Screening for Peripheral Neuropathies in Children With Diabetes: A Systematic Review

abstract

BACKGROUND AND OBJECTIVE: Although guidelines for the management of children with type 1 diabetes include recommendations to screen for diabetic peripheral neuropathies (DPN), the research into the diagnostic utility of screening methods has not been systematically reviewed. The goal of this study was to summarize the findings with regard to the diagnostic accuracy of the Semmes-Weinstein monofilament and the Rydel-Seiffer tuning fork in detecting DPN in children and adolescents compared with the gold standard nerve conduction studies.

METHODS: Based on a PubMed search (conducted on April 26, 2013) and secondary searching, we identified 72 articles for review. We included studies that: (1) assessed DPN with the gold standard nerve conduction studies; (2) used noninvasive screening for DPN (monofilament, tuning fork, or biothesiometer); and (3) were performed in the relevant population (children with diabetes). Five articles met these criteria. Study quality was assessed by using the revised Quality Assessment of Diagnostic Accuracy Studies criteria. Heterogeneous methods precluded a formal meta-analysis of effects.

RESULTS: Diagnostic accuracies were heterogeneous for the different screening methods. Sensitivities ranged from 1% to 19% for the tuning fork (3 studies); from 61% to 80% for the biothesiometer (2 studies); and from 19% to 73% for the monofilament (2 studies).

CONCLUSIONS: Data show extremely low diagnostic utility for standard screening methods (tuning fork and 10-g monofilament) but acceptable utilities for biothesiometry and finer (1 g) monofilaments. Data on the diagnostic utility should be used to inform national and international guidelines on diabetes management. Pediatrics 2014;133:e1324–e1330
Diabetic peripheral neuropathies (DPN) are major complications of insulin-dependent diabetes mellitus (IDDM). Studies using objective measures of nerve dysfunction (e.g., nerve conduction studies [NCS]) demonstrated abnormalities in 28% to 58% of children with diabetes.1–4 and categorized them as high risk for developing severe secondary complications such as foot ulcerations.5 These severe complications can only be averted when signs of the disease are detected early. The highest prevalence for DPN is found in children and adolescents with poor glycemic control and longer duration of diabetes.6–8 Accordingly, national and international guidelines suggest annual screening for DPN in children and adolescents.9–11 These guidelines differ slightly in the recommended methods (monofilament, biothesiometer, or tuning fork) and indications (from what age or diabetes duration onward). For example, the American Diabetes Association recommends annual screening by using a 10-g monofilament beginning in puberty.9 Given the high prevalence of diabetes in US children (which affects ∼215,000 children and adolescents aged <20 years),12 several hundred children with diabetes are tested every day in the United States alone. In contrast to the clinical importance and frequency at which these tests are conducted, very little is known about the diagnostic utility of the screening methods in children with IDDM.

In pediatric practices, screening is performed by using noninvasive screening methods because the gold standard for DPN diagnoses, the measurement of the nerve conduction velocity,13,14 is painful and not available to all pediatricians. The 2 most frequently used methods to screen for DPN are the Semmes-Weinstein monofilament and the Rydel-Seiffer tuning fork.15 Recent research into quantitative sensory testing has shown that interpretation of these tests needs to account for patient age because the normal ranges of these tests vary systematically according to age. For example, although a vibration detection threshold of 5 (x/8) is still in the range for normal vibration perception in adults aged >40 years,16 the same value is highly abnormal in children.17 These differences in general nerve functioning make it difficult to adopt the same methods and cutoff values developed for adults for children and adolescents, and they afford an independent evaluation of the screening methods in this group of patients.

We systematically reviewed the literature on the diagnostic utility of noninvasive screening tests for DPN in children and adolescents. We focused on both vibration perception–based and mechanical perception–based screenings and compared them with the gold standard NCS to assess their diagnostic utility.

