

There is need for outcome measures that reflect changes in neuropathic pain in response to treatment. However, the LANSS did not meet all requirements for reliable and valid measurement of PU-related neuropathic pain. Reliability in this context does not support use of the LANSS as an outcome measure at the individual level or as a measurement scale across PU severity groups. If the aim is to use a version of the LANSS as an outcome measure then expanding the number of items to aid differentiation between neuropathic pain levels and improving scale reliability is recommended.

Authors' contributions

CR contributed to study concept and design, data analysis and interpretation, and preparation of manuscript. JN, JB and MB contributed to study design, interpretation of data and preparation of manuscript. MH contributed to data analysis and interpretation of data, and preparation of manuscript. All authors discussed the results and commented on the manuscript.

Ethical considerations

This is a secondary analysis therefore ethical approval was not required nor obtained. Ethical approval was obtained for collecting data for the original studies.

References

Andrich, D. (1985) An elaboration of Guttman scaling with Rasch models for measurement. In Brandon-Tuma, N. (ed) *Sociological Methodology*. Jossey-Bass, San Fransisco, pp. 33-80.

Andrich, D. (1988) *Rasch models for measurement*. Sage Publications, Beverly Hills.

Andrich, D. (1989) *Distinctions between assumptions and requirements in measurements in the social sciences*. Elsevier Science Publications BV, North Holland.

Andrich, D., Sheridan, B. & Luo, G. (2010) RUMM 2030. RUMM Laboratory PTY LTD, Perth, WA.

Bennett, M. (2001) The LANSS Pain Scale: the Leeds assessment of neuropathic symptoms and signs. *Pain*, **92**, 147-157.

Briggs, M., Bennett, M.I., Closs, S.J. & Cocks, K. (2007) Painful leg ulceration: a prospective, longitudinal cohort study. *Wound Repair & Regeneration.*, **15**, 186-191.

Briggs, M., Collinson, M., Wilson, L., Rivers, C., McGinnis, E., Dealey, C., Brown, J., Coleman, S., Stubbs, N., Stevenson, R., Nelson, E.A. & Nixon, J. (2013) The prevalence of pain at pressure areas and pressure ulcers in hospitalised patients. *BMC Nurs*, **12**, 1472-6955.

Burnand, B., Kernan, W.N. & Feinstein, A.R. (1990) Indexes and boundaries for "quantitative significance" in statistical decisions. *Journal of Clinical Epidemiology*, **143**, 1278-1284.

Campbell, D.T. & Fiske, D.W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, **56**, 81-105.

Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, 37-46.

Coleman, S., Gorecki, C., Nelson, E.A., Closs, J., Defloor, T., Halfens, R., Farrin, A., Brown, J., Schoonhoven, L. & Nixon, J. (2013) Patient risk factors for pressure ulcer development: Systematic review. *International Journal of Nursing Studies*, **50** 974-1003.

Cronbach, L. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297-334.

Dworkin, R.H., Turk, D.C., Wyrwich, K.W., Beaton, D., Cleeland, C.S., Farrar, J.T., Haythornthwaite, J.A., Jensen, M.P. & Kerns, R.D. (2008) Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *The Journal of Pain*, **9**, 105-121.

FDA., U.D.o.H.H.S. (2009) Patient reported outcome measures: Use in medical product development to support labelling claims. US Department of Health & Human Support Food & Drug Administration, MD.

Francis, G. (2001) *Introduction to SPSS for Windows*. Pearson Education Australia, Sydney.

Gorecki, C., Brown, J.M., Nelson, E.A., Briggs, M., Schoonhoven, L., Dealey, C., Defloor, T., Nixon, J. & and on behalf of the European Quality of Life Pressure Ulcer, P. (2009) Impact of pressure ulcers on quality of life in older patients: A systematic review. *Journal of the American Geriatrics Society*, **57**, 1175-1183.

Gorecki, C., Closs, J.S., Nixon, J. & Briggs, M. (2011) Patient-reported pressure ulcer pain: A mixed methods systematic review. *Journal of Pain & Symptom Management*, **42**, 443-459.

Gorecki, C., Lamping, D.L., Brown, J.M., Madill, A., Firth, J. & Nixon, J. (2010) Development of a conceptual framework of health-related quality of life in pressure ulcers: a patient-focused approach. *International Journal of Nursing Studies*, **47**, 1525-1534.

Gorecki, C., Nixon, J., Madill, A. & Firth, J.J.M. (2012) What influences the impact of pressure ulcers on health-related qualityof life? A qualitative patient-focused exploration of contributory factors. *Journal of Tissue Viability*, **21**, 3-12.

Haanpaa, M., Attal, N., Backonja, M., Baron, R., Bennett, M., Bouhassira, D., Cruccu, G., Hansson, P., Haythornthwaite, J.A., Iannetti, G.D., Jensen, T.S., Kauppila, T., Nurmikko, T.J., Rice, A.S., Rowbotham, M., Serra, J., Sommer, C., Smith, B.H. & Treede, R.D. (2011) NeuPSIG guidelines on neuropathic pain assessment. *Pain*, **152**, 14-27.

