


RESEARCH

Open Access



# Non-invasive assessment of vibration perception and protective sensation in people with diabetes mellitus: inter- and intra-rater reliability

Sean Michael Lanting<sup>1\*</sup> , Martin Jeremy Spink<sup>1</sup>, Peta Ellen Tehan<sup>1,2</sup>, Stephanie Vickers<sup>1</sup>, Sarah Louise Casey<sup>1</sup> and Vivienne Helaine Chuter<sup>1,3</sup>

## Abstract

**Background:** Testing of protective sensation and vibration perception are two of the most commonly used non-invasive methods of screening for diabetes-related peripheral neuropathy (DPN). However, there is limited research investigating the reliability of these tests in people with diabetes. The aim of this study was to determine the inter- and intra-rater reliability of methods used to test vibration perception and protective sensation in a community-based population of adults with type 2 diabetes.

**Methods:** Three podiatrists with varying clinical experience tested four- and 10-site, 10 g monofilament and vibration perception threshold (VPT). In a separate cohort, the reliability of a graduated tuning fork as well as two methods of conventional tuning fork (on/off method and dampening method) was undertaken by a new graduate podiatrist and podiatrist with one-year's clinical experience. The intra- (Cohen's K) and inter-rater (Cohen's or Fleiss' K) reliability of each test was determined.

**Results:** Fifty participants (66% male, 100% type 2, 32% with DPN) underwent monofilament and neurothesiometer testing with 44 returning for the retest. Twenty-four participants (63% male, 100% type 2, 4% with DPN) underwent tuning fork testing and returned for retest. All tests demonstrated acceptable inter-rater reliability ranging from moderate (10-site monofilament, K: 0.54, CI: 0.38–0.70,  $p = 0.02$ ) to substantial (graduated tuning fork, K: 0.68, CI: 0.41–0.95,  $p < 0.01$ ). The 10-site monofilament (K: 0.44–0.77) outperformed the 4-site test (K: 0.34–0.67) and the dampened tuning fork method (K: 0.41–0.49) showed lower intra-rater reliability compared to both conventional (K: 0.52–0.57) and graduated methods (K: 0.50–0.57).

**Conclusion:** We support the current recommendations of using more than one test to screen and monitor progression of DPN. Four- and 10-site 10 g monofilament testing have similarly acceptable levels of reliability and the neurothesiometer is the most reliable method of assessing vibration perception function. Use of a graduated tuning fork was slightly more reliable than other methods of tuning fork application however all had substantial reliability. Years of clinical experience only marginally affected test reliability overall and due to subjective nature of the tests we suggest that testing should be performed regularly and repetitively.

**Keywords:** Diabetes, Neuropathy, Reliability, Monofilament, Vibration, Tuning fork, Neurothesiometer

\* Correspondence: [Sean.lanting@newcastle.edu.au](mailto:Sean.lanting@newcastle.edu.au)

<sup>1</sup>School of Health Sciences, University of Newcastle, Newcastle, Australia  
Full list of author information is available at the end of the article



## Background

Diabetes is a significant health problem and was recently estimated to affect approximately 451 million people worldwide [1]. Up to 50% of persons with diabetes are affected by diabetic peripheral neuropathy (DPN), which causes widespread sensory loss, primarily affecting the feet and legs [2–5]. DPN is associated with lower limb complications such as foot deformity [6], increased plantar pressures [7], ulceration and infection and, is implicated in 50–75% of all non-traumatic lower limb amputations [8]. Prophylactic care in people with diabetes has been shown to prevent or delay development of DPN. For example, intensive glycaemic control has demonstrated a reduction of neuropathy incidence of between 25% [9] and 57% [10]. Additionally education and routine foot care in those with DPN have been shown to reduce risk of associated foot complications [11, 12]. Therefore, early and accurate diagnosis of DPN is paramount to mitigating the risk of associated foot complications.

Methods for conducting clinical chairside neurological tests to establish the presence and monitor the progression of DPN are varied, and assess different nerve fibre types. Current international guidelines recommend testing of protective sensation using monofilament, as well as additional tests such as vibration perception, reflexes, pain perception and asking about neurological symptoms [13, 14]. Diminished vibration perception and ability to detect 10 g monofilament have demonstrated predictive capacity for future foot ulceration [8, 15–18], and are widely used both clinically and in research. Several techniques are available for testing vibration perception, including use of a neurothesiometer or similar instrument, as well as graduated and non-graduated tuning forks. Similarly, methods for testing protective sensation testing using monofilament examination can vary clinically in terms of location and number of sites tested. However there are limited data available comparing the reliability of different testing methods. Reliability refers to the level of consistency of measurement results between different clinicians (inter-rater) and the same clinician on multiple occasions (intra-rater). While there have been several small studies investigating inter- and/or intra-rater reliability of monofilament [19–21] and vibration perception testing [21–25] results of these studies are variable, and generalisability of these findings limited by inconsistency of testing methods. One larger study recently compared effectiveness of three, 4 and 10 site monofilament for identifying DPN in 1915 people with diabetes, and in doing so, reported high level of agreement between testing methods (K: 0.797 to 0.925) [26], but did not report reliability on individual tests.