METHODS
Criteria for Inclusion and Exclusion
The following criteria had to be met by the individual study to be included: (1) study patients were diagnosed clinically with diabetes; (2) included patients were aged <18 years; (3) ≥1 noninvasive screening tool for neuropathy was used; (4) screening results were compared with NCS; and (5) were written in English or German. We also excluded studies that were: (1) reviews of empirical studies; (2) case studies; and (3) written in languages other than German or English.

Search Strategy
The initial search on PubMed (conducted on April 26, 2013) was inspired by the strategy used in the systematic review of screening methods in adults.18 The search used 4 groups of key words. The first group defined the target condition (i.e., diabetic peripheral neuropathy), and related diseases were included. The second group defined the investigated population (i.e., children and adolescents). The third group defined the clinical condition (i.e., type 1 diabetes mellitus). The fourth group defined the diagnostic methods (i.e., tuning fork). The complete strategy is available in Supplemental Table 3. This search yielded 53 studies. Furthermore, we used forward (within the reference sections) and backward (via PubMed’s “cited-by” feature) secondary search strategies to identify additional articles. This method was necessary because several studies that met inclusion criteria were conducted for a purpose other than to assess the diagnostic accuracy of screening methods for DPN in children (e.g., to describe the relation of objective DPN measures to metabolic control). These secondary search strategies yielded 19 new articles for review (Fig 1).

The title and abstract of all articles identified by using the search strategy were independently screened for inclusion by 2 researchers (Dr Hirschfeld and Mr von Glischinski), using the criteria described earlier. Discrepancies were resolved by discussion. The majority of articles were excluded because they did not screen for DPN and only assessed objective measures from the NCS. The full texts of studies that met these criteria (55 studies) were retrieved and reviewed in greater detail. Two researchers (Dr Hirschfeld and Mr von Glischinski) reviewed the final set of studies for information necessary to construct the 2 × 2 tables from which the sensitivity and specificity data were computed. For studies that reported collecting relevant data (i.e., the methods section described performance of NCS and a noninvasive screening for DPN) but did not report the sensitivity and specificity or statistics from which...
these could be computed, the authors were contacted. Because some studies had been published >20 years earlier, the most recent e-mail address was sought. Twelve authors were contacted, and 3 responded to our request. One author provided the missing information, and 2 authors responded that the raw data were no longer accessible because the study was performed >20 years ago.

Quality Assessment of Diagnostic Accuracy Studies

The methodologic quality of the studies was assessed by using the revised Quality Assessment of Diagnostic Accuracy Studies framework. The aim of this instrument is twofold. First, it facilitates unraveling design problems, such as the lack of blinding that may prohibit the interpretation of the studies. Second, it highlights methodologic differences between studies, such as differences in the measurement of both gold standard methods and screening methods. This latter information is central to the question of whether an integration of findings, in terms of meta-analysis, is possible.

Diagnostic Accuracy

Sensitivity, specificity, positive and negative likelihood ratios, and diagnostic odds ratios, along with their 95% confidence intervals, were calculated for each study from the extracted data on true-positive, false-negative, true-negative, and false-positive results. For all studies, a continuity correction was applied. For studies that only reported derived measures (sensitivity and specificity), but not the 2×2 tables, these tables were calculated from the respective sample sizes, and the derived measures were re-calculated by using the same methods used for the studies (notably, continuity correction). We also included data from trials that only indirectly provided data necessary to construct a 2×2 table from which sensitivity and specificity data could be computed. For example, the study by Gallai et al notes that “normal responses were reported in these patients with regard to […] vibration perception detected with a tuning fork.”

RESULTS

Overall, the search strategy identified 72 articles as potentially relevant. Of these, 17 were eliminated in the screening phase because they did not focus on detection of DPN in children with IDDM. Assessment of the remaining 55 articles yielded 5 articles that encompassed the information necessary to compute sensitivity and specificity. Core information about the studies is given in Table 1.

