

University of Groningen

## Investigating the validity of the DN4 in a consecutive population of patients with chronic pain

Timmerman, Hans; Steegers, Monique A. H.; Huygen, Frank J. P. M.; Goeman, Jelle J.; van Dasselaar, Nick T.; Schenkels, Marcel J.; Wilder-Smith, Oliver H. G.; Wolff, André; Vissers, Kris C. P.

*Published in:*  
 Plos one

*DOI:*  
[10.1371/journal.pone.0187961](https://doi.org/10.1371/journal.pone.0187961)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
 Publisher's PDF, also known as Version of record

*Publication date:*  
 2017

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Timmerman, H., Steegers, M. A. H., Huygen, F. J. P. M., Goeman, J. J., van Dasselaar, N. T., Schenkels, M. J., ... Vissers, K. C. P. (2017). Investigating the validity of the DN4 in a consecutive population of patients with chronic pain. *Plos one*, 12(11), [0187961]. DOI: 10.1371/journal.pone.0187961

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

RESEARCH ARTICLE

# Investigating the validity of the DN4 in a consecutive population of patients with chronic pain

Hans Timmerman<sup>1\*</sup>, Monique A. H. Steegers<sup>1</sup>, Frank J. P. M. Huygen<sup>2</sup>, Jelle J. Goeman<sup>1,3</sup>, Nick T. van Dasselaar<sup>4</sup>, Marcel J. Schenkels<sup>5</sup>, Oliver H. G. Wilder-Smith<sup>1,6</sup>, André P. Wolff<sup>7</sup>, Kris C. P. Vissers<sup>1</sup>

**1** Radboud university medical center, Department of Anesthesiology, Pain and Palliative Medicine, Nijmegen, the Netherlands, **2** ErasmusMC, Department of Anesthesiology, University Center of Pain Medicine, Rotterdam, the Netherlands, **3** Leiden University Medical Center, Department of Medical Statistics and Bioinformatics, Leiden, the Netherlands, **4** Reinier de Graaf Gasthuis, Department of Anesthesiology, Pain Medicine and Palliative Care, Delft, the Netherlands, **5** Bernhoven Ziekenhuis, department of Anesthesiology, Uden, the Netherlands, **6** Aalborg University, Center for Sensory-Motor Interaction, Aalborg, Denmark, **7** University of Groningen, University Medical Center Groningen, Department of Anesthesiology, Pain Center, Groningen, the Netherlands

\* [hans.timmerman@radboudumc.nl](mailto:hans.timmerman@radboudumc.nl)



**OPEN ACCESS**

**Citation:** Timmerman H, Steegers MAH, Huygen FJPM, Goeman JJ, van Dasselaar NT, Schenkels MJ, et al. (2017) Investigating the validity of the DN4 in a consecutive population of patients with chronic pain. PLoS ONE 12(11): e0187961. <https://doi.org/10.1371/journal.pone.0187961>

**Editor:** Claudia Sommer, University of Würzburg, GERMANY

**Received:** May 31, 2017

**Accepted:** October 30, 2017

**Published:** November 30, 2017

**Copyright:** © 2017 Timmerman et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This study was performed within DALI for PAIN, a national program that focuses on neuropathic pain care optimisation to APW. DALI for PAIN is an initiative of Pfizer. This project is supported by an unrestricted grant from Pfizer. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Neuropathic pain is clinically described as pain caused by a lesion or disease of the somato-sensory nervous system. The aim of this study was to assess the validity of the Dutch version of the DN4, in a cross-sectional multicentre design, as a screening tool for detecting a neuropathic pain component in a large consecutive, not pre-stratified on basis of the target outcome, population of patients with chronic pain. Patients' pain was classified by two independent (pain-)physicians as the gold standard. The analysis was initially performed on the outcomes of those patients (n = 228 out of 291) in whom both physicians agreed in their pain classification. Compared to the gold standard the DN4 had a sensitivity of 75% and specificity of 76%. The DN4-symptoms (seven interview items) solely resulted in a sensitivity of 70% and a specificity of 67%. For the DN4-signs (three examination items) it was respectively 75% and 75%. In conclusion, because it seems that the DN4 helps to identify a neuropathic pain component in a consecutive population of patients with chronic pain in a moderate way, a comprehensive (physical-) examination by the physician is still obligate.

## Introduction

Neuropathic pain is described as pain caused by a lesion or disease of the somatosensory nervous system and requires a demonstrable lesion or a disease that satisfies established neurological diagnostic criteria[1]. Moreover, neuropathic pain is a clinical description and not a diagnosis[1]. In daily clinical practice it is to our opinion more appropriate to speak of a present neuropathic pain component (present NePC) or absent neuropathic pain component (absent NePC)[2, 3]. This because the pain experienced by the patient in the clinical context

**Competing interests:** This project is supported by an unrestricted grant from Pfizer. This unrestricted grant does not alter our adherence to PLOS ONE policies on sharing data and materials. The authors declare that they have no other competing interests.

may be caused by both neuropathic- as well as nociceptive mechanisms (also known as ‘mixed pain’)[2, 4–6]. The main features of neuropathic pain components are, in clinical practice, the painful signs and symptoms in a region of altered sensations (numbness or increased sensitivity)[6]. The assessment of neuropathic pain is nowadays primarily based on history and physical examination including (bedside-)sensory testing[7–9] to assess patients’ pain.

Since current pharmacological treatment of patients with and without a NePC differs strongly, a correct pain classification is imperative[7, 10]. The availability of a simple and validated screening tool to determine the presence of NePC for clinical triage and epidemiological purposes can assist in detection of NePC[7, 8, 11–16]. This is especially true when this tool can be used by non pain specialists.

The original French validation study of the ‘Douleur Neuropathique en 4 Questions’ (DN4) [17] was performed in patients with neuropathic pain resulting from, for example, nerve trauma or post herpetic neuralgia. Patients with non-neuropathic pain were, amongst other diagnoses, suffering from osteoarthritis. All included patients had pain of at least a moderate severity ( $\geq 40$  on a 100mm visual analogue scale). Pain classification in this study was based on medical history, physical examination, electromyography and/or imaging by two independently working physicians. DN4 application resulted in a sensitivity of 83% and a specificity of 90%[17]. As indicated in a systematic review by Mathieson et al[16] the classification of a NePC may differ between clinicians and may be more difficult when there are patients included with mixed pain and with all levels of pain. This reflects the patient population in a daily clinical practice, but might have an influence on the validity. Moreover, the accuracy of screening tools is dependent on the standardization of the assessment strategy[18]. Translation/ cross-cultural adaptation and/or validation of the DN4 was performed in more than 75 languages[19–31].

The neuropathic pain special interest group (NeuPSIG) grading system[32] is developed by Treede et al in 2008 and updated in 2016[33]. It is a system to help the clinician to determine the certainty of the pain classification for the existence of a NePC in an individual patient: non-neuropathic pain; possible, probable or definite neuropathic pain. The grading system is suggested to be helpful in the assessment of the pain classification in clinical practice[34–38].

The aim of this study was to assess the validity and reliability of the DN4 as a screening tool for use in daily outpatient practices to detect a NePC in a, not pre-stratified on the target outcome, consecutive patient population having chronic pain syndromes due to low back and leg pain (LBLEP), neck-shoulder-arm-pain (NSAP) or pain of suspected neuropathic origin (PSNO).

## Methods

This validity and reliability study had a cross-sectional, longitudinal, research design with a 2-weeks and 3-months follow-up period. Comparisons were made between the DN4 (as a whole and for the symptom questions and signs tests separately) and the classification of patients’ pain by two, independently working, physicians (the gold standard) as well as with the grading system.

The study was approved by the Committee on Research Involving Human Subjects region Arnhem-Nijmegen, Nijmegen, the Netherlands, (dossier number: 2008/348; NL 25343.091.08) which counts for participation of the Dutch academic pain centers (Radboud University Medical Center, Nijmegen; Utrecht University Medical Center, Utrecht; Erasmus Medical Center, Rotterdam), Dutch non-academic pain centers (Bernhoven Ziekenhuis, Oss; St. Anna Ziekenhuis, Geldrop) and a Dutch non-academic department of neurology (Rijnstate Ziekenhuis, Arnhem). Participation of Dutch non-academic pain center in Delft, the Netherlands (Reinier

de Graaf gasthuis) was approved by Medisch Ethische Toestings Commissie Zuidwest Holland (dossier number: 10–145). The study protocol was registered in the Dutch National Trial Register (NTR3030).

We used the same methodology as in the published protocol [39] and as employed in a simultaneous study regarding the validity of the PainDETECT (Timmerman et.al / Under review by BMC Neurology).

## Participants

Consecutive patients (first time visitors of the participating centers) without pre-stratification based on the target outcome[40] were included in the study between October 2009 until July 2013. Patients were asked to participate by their doctor. Each patient signed informed consent before participation in the study.

At that time, there was only a rough diagnosis: LBLP, NSAP or PSNO. Inclusion criteria: Male and female adult patients ( $\geq 18$  years of age) with chronic ( $\geq 3$  months) LBLP or NSAP radiating into respectively leg(s) or arm(s) or patients with chronic pain due to a PSNO (pain associated with a lesion or disease of the peripheral somatosensory system). Exclusion criteria: Patients diagnosed with malignancy; compression fractures; patients with diffuse pains (such as fibromyalgia or ankylosing spondylitis); severe mental illness; chronic alcoholism or substance abuse; inability to fill in the questionnaire adequately or incapable of understanding the Dutch language.

## Physicians

The physicians (pain specialists, pain specialist fellows or neurologists always operating in differently composed pairs) participating in this study were not selected on basis of age, experience as a physician or any other criteria. Classification of patients' pain was based on the NeuPSIG guidelines on neuropathic pain assessment[7] and recorded as absent NePC or present NePC. Pain classification was performed consecutively on the same patient by two physicians and categorized afterwards in three groups: absent NePC, present NePC or 'undetermined' (i.e. the pain classification of the two physicians was not the same). A full medical history and clinical examination including sensory bedside examination (touch, pinprick, pressure, cold, heat and temporal summation) was taken[7, 8, 39, 41, 42] and was considered to be the gold standard when assessed by two physicians. The NeuPSIG grading system[32, 33] was used as a secondary comparison with the outcome of the DN4 and was assessed by both the physicians separately. The outcomes "probable" and "definite" were regarded as present NePC. "unlikely" and "possible" as absent NePC[38, 43, 44]. The physicians worked independently of each other and were blinded to the pain classification of the other physician. Each physician was allowed to perform the clinical examination in the way he or she is used to do but were supported by a standardized assessment form[39]. In this form, the pain score, a body map to indicate the localization of patients pain, the sensory examination and the four questions of the grading system had to be filled in by the physician. The participating physicians were trained in a standardized way (presentation about the study and the outcome parameters and a practical training on how to use the (measurement) instruments), by the investigator (HT) or by a designated person on location before participation in the study. Practical training was focused on the classification of NePC, the assessment of the grading system, the performance of bedside examination tests and the performance and assessment of the examination items of the DN4.

