The clinical utility of pain classification in non-specific arm pain

Authors:

Niamh A Moloney ¹ ² PhD, M.Man.Th, B.Physio.

Current appointment: Lecturer, The University of Sydney.

Toby M Hall ³ ⁴ PhD, MSc, Post Grad Dip Manip Ther.

Current appointment: Adjunct Associate Professor, School of Physiotherapy and Curtin Health Innovation Research, Curtin University.

Andrew M Leaver ² PhD, GradDipAppSc(ManipPhty), BAppSc(Phty),GradCertEdSt

Current appointment: Lecturer, The University of Sydney

Catherine M Doody ¹ PhD, MSc, BSc.

Current appointment: Lecturer, University College Dublin.

Affiliations:

¹UCD School of Public Health, Physiotherapy and Population Science, University College Dublin, Belfield, Dublin 4, Ireland

² Discipline of Physiotherapy, Faculty of Health Sciences, The University of Sydney, PO Box 170, Lidcombe, NSW 1825, Australia.

³School of Physiotherapy, Curtin University of Technology, G.P.O. Box U1987, Perth, WA 6845, Australia and ⁴Manual Concepts, P.O. Box 1236, Booragoon, WA 6954, Australia.

Corresponding author: Niamh Moloney, Discipline of Physiotherapy, Faculty of Health Sciences, The University of Sydney, PO Box 170, Lidcombe, NSW 1825, Australia.

Email: niamh.moloney@sydney.edu.au

Tel: +61 2 9351 9266 Fax: +61 293519278
Abstract

Pain classification according to pathophysiological mechanisms has received considerable attention over recent years with potential use for clinical decision making. A number of algorithms for pain classification have been proposed. Non-specific arm pain (NSAP) is a poorly defined condition, which could benefit from classification according to pain mechanisms to improve treatment selection. This study used three previously published classification algorithms (heretofore called NeuPSIG, Smart, Schafer) to investigate the frequency of different pain classifications in NSAP and the clinical utility of these systems in assessing NSAP.

Forty people with NSAP underwent a clinical examination and quantitative sensory testing. Findings were used to classify participants according to three algorithms. Frequency of pain classification including number unclassified was analysed using descriptive statistics. Inter-observer agreement was analysed using and kappa coefficients.

NSAP was classified as ‘unlikely neuropathic pain’ using NeuPSIG criteria, ‘peripheral neuropathic pain’ using Smart’s criteria and ‘peripheral nerve sensitisation’ using the Schafer algorithm. Two of the three algorithms allowed classification of all but one participant; up to 45% of participants (n=18) were categorised as mixed by the ‘Smart algorithm’. Inter-observer agreement was good for the ‘Schafer algorithm’ (κ=0.78) and moderate for the ‘Smart algorithm’ (κ=0.40). A kappa value was unattainable for the NeuPSIG algorithm but agreement was high.

Pain classification was achievable with high inter-observer agreement for two of the three algorithms assessed. The third classification may be useful but requires further direction regarding the use of clinical criteria included. The impact on outcomes of adding a pain classification to clinical assessment needs to be evaluated in this cohort.
INTRODUCTION

Classification of pain based on pathophysiological mechanisms has received considerable attention [1-9] and is increasingly used in diagnosis and management of musculoskeletal conditions. In musculoskeletal conditions, there is often a poor relationship between pathology, pain and disability [10], as well as high prevalence of undiagnosed disorders [11, 12], suggesting the potential clinical value of mechanisms-based pain classification [1].

In the absence of a gold standard classification, a number of mechanisms based algorithms have been proposed [8, 13, 14]. The classification algorithm endorsed by the Neuropathic Pain Special Interest Group (NeuPSiG) of the International Association for the Study of Pain [13, 15] classifies patients into one of four groups; Definite-, Probable-, Possible-, and Unconfirmed-Neuropathic pain, based on the number of corroborative signs (Figure 1). The ‘NeuPSiG algorithm’ is a consensus document of NeuPSiG and its reliability or validity have not been formally tested.

Classification criteria outlined by [8], for chronic low back pain (+/- leg pain), classifies patients into three groups (nociceptive-, peripheral neuropathic- and central sensitisation pain) (Figure 2). There is preliminary evidence for the validity of the ‘Smart classification’ when used in a low back pain population [8].

The algorithm of [14] for classification of low-back related leg pain classifies patients into four groups (neuropathic sensitisation, denervation, peripheral nerve sensitisation and musculoskeletal pain) (Figure 3). The ‘Schafer algorithm’ has good inter-rater reliability [16] and good discriminative validity for the group ‘peripheral nerve sensitization’ [6].