With regard to study quality, we found the following positive aspects throughout: all studies included patients in whom it was clinically sensible to suspect the target disease; withdrawals from study participation were always explained in detail; participant demographic characteristics (eg, age, height) and disease parameters (eg, duration of diabetes, glycohemoglobin levels) were assessed in sufficient detail; assessment of screening and index tests was consistent within each study; and the nerve conduction velocity of both sensory and motor nerves were examined (except for 1 study, which only tested the sensory nervus suralis). Although the studies varied greatly with regard to methods, there were also several consistent methodologic limitations: most studies omitted information about the order of testing.
(screening and index test first) and the individual testing procedures; screening results might not have been interpreted without knowledge of the NCS; interpretation of the clinical data was described as insufficient; and not all participants received the same criterion test. Furthermore, there were many procedural differences between studies that make an integration difficult, including wide-ranging sample sizes (range: 21–75 participants), varied localization of the screening and/or NCS testing area, and differing thresholds used to identify the clinical results as abnormal. Detailed information and revised Quality Assessment of Diagnostic Accuracy Studies criteria are listed in Table 2.

All 5 studies assessed tests of vibration detection (tuning fork or biothesiometer) to screen for DPN. Only 2 studies assessed sensitivity to light touch (monofilaments), but 1 study relied on individual filaments and the other used a staircase procedure to estimate the mechanical perception threshold. These were the only studies that were specifically aimed at describing the diagnostic utility of noninvasive screening methods. The remaining studies reported data on diagnostic utility as a byline. Because of the large heterogeneity with regard to the gold standard and screening assessments, we did not perform a formal meta-analysis but instead present the results of the individual studies (Table 1) and provide a graphical overview (Fig 2).

In the 2 studies that used a biothesiometer, vibration sense achieved acceptable sensitivity (80% and 61%, respectively) and specificity (76% and 64%). For the remaining studies that used a tuning fork, the classification of the screening with regard to the gold standard was near-chance. Specifically, these studies resulted in small sensitivities (between 1% and 19%) and high specificities (between 99% and 87%). Sensitivity of light touch yielded heterogeneous results in 2 studies. One of these studies generated unacceptable results (sensitivity: 19%; specificity: 64%) by using relatively coarse (17 MN) monofilaments. The other study generated acceptable results (sensitivity: 73%; specificity: 87%) by using very fine monofilaments (1 MN).

**DISCUSSION**

The aim of the present review was to evaluate evidence-based screening methods for DPN in children and adolescents. Our review yielded 2 main results. First, despite the high clinical relevance and frequent use of noninvasive screening for DPN, a small number of studies address the diagnostic utility of screenings for DPN in children with IDDM. Second, the studies that were identified only provided limited support for the DPN screening procedures that are routinely advocated for in guidelines.

The paucity of studies on the diagnostic utility of screening procedures is notable. We were able to identify only 5 studies from which we were able to extract data, and only 2 studies were conducted with the aim of describing the diagnostic utility. One reason for this small number may be the implicit belief that findings from adults can be transferred to children, but especially in developmentally sensitive domains such as peripheral nerve functioning, this assessment cannot be taken for granted. Another reason for the few studies into DPN screening in children is that curative therapies for DPN are currently not available. However, there are many research initiatives exploring several pathways to develop such treatments. Any treatment would profit from early detection based on reliable screening tools. Furthermore, although there are several studies comparing various screening methods in adults with diabetes, a recent review of the diagnostic utility of monofilament tests also identified 3 studies that have methodologic problems similar to those that were reported here. The lack of standardization of screening methods seems to be a general problem in diagnostic research on DPN screening.