In this study, 62 physicians (pain specialist, pain specialist-fellow or neurologist) participated. The physicians who were classifying patients' pain at the first session were called

'Physicians A'. The physicians who performed the classification at the second session, were called 'Physicians B'.

## Measurements

**Douleur Neuropathique en 4 questions (DN4).** The DN4 [17, 20, 25] (Pfizer bv. Capelle a/d IJssel, the Netherlands) consists of 10 items in total and is developed to screen for symptoms and signs of neuropathic pain resulting in a yes/no answer for the presence of neuropathic pain. This instrument is divided into two questions (seven answers, DN4-symptoms: score range 0–7) and two physical examination tests (three answers, DN4-signs: score range 0–3). The examination items of the DN4 regarding the signs (hypoesthesia to touch, hypoesthesia to prick and brushing) were incorporated in the sensory examination part of the standardized assessment form and were carried out according the original publication by Bouhassira et al [17]. This assessment form was filled in by both physicians separately. The seven symptom items are consisting of characteristics (Burning, painful cold, electric shocks) and symptoms (Tingling, pins and needles, numbness, and itching). The patient completed the DN4-symptoms directly after the clinical assessments by the physicians but without interference. The researcher (HT) or a nurse was available for help in person or via telephone when it was not clear for the patient how to fill in the questionnaires.

The items of the DN4 are scored based on a yes (1 point) /no (0 points) answer. This leads to a score range of 0–10 when the symptoms (range 0–7 points) as well as the signs (range 0–3 points) items are included. Values in the DN4 who were not filled in were considered as 'no' (0 points). However, in the reliability analysis these data were not incorporated.

**Patient global impression of change (PGIC).** The Patients Global Impression of Change (PGIC)[11, 45–47] was used to assess the change of pain complaints, based on the patients' own impression of change over time, during the follow-up period (7-points scale: Very much improved-very much worse). Follow-up took place two weeks and three months after the initial visit. To compare the outcome of the DN4 in the follow-up period the pain complaints as addressed by the patient had to be unchanged.

## Time-line

All baseline measurements (the assessment by the physicians, the grading system by both physicians as well as filling in the questionnaires by the patient) took place on preferably the same day. The PGIC [45–47] and the DN4-symptoms (sensory testing for the DN4-signs was not performed) were sent to the patient after two weeks and three months with instructions how to fill them in by mail. Also for the follow-up measurements help was available in person or via telephone when it was not clear how to fill in the questionnaires.

## Data

All data was collected on paper and stored by Radboudumc, Nijmegen, The Netherlands. Data management and monitoring were performed within MACRO (MACRO, version 4.1.1.3720, Infermed, London, United Kingdom). Data analysis and statistics was performed by use of Statistical Package for the Social Sciences (IBM SPSS statistics 22, SPSS Inc., Chicago, Illinois, USA).

## Statistical analysis

According to the power-calculation in the protocol 132 patients with LBLP, NSAP or PSNO were needed such that the sample size contains adequate numbers of cases and controls[39].

Qualitative variables are presented as frequencies and percentages. The quantitative variables are presented as mean and standard deviation (SD) or as median and inter quartile range (IQR).

The agreement between any of the two combinations of the two observers (pain classification by the physician and the outcome of the grading system) to establish a present NePC or absent NePC, and of the DN4 (DN4 / DN4-symptoms / DN4-signs outcome) was evaluated by use of Cohen's kappa (K), prevalence index (Pi) and percentage of pair wise agreement (PA). The categorization of the kappa values are, according to the categorization of observer agreement by Landis and Koch[48], none beyond chance ( $K \leq 0.00$ ); slight ( $K = 0.01-0.20$ ); fair ( $K = 0.21-0.40$ ); moderate ( $K = 0.41-0.60$ ); substantial ( $K = 0.61-0.80$ ) and (almost) perfect agreement ( $K = 0.81-1.00$ ). A  $K \geq 0.40$  and a  $PA \geq 70\%$  is considered indicative of interobserver reliability acceptable for use in clinical practice[48]. Moreover, also the interobserver reliability of the examination items in the DN4-signs were tested.

Based on the classifications of the two physicians, all patients were categorized as absent NePC, present NePC or 'undetermined' (i.e. the pain classification of the two physicians was not the same).

Statistical significant differences between absent NePC and present NePC were determined by use of students t-test (Interval scales), Mann-Whitney U-test groups (ordinal scales) or via Chi<sup>2</sup>-test (nominal scale). The statistical significant differences between present NePC, absent NePC and the Undetermined group was assessed by use of One-way ANOVA (with additional Tukey's studentized range post-hoc test) or Kruskal-Wallis test. Chi<sup>2</sup> test was also used to analyze the nominal outcome scale of the DN4 regarding the three groups.

A factor analysis was used to study the structure of the DN4 in such a way that variables that were thought to reflect a smaller number of underlying variables were observed. This method was performed for all three versions of the DN4 (DN4; DN4-symptoms and DN4-signs). Principal axis factoring was used as the extraction method. The varimax rotation with Kaiser normalization was used. Extraction of the factors was based on Eigenvalues being greater than 1.0. Cronbach's alpha was used to calculate the internal consistency of the factors constructed. The results are only shown for the Physicians A (the assessment of the patient by the first physician). The outcomes by the Physicians B (the assessment of the patient by the second physician) are shown in [S1 Table](#). However, the conclusions, which are drawn, are identical for physicians A and for physicians B.

A receiver operating characteristic (ROC) curve was calculated for the DN4 and the DN4 signs by both the physicians A and B and for the DN4-symptoms as filled in by the patient. The area under the curve (AUC) with 95% confidence interval was presented to indicate the discriminatory power of the DN4 to discriminate patients by present NePC or absent NePC. This dichotomy was based on the physicians' assessment outcome or based on the grading system outcome, respectively. The theoretical maximum of the AUC is 100%, indicating a perfect discrimination and 50% is equal to tossing a coin. An AUC between 0.9 and 1 is considered to be excellent, an AUC between 0.8 and 0.9 is good and between 0.7 and 0.8 is fair. An AUC between 0.6 and 0.7 is considered to be poor. Between 0.5 and 0.6 the AUC is considered to be failed[49–52]. The optimal cut-off point of the DN4 was calculated under the condition of equal-costs of misclassification using the Youden-index. Sensitivity, specificity, positive and negative predictive values and the likelihood ratio in the population in this study was calculated at this cut-off point. The outcome results were averaged between both physicians and the 95% confidence intervals were noted with respect to the lowest and highest level.

Clinimetrics of the DN4 based on both the physicians assessment and/or both the grading system outcome were assessed for the DN4, the DN4-symptoms and for the DN4-signs items.

A screening tool for the presence of a NePC is considered valid if it has a high sensitivity, specificity, high positive predictive value and a high positive likelihood ratio [53].

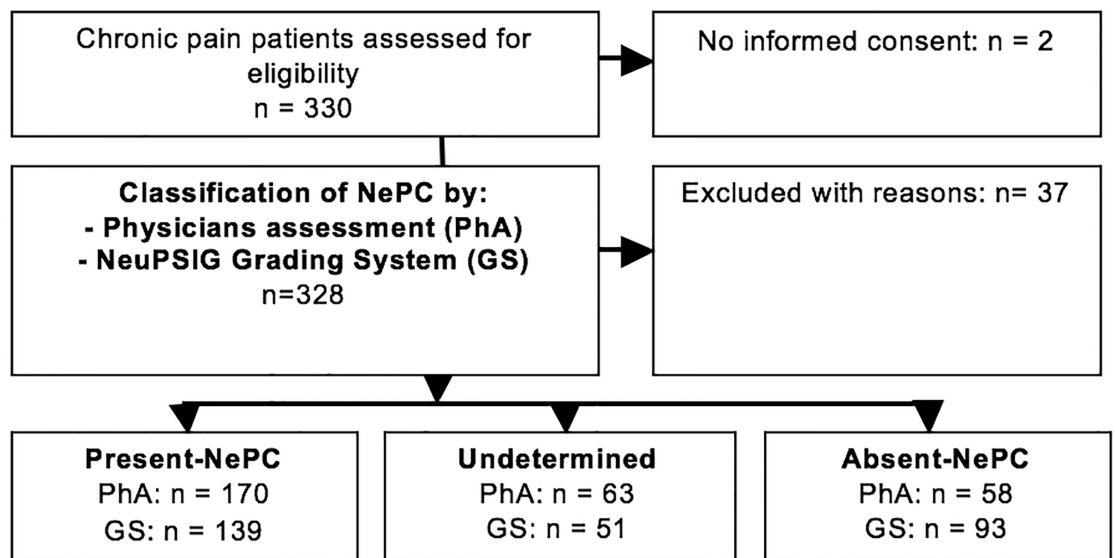
Intraclass correlation (ICC) was used to assess reproducibility ('test-retest reliability') of the DN4-symptoms between the predetermined time points (baseline versus two weeks & baseline versus three months). Based on the guidelines by Cicchetti et al.[54, 55] an ICC <0.40 indicates poor level of clinical significance. The level is fair when the ICC is between 0.40 and 0.59, good between 0.60 and 0.74 and excellent when the ICC is between 0.75 and 1.00. To assess the test-retest reliability patients' pain should not have changed (outcome based on the PGIC) because otherwise the ICC would not reflect the consistency of the DN4. Test-retest reliability was assessed for those questionnaires returned within 7–21 days for the two weeks test-retest reliability and 60–120 days for the three months test—retest reliability. The ICC and responsiveness of the DN4-symptoms was assessed at each point of measurement.

Two-tailed p-value below 0.05 was considered statistically significant.

## Results

### Patients

In this study 330 consecutive patients were assessed for eligibility (Fig 1). Of these, 291 participated in the study between October 2009 and July 2013. Two patients did not give their informed consent. Exclusion (n = 37) was because of not fulfilling the in- and exclusion criteria (n = 13): patients with LBLP or NSAP without radiating pain: n = 1; patients with less than 3 months pain complaints: n = 2; patients with pain with an oncological cause: n = 2; patients with painful syndromes of unknown origin or associated with diffuse pains: n = 7; patients with severe mental illness: n = 1; missing baseline measurements due to not returning questionnaires by the patient: n = 16; missing pain classification based on the grading system by one physician (n = 5) or both the physicians (n = 3). 132 patients had LBLP with radiation in



**Fig 1. Flow diagram for the outcome of the physicians assessment and the NeuPSIG grading system.** Present NePC: present neuropathic pain component; Undetermined: Both physicians disagree with each other about the existence of a neuropathic pain component; Absent NePC: absent neuropathic pain component; n = total number of patients in analysis PhA: Physicians assessment; GS: Neuropathic pain special interest group grading system (missing pain classification based on the grading system: n = 8).