All three algorithms aim to distinguish patients who have pain with demonstrated painful neuropathy from those with non-neuropathic pain. The Smart classification and Schafer algorithm aim to further distinguish patients who have widespread sensitisation from those
with nociceptive pain or peripheral nerve mechanosensitivity. The NeuPSIG algorithm does not make this distinction.

Clinically, differentiation of pain mechanisms may influence decision making about potential interventions, for example, musculoskeletal pain and neuropathic pain would warrant different treatment approaches with manual therapy and therapeutic exercise likely more useful in musculoskeletal pain than in people with neuropathic pain. The clinical utility of classification algorithms is dependent on the ability of the algorithm to influence clinical decision making. To do this an algorithm must have the capacity to reliably and correctly assign patients without there being too many ‘unclassifiable’ cases [17].

The use of a mechanisms-based pain classification system holds potential for non-specific arm pain (NSAP). NSAP is a common upper limb disorder [18], that is frequently associated with poor outcome [19]. Whilst NSAP is a diagnosis of exclusion [20], the high prevalence of weakness and paraesthesia in NSAP [21] suggests a neural tissue disorder might underpin some presentations of NSAP. This hypothesis is supported by findings of altered vibration thresholds [22, 23] and neural tissue sensitivity [22, 24]. However, data also exist suggesting that a muscle tissue disorder [25, 26] might be a prevalent pathology in NSAP. Recently, we presented data that widespread sensory hypersensitivity along with localised neural tissue sensitivity were characteristic features in this condition [27]. Given these findings, it is not clear whether identification of a single specific pain classification can be achieved in NSAP. The purpose of this study was to (1) investigate the frequency of different pain classification in NSAP and (2) to investigate the clinical utility of three pain mechanism classification algorithms for NSAP. Specifically, we aimed to examine the completeness of classification and the inter-rater agreement for each classification algorithm.
METHODS

Study Design

A cross-sectional observational study was conducted. Participants with NSAP underwent standardized assessment by a physiotherapist (NM). Assessment findings were then used by two physiotherapists (NM and TH) to classify participants according to three pain classification algorithms. The clinical utility of each algorithm was evaluated according to its capacity to completely classify all participants and by assessment of inter-rater agreement.

Setting

This study was set in a university laboratory. Participants were recruited from metropolitan hospitals, medical and physiotherapy practices and the general population. The study was approved by the Human Research Ethics Committee for Life Sciences, University College Dublin, and participating hospitals. All participants were unpaid volunteers who provided written informed consent.

Participants

Forty volunteers with arm pain, aged between 18-65 years were recruited. Participants were included if they had pain ≥3/10 of >3 months duration, who used desktop equipment for more than 40% of their working week [22] and had been employed using desk-top equipment for at least two years [28]. Participants were excluded if they had a diagnosed specific musculoskeletal condition or any of the following: generalized neurological or musculoskeletal disorders, low back pain or migraine over the previous 6 months, significant upper quadrant trauma, diabetes, endocrine disorders, epilepsy or significant mental health disorders.

Investigators
Clinical assessment of the participants and collection of baseline data and allocation of participants to pain classification groups was conducted by the lead author (NM). A second investigator (TH) independently verified assessment findings and independently allocated participants to pain classification groups. Data entry and analysis was conducted by a third investigator (AL). All investigators were physiotherapists with post-graduate qualification in musculoskeletal physiotherapy, with more than 11 years clinical experience and training in quantitative sensory testing (QST).

**Procedure**

Participants attended a one-off clinical assessment that included quantitative sensory testing (Figure 4). The diagnosis of ‘NSAP’ was confirmed by excluding alternative diagnoses including cervical referred pain and upper limb neuropathies (see Appendix 1). Neurological function and nervous system sensitisation (localised and widespread) was also assessed. Participants completed questionnaires to assess pain and disability. Assessment findings were evaluated by two independent investigators. Participants were classified according to the three algorithms by two independent investigators (Figures 1-3).

**Variables**

Variables used to classify participants for each algorithm were extracted from the patient history, written questionnaires, physical examination and QST findings. The interpretation of each variable within each classification algorithm is detailed in Table 1.

**History**

Variables extracted from the patient history included the presenting complaint, pain features, pattern of aggravating and easing factors, and signs and symptoms of nerve injury or compromise.