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>No. of Participants (% Female)</th>
<th>Mean ± SD Age, y</th>
<th>Mean ± SD Duration of IDDM, y</th>
<th>Mean ± SD Glycohemoglobin Level, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blankenburg et al, 2012</td>
<td>IDDM patients</td>
<td>45 (51.1)</td>
<td>13.2 ± 2.8</td>
<td>6.7 ± 2.5</td>
<td>8.2 ± 1.1</td>
</tr>
<tr>
<td>Davis et al, 1997</td>
<td>IDDM patients</td>
<td>21a (53.0)</td>
<td>13.3 ± 4.6</td>
<td>5.1 ± 4.8</td>
<td>10.7 ± 3.5</td>
</tr>
<tr>
<td>Gallai et al, 1988</td>
<td>IDDM patients</td>
<td>50 (54.0)</td>
<td>13.0 ± 1.3</td>
<td>2.3 ± 1.4</td>
<td>9.1 ± 0.5</td>
</tr>
<tr>
<td>Hyllienmark et al, 1995</td>
<td>IDDM patients</td>
<td>75 (38.6)</td>
<td>15.4 ± 3.6</td>
<td>8.2 ± 3.5</td>
<td>7.0 ± 1.1</td>
</tr>
<tr>
<td>Nelson et al, 2006</td>
<td>IDDM patients</td>
<td>73 (47.9)</td>
<td>13.7 ± 2.6</td>
<td>8.1 ± 2.6</td>
<td>9.0 ± 1.0</td>
</tr>
</tbody>
</table>

NR, not reported.

* Even though 307 subjects were screened, only 21 were tested with NCS.
## TABLE 2 Study Characteristics, Methods, and Results

<table>
<thead>
<tr>
<th>Study</th>
<th>Device</th>
<th>Testing Sites</th>
<th>Threshold</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR− (95% CI)</th>
<th>DOR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blankenburg et al,14 2012b</td>
<td>Tuning fork</td>
<td>Both feet</td>
<td>Age- and gender-specific reference values</td>
<td>Right n. suralis</td>
<td>&lt;41 m/s &lt;3 μV</td>
<td>0.19 (0.06–0.46)</td>
<td>0.87 (0.70–0.95)</td>
<td>1.53 (0.34–6.78)</td>
</tr>
<tr>
<td></td>
<td>Monofilament</td>
<td>Both feet</td>
<td>Age- and gender-specific reference values</td>
<td>Right n. suralis</td>
<td>&lt;41 m/s &lt;3 μV</td>
<td>0.73 (0.46–0.98)</td>
<td>0.87 (0.70–0.95)</td>
<td>5.84 (2.07–16.44)</td>
</tr>
<tr>
<td>Davis et al,21 1997b,c</td>
<td>Biothesiometer</td>
<td>Medial malleolus, plantar surface of both great toes (&gt;97th percentile of healthy control subjects)</td>
<td>SNCV: right n. medianus, bil. n. suralis; MNCV: right n. medianus, bil. n. tibialis, bil. n. peroneus</td>
<td>&gt;3 SDs from the mean of healthy control subjects</td>
<td>0.80 (0.60–0.91)</td>
<td>0.76 (0.51–0.91)</td>
<td>3.44 (1.34–8.81)</td>
<td>0.25 (0.10–0.61)</td>
</tr>
<tr>
<td>Gallai et al,6 1988c,d</td>
<td>Tuning fork</td>
<td>NR</td>
<td>NR</td>
<td>SNCV: n. radialis, n. suralis; MNCV: median and posterior n. tibialis</td>
<td>&gt;2 SDs from the mean of healthy control subjects</td>
<td>0.03 (0.00–0.24)</td>
<td>0.98 (0.87–0.99)</td>
<td>2.25 (0.04–108.45)</td>
</tr>
<tr>
<td>Hyllienmark et al,8 1995</td>
<td>Tuning fork</td>
<td>Big toe bilaterally</td>
<td>NR</td>
<td>SNCV: bil. n. medianus, bil. n. suralis; MNCV: bil. n. medianus, bil. n. peroneus</td>
<td>&gt;2 SDs from the mean of healthy control subjects</td>
<td>0.01 (0.00–0.10)</td>
<td>0.99 (0.87–0.99)</td>
<td>0.75 (0.01–36.83)</td>
</tr>
<tr>
<td>Nelson et al,4 2006</td>
<td>Monofilament</td>
<td>Distal hallux</td>
<td>Inability to sense the 4.17-U filament</td>
<td>SNCV: right n. suralis; MNCV: right n. medianus, right n. peroneus</td>
<td>&gt;2 SDs from the mean of reference data</td>
<td>0.19 (0.10–0.33)</td>
<td>0.64 (0.46–0.78)</td>
<td>0.55 (0.25–1.17)</td>
</tr>
<tr>
<td></td>
<td>Biothesiometer</td>
<td>Distal hallux bil. &gt;0.5 μm</td>
<td>SNCV: right n. suralis; MNCV: right n. medianus, right n. peroneus</td>
<td>Values &lt;2 SDs below mean of reference data</td>
<td>0.61 (0.46–0.74)</td>
<td>0.64 (0.46–0.78)</td>
<td>1.71 (1.02–2.88)</td>
<td>0.59 (0.37–0.94)</td>
</tr>
</tbody>
</table>