<https://doi.org/10.1371/journal.pone.0187961.g001>

one or two legs (45.4%), 51 NSAP with radiation in one or both arms (17.5%) and 108 patients (37.1%) had PSNO: 86 patients with pain after treatment for breast cancer (surgery and chemotherapy and/or radiation therapy and/or hormonal therapy). Twenty-two patients had pain for various reasons: peripheral nerve damage (n = 12), radicular pain (n = 3), polyneuropathy (n = 3), CRPS (n = 2) and post stroke pain (n = 2). The gold standard for presence of the NePC in this study was the concordant clinical opinion of both physicians. After pain classification by two physicians, 170 patients were classified as present NePC, 58 as absent NePC and in 63 patients the two physicians made a different pain classification: ‘undetermined’. Using the grading system, 139 patients were assigned as having a present NePC, 93 patients as absent NePC and 51 patients were assigned as undetermined. The DN4 was full filled by the patients at a median of one day (IQR 0–5 days) following the assessments by the physicians.

Clinical and social-demographic details of the 291 patients were analyzed based on their pain classification. No statistically significant differences were found between present NePC and absent NePC for gender, age, height, weight, BMI, medication and duration of pain. Also no statistically significant difference was observed between absent NePC and present NePC regarding current- worst and average pain (Table 1).

**Table 1. Clinical and socio-demographic characteristics of the patients related to physicians agreement for the existence of a neuropathic pain component.**

	NePC Absent		Present		P	Undetermined		P	
	N		N			N			
<b>Total number of patients</b>	58		170			63			
<b>Gender</b>					0.163 <sup>c</sup>			0.164 <sup>c</sup>	
	Male	25 (43%)		56 (33%)			17 (27%)		
	Female	33 (57%)		114 (67%)			46 (73%)		
<b>Age (Years) #</b>	58	55 ± 12	170	56 ± 11	0.594 <sup>a</sup>	63	58 ± 13	0.522 <sup>d</sup>	
<b>Height (cm) #</b>	55	172 ± 9	164	172 ± 8	0.845 <sup>a</sup>	6	170 ± 9	0.250 <sup>d</sup>	
<b>Weight (kg) #</b>	55	84 ± 2	167	80 ± 17	0.382 <sup>a</sup>	6	80 ± 16	0.461 <sup>d</sup>	
<b>BMI (kg/m<sup>2</sup>) #</b>	54	28 ± 8	164	27 ± 5	0.436 <sup>a</sup>	6	27 ± 5	0.593 <sup>d</sup>	
<b>Medication use ^</b>	55	56.9%	168	66.1%	0.414 <sup>c</sup>	61	57.4%	0.423 <sup>c</sup>	
<b>Duration of pain (months) #</b>	57	72 ± 90	169	60 ± 76	0.327 <sup>a</sup>	62	49 ± 46	0.247 <sup>d</sup>	
<b>Pain* (NRS; 0–10)</b>									
	Current pain	57	5 (3–7)	167	6 (3–7)	0.577 <sup>b</sup>	6	4 (1–7)	0.084 <sup>e</sup>
	Worst pain during the past four weeks	57	8 (5–9)	167	8 (7–9)	0.371 <sup>b</sup>	6	7 (5–8)	0.053 <sup>e</sup>
	Average pain during the past four weeks	57	6 (3.5–7)	167	6 (5–8)	0.233 <sup>b</sup>	6	6 (3–7)	<b>0.018<sup>e</sup></b>

Classification for the existence of NePC is based on physicians assessment of the patients. NePC: neuropathic pain component; Absent: NePC is absent; Present: NePC is present; Undetermined: both physicians disagree with each other about the existence of a neuropathic pain component; N: total number of patients in analysis; n: number of patients;

^percentage;

#Standard deviation;

\* Inter quartile range. A: physicians A; B: Physicians B; P value for significant difference between groups (P ≤ 0.05) by use of different analyse methods:

<sup>a</sup>: Students t-test;

<sup>b</sup>: Mann-Whitney U test;

<sup>c</sup>: Chi-square;

<sup>d</sup>: One-Way ANOVA;

<sup>e</sup>: Kruskal-Wallis test.

<https://doi.org/10.1371/journal.pone.0187961.t001>



## Reliability

The proportion of agreement after chance agreement is removed (Cohen's Kappa, K) for the classification of patients' pain (absent NePC or present NePC) by the physicians was 0.49 (moderate), with a PA of 78.4% ( $P_i = 0.38$ ;  $n = 291$ ). For the classification of patients' pain on basis of the grading system K was 0.63 (good) and PA was 82% ( $P_i = 0.16$ ;  $n = 283$ ). The outcome of K and PA regarding the DN4 compared to the outcome of the assessment by physicians A was respectively 0.34 (fair) and 69.8% ( $P_i = 0.33$ ;  $n = 275$ ). Compared to the outcome of the assessment by physicians B it was 0.33 (fair) and 69.2% ( $P_i = 0.30$ ;  $n = 263$ ). Comparing the outcome of the DN4 to the outcome of the grading system, it was 0.35 (fair) and 69.1% ( $P_i = 0.22$ ;  $n = 272$ ) for physicians A, and 0.32 (fair) and 67.3% ( $P_i = 0.19$ ;  $n = 260$ ) for physicians B (Table 2). The interobserver reliability for 'hypoesthesia to touch' as well as for 'brushing' was respectively  $K = 0.59$  (moderate) (PA = 79.7%) and  $K = 0.53$  (moderate) (PA = 76.6%). The interobserver reliability for 'hypoesthesia to prick' was  $K = 0.21$  (fair); PA = 87% (Table 3).

In 253 patients all the six outcome variables (two times the physicians' assessment, two times the grading system and The DN4 by physician A and DN4 by physician B) was available. In 83 patients (32.8%), the pain was classified as present NePC in all outcomes and in 22 patients (8.7%) it was six times negative, indicating absent NePC, so the agreement on all the six measures was 41.5% (the percentage of agreement based on both the gold standards and both the grading systems only was 56.9%).

## Factor analysis

Table 4 shows the loading factor of the items of the DN4 according to the rotated component matrix factor analysis with Kaiser normalization. The analysis was performed by use of the 10 questions in the DN4 and revealed a 4-factor solution explaining 59.3% of the variance for the first physicians' assessment (physicians A): Factor 1 included two items (hypoesthesia to touch, brushing) indicating that there was an inter-relation between those items (Cronbach's  $\alpha$ : 0.87). Factor 2 included three items (painful cold, tingling, hypoesthesia to prick) (Cronbach's  $\alpha$ : 0.37). Factor 3, consisted of four items (burning, electric shocks, pins and needles, numbness); Cronbach's  $\alpha$ : 0.51). Factor 4 consisted of one item (itching) (Table 4). In the S1 Table we provided the factor analysis for both the physicians assessments (A & B), the DN4 symptoms solely and the DN4 signs for both physicians' assessments (A & B). Internal consistency of all the components of the DN4 for the physicians A at baseline was assessed via Cronbach's  $\alpha$ : 0.57; for the physicians B it was 0.55. Cronbach's  $\alpha$  for DN4-symptoms was 0.52. Cronbach's  $\alpha$  for the DN4-signs for A and B were respectively 0.68 and 0.66.

## Items of the DN4

The DN4-symptoms (pain descriptors) burning, electric shocks, tingling, pins and needles, and numbness were statistically significant associated ( $\chi^2$ ) with the classification by the physicians (absent NePC, present NePC or undetermined),  $p < 0.05$ . The descriptors 'painful cold' ( $p = 0.210$ ) and 'itching' ( $p = 0.409$ ) were not associated with the outcome of the classification. The DN4-signs (examination items) hypoesthesia to touch, pricking and brushing were statistically significant associated ( $\chi^2$ ) with the classification by the physicians (absent NePC, present NePC or undetermined),  $p < 0.05$ .

The median of the total sum score of the DN4 for patients classified as absent NePC was 2, the median for the DN4-symptoms items was 2 and for the DN4-signs items the median was 0; for patients classified as present NePC it was at median 5, 3 and 2, respectively. As calculated based on the Kruskal-Wallis test there was for the sum scores of the DN4, the DN4-symptoms items and the DN4-signs items a statistical significant difference between absent NePC and

**Table 2. The kappa coefficient between the classification on basis of the assessment by the physicians, the grading systems, the DN4 and the kappa coefficient between both physicians regarding the DN4-signs.**

		Classification physician B	Grading A	Grading B	DN4 A	DN4 B	DN4 Symptoms	DN4-Signs A	DN4-Signs B
<b>Classification physician A</b>	n	<b>291</b>	<b>286</b>	288	275	263	288	279	266
	K	<b>0.49</b>	<b>0.48</b>	0.32	0.34	0.34	0.32	0.37	0.26
	PA	<b>78.4</b>	<b>76.2</b>	67.4	69.8	70.0	67.4	70.3	64.4
	Pi	<b>0.38</b>	<b>0.32</b>	0.26	0.33	0.31	0.26	0.30	0.30
<b>Classification physician B</b>	n		286	<b>288</b>	275	263	288	279	266
	K		0.38	<b>0.48</b>	0.33	0.33	0.21	0.39	0.37
	PA		71.0	<b>75.0</b>	69.1	69.2	62.8	71.0	70.7
	Pi		0.28	<b>0.22</b>	0.29	0.30	0.25	0.26	0.28
<b>Grading A</b>	n			<b>283</b>	272	259	283	<b>276</b>	262
	K			<b>0.63</b>	0.35	0.31	0.14	<b>0.54</b>	0.31
	PA			<b>82.0</b>	69.1	67.2	58.6	<b>77.5</b>	67.2
	Pi			<b>0.16</b>	0.22	0.23	0.19	<b>0.19</b>	0.21
<b>Grading B</b>	n				272	260	285	<b>276</b>	<b>263</b>
	K				0.29	0.32	0.13	<b>0.53</b>	<b>0.45</b>
	PA				65.4	67.3	57.2	<b>76.8</b>	<b>73.4</b>
	Pi				0.18	0.19	0.14	<b>0.14</b>	<b>0.16</b>
<b>DN4 A</b>	n					<b>257</b>	<b>275</b>	<b>275</b>	257
	K					<b>0.76</b>	<b>0.62</b>	<b>0.52</b>	0.29
	PA					<b>88.7</b>	<b>81.8</b>	<b>76.7</b>	65.8
	Pi					<b>0.21</b>	<b>0.19</b>	<b>0.17</b>	0.19
<b>DN4 B</b>	n						<b>263</b>	<b>257</b>	<b>263</b>
	K						<b>0.65</b>	<b>0.40</b>	<b>0.45</b>
	PA						<b>82.9</b>	<b>71.2</b>	<b>73.4</b>
	Pi						<b>0.18</b>	<b>0.21</b>	<b>0.17</b>
<b>DN4 symptoms</b>	n							276	263
	K							0.15	0.10
	PA							58.7	56.3
	Pi							0.17	0.16
<b>DN4-Signs A</b>	n								<b>260</b>
	K								<b>0.55</b>
	PA								<b>78.4</b>
	Pi								<b>0.18</b>

n = total number of patients in the analysis; K = Cohen's kappa value; PA (%) = percentage of agreement between two outcome variables; Pi = Prevalence index

<https://doi.org/10.1371/journal.pone.0187961.t002>

present NePC (P<0.001), between present NePC and undetermined (P<0.001) and between absent NePC and undetermined (P<0.001). In Table 5 the outcomes for all individual items and the three DN4 scales (for physicians A as well as for physicians B) are presented according to the pain classification by the physicians (Table 5).

