Chapter 5

Diabetic Neuropathy Examination:
a hierarchical scoring system to diagnose
distal polyneuropathy in diabetes

J.W.G. Meijer, E. van Sonderen, E.E. Blaauwwiekel,

*Diabetes Care* 2000; 23 (6): 750-753.
Abstract

Objective Existing physical examination scoring systems for distal diabetic polyneuropathy (PNP) do not fulfil all of the following criteria: validity, manageability, predictive value, and hierarchy. The aim of this study was to adapt the Neuropathy Disability Score (NDS) to diagnose PNP in diabetes mellitus (DM) so that it fulfils these criteria.

Methods A total of 73 patients with DM were examined with the NDS. Monofilaments and biothesiometry were used as clinical standards for PNP to modify the NDS.

Results A total of 43 men and 30 women were studied; the mean duration of DM was 15 years (1-43), and the mean age was 57 years (19-90). Twenty-four patients had DM type 1 and 49 had type 2 DM. Clinically relevant items were selected from the original 35 NDS items (specific item scored positive score in > 3 patients). The resulting 8-item Diabetic Neuropathy Examination score (DNE) could accurately predict the results of the clinical standards and is strongly hierarchical (H-value 0.53). The sensitivity and specificity of the DNE at a cut-off level of 3 to 4 were 0.96 and 0.51 for abnormal monofilament scores, respectively. For abnormal biothesiometry scores, these values were 0.97 and 0.59, respectively. Reproducibility, as assessed by inter- and intrarater agreement, was good.

Conclusions The DNE is a sensitive and well-validated hierarchic scoring system that is fast and easy to perform in clinical practice.
5.1 Introduction

Early detection of symmetric distal sensori-motor polyneuropathy (PNP) is important in patients with diabetes mellitus (DM), because preventive interventions can be applied to decrease morbidity. Unfortunately, no "gold standard" exists for diagnosing PNP, but a consensus panel has recommended that at least 1 measurement should be performed in 5 different diagnostic categories. One of these categories is a standardised physical examination. In our opinion, diagnostic tests should fulfil the following criteria: validation (presence of independent reference standard, adequate spectrum and number of patients, standardisation, soundly based item selection), predictive value, manageability (reproducibility, performance in clinical practice) and hierarchy. Frequently used and accepted examination scores for diabetic neuropathy are the Neuropathy Disability Score (NDS), the Neuropathy Impairment Score in the Lower Limbs (NIS-LL), various modified NDS scores, the Neuropathy Deficit Score, the Michigan Neuropathy Screening Instrument (MNSI) and the Clinical Examination score of Valk (CE-V).

The NDS was designed for neuropathy in general. Although the score is well founded and complete, it is difficult to perform in clinical practice on patients with diabetic foot problems. Precise descriptions of how the tests should be performed and how items should be scored are lacking. The NIS-LL is a modification of the NDS specific for distal PNP, although motor activity grading is the focus and involves 64 of a maximum of 88 points. The NIS-LL has not been validated. Various other modified NDS scoring systems have been used, such as those of Veves et al. and Young et al. However, these instruments also have not been validated and no information is available on their predictive value regarding the results of clinical standards. The Neuropathy Deficit Score is a neurological examination score aimed at anatomical levels in the legs and arms. It has not been validated and no information is available about how to interpret modifications, which is also the case for the other modified NDS scoring systems. Feldman et al. developed a combination of two scoring systems: the Michigan Neuropathy Screening Instrument (symptom and examination score) and the Michigan Diabetic Neuropathy Score (neurological examination and nerve conduction studies). These scores do not have a separate examination score, as advised by consensus reports. The CE-V can be used to examine sensory functions, tendon reflexes and muscle strength in the lower extremities. The scoring systems of Feldman et al. and Valk et al. have been validated and are easy to perform in clinical practice. None of the afore mentioned scores is known to be hierarchical.
The aim of this study was to adapt the NDS into a valid, easily managed, graded and accurate scoring system for diagnosing PNP, the Diabetic Neuropathy Examination (DNE) score.

5.2 Research Design and Methods

Patients:
Our study group consisted of 73 patients with DM. Exclusion criteria were factors that may interfere with the neurological condition of the subjects other than PNP. Fifty patients were randomly selected from the diabetes outpatient clinic of the University Hospital Groningen. Twenty-three positive control patients with obvious diabetic foot complications or clinical neuropathy were selected from the Department of Diabetes of the Rehabilitation Centre Beatrixoord. The characteristics of these 73 patients are shown in Table 1.