**Questionnaires**
Participants completed the Leeds Assessment for Neuropathic Symptoms and Signs (LANSS) questionnaire, Tampa Scale for Kinesiophobia (TSK) [29] [30], and Short-Form McGill Pain Questionnaire (SFMPQ) [31]. A LANSS of ≥12 was considered an indicator of possible neuropathic pain [32]. A TSK score of ≥37 was considered an indicator of significant fear avoidance beliefs [30]. In addition, the disability of arm, shoulder and hand (DASH) questionnaire was administered [34]. Pain descriptors in the MPQ, in addition to descriptors nominated by participants, and results from these questionnaires were used for completion of clinical algorithms (Table 1).

**Physical examination**

Physical examination included:

- Neurological examination (reflexes, myotomes, dermatomes) [35]
- Neural tissue sensitivity tests including the upper limb neurodynamic test 1 and palpation of the median, ulnar and radial nerves. The neurodynamic test was considered positive if arm symptoms (at least in part) were reproduced along with positive structural differentiation tests [36, 37].
- Assessment of pin-prick hyperalgesia and brush-stroke allodynia over symptomatic and remote asymptomatic regions
- Assessment of general tenderness to light touch.

**Quantitative sensory testing**

A previously published QST protocol [38] was implemented to test the following: cold, warm and vibration detection thresholds; cold, heat and pressure pain thresholds. All measures were recorded bilaterally over three upper limb sites. The tibialis anterior region was used as a remote reference point for thermal testing and pressure pain thresholds.
Sensory phenotypes were determined for each participant by generating z-score profiles from QST data [39]. QST data were log transformed before calculation of z-scores. Z-transformation for each variable was generated using the formula: 
\[ Z\text{-score} = \frac{X \text{ single participant} - \text{Mean controls}}{\text{SD controls}} \] 
Z-scores of > 2 z-scores from the mean were considered indicative of sensory gain to the tested stimuli, while z-scores of < 2 from the mean were considered indicative of sensory loss [40]. If a sensory abnormality was detected, it was sub-grouped as (a) localised, (b) localised within a neuroanatomical distribution or (c) widespread.

**Analyses**

Sample size was calculated using a method reported by Sim and Wright [41]. We considered that \( \kappa = 0.40 \) would be the least clinically acceptable agreement measurement. Using the null hypothesis of \( \kappa = 0.40 \) and assuming a 5% level of significance and a power of 0.8, a sample size of 39 was required.

Participant characteristics and classification frequencies were analysed using descriptive statistics. Clinical utility of each algorithm was assessed according to the completeness of the classification according to each algorithm i.e. the proportion of participants classified versus unclassified. The Kappa coefficient was used to assess inter-rater agreement of classification according to each algorithm and was interpreted according to published guidelines [42].

**RESULTS**

Participant characteristics are presented in Table 2.

**Frequency of pain classification**
The frequency of assignment to each pain group and inter-rater agreement for each algorithm are presented in Tables 3, 4 & 5. The dominant pain classification was ‘unlikely neuropathic pain’ using the NeuPSIG algorithm, ‘peripheral nerve sensitisation’ using the Schafer algorithm, and ‘peripheral neuropathic pain’ using the Smart classification.

**Clinical Utility**

For two classification systems (NeuPSIG and Schafer algorithms) all but one participant could be classified by both observers. Using the Smart classification, 20% and 45% of participants were deemed to have mixed pain types by rater 1 and 2 respectively. A statistical coefficient of agreement was unattainable for the NeuPSIG algorithm; however, visual inspection of the data revealed excellent agreement between the observers (92% agreement). The inter-rater agreement was moderate at 40% for the Smart classification, which was set a priori as the minimal clinically-relevant acceptable agreement. Classification using the Schafer algorithm demonstrated good agreement at 78.3%.

**DISCUSSION**

This study demonstrates that mechanisms-based pain classification algorithms developed for back pain +/- leg pain [8], low back related leg pain [14] and neuropathic pain [13] have potential clinical utility in the assessment of NSAP. Using the NeuPSIG algorithm, the most frequently allocated classification was ‘unlikely neuropathic pain’, while using the Smart and Schafer algorithms this was ‘peripheral neuropathic pain’ and ‘peripheral nerve sensitisation’ respectively. Between five and 11 participants were considered to have musculoskeletal/nociceptive pain using these latter algorithms, while few participants were deemed to have central or neuropathic sensitisation. The best agreement between raters was achieved using the Schafer algorithm.
Pain Classifications in NSAP