### Validity

We constructed ROC-curves for the DN4, the DN4-symptoms and the DN4-signs with respect to the classification by physician A or B and according to the neuropathic pain grading system

**Table 3. The kappa coefficient between both physicians regarding the DN4-signs.**

		Hypoesthesia to touch DN4-signs B	Hypoesthesia to prick DN4-signs B	Brushing DN4-signs B
<b>Hypoesthesia to touch</b> DN4-signs A	n	<b>222</b>		
	K	<b>0.59</b>		
	PA	<b>79.7</b>		
	Pi	<b>0.10</b>		
<b>Hypoesthesia to prick</b> DN4-signs A	n		244	
	K		0.21	
	PA		87.3	
	Pi		-0.82	
<b>Brushing</b> DN4-signs A	n			<b>222</b>
	K			<b>0.53</b>
	PA			<b>76.6</b>
	Pi			<b>0.11</b>

n = total number of patients in the analysis; K = Cohen's kappa value; PA (%) = percentage of agreement between two outcome variables; Pi = Prevalence index

<https://doi.org/10.1371/journal.pone.0187961.t003>

by physician A or B and all the combinations (Concordant assessment by physicians A and B together, concordant grading system by Physicians A and B together and concordant grading system for Physicians A and B together with the concordant grading system by physicians A and B). This because of the chosen gold standard and the grading system in which patients were classified by two different physicians. This might have lead to differences in the outcomes relative to the individual outcome by the physician. In Fig 2 the ROC-curve is displayed for the DN4 (physicians A and physicians B), DN4-symptoms and the DN4-signs (physicians A and physicians B) (Fig 2).

**Table 4. Loading factors of the items of the DN4 according to the rotated component matrix factor analysis.**

DN4	Component (Physicians A)			
	1	2	3	4
<b>Burning</b>	0.25	0.28	<b>0.29</b>	0.15
<b>Painful cold</b>		<b>0.62</b>		
<b>Electric shocks</b>			<b>0.72</b>	
<b>Tingling</b>		<b>0.68</b>		
<b>Pins and needles</b>		0.35	<b>0.45</b>	0.27
<b>Numbness</b>			<b>0.71</b>	
<b>Itching</b>				<b>0.86</b>
<b>Hypoesthesia to touch</b>	<b>0.87</b>			
<b>Hypoesthesia to prick</b>	0.38	<b>0.63</b>		
<b>Brushing</b>	<b>0.90</b>			
<b>Cronbach's alpha</b>	<i>0.81</i>	<i>0.37</i>	<i>0.51</i>	

Loading factors < 0.25 are omitted to improve readability

<https://doi.org/10.1371/journal.pone.0187961.t004>

**Table 5. The median (IQR) and percentages of the items of the DN4 by physicians agreement of a NePC.**

NePC		Absent		Present		Undetermined				
		N		N	P	N	P			
<b>Total number of patient</b>		58		170		63				
<b>DN4-Symptoms<sup>^</sup></b>	Burning	56	12 (21%)	161	77 (48%)	0.001 <sup>a</sup>	57	22 (39%)	0.002 <sup>a</sup>	
	Painful Cold	54	6 (11%)	154	34 (22%)	0.078 <sup>a</sup>	53	11 (21%)	0.210 <sup>a</sup>	
	Electric Shocks	55	18 (33%)	162	87 (54%)	0.007 <sup>a</sup>	55	19 (35%)	0.005 <sup>a</sup>	
	Tingling	55	29 (53%)	160	110 (69%)	0.032 <sup>a</sup>	57	28 (49%)	0.011 <sup>a</sup>	
	Pins and Needles	52	19 (37%)	157	101 (60%)	0.000 <sup>a</sup>	58	27 (47%)	0.001 <sup>a</sup>	
	Numbness	54	29 (54%)	165	131 (79%)	0.000 <sup>a</sup>	59	42 (71%)	0.001 <sup>a</sup>	
<b>DN4-signs<sup>^</sup></b>	Itching	51	10 (20%)	149	25 (17%)	0.646 <sup>a</sup>	56	14 (25%)	0.409 <sup>a</sup>	
	Hypoesthesia to touch A	A	42	9 (21%)	153	102 (67%)	0.000 <sup>a</sup>	60	16 (27%)	0.000 <sup>a</sup>
		B	41	11 (27%)	151	101 (67%)	0.000 <sup>a</sup>	49	18 (37%)	0.000 <sup>a</sup>
	Hypoesthesia to prick A	A	47	0 (0%)	162	20 (12%)	0.011 <sup>a</sup>	58	3 (5%)	0.017 <sup>a</sup>
		B	48	0 (0%)	159	21 (12%)	0.008 <sup>a</sup>	53	1 (2%)	0.002 <sup>a</sup>
	Brushing A	A	43	9 (21%)	157	110 (70%)	0.000 <sup>a</sup>	55	14 (25%)	0.000 <sup>a</sup>
		B	43	13 (30%)	151	99 (66%)	0.000 <sup>a</sup>	52	19 (37%)	0.000 <sup>a</sup>
	<b>Total sum score DN4 A* (0–10)</b>		47	2 (1–3)	166	5 (3–6)	0.000 <sup>b</sup>	62	3 (2–4)	0.000 <sup>c</sup>
<b>Total sum score DN4 B* (0–10)</b>		48	2 (2–3,75)	159	5 (3–6)	0.000 <sup>b</sup>	56	3 (2–4.75)	0.000 <sup>c</sup>	
<b>Total sum score DN4 symptoms* (0–7)</b>		57	2 (1–3)	168	3 (2–5)	0.000 <sup>b</sup>	63	2 (2–4)	0.000 <sup>c</sup>	
<b>Total sum score DN4 signs A* (0–3)</b>		49	0 (0–0)	168	2 (1–2)	0.000 <sup>b</sup>	62	0 (0–1)	0.000 <sup>c</sup>	
<b>Total sum score DN4 signs B* (0–3)</b>		49	0 (0–1)	161	2 (0–2)	0.000 <sup>b</sup>	56	0 (0–2)	0.000 <sup>c</sup>	

Classification for the existence of NePC is based on physicians assessment of the patients. NePC: neuropathic pain component; Absent: NePC is absent; Present: NePC is present; Undetermined: both physicians disagree with each other about the existence of a neuropathic pain component; N: total number of patients in analysis; n: number of patients;

<sup>^</sup>percentage;

\* Inter quartile range. A: physicians A; B: Physicians B; P value for significant difference between groups ( $P \leq 0.05$ ) by use of different analyse methods:

<sup>a</sup>: Chi-Squared;

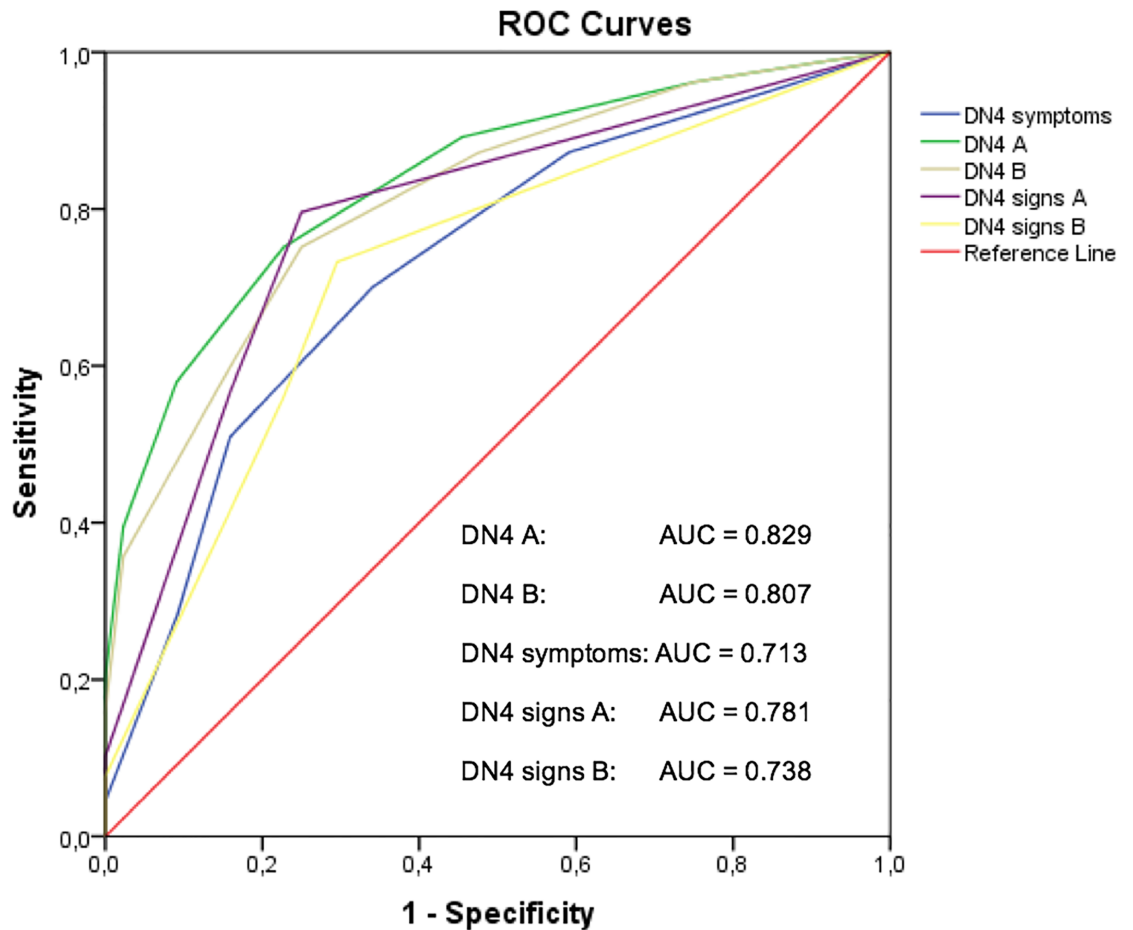
<sup>b</sup>: Mann-Whitney U test;

<sup>c</sup>: Kruskal-Wallis test.