Table 1: Patient characteristics

<table>
<thead>
<tr>
<th>N</th>
<th>73</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (years)(SD)</td>
<td>56.9 (16.1)</td>
</tr>
<tr>
<td>Min – max (years)</td>
<td>19 – 90</td>
</tr>
<tr>
<td>Mean duration DM (years) (SD)</td>
<td>14.9 (9.9)</td>
</tr>
<tr>
<td>Min – max (years)</td>
<td>1 – 43</td>
</tr>
<tr>
<td>Sex</td>
<td>male – female</td>
</tr>
<tr>
<td></td>
<td>43 – 30</td>
</tr>
<tr>
<td>Type DM</td>
<td>1 - 2</td>
</tr>
<tr>
<td></td>
<td>24 – 49</td>
</tr>
<tr>
<td>Mean HbA1c (%) (SD)</td>
<td>8.7 (1.4)</td>
</tr>
<tr>
<td>Min – max</td>
<td>6.6 – 13.5</td>
</tr>
</tbody>
</table>

Methods:
The same researcher (JWGM) examined all 73 patients. First, the NDS and NIS-LL were performed, followed by quantitative sensory tests that acted as a clinical standard.

1 NDS and NIS-LL
The NDS is the most widely used and widely accepted scoring system for diabetic neuropathy; it has also been recommended in consensus reports. The instrument examines cranial nerves, muscle weakness, reflexes and sensation. The scale consists of 35 items for testing the left and right sides of
the body; scores range from 0 to 4. A sum score is obtained with a maximum of 280 points. The NIS-LL is a modified version of the NDS to quantify diabetic PNP. The lower limb items of the NDS are used complemented with two muscle power items (toe extension and toe flexion). The NIS-LL has 14 items: 8 items evaluate muscle power (0-4 points), 2 items evaluate reflexes (0-2 points), and 4 items evaluate sensory modalities (0-2 points). All items are tested on both sides. The maximum score is 88 points. The NDS, as the most complete and accepted score, was used for item selection to develop the DNE-score.

2 Clinical Standards
Semmes Weinstein Monofilaments (SWMF) and Vibration Perception Threshold (VPT) were chosen as clinical standards to study the construct validity of the scoring system for PNP. SWMF were tested on the plantar surface of the hallux and centrally at the heel (when necessary after removal of excessive calluses). This method was standardised according to generally accepted guidelines. The "yes-no" method was used, which means that the patient says yes each time he or she senses the application of a monofilament. Six trials were administered, when the patient was unable to respond correctly in more than 1 trial, a heavier monofilament was used. The 1, 10 and 75 gram monofilaments have been used. We present the results in four categories: category 1: 1 gram monofilament felt; category 2: 10 gram monofilament felt, 1 gram monofilament not felt; category 3: 75 gram monofilament felt, 10 gram monofilament not felt; category 4: 75 gram monofilament not felt.

Vibration Perception Thresholds (VPTs) were determined using a hand-held biothesiometer (Biomedical Instruments Inc., Ohio, USA). VPT was tested at the dorsum of the hallux on the interphalangeal joint. It was performed in a standardised way. The voltage of vibration was increased until the patient could perceive a vibration. This was done three times. The mean of these three trials was used to determine the VPT.

Reproducibility
To test reproducibility, inter- and intrarater agreement were assessed in a separate study of 10 patients. The 6 women and 4 men, with a mean age of 50.0 years (SD 15.9) had a wide range of neuropathy severity. The mean duration of DM was 11.5 years (SD 10.5); 3 participants had type 1 DM and 7 participants had type 2 DM. Two experienced physicians, an endocrinologist (EEB) and a physician for rehabilitation medicine (JWGM), both experienced in diagnosing diabetic neuropathies, rated these patients twice within one week.
Statistical Analyses

Internal consistency of the DNE-score was assessed by calculating Cronbach's alpha, and reliability coefficient Rho, which is comparable to alpha. In addition to internal consistency, scalability coefficient H was computed with the probabilistic scaling programme MSP (Mokken Scaling Polychotomous items) to assess the hierarchical structure of the items. High values of H increase the likelihood that patients with the same scale score have difficulties or problems with the same items.