The results of this study indicate that NSAP may be considered a ‘peripheral neuropathic pain’ condition when using the \textit{Smart algorithm} but was considered ‘unlikely neuropathic pain’ using the \textit{NeuPSIG algorithm} \cite{13}. This highlights discord in the clinical criteria used to classify neuropathic pain. The \textit{NeuPSIG algorithm} requires that evidence of a lesion or disease of the nervous system is obtained through the history, pain distribution and further testing and provides a stepwise approach to the classification of neuropathic pain, from probable to definite. There are two key differences to the way we used the criteria outlined by \cite{8} Firstly, because a hierarchy of the clinical criteria has not been indicated, all of the criteria were treated with equal weight when classifying patients. Secondly, signs and symptoms of neural tissue mechanosensitivity, even in the absence of neurological deficit, were assigned to peripheral neuropathic pain. This may explain the discrepancy between the results of the \textit{Smart classification} and the \textit{NeuPSIG algorithm}. Furthermore, the \textit{Schafer algorithm} allows for separation of those with neurological deficit and those with neural tissue mechanosensitivity under the headings ‘denervation’ and ‘peripheral nerve sensitisation’; this resulted in 65% of participants falling into the latter category. The presence of peripheral nerve sensitisation in NSAP is consistent with our previous report \cite{27} and other studies \cite{22, 24, 43, 44}. These findings hold relevance for clinical decision making, discussed later in this paper.

With respect to other classifications of pain, up to 28% of participants were classified as having nociceptive/musculoskeletal pain, but few were identified with widespread sensitisation (central or neuropathic sensitisation) (≤10%). The infrequent identification of
widespread sensitisation in NSAP using these algorithms is surprising and contrasts with data from QST in this group that highlighted widespread sensitisation as a key characteristic [27]. To our knowledge, correlations between the clinical appraisal of the presence of widespread sensitisation and results from QST have not been yet been reported; however, quantitative measures of pain sensitivity have previously been shown to correlate poorly with self-reported pain and disability [45]. Hübscher et al. (2013) proposed one reason for this poor correlation may be the fact that none of the included studies in that review specifically recruited people deemed clinically to present with widespread sensitisation and as such, correlations between QST and clinical presentation of pain may have been diluted. Interestingly, while widespread hypersensitivity is acknowledged as an important mechanism underpinning some chronic pain states e.g. whiplash and fibromyalgia [46-48], both its status as a distinct pain classification and criteria for its classification are not yet fully established.

**Appraisal of each algorithm**

In consideration of the *NeuPSIG algorithm*, it is unsurprising that the majority of the participants were deemed unlikely to have neuropathic pain, with only three participants allocated to the ‘possible’ category by one observer. No participant was classified as probable or definite in this group. Our participants were screened for the presence of specific diagnoses prior to inclusion in this study, and neurological injury would have meant exclusion. The *NeuPSIG algorithm* was therefore accurate in this respect; however, its use may have been redundant in this group. The purpose of the *NeuPSIG algorithm* is to provide a framework for classification of neuropathic pain, which relies on the diagnosis of neuropathy [13]. There has been some debate about this approach and a view that neuropathic pain may be a broader entity [49]. Nonetheless, the benefit of this algorithm
for guiding clinical decision making in primary care lies in its ability to identify patients who may warrant further investigation or more specific management for neuropathy [50]. The limitation of this algorithm is that in the absence of neuropathic pain, other guidelines are required for further classification and direction for management.

The *Smart classification* resulted in the highest number of participants classified with mixed pain presentations and the lowest consistency between raters. This was largely influenced by the number of participants classified as mixed by one rater. Improvement in inter-rater agreement might be achieved with more specific instruction about interpreting assessment findings in light of the classification criteria. Inter-rater agreement in the current study are considerably lower than previously reported [51]; however, that study used different methodology. [51] assessed agreement of a larger suite of clinical criteria items that were used in an earlier phase of their study. Further, [51] tested decisions made immediately following patient assessment rather than post-hoc as they were in this study. The clinical criteria for selection in each category incorporates a combination of subjective and objective features, the absence of a hierarchical model meant that many participants were classified as ‘mixed’. Whilst this makes clinical decision making difficult, it may reflect the real-life clinical situation, where patients are likely to present with a mix of pain types and where pain states are a continuum rather than discrete entities. However, developing this classification system into a stepwise clinical algorithm might be beneficial.

The dominant classification of the *Smart classification* was peripheral neuropathic pain despite this cohort being screened and cleared of neuropathy. The inclusion of both neurological deficit and signs and symptoms of neural tissue mechanosensitivity within the same category in this algorithm diminishes the capacity of this algorithm to guide clinical
decision making as it is necessary to segregate those with neurological deficit from those without for the purpose of guiding further investigation and possible treatment.