<https://doi.org/10.1371/journal.pone.0187961.t005>

Based on the gold standard the sensitivity of the DN4 was on average (at maximal Youden-index, cut off point: 4/10) 75% (95% CI 0.68–0.81), specificity 76% (95% CI 0.61–0.86), positive predictive value 92% and the positive likelihood ratio was 3.09 (95% CI 1.82–5.39) (Table 5; S2 Table). For patients with LBLP the sensitivity was on average 75% and specificity was on average 81%. For patients with NSAP the averaged sensitivity was 73% and the specificity was on average 72%. For patients with pain due to a PSNO it was respectively, on average, 70% and 78%. The sensitivity of the DN4-symptoms was, in respect to the gold standard, 70% (95% CI 0.63–0.77) and the specificity was 67% (95% CI 0.54–0.78) (at maximal Youden-index, cut off point 3/7). Analysis of the DN4-signs solely resulted in an average sensitivity of 75% (95% CI 0.66–0.82) and an average specificity of 75% (95% CI 0.58–0.87) (at maximal Youden-index, cut off point 1/3). With the outcome based on the grading system the sensitivity was on average 76% (95% CI 0.68–0.82) and the specificity was 64% (95% CI 0.51–0.74) (at maximal Youden-index, cut off point 4/10). (Table 6; S2 Table).

In Table 6 and S2 Table we present the number of patients per group, values of the AUC, Youden index, cut-off score, true positives, false positives, false negatives, true negatives, sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, the diagnostic odds ratio, the a-priori chance for the existence (or not) of a NePC and



**Fig 2. The ROC curve of the DN4, DN4 symptoms and the DN4 signs to the probability of the presence of NePC as classified based on the assessment by the physicians (A and B).** DN4: Doleur Neuropathique en 4 questions; DN4-symptoms: the items filled in by the patient; DN4 A: DN4-symptoms filled in by the patient and DN4-signs as assessed by physicians A; DN4 B: DN4-symptoms filled in by the patient and DN4-signs as assessed by physicians B; DN4 signs A: DN4-signs as assessed by physicians A; DN4 signs B: DN4-signs as assessed by physicians B.

<https://doi.org/10.1371/journal.pone.0187961.g002>

false positive and negative ratios for all validity outcomes (DN4 A & B, DN4-symptoms, DN4-signs A & B) divided according to the pain classification and divided into LBLP, NSAP and PSNO (Table 6 and S2 Table).

### Test-retest reliability

Stability and responsiveness of the DN4-symptoms over time was assessed over a period of two weeks. The median sum score (IQR) of the DN4 at baseline for the total group was 3 (2–4), after two weeks it was 3 (2–4). Taking into consideration the fact that patients’ pain should not have changed (outcome based on the PGIC) because otherwise the ICC would not reflect the consistency of the DN4, test-retest reliability via ICC was 0.84 (excellent) (95%CI 0.80–0.87; n = 265). For the time gap of 7–21 days (to rule out the early or delayed return of questionnaires) between the first and second DN4-symptoms the ICC was 0.85 (excellent) (95% CI 0.79–0.90; n = 122). After three months, with no change in patients pain

**Table 6. The area under the curve and the sensitivity / specificity at the optimal cut-off point of the DN4 under the condition of equal costs of misclassification to classify a neuropathic pain component by the classification and the grading system of the physicians.**

	Present NePC	Absent NePC	AUC	(95%CI)	Youden index	Cut-off	Sens %	95% CI	Spec %	95% CI	PPV %	NPV	PLR	95% CI
<b>Classification A = B</b>														
<b>DN4 A</b>	<b>166</b>	<b>47</b>	<b>0.829</b>	<b>0.767–0.890</b>	<b>0.513</b>	<b>4</b>	<b>75</b>	<b>0.676–0.807</b>	<b>77</b>	<b>0.628–0.864</b>	<b>92</b>	<b>46</b>	<b>3.19</b>	<b>1.889–5.394</b>
LBP	72	26	0.823	0.738–0.90	0.544	4	74	0.624–0.824	81	0.621–0.915	91	53	3.83	1.72–8.517
NSAP	23	10	0.763	0.576–0.950	0.439	3	74	0.535–0.875	70	0.397–0.892	85	54	2.46	0.927–6.547
PSNO	71	11	0.836	0.713–0.959	0.543	5	63	0.518–0.736	91	0.623–0.984	98	28	6.97	1.067–45.558
<b>DN4 B</b>	<b>159</b>	<b>48</b>	<b>0.807</b>	<b>0.742–0.872</b>	<b>0.498</b>	<b>4</b>	<b>75</b>	<b>0.678–0.81</b>	<b>75</b>	<b>0.612–0.851</b>	<b>91</b>	<b>47</b>	<b>2.99</b>	<b>1.819–4.927</b>
LBP	67	26	0.821	0.736–0.906	0.554	4	75	0.631–0.835	81	0.621–0.915	91	55	3.88	1.744–8.637
NSP	21	11	0.725	0.529–0.921	0.442	4	71	0.5–0.862	73	0.434–0.903	83	57	2.62	0.961–7.135
PSNO	71	11	0.777	0.644–0.910	0.397	4	76	0.65–0.845	64	0.354–0.848	93	29	2.09	0.947–4.62
<b>DN4-symptoms</b>	<b>168</b>	<b>57</b>	<b>0.713</b>	<b>0.634–0.791</b>	<b>0.369</b>	<b>3</b>	<b>70</b>	<b>0.629–0.766</b>	<b>67</b>	<b>0.537–0.775</b>	<b>86</b>	<b>43</b>	<b>2.11</b>	<b>1.441–3.082</b>
LBP	74	28	0.716	0.606–0.826	0.348	3	72	0.605–0.806	61	0.424–0.764	83	45	1.82	1.126–2.953
					0.348	4	53	0.415–0.637	82	0.644–0.921	89	40	2.95	1.296–6.723
NSAP	23	18	0.661	0.484–0.837	0.374	3	65	0.449–0.812	72	0.491–0.875	75	62	2.35	1.052–5.238
PSNO	71	11	0.764	0.611–0.918	0.431	3	70	0.59–0.798	73	0.434–0.903	94	28	2.58	0.972–6.858
<b>DN4-signs A</b>	<b>168</b>	<b>49</b>	<b>0.781</b>	<b>0.709–0.852</b>	<b>0.537</b>	<b>1</b>	<b>76</b>	<b>0.692–0.82</b>	<b>78</b>	<b>0.641–0.87</b>	<b>92</b>	<b>49</b>	<b>3.39</b>	<b>2.003–5.75</b>
LBP	73	26	0.744	0.637–0.850	0.479	1	67	0.557–0.768	81	0.621–0.942	91	47	3.49	1.562–7.799
NSAP	23	11	0.783	0.632–0.933	0.565	1	57	0.368–0.744	100	0.741–1.000	100	52	---	---
PSNO	72	12	0.763	0.608–0.917	0.417	1	92	0.83–0.961	50	0.254–0.746	92	50	1.83	1.037–3.242
<b>DN4-signs B</b>	<b>161</b>	<b>49</b>	<b>0.738</b>	<b>0.660–0.816</b>	<b>0.447</b>	<b>1</b>	<b>73</b>	<b>0.66–0.795</b>	<b>71</b>	<b>0.576–0.822</b>	<b>89</b>	<b>45</b>	<b>2.57</b>	<b>1.632–4.033</b>
LBP	68	26	0.742	0.636–0.847	0.455	1	65	0.528–0.75	81	0.621–0.915	90	47	3.36	1.501–7.541
NSAP	21	11	0.777	0.606–0.948	0.576	1	67	0.454–0.828	91	0.623–0.984	93	59	7.33	1.104–48.691
PSNO	72	12	0.628	0.463–0.794	0.167	1	83	0.731–0.902	33	0.138–0.609	88	25	1.25	0.827–1.89
					0.167	2	67	0.552–0.765	50	0.254–0.746	89	20	1.33	0.74–2.403
<b>Grading A = B</b>														
<b>DN4 A</b>	<b>138</b>	<b>81</b>	<b>0.771</b>	<b>0.709–0.833</b>	<b>0.396</b>	<b>4</b>	<b>75</b>	<b>0.676–0.818</b>	<b>64</b>	<b>0.533–0.738</b>	<b>78</b>	<b>60</b>	<b>2.10</b>	<b>1.549–2.861</b>
<b>DN4 B</b>	<b>135</b>	<b>75</b>	<b>0.744</b>	<b>0.673–0.814</b>	<b>0.382</b>	<b>4</b>	<b>76</b>	<b>0.677–0.82</b>	<b>63</b>	<b>0.514–0.727</b>	<b>7</b>	<b>59</b>	<b>2.02</b>	<b>1.487–2.755</b>
<b>DN4-symptoms</b>	<b>139</b>	<b>91</b>	<b>0.610</b>	<b>0.537–0.684</b>	<b>0.179</b>	<b>4</b>	<b>45</b>	<b>0.373–0.536</b>	<b>73</b>	<b>0.626–0.806</b>	<b>72</b>	<b>46</b>	<b>1.65</b>	<b>1.128–2.414</b>
<b>DN4-signs A</b>	<b>138</b>	<b>84</b>	<b>0.855</b>	<b>0.803–0.908</b>	<b>0.653</b>	<b>1</b>	<b>86</b>	<b>0.787–0.904</b>	<b>80</b>	<b>0.7–0.87</b>	<b>87</b>	<b>77</b>	<b>4.23</b>	<b>2.748–6.495</b>
<b>DN4-signs B</b>	<b>135</b>	<b>77</b>	<b>0.759</b>	<b>0.691–0.827</b>	<b>0.466</b>	<b>1</b>	<b>78</b>	<b>0.701–0.84</b>	<b>69</b>	<b>0.578–0.781</b>	<b>81</b>	<b>64</b>	<b>2.50</b>	<b>1.769–3.52</b>

DN4: Douleur neuropathique en 4 questions; present NePC: Neuropathic pain component existing; Absent NePC: Neuropathic pain component not existing; AUC: Area under curve; 95%CI: 95% confidence interval; Sens.: Sensitivity; Spec.: Specificity; PPV: Positive predictive value; A: Physicians A; B: Physicians B; LBP: Patients suffering from low back and leg pain; NSAP: Patients suffering from neck shoulder arm pain; PSNO: Patients suffering from pain due to a suspected neuropathic origin.

<https://doi.org/10.1371/journal.pone.0187961.t006>

and a time gap of 60–120 days between the first and third DN4-symptoms, ICC was 0.79 (excellent) (95% CI 0.70–0.86; n = 102).

## Discussion

The DN4 seems, in this study, to help to identify a neuropathic pain component in a consecutive population of patients with chronic pain in a moderate way.

## Reliability

We used the concordant opinion about the classification of patients' pain by two physicians as the gold standard. It is disputable if the term gold standard is practicable. However, as written by Versi[56] [57] "the gold standard is not the perfect test but merely the best available test. ... Against which newer tests can be compared". There are studies regarding the validity of the DN4 using only one physician's opinion[21, 30]. To our opinion it is preferable to use two

physicians as the gold standard, which is also performed in the original validation study of the DN4[17]. This might lead to less false positive or false negative outcomes which, of course, will lead to a more accurate validity outcome. The physicians in this study agreed on pain classification in 78% of the patients. In other studies without pre-stratification of patients on the target outcome the results for the physicians agreement were 53%[25] and 89%[27]. The kappa coefficient between the DN4 as filled in by physician A compared to the DN4 by physician B was 'good' with a high percentage of agreement. Test-retest reliability of the DN4-symptoms in this study was excellent. Based on these results DN4 seems to be reliable. However, it is possible that an instrument is reliable without being valid[58].