The statistical package SPSS-PC (Chicago) was used to compute the descriptive statistics, factor analysis, reliability coefficient Cronbach’s alpha, Pearson's correlation coefficient r and Student's t-test. Inter- and intrarater agreement were assessed on a scale level by computing Pearson's correlation coefficients and t-test values for differences in means.

5.3 Results

Items were excluded from the original NDS if they conformed to the following definition of clinical irrelevance: specific item scored positive in 3 patients or less. After examining the patients, 9 of the original 35 items remained. No relevant differences were found between the measurements made on the left and right side, so only the right-side items were used in the analyses.

Factor analysis was performed on the 9 items to investigate coherence. The coherence of the 8 items was good; only item 22 (muscle strength triceps surae) had poor coherence compared to the other items.

Calculation of hierarchy was performed using the MSP items. This resulted in a hierarchical scale of 8 items. Item 22 disturbed the hierarchy severely.

Logistic regression analysis was performed to study whether item 22, in addition to the 8-item hierarchical scale, could predict the results of the clinical standards SWMF and VPT. Item 22 did not make any significant contribution, so it was excluded.

Modification of the NDS resulted in an 8-item scale: the Diabetic Neuropathy Examination (DNE) score. The DNE-score is shown in Appendix 1.

Reliability of the scale was assessed by measuring the internal consistency. According to both Cronbach's alpha (.78) and reliability coefficient Rho (.81) the scale appears to be reliable. The H-value for hierarchy was 0.53, which indicates the presence of a strong hierarchical scale.

Table 2 shows the characteristics of the DNE-score, NDS and NIS-LL. As expected, the correlation between the DNE-score and respectively the NDS (Pearson's r .96, p<.001) and NISS-LL (Pearson's r .92, p<.001) were both high. The reliability of the scoring systems was good.
The DNE is fast and easy to perform in clinical practice; application takes about 5 minutes.

Table 2: Characteristics of the NDS, NIS-LL and the DNE-score in our study population

<table>
<thead>
<tr>
<th></th>
<th>NDS</th>
<th>NIS-LL</th>
<th>DNE-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>n= 73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean score (SD)</td>
<td>19.7 (14.5)</td>
<td>9.7 (7.9)</td>
<td>5.0 (3.6)</td>
</tr>
<tr>
<td>Reliability (alpha)</td>
<td>.88</td>
<td>.87</td>
<td>.78</td>
</tr>
<tr>
<td>Number of items</td>
<td>70</td>
<td>28</td>
<td>8</td>
</tr>
<tr>
<td>Maximum score</td>
<td>280</td>
<td>88</td>
<td>16</td>
</tr>
<tr>
<td>Maximum scored</td>
<td>56</td>
<td>32</td>
<td>13</td>
</tr>
<tr>
<td>Items not scored</td>
<td>44</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Less than three scores</td>
<td>8</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

Relationship of the NDS, NIS-LL and DNE-score with the clinical standards

Pearson's correlation coefficient r of SWMF with NDS, NIS-LL and DNE-score was similar with values of .76 (p<.001), .74 (p<.001) and .75 (p<.001), respectively. Pearson's correlation coefficient r for VPT with NDS, NIS-LL and DNE-score was similar with values of .73 (p<.001), .71 (p<.001) and .75 (p<.001), respectively. The NDS, NISS-LL and DNE-score predicted the results of the clinical standards very accurately (p < .001).

At a cutoff point of 3 to 4, sensitivity and specificity of the DNE-score were 0.96 and 0.51, respectively, for an abnormal result using SWMF. For an abnormal result using the VPT, these values were 0.97 and 0.59, respectively.

Reproducibility

Comparing the scores of 2 raters obtained on 2 occasions (interval 1 week) assessed reproducibility of the DNE-score. The interrater correlation was .97 at t1 and .92 at t2, respectively. Differences in mean scores were less than 10% and not significant (p=.08 and .55, respectively). The intrarater correlation was .89 for one rater and .99 for the other. The mean scores of the two raters did not differ significantly at t1 and t2 (p=.17 and .60, respectively).
5.4 Discussion

The NDS is a widely accepted and validated physical examination scoring system used to diagnose neuropathy. Its predictive value and reproducibility are high. It is well correlated with neurophysiological and sural nerve morphometric abnormalities in patients with diabetes mellitus. Because the aim of the NDS is to evaluate neuropathy in general, it is not completely suitable for use at an outpatient diabetic foot clinic. Consequently, several other scoring systems have been developed, but they do not sufficiently fulfil all of the criteria necessary for adequate diagnostic tests. One of these is NIS-LL, a score for distal diabetic polyneuropathy with 14 items. The score has not been validated, and focuses more on motor problems than on sensory problems.