Hardy, J., Quinn, S., Fazekas, B., Agar, M. & Currow, D. (2013) Can the LANSS scale be used to classify pain in chronic cancer pain trials? *Support Care Cancer*, **21**, 3387-3391.

Hobart, J., Cano, S., Zajicek, J. & Thompson, A. (2007) Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurology*, **6**, 1094-1095.

Kerlinger, F.N. (1973) *Foundations of behavioural research*. Holt, Rinehart and Winston, New York.

Krasner, D. (1995) The chronic wound pain experience: a conceptual model. *Ostomy Wound Management*, **41**, 20-25.

Mathieson, S., Maher, C.G., Terwee, C.B., Folly de Campos, T. & Lin, C.W. (2015) Neuropathic pain screening questionnaires have limited measurement properties. A systematic review. *J Clin Epidemiol*, **23**, 00155-00159.

McGinnis, E., Briggs, M., Collinson, M., Wilson, L., Dealey, C., Brown, J., Coleman, S., Stubbs, N., Stevenson, R., Nelson, E.A. & Nixon, J. (2014) Pressure ulcer related pain in community populations: a prevalence survey. *BMC Nurs*, **13**, 1472-6955.

Muthén, L. & Muthén, B. (2011) *Mplus User's Guide, 6th ed.* Muthén & Muthén, Los Angeles, CA.

National Pressure Ulcer Advisory, P. & the European Pressure Ulcer Advisory, P. (2009) Prevention and Treatment of Pressure Ulcers: Clinical Practice Guideline. NPUAP, Washington CD, pp. 16-20.

Nixon, J., Nelson, E., Rutherford, C., Coleman, S., Muir, D., Keen, J., McCabe, J., Dealey, C., Briggs, M., Brown, S., Collinson, M., Hulme, C., Meads, D., McGinnis, E., Patterson, M., Czoski-Murray, C., Pinkney, L., Smith, I., Stevenson, R., Stubbs, N., Wilson, L. & Brown, J. (2015) *Pressure UlceR Programme Of ReSEarch: using mixed methods (systematic reviews, prospective cohort, case-study, consensus and psychometrics) to identify patient and organisational risk and develop a risk assessment tool and patient reported outcome Quality of Life and Health Utility measures.* National Institute of Health Research UK, London, UK.

Nixon, J., Nelson, E.A., Cranny, G., Iglesias, C.P., Hawkins, K., Cullum, N.A., Phillips, A., Spilsbury, K., Torgerson, D.J. & Mason, S. (2006) Pressure relieving support surfaces: A randomised evaluation. *Health Technology Assessment.*, **10**.

Nunnally, J.C. & Bernstein, I.H. (1994) *Psychometric Theory*. McGraw-Hill, New York.

Pallant, J. & Tennant, A. (2007) An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*, **46**, 1-18.

Panel, E.P.U.A. (1998) Pressure Ulcer Prevention Guidelines. European Pressure Ulcer Advisory Panel: URL: http://www.epuap.org/glprevention.html:, Oxford.

Rasch, G. (1960) *Probabilistic models for some intelligence and attainment tests*. University of Chicago, Chicago.

Rastinehad, D. (2006) Pressure ulcer pain. *Journal of WOCN.*, **33**, 252-257.

Reddy, M., Kohr, R., Queen, D., Keast, D. & Sibbald, R.G. (2003) Practical treatment of wound pain and trauma: a patient-centered approach. An overview. *Ostomy Wound Management.*, **49**, 2-15.

Royal College of Physicians, British Geriatrics Society, British Pain Society. (2007) The assessment of pain in older people: national guidelines. Report 8, Consice guidance to good practice series. RCP, London.

Scientific Advisory Committee of the Medical Outcomes Trust (2002) Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research*, **11**, 193-205.

Searle, R.D., Bennett, M.I. & Tennant, A. (2011) Can neuropathic screening tools be used as outcome measures? *Pain Medicine*, **12**, 276-281.

Smith, R.M. (2000) Fit anlysis in latent trait measurement models. *Journal of Applied Measurement*, **2**, 199-218.

Szor, J.K. & Bourguignon, C. (1999) Description of pressure ulcer pain at rest and at dressing change. *Journal of Wound, Ostomy, & Continence Nursing.*, **26**, 115-120.

Tennant, A. & Conaghan, P.G. (2007) The measurement model model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism*, **57**, 1358-1362.

Teresi, J.A., Ramirez, M., Lai, J.S. & Silver, S. (2008) Occurences and sources of Differential Item Fuctioning (DIF) in patient-reported outcome measures: Description of DIF methods, and review of measures of depression, quality of life and general health. *Psychol Sci Q*, **50**, 538.

Wright, B.D. & Masters, G. (1982) *Rating scale analysis: Rasch measurement*. MESA, Chicago.