## Validity

To quantify the screening ability of the DN4, for the existence of a NePC, sensitivity and specificity can be used[59]. However, in clinical practice we want to know how many patients with a positive score on the DN4 really does have a NePC. To report this, the positive and negative predictive values are important because they give the proportion of patients with positive or negative test results which are correctly diagnosed[60]. The predictive value depends on the prevalence of NePC in the group of patients under study[60]. In our study the prevalence of NePC was high, 75%. The higher the prevalence of NePC in the group under study the more sure it is that a positive outcome of the DN4 indicates the presence of a NePC, but the less sure it is that a negative DN4 outcome indicates absent NePC[60]. The likelihood ratio gives an indication of the value of the DN4 for increasing certainty about a positive diagnosis[60]. A higher likelihood ratio might indicate that the DN4 is useful, but is still not sure that a positive outcome of the DN4 is a good indicator for the presence of a NePC[60]. In the literature there are, as far as we know, no 'cut-off' scores for the validity indices. In our study we found a sensitivity of 75% (DN4-symptoms 70%), a specificity of 76% (DN4-symptoms 67%), positive predictive value of 92%, negative predictive value of 46% and the positive and negative likelihood ratios were respectively 3.09 and 0.34. In the original study by Bouhassira et al.[17] patients with only 'typical' neuropathic or nociceptive entities and a VAS of  $\geq 40$  mm (0–100mm) were included. They found a sensitivity of 83% and a specificity of 90%. For the DN4-symptoms the sensitivity was 78% and the specificity 81%. The Dutch version of the DN4[20] was validated before in a consecutive group of patients suffering from chronic pain for more than three months with a pain score of 5 or higher on a 0–10 numeric rating scale (NRS)[25]. For the DN4 a sensitivity of 75% and a specificity of 79% was found. For the DN4-symptoms version sensitivity was 74% and the specificity 79%. Van Seventer et al. concluded that the DN4 was a diagnostic tool with a good ability to discriminate between neuropathic pain and nociceptive pain[25]. However, the paper by Bouhassira et al.[17] and the paper by Van seventer et al.[25] both didn't report the predictive values and likelihood ratios. Inappropriate screening might result in higher health care costs due to more diagnostic testing or even lead to a harmful treatment for the patient[61]. It seems that the validity indices in our study are resulting in a lower score for the DN4 as in the original publication[17] and than in other studies[4, 21, 23–28, 30, 31, 62–67]. This might have several reasons. At first, we did not pre-stratify on the target outcome. In studies, besides the original validation study[17] with pre-stratification on the target outcome[23, 24, 26, 28, 31] (neuropathic or non-neuropathic pain), the sensitivity of the DN4 was ranging from 90%[26] till 100%[24], the specificity from 93%[24]-97%[23, 28]. In studies where there was no pre-stratification on the target outcome (neuropathic or non-neuropathic pain), the sensitivity of the DN4 was ranging from 80%[21] till 100%[30], the specificity ranges from 78%[21, 27] till 87%[30]. These results are showing that the validity of the DN4 is lower in studies without pre-stratification than in studies where patients were stratified based on their

pain classification before entering the study. In studies with specified diseases as spinal cord injury[64]; diabetes[63, 64]; leprosy[65, 66]; FBSS[67], chronic low back pain[4] and in patients with cancer before starting with chemotherapy[68], the sensitivity (62%-100%) and specificity (44%-93%) ranges were much wider. Our results, also when separated into results for LBLP, NSAP and PSNO, falls within these ranges. This indicates that the neuropathic pain component is not always clear and/or easy to classify by use of the DN4 in the different medical conditions. Secondly, in our study we did not have a minimum level of pain as an inclusion criteria. In seven studies a minimal level of pain (on a rating scale of 0–10) was not an inclusion criteria [21, 23, 31, 62, 63, 65, 66]. In other studies a level  $\geq$  three[64, 67],  $\geq$  four[4, 17, 24, 26, 28, 30] or  $\geq$  five[25, 27] is set as an inclusion criterium. As shown by Perez et al[21], pain severity has a major influence on the sensitivity and specificity of the DN4. A severity of  $< 40$  mm on a 0-100mm VAS resulted in a sensitivity of 56% and a specificity of 67%. For moderate pain (between 40mm en 70mm on a 0-100mm VAS) it was 85% and 84% respectively, and  $>70$  mm sensitivity was 80% and specificity was 74%[21]. In a study by Marksman[67] in patients after FBSS it was showed that the presence of neuropathic characteristics, as determined by the DN4, was associated with a higher pain intensity. These facts are crucial for the validation of a screening instrument because such a tool must be valid for use in daily clinical out-patient practice and/or for epidemiological purposes.

As a second comparison, we validated the DN4 in comparison with the grading system[32] [21]. In this study, we combined 'unlikely' and possible neuropathic pain as absent NePC and probable and definite as present NePC, which resulted in an average sensitivity for the DN4 of 76% and an average specificity of 64%. In patients with a failed back surgery syndrome[67], the validation of the DN4 resulted in a sensitivity of 62% and a specificity of 44%. In a study by Sadler et al [69] where patients with neuropathic pain were compared to musculoskeletal pain the sensitivity was 76% and the specificity was 70%. However, in patients with a more mixed pain the sensitivity and specificity descended to 59% and 70% respectively. Abdallah et al [36] compared the DN4 with the grading system in patients after breast tumor resection with and without paravertebral blocks. This resulted in a sensitivity of 90% and a specificity of 60% to identify patients with chronic neuropathic pain based on the outcome of the grading system. However, this outcome was not validated by (expert) physicians. The distinction between possible neuropathic pain and probable or definite neuropathic pain is of high importance because the outcome forms the basis for selecting a different treatment strategy[34]. The combination of outcomes in our study might have resulted in a lower sensitivity and a bit higher specificity in comparison with the classification in the study of Abdallah et al[36].

**Douleur Neuropathique en 4 questions (DN4).** Bouhassira [17] presented the DN4 as a clinician-administered questionnaire. In different studies not a physician but a research coordinator[30], a nurse[25] or the patient self [25, 70] filled in the DN4. In our study we gave the patient the questionnaire with the 7-items (DN4-symptoms) to fill them in after the physical examinations. The three examination-items (DN4-signs) were incorporated in the standardized assessment form which should be filled in by the physician. We presented the DN4 total sum score as well as the DN4-signs score separately for physicians A and B. This is due to the fact that it is only possible to have one outcome when the sign-items were performed by one physician.

## Strength and weaknesses

There are several strengths in this study. At first, this study reflects daily clinical practice. In this study, we included a large cohort of patients irrespective of the predominant origin of the pain and level of pain which corresponds to a typical daily clinical patient population. These



patients were associated with the most common specified medical conditions for pain (i.e. LBLP or NSAP or PSNO) and classified by two, independently working, physicians. Moreover, patients were referred from primary care to secondary and tertiary pain clinics and were assessed for their complaints for the first time at the time of inclusion in this study. This limits the risk of systematic bias and also reflects daily clinical practice. Secondly, we used a standardized assessment form in which the bedside examination and the grading system[32, 33] and the DN4-signs were incorporated. This might, however, have led to an influence on each other which made the physician more sure about the final classification of patients pain and thus made the gold standard stronger. There are also some weaknesses in this study. As said before, we have not used the DN4-symptoms as a interview by a physician but as a questionnaire which has to be filled in by the patient. This might have had an influence on the reliability and validity. In the revised EFNS guidelines on neuropathic pain assessment[42] it is suggested that “The seven sensory descriptors can be used as a self-report questionnaire with similar results”. Moreover, above the official Dutch version[20, 25] of the DN4 is written in Dutch: “To be completed by the patient”. In the paper by van Seventer et al the agreement between the patient administered and a nurse administered was good till very good for the first seven items[25]. It would be of interest to see if there are differences in the outcome when the DN4 is filled in by the patient himself or as an interview by the pain physician. Questions by the patient to the nurse or via telephone to the researcher regarding the DN4 were very rare. However, we didn't keep track of the questions. Another limitation is the fact that we only tested the test-retest reliability regarding the DN4-symptoms and not the DN4-signs to prevent the patient to come back to the hospital only for these test-items. Another weakness is the gold standard which is, for now, the best measure for the existence of a neuropathic pain component but the result is still open for discussion.

### Suggestions for the validation of neuropathic pain screening tools

Validation of screening tools should be performed in a standardized manner and described in detail, but performed in a setting which is comparable to a daily clinical practice. A research setting might be different from a clinical setting and thus might have influence on the patient and on the study results. The group of patients as well as the physicians under study should be comparable to the patients/physicians for who the tool is intended. Pre-stratification on the target outcome must be avoided (especially the exclusion of the so called mixed pain), because this will lead to a non-clinical situation and thus decreases the validity and generalizability of the instrument[16, 71].

### Conclusion

The validity of DN4-signs is equal to the DN4 outcome and, importantly, both are more valid than the DN4-symptoms alone. It seems that the patients' symptoms and signs doesn't reliably reflect the underlying mechanisms, indicating there is a need for a more objective way to assess patients' pain to facilitate improvement in the treatment of patients with pain. The physicians' assessment cannot be replaced by a screening tool as the DN4, but gives the physician a little hint towards the (non-)existence of neuropathic pain component.

### Supporting information

**S1 Table. Loading factors of the three versions of the DN4 according to the rotated component matrix factor analysis.**

(DOCX)

**S2 Table. The area under the curve and the sensitivity / specificity at the optimal cut-off point of the DN4 under the condition of equal costs of misclassification to classify a neuropathic pain component by the classification and the grading system of the physicians.**

PD-Q: PainDETECT questionnaire; Present NePC: Neuropathic pain component existing; Absent NePC: Neuropathic pain component not existing; AUC: Area under curve; Std.Error: Standard error; Asymp. Sig.: Asymptotic Significance; 95%CI: 95% confidence interval; Sens.: Sensitivity; Spec.: Specificity; +DV: Positive diagnostic value; -DV: Negative diagnostic value; +LR: Positive likelihood ratio; -LR: Negative likelihood ratio; P[Z+]: a-priori chance for the existence of a NePC; P[Z-]: a-priori chance for no existence of NePC; FPR: False positive ratio; FNR: False negative ratio; A: Physician A; B: Physician B; LBLP: Low back and leg pain; NSAP: Neck shoulder arm pain; PSNO: Pain of suspected neuropathic origin.

(DOCX)

## Acknowledgments

Thanks to all the participating patients, physicians and assistants for their invaluable work to this study.