bil., bilateral; CI, confidence interval; DOR, diagnostic odds ratio; LR+, positive likelihood ratio; LR−, negative likelihood ratio; MNCV, motor nerve conduction velocity; n., nervus; NR, not reported; SNCV, sensory nerve conduction velocity.

* Diagnostic accuracies differ from those reported in primary study due to continuity correction.

† Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) limitation: tests results possibly not interpreted without knowledge of the other test results.

‡ QUADAS limitation: execution of index test and reference standard not described in sufficient detail.

§ QUADAS limitation: interpretation of screening and/or index results not clearly described.
The results concerning the diagnostic utility of methods routinely used for DPN screening (tuning fork and monofilament) were discouraging. With the exception of biothesiometry, which showed above-chance sensitivity and specificity in both studies in which it was used, the screening methods did not demonstrate a reliable positive diagnostic utility. However, 1 of the 2 studies favoring biothesiometry had a serious flaw (only a subset of patients received the gold standard). Furthermore, the finding that the use of finer monofilaments which are near the age- and gender-specific reference values results in higher diagnostic utility points to a possible and practical avenue for improvement.

Several limitations have to be noted when interpreting the results of the present review. First, although NCS is an objective measure for nerve dysfunction, it is unknown whether significant deviations among the control subjects constitute clinically relevant results. It is hoped that early detection of abnormalities improves patient-relevant outcomes, possibly via improved compliance for glycemic control, but there is no evidence (eg, diagnostic Phase IV studies) supporting such claims. It may be that higher rates of detecting abnormalities do not translate into meaningful improvements for patients. Second, methodologic differences between studies (eg, use of study-specific thresholds) prohibited a formal integration of the results from different studies. Although some of these differences have positive aspects by controlling for factors that strongly affect NCS results (eg, skin temperature), the differences in how the screenings were performed have no positive aspects. Third, because the majority of studies were conducted with a different objective, we had to calculate the relevant indices from information given in the text, which was mainly possible in cases in which it was mentioned that none of the patients had clinical signs. As a result, our findings are biased toward overly negative results. However, even when discarding these imputed studies, the results regarding the low utility of noninvasive screening methods remain valid.

The present systematic review revealed that despite the widespread use of screening methods for DPN in children, the diagnostic utility of these screening methods is largely unknown. From the 2 high-quality trials reviewed, it seems plausible that biothesiometry and the use of finer monofilaments may improve detection rates. Using more sensitive measures would result in a marked increase in children with positive screening results. Changes to guidelines that pertain to so many children and adolescents need to be informed by evidence that also considers the consequences of the screening methods.
REFERENCES


Screening for Peripheral Neuropathies in Children With Diabetes: A Systematic Review
Gerrit Hirschfeld, Michael von Glischinski, Markus Blankenburg and Boris Zernikow
Pediatrics 2014;133:e1324; originally published online April 7, 2014; DOI: 10.1542/peds.2013-3645
Screening for Peripheral Neuropathies in Children With Diabetes: A Systematic Review
Gerrit Hirschfeld, Michael von Glischinski, Markus Blankenburg and Boris Zernikow

*Pediatrics* 2014;133:e1324; originally published online April 7, 2014;
DOI: 10.1542/peds.2013-3645

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/133/5/e1324.full.html