In this study, the NDS was modified once again with the aim of achieving a new physical examination scoring system for diagnosing distal symmetric polyneuropathy in diabetes mellitus. The new instrument is the DNE-score, a scoring system with 8 items. It was validated in DM patients with a wide spectrum of complications. The DNE-score is hierarchical, sensitive, fast, and easy to perform in clinical practice (application takes about 5 minutes). Hierarchy implies that patients with the same scale score have difficulties or problems with the same items, which makes this scoring system able to differentiate between severity levels of PNP and to compare groups or individuals over time. The NDS, NIS-LL and the other instruments for evaluating PNP have not been documented to represent a hierarchical scale.

Our modifications were validated with SWMF measurements and VPTs. These are both semiquantitative, reliable measurements with proven predictive value for the development of clinical problems, such as foot ulcers and amputations. They are non-invasive, patient-friendly, independent and complementary. SWMF and VPT only assess large fiber function, no small fiber tests have been used in this study. Testing the DNE-score on a random sample from the outpatient clinic in addition to a set of patients with definite neuropathy means that the results are generalisable to the complete range of patients with DM.

Many clinicians prefer using electro-diagnostic techniques to diagnose diabetic PNP. Although neurophysiological examination is sensitive, specific and reproducible regarding the presence and severity of peripheral nerve involvement in patients with diabetes, it is not suitable for making a quick preliminary diagnosis at a diabetes outpatient clinic. No data are available on the predictive value of these techniques in relation to the development of clinical problems, such as diabetic foot disease.
Because the aim of this study was to develop a screening instrument as a tool in the detection and prevention of patients at risk for diabetic foot complications, the observed sensitivity and specificity of the DNE-score are satisfactory. Because sensitivity is of greater importance than specificity for screening instruments, the chosen cut off value results in the desired high sensitivity with an acceptable specificity. A low specificity might burden prevention education programmes. The combined use of different diagnostic tools, as advised in consensus reports, will enhance specificity.

The selection of the item muscle strength of the quadriceps femoris in the DNE-score is surprising and suggests the presence of mononeuropathy. Nevertheless, all patients with quadriceps dysfunction also showed other abnormalities regarding sensation in the feet, that were not related to the same peripheral nerves, which makes mononeuropathy less probable. The ankle dorsiflexion item was excluded because of poor coherence and disturbance of hierarchy. It did not contribute to the 8 definite items. Perhaps this discrepancy in muscle strength and its assessment is because of other factors, such as limited joint mobility.

The results of validation and the predictive value of the NDS, NIS-LL and DNE-score were very satisfactory. The strengths of the DNE-score are its manageability in clinical practice and its hierarchy. The DNE-score is the most efficient according to the criteria shown in Table 2.

In conclusion, the DNE-score as modified from the NDS is fast and easy to perform, hierarchical, and sensitive for PNP, and patient scores are more differentiated.
5.5 References


<table>
<thead>
<tr>
<th><strong>DNE-score</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>muscle strength</strong></td>
</tr>
<tr>
<td>1 quadriceps femoris: extension of the knee</td>
</tr>
<tr>
<td>2 tibialis anterior: dorsiflexion of the foot</td>
</tr>
<tr>
<td><strong>reflex</strong></td>
</tr>
<tr>
<td>3 triceps surae</td>
</tr>
<tr>
<td><strong>sensation index finger</strong></td>
</tr>
<tr>
<td>4 sensitivity to pin pricks</td>
</tr>
<tr>
<td><strong>sensation big toe</strong></td>
</tr>
<tr>
<td>5 sensitivity to pin pricks</td>
</tr>
<tr>
<td>6 sensitivity to touch</td>
</tr>
<tr>
<td>7 vibration perception</td>
</tr>
<tr>
<td>8 sensitivity to joint position</td>
</tr>
</tbody>
</table>

- only the right leg and foot are tested
- scoring from 0-2:
  - 0= normal
  - 1= mild/moderate deficit:
    - muscle strength: MRC 3-4
    - reflex: decreased, but present
    - sensation: decreased, but present
  - 2= severely disturbed/absent:
    - muscle strength: MRC 0-2
    - reflex: absent
    - sensation: absent
- maximum score: 16 points