## Author Contributions

**Conceptualization:** Hans Timmerman, Monique A. H. Steegers, André P. Wolff, Kris C. P. Vissers.

**Data curation:** Hans Timmerman, Frank J. P. M. Huygen, Nick T. van Dasselaar, Marcel J. Schenkels, André P. Wolff, Kris C. P. Vissers.

**Formal analysis:** Hans Timmerman, Jelle J. Goeman, Oliver H. G. Wilder-Smith, André P. Wolff, Kris C. P. Vissers.

**Funding acquisition:** André P. Wolff, Kris C. P. Vissers.

**Investigation:** Frank J. P. M. Huygen, Nick T. van Dasselaar, Marcel J. Schenkels, André P. Wolff, Kris C. P. Vissers.

**Methodology:** Hans Timmerman, Monique A. H. Steegers, Jelle J. Goeman, Oliver H. G. Wilder-Smith, André P. Wolff, Kris C. P. Vissers.

**Project administration:** Hans Timmerman, Nick T. van Dasselaar, André P. Wolff.

**Resources:** André P. Wolff, Kris C. P. Vissers.

**Software:** Kris C. P. Vissers.

**Supervision:** Hans Timmerman, Nick T. van Dasselaar, Marcel J. Schenkels, Oliver H. G. Wilder-Smith, André P. Wolff, Kris C. P. Vissers.

**Validation:** Hans Timmerman, Monique A. H. Steegers, Jelle J. Goeman, Oliver H. G. Wilder-Smith, André P. Wolff, Kris C. P. Vissers.

**Visualization:** Jelle J. Goeman.

**Writing – original draft:** Hans Timmerman, Monique A. H. Steegers, Frank J. P. M. Huygen, Jelle J. Goeman, Nick T. van Dasselaar, Marcel J. Schenkels, Oliver H. G. Wilder-Smith, André P. Wolff, Kris C. P. Vissers.

**Writing – review & editing:** Hans Timmerman, Monique A. H. Steegers, Frank J. P. M. Huygen, André P. Wolff, Kris C. P. Vissers.

## References

1. IASP. IASP Taxonomy Neuropathic Pain: International Association for the Study of Pain; 2015 [cited 2015 May 19, 2015]. <http://www.iasp-pain.org/Taxonomy#Neuropathicpain>.
2. 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(10):1911–20. Epub 2006/10/07. <https://doi.org/10.1185/030079906X132488> PMID: 17022849.
3. La Cesa S, Tamburin S, Tugnoli V, Sandrini G, Paolucci S, Lacerenza M, et al. How to diagnose neuropathic pain? The contribution from clinical examination, pain questionnaires and diagnostic tests. *Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*. 2015; 36(12):2169–75. <https://doi.org/10.1007/s10072-015-2382-z> PMID: 26410087.
4. Attal N, Perrot S, Fermanian J, Bouhassira D. The neuropathic components of chronic low back pain: a prospective multicenter study using the DN4 Questionnaire. *The journal of pain: official journal of the American Pain Society*. 2011; 12(10):1080–7. Epub 2011/07/26. <https://doi.org/10.1016/j.jpain.2011.05.006> PMID: 21783428.
5. Baron R, Binder A, Attal N, Casale R, Dickenson AH, Treede RD. Neuropathic low back pain in clinical practice. *European journal of pain*. 2016; 20(6):861–73. <https://doi.org/10.1002/ejp.838> PMID: 26935254.
6. Freynhagen R, Baron R. The evaluation of neuropathic components in low back pain. *Curr Pain Headache Rep*. 2009; 13(3):185–90. Epub 2009/05/22. PMID: 19457278.
7. Haanpaa M, Attal N, Backonja M, Baron R, Bennett M, Bouhassira D, et al. NeuPSIG guidelines on neuropathic pain assessment. *Pain*. 2011; 152(1):14–27. Epub 2010/09/21. <https://doi.org/10.1016/j.pain.2010.07.031> PMID: 20851519.
8. Haanpaa ML, Backonja MM, Bennett MI, Bouhassira D, Cruccu G, Hansson PT, et al. Assessment of neuropathic pain in primary care. *Am J Med*. 2009; 122(10 Suppl):S13–21. Epub 2009/10/07. <https://doi.org/10.1016/j.amjmed.2009.04.006> PMID: 19801048.
9. Truini A, Cruccu G. How diagnostic tests help to disentangle the mechanisms underlying neuropathic pain symptoms in painful neuropathies. *Pain*. 2016; 157 Suppl 1:S53–9. <https://doi.org/10.1097/j.pain.0000000000000367> PMID: 26785156.
10. Dworkin RH, O'Connor AB, Kent J, Mackey SC, Raja SN, Stacey BR, et al. Interventional management of neuropathic pain: NeuPSIG recommendations. *Pain*. 2013; 154(11):2249–61. Epub 2013/06/12. <https://doi.org/10.1016/j.pain.2013.06.004> PMID: 23748119.
11. Bennett MI, Attal N, Backonja MM, Baron R, Bouhassira D, Freynhagen R, et al. Using screening tools to identify neuropathic pain. *Pain*. 2007; 127(3):199–203. Epub 2006/12/22. <https://doi.org/10.1016/j.pain.2006.10.034> PMID: 17182186.
12. Bouhassira D, Attal N. Diagnosis and assessment of neuropathic pain: the saga of clinical tools. *Pain*. 2011; 152(3 Suppl):S74–83. <https://doi.org/10.1016/j.pain.2010.11.027> PMID: 21185120.
13. Andermann A, Blancquaert I, Beauchamp S, Dery V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. *Bull World Health Organ*. 2008; 86(4):317–9. <https://doi.org/10.2471/BLT.07.050112> PMID: 18438522
14. Wilson JM, Jungner YG. [Principles and practice of mass screening for disease]. *Bol Oficina Sanit Panam*. 1968; 65(4):281–393. PMID: 4234760.
15. Dieleman JP, Kerklaan J, Huygen FJ, Bouma PA, Sturkenboom MC. Incidence rates and treatment of neuropathic pain conditions in the general population. *Pain*. 2008; 137(3):681–8. Epub 2008/04/29. <https://doi.org/10.1016/j.pain.2008.03.002> PMID: 18439759.
16. Mathieson S, Maher CG, Terwee CB, Folly de Campos T, Lin CW. Neuropathic pain screening questionnaires have limited measurement properties. A systematic review. *Journal of clinical epidemiology*. 2015; 68(8):957–66. <https://doi.org/10.1016/j.jclinepi.2015.03.010> PMID: 25895961.
17. Bouhassira D, Attal N, Alchaar H, Boureau F, Brochet B, Bruxelle J, et al. Comparison of pain syndromes associated with nervous or somatic lesions and development of a new neuropathic pain diagnostic questionnaire (DN4). *Pain*. 2005; 114(1–2):29–36. Epub 2005/03/01. PMID: 15733628.
18. CADTH. Diagnostic methods for neuropathic pain: A review of diagnostic accuracy. Canadian Agency for Drugs and Technologies in Health [Internet]. 2015 10-28-2015 [cited 2015]. [http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0078647/pdf/PubMedHealth\\_PMH0078647.pdf](http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0078647/pdf/PubMedHealth_PMH0078647.pdf).
19. Proqolid. neuropathic pain 4 questions: MAPI Research trust; 2015 [cited 2015]. [http://www.proqolid.org/instruments/neuropathic\\_pain\\_4\\_questions\\_dn](http://www.proqolid.org/instruments/neuropathic_pain_4_questions_dn).
20. Van Seventer R, Vos C, Meerding W, Mear I, Le Gal M, Bouhassira D, et al. Linguistic validation of the DN4 for use in international studies. *European journal of pain*. 2010; 14(1):58–63. Epub 2009/03/14. <https://doi.org/10.1016/j.ejpain.2009.01.005> PMID: 19282208.

21. Perez C, Galvez R, Huelbes S, Insausti J, Bouhassira D, Diaz S, et al. Validity and reliability of the Spanish version of the DN4 (Douleur Neuropathique 4 questions) questionnaire for differential diagnosis of pain syndromes associated to a neuropathic or somatic component. *Health Qual Life Outcomes*. 2007; 5:66. <https://doi.org/10.1186/1477-7525-5-66> PMID: 18053212
22. Harifi G, Ouilki I, El Bouchti I, Ouazar MA, Belkhou A, Younsi R, et al. Validity and reliability of the Arabic adapted version of the DN4 questionnaire (Douleur Neuropathique 4 Questions) for differential diagnosis of pain syndromes with a neuropathic or somatic component. *Pain practice: the official journal of World Institute of Pain*. 2011; 11(2):139–47. <https://doi.org/10.1111/j.1533-2500.2010.00399.x> PMID: 20602717.
23. Unal-Cevik I, Sarioglu-Ay S, Evcik D. A comparison of the DN4 and LANSS questionnaires in the assessment of neuropathic pain: validity and reliability of the Turkish version of DN4. *The journal of pain: official journal of the American Pain Society*. 2010; 11(11):1129–35. <https://doi.org/10.1016/j.jpain.2010.02.003> PMID: 20418179.
24. Santos JG, Brito JO, de Andrade DC, Kaziyama VM, Ferreira KA, Souza I, et al. Translation to Portuguese and validation of the Douleur Neuropathique 4 questionnaire. *The journal of pain: official journal of the American Pain Society*. 2010; 11(5):484–90. <https://doi.org/10.1016/j.jpain.2009.09.014> PMID: 20015708.
25. van Seventer R, Vos C, Giezeman M, Meerding WJ, Arnould B, Regnault A, et al. Validation of the Dutch version of the DN4 diagnostic questionnaire for neuropathic pain. *Pain practice: the official journal of World Institute of Pain*. 2013; 13(5):390–8. <https://doi.org/10.1111/papr.12006> PMID: 23113981.
26. Madani SP, Fateh HR, Forogh B, Fereshtehnejad SM, Ahadi T, Ghaboussi P, et al. Validity and reliability of the persian (Farsi) version of the DN4 (Douleur Neuropathique 4 Questions) questionnaire for differential diagnosis of neuropathic from non-neuropathic pains. *Pain practice: the official journal of World Institute of Pain*. 2014; 14(5):427–36. <https://doi.org/10.1111/papr.12088> PMID: 23763722.
27. Sykioti P, Zis P, Vadalouca A, Siafaka I, Argyra E, Bouhassira D, et al. Validation of the Greek Version of the DN4 Diagnostic Questionnaire for Neuropathic Pain. *Pain practice: the official journal of World Institute of Pain*. 2014. <https://doi.org/10.1111/papr.12221> PMID: 24796220.
28. Hamdan A, Luna JD, Del Pozo E, Galvez R. Diagnostic accuracy of two questionnaires for the detection of neuropathic pain in the Spanish population. *European journal of pain*. 2014; 18(1):101–9. <https://doi.org/10.1002/j.1532-2149.2013.00350.x> PMID: 23776139.
29. Chaudakshetrin P, Prateepavanich P, Chira-Adisai W, Tassanawipras W, Leechavengvongs S, Kitisomprayoonkul W. Cross-cultural adaptation to the Thai language of the neuropathic pain diagnostic questionnaire (DN4). *J Med Assoc Thai*. 2007; 90(9):1860–5. PMID: 17957931.
30. Kim HJ, Park JH, Bouhassira D, Shin JH, Chang BS, Lee CK, et al. Validation of the Korean Version of the DN4 Diagnostic Questionnaire for Neuropathic Pain in Patients with Lumbar or Lumbar-Radicular Pain. *Yonsei Med J*. 2016; 57(2):449–54. <https://doi.org/10.3349/ymj.2016.57.2.449> PMID: 26847299
31. Chatila N, Pereira B, Maarrawi J, Dallel R. Validation of a New Arabic Version of the Neuropathic Pain Diagnostic Questionnaire (DN4). *Pain practice: the official journal of World Institute of Pain*. 2016. <https://doi.org/10.1111/papr.12419> PMID: 26895970.
32. Treede RD, Jensen TS, Campbell JN, Cruccu G, Dostrovsky JO, Griffin JW, et al. Neuropathic pain: redefinition and a grading system for clinical and research purposes. *Neurology*. 2008; 70(18):1630–5. Epub 2007/11/16. <https://doi.org/10.1212/01.wnl.0000282763.29778.59> PMID: 18003941.
33. Finnerup NB, Haroutounian S, Kamerman P, Baron R, Bennett DL, Bouhassira D, et al. Neuropathic pain: an updated grading system for research and clinical practice. *Pain*. 2016. PMID: 27115670.
34. Geber C, Baumgartner U, Schwab R, Muller H, Stoeter P, Dieterich M, et al. Revised definition of neuropathic pain and its grading system: an open case series illustrating its use in clinical practice. *Am J Med*. 2009; 122(10 Suppl):S3–12. <https://doi.org/10.1016/j.amjmed.2009.04.005> PMID: 19801050.
35. Mulvey MR, Rolke R, Klepstad P, Caraceni A, Fallon M, Colvin L, et al. Confirming neuropathic pain in cancer patients: applying the NeuPSIG grading system in clinical practice and clinical research. *Pain*. 2014; 155(5):859–63. <https://doi.org/10.1016/j.pain.2013.11.010> PMID: 24275256.
36. Abdallah FW, Morgan PJ, Cil T, Escallon JM, Semple JL, Chan VW. Comparing the DN4 tool with the IASP grading system for chronic neuropathic pain screening after breast tumor resection with and without paravertebral blocks: a prospective 6-month validation study. *Pain*. 2015; 156(4):740–9. PMID: 25719620.
37. Guastella V, Mick G, Soriano C, Vallet L, Escande G, Dubray C, et al. A prospective study of neuropathic pain induced by thoracotomy: incidence, clinical description, and diagnosis. *Pain*. 2011; 152(1):74–81. <https://doi.org/10.1016/j.pain.2010.09.004> PMID: 21075523.
38. Timmerman H, Heemstra I, Schalkwijk A, Verhagen C, Vissers K, Engels Y. Assessment of Neuropathic Pain in Patients with Cancer: The Interobserver Reliability. *An Observational Study in Daily Practice*. *Pain physician*. 2013;( 16):11.