The Schafer algorithm was considered the most clinically useful algorithm for NSAP, both due to the high number of people classified and the good inter-rater agreement. This is likely due to two factors; firstly the hierarchical nature of the algorithm forces a category selection that is based on limited but key information. Secondly, the presence of a separate category for neural tissue mechanosensitivity i.e. peripheral nerve sensitisation was useful in NSAP. The identification of peripheral nerve sensitisation as a distinct category may be controversial from a mechanisms-based pain classification perspective. As is apparent from the Smart classification, it has become common to describe neural tissue mechanosensitivity as a neuropathic pain condition. This description reflects the inconsistency across some disciplines in the semantic use of the term 'neuropathic'. Whilst neural tissue provocation tests have demonstrated validity [52, 53], whether peripheral nerve sensitisation is a discrete neuropathic pain condition is debatable. Certainly, neural tissue sensitisation frequently occurs with painful neuropathies [54]; however, in the absence of a distinct neuropathy, a number of mechanisms to explain sensitisation of neural tissue include neuritis [55, 56], sensitisation of nervi nervorum [57] and minor neuropathy [22]. Regardless of the semantic arguments, arm pain with neuropathy and arm pain without neuropathy but with neural tissue sensitivity usually require different treatment approaches so distinguishing between these conditions is important.

Clinical value of pain classification algorithms
It is important to consider how pain classification might influence clinical decision-making.

From the algorithms outlined, four sub-groups are evident: (1) Neuropathic pain; (2) Musculoskeletal pain; (3) Widespread sensitisation and (4) Peripheral nerve sensitisation (i.e. localised neural tissue mechanosensitivity). An important caveat in discussing sub-grouping is whether it results in improved outcomes. Whilst sub-grouping is attractive for the stratification of healthcare and appears logical, as will be outlined, the impact of sub-grouping on treatment outcomes is, as yet, inconsistent [58-60].

**Neuropathic pain**: A key task when triaging a patient is the identification of specific pathologies including neuropathy ± neuropathic pain. Such patients often require specific investigations and sometimes surgical intervention [50]. Conservative management in this group including physical therapies [61], and anti-convulsant medications [62] might also be considered.

**Musculoskeletal pain**: The identification of musculoskeletal (nociceptive) pain, in the absence of neuropathy or central sensitisation, suggests that clinicians should focus on conservative interventions, such as, education, simple analgesic/anti-inflammatory medication and physical treatments such as exercise and manual therapy [63-66].

**Widespread sensitisation**: Widespread sensitisation involves peripheral and central nervous system sensitisation which poses a particular challenge in clinical practice. There is growing research evidence that many people with chronic musculoskeletal conditions display signs of widespread sensitisation [47, 48, 67-70]. Evidence based management approaches remain elusive in this group [47, 59]; however, comprehensive, multi-disciplinary approaches incorporating pain science education, are likely warranted [70, 71].

**Peripheral nerve sensitisation**: The mechanisms underpinning peripheral nerve sensitisation are still open to debate. Nonetheless, identification of peripheral nerve sensitisation as a
distinct entity is potentially beneficial. Emerging data demonstrate positive responses to therapies specifically targeting neural tissue sensitivity i.e. non-provocative neural mobilisation [6, 72-74] in those classified with peripheral nerve sensitisation.

Study Limitations

Further studies on a larger sample size and by researchers who have not been involved in the development of the classification systems is warranted.

CONCLUSION

This study explored the clinical utility for NSAP of three pain algorithms that classify patients according to pain mechanisms. The results indicate that the NeuPSIG algorithm is effective in identifying those with neuropathic pain resulting from an identifiable neurological lesion. The Schafer algorithm was demonstrated to have the best clinical utility in terms of number of participants classified and inter-rater agreement. Finally the Smart classification resulted in the most participants classified with ‘mixed pain’ in this cohort, which also largely accounted for the lower inter-rater rates using this method. The results from this study support previous reports of peripheral nerve sensitisation as a key characteristic of NSAP.
ACKNOWLEDGEMENTS

The authors would like to acknowledge the late Max Zusman for his contribution to the body of research from which this paper is derived. This research was funded by the Irish Research Council for Science, Engineering and Technology: Embase PhD Scholarship; Chartered Physiotherapists in Musculoskeletal Therapy, Ireland: Research Bursary; Chartered Physiotherapists in Occupational Health and Ergonomics, Ireland: Research Bursary.

Conflicts of Interest: None declared
REFERENCES