39. Timmerman H, Wilder-Smith O, van Weel C, Wolff A, Vissers K. Detecting the neuropathic pain component in the clinical setting: a study protocol for validation of screening instruments for the presence of a neuropathic pain component. *BMC Neurol*. 2014; 14(1):94. Epub 2014/06/03. <https://doi.org/10.1186/1471-2377-14-94> PMID: 24885108
40. Lavand'homme P, Thienpont E. Pain after total knee arthroplasty: a narrative review focusing on the stratification of patients at risk for persistent pain. *Bone Joint J*. 2015; 97-B(10 Suppl A):45–8. <https://doi.org/10.1302/0301-620X.97B10.36524> PMID: 26430086.
41. Cruccu G, Anand P, Attal N, Garcia-Larrea L, Haanpaa M, Jorum E, et al. EFNS guidelines on neuropathic pain assessment. *European journal of neurology*. 2004; 11(3):153–62. PMID: 15009162.
42. Cruccu G, Sommer C, Anand P, Attal N, Baron R, Garcia-Larrea L, et al. EFNS guidelines on neuropathic pain assessment: revised 2009. *European journal of neurology*. 2010; 17(8):1010–8. Epub 2010/03/20. <https://doi.org/10.1111/j.1468-1331.2010.02969.x> PMID: 20298428.
43. Tampin B, Briffa NK, Goucke R, Slater H. Identification of neuropathic pain in patients with neck/upper limb pain: application of a grading system and screening tools. *Pain*. 2013; 154(12):2813–22. Epub 2013/08/27. <https://doi.org/10.1016/j.pain.2013.08.018> PMID: 23973362.
44. Vaegter HB, Andersen PG, Madsen MF, Handberg G, Enggaard TP. Prevalence of neuropathic pain according to the IASP grading system in patients with chronic non-malignant pain. *Pain Med*. 2014; 15(1):120–7. <https://doi.org/10.1111/pme.12273> PMID: 24165161.
45. Collins SL, Edwards J, Moore RA, Smith LA, McQuay HJ. Seeking a simple measure of analgesia for mega-trials: is a single global assessment good enough? *Pain*. 2001; 91(1–2):189–94. Epub 2001/03/10. PMID: 11240091.
46. Farrar JT, Young JP Jr., LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001; 94(2):149–58. Epub 2001/11/03. PMID: 11690728.
47. Fischer D, Stewart AL, Bloch DA, Lorig K, Laurent D, Holman H. Capturing the patient's view of change as a clinical outcome measure. *Jama*. 1999; 282(12):1157–62. Epub 1999/09/29. PMID: 10501119.
48. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159–74. Epub 1977/03/01. PMID: 843571.
49. El Khouli RH, Macura KJ, Barker PB, Habba MR, Jacobs MA, Bluemke DA. Relationship of temporal resolution to diagnostic performance for dynamic contrast enhanced MRI of the breast. *J Magn Reson Imaging*. 2009; 30(5):999–1004. <https://doi.org/10.1002/jmri.21947> PMID: 19856413
50. Obuchowski NA. Receiver operating characteristic curves and their use in radiology. *Radiology*. 2003; 229(1):3–8. <https://doi.org/10.1148/radiol.2291010898> PMID: 14519861.
51. Ludemann L, Grieger W, Wurm R, Wust P, Zimmer C. Glioma assessment using quantitative blood volume maps generated by T1-weighted dynamic contrast-enhanced magnetic resonance imaging: a receiver operating characteristic study. *Acta Radiol*. 2006; 47(3):303–10. PMID: 16613313.
52. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med*. 1978; 8(4):283–98. PMID: 112681.
53. Greenhalgh T. How to read a paper: Papers that report diagnostic or screening tests (vol 315, pg 540, 1997). *British Medical Journal*. 1997; 315(7113):942–.
54. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*. 1994; 6(4):6.
55. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American journal of mental deficiency*. 1981; 86(2):127–37. PMID: 7315877.
56. Versi E. "Gold standard" is an appropriate term. *British Medical Journal*. 1992; 305:187.
57. Versi E. Discriminant analysis of urethral pressure profilometry data for the diagnosis of genuine stress incontinence. *British journal of obstetrics and gynaecology*. 1990; 97(3):251–9. PMID: 2334653.
58. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. *American journal of health-system pharmacy: AJHP: official journal of the American Society of Health-System Pharmacists*. 2008; 65(23):2276–84. <https://doi.org/10.2146/ajhp070364> PMID: 19020196.
59. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. *Bmj*. 1994; 308(6943):1552. PMID: 8019315
60. Altman DG, Bland JM. Diagnostic tests 2: Predictive values. *Bmj*. 1994; 309(6947):102. PMID: 8038641
61. Grimes DA, Schulz KF. Uses and abuses of screening tests. *Lancet*. 2002; 359(9309):881–4. [https://doi.org/10.1016/S0140-6736\(02\)07948-5](https://doi.org/10.1016/S0140-6736(02)07948-5) PMID: 11897304.

62. Themistocleous AC, Ramirez JD, Shillo PR, Lees JG, Selvarajah D, Orengo C, et al. The Pain in Neuropathy Study (PiNS): a cross-sectional observational study determining the somatosensory phenotype of painful and painless diabetic neuropathy. *Pain*. 2016; 157(5):1132–45. PMID: [27088890](#)
63. Spallone V, Morganti R, D'Amato C, Greco C, Cacciotti L, Marfia GA. Validation of DN4 as a screening tool for neuropathic pain in painful diabetic polyneuropathy. *Diabet Med*. 2012; 29(5):578–85. <https://doi.org/10.1111/j.1464-5491.2011.03500.x> PMID: [22023377](#).
64. Hallstrom H, Norrbrink C. Screening tools for neuropathic pain: can they be of use in individuals with spinal cord injury? *Pain*. 2011; 152(4):772–9. Epub 2011/01/29. <https://doi.org/10.1016/j.pain.2010.11.019> PMID: [21272997](#).
65. Lasry-Levy E, Hietaharju A, Pai V, Ganapati R, Rice AS, Haanpaa M, et al. Neuropathic pain and psychological morbidity in patients with treated leprosy: a cross-sectional prevalence study in Mumbai. *PLoS Negl Trop Dis*. 2011; 5(3):e981. <https://doi.org/10.1371/journal.pntd.0000981> PMID: [21408111](#)
66. Haroun OM, Hietaharju A, Bizuneh E, Tesfaye F, Brandsma JW, Haanpaa M, et al. Investigation of neuropathic pain in treated leprosy patients in Ethiopia: a cross-sectional study. *Pain*. 2012; 153(8):1620–4. <https://doi.org/10.1016/j.pain.2012.04.007> PMID: [22727538](#).
67. Markman JD, Kress BT, Frazer M, Hanson R, Kogan V, Huang JH. Screening for neuropathic characteristics in failed back surgery syndromes: challenges for guiding treatment. *Pain Med*. 2015; 16(3):520–30. <https://doi.org/10.1111/pme.12612> PMID: [25530081](#).
68. Geber C, Breimhorst M, Burbach B, Egenolf C, Baier B, Fechir M, et al. Pain in chemotherapy-induced neuropathy—more than neuropathic? *Pain*. 2013; 154(12):2877–87. PMID: [23999056](#).
69. Sadler A, Wilson J, Colvin L. Acute and chronic neuropathic pain in the hospital setting: use of screening tools. *The Clinical journal of pain*. 2013; 29(6):507–11. PMID: [23247003](#).
70. VanDenKerkhof EG, Mann EG, Torrance N, Smith BH, Johnson A, Gilron I. An Epidemiological Study of Neuropathic Pain Symptoms in Canadian Adults. *Pain research & management*. 2016; 2016:9815750. <https://doi.org/10.1155/2016/9815750> PMID: [27445636](#)
71. Deeks JJ. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. *Bmj*. 2001; 323(7305):157–62. PMID: [11463691](#)