The aim of this study was to determine the inter- and intra-rater reliability of commonly used testing methods

of protective sensation and vibration perception, performed by podiatrists with varying amounts of clinical experience, in people with diabetes. Specifically, a four-site and a 10-site monofilament test, as well as vibration perception as determined by neurothesiometer, graduated tuning fork and non-graduated (dampened and conventional methods) tuning fork.

## Methods

This study was conducted at the University of Newcastle Podiatry clinics in New South Wales, Australia. Ethics approval was obtained from the University of Newcastle Human Research Ethics Committee prior to undertaking this study, protocol code H-2012-0141. All participants involved in this study provided written informed consent prior to study commencement.

## Participants

Participants were recruited on a volunteer basis, with flyers posted up in university clinic consultation rooms and the waiting room, directing potential recruits to register their interest. Recruitment was performed by people who were not involved in test performance thereby ensuring blinding of raters to participant health status. Participants included in the study were required to be representative of the population in which screening for DPN is recommended [14]. Therefore, inclusion criteria were Type-1 diabetes of five years or more or Type-2 diabetes of any duration with and without history of diagnosed DPN, confirmed by medical records. Participants were required to be fluent in English language to satisfy consent for the study. Exclusion criteria included active foot ulceration, visual evidence of recently healed foot ulceration, lower limb amputation of any kind or diagnosed peripheral neuropathy of an origin other than diabetes.

The inter- and intra-rater reliability of 10 g monofilament testing using four-site and 10-site testing techniques as well as vibration perception threshold (VPT) using a neurothesiometer were determined across three raters [a new graduate podiatrist (R1); a podiatrist with five years of clinical experience (R2); and a podiatrist with 10 years of clinical experience (R3)]. In addition, inter- and intra-rater reliability of a graduated tuning fork as well as an on/off and a dampened method of a conventional tuning fork were tested in a podiatrist with one year's clinical experience (R4) and a new graduate podiatrist (R5).

## Testing methods

### Monofilament testing

Semmes-Weinstein 10 g monofilaments (North Coast Medical, California) were used to conduct all monofilament testing. A four-site [27] and a ten-site

monofilament [8] test were used. For the four-site test, site application was plantar surface of the hallux as well as first, third and fifth metatarsal heads, while the 10-site test also included the plantar surface of the third and fifth digits, heel, medial arch, lateral arch and the dorsal surface of the mid foot. Perception of six or less sites in the 10-site test [28] and three or less sites in the four-site test [8] were considered abnormal. Monofilaments were applied perpendicular to the skin until buckling and held in place for 1–2 s. The participants were asked to respond with a “yes” on each occasion where they could perceive the 10 g force. Monofilaments used in this study were discarded following use on nine consecutive participants ensuring they were not used more than 100 times within 24 h in order to maintain the force applied at 10 g [29].

### Neurothesiometer testing

Horwell neurothesiometers (Wilford Industrial, Nottingham) were used to determine vibration perception threshold (VPT). The stylus of the device was applied to the apex of the right hallux and the amplitude of vibration of the device was then gradually increased until the participant could perceive the vibration. The corresponding VPT value was immediately written on the assessment form and the process repeated until three values were recorded. The mean of the three values was calculated, with a mean VPT value > 25 v considered an abnormal response [27].

### Tuning fork testing

The vibration perception tests were performed using graduated C64-Hz Ragg Rydel-Seiffer (Granton Medical, Sheffield) and conventional (non-graduated) Ragg Gardner Brown C128-Hz (Granton Medical, Sheffield) tuning forks. For each tuning fork test a manually applied force to induce vibration was applied to the tines of the tuning fork before placing the device on the apex of the right hallux. Participants were instructed to indicate verbally when they felt vibration and then when they perceived the vibration had stopped (C128-Hz). If the participant could not detect the vibration at all then it was considered abnormal (on/off method) [17]. Once the vibration was perceived, the rater would randomly dampen the tuning fork (C128-Hz) with their other hand and if the participant could not perceive that the vibration had stopped then this was considered an abnormal response (dampening method) [23]. Lastly, perception of less than four octals as quantified by the graduated tuning fork (C64 Hz) constituted an abnormal response [30].

### Testing protocol

In both the initial testing session and retest for all testing conducted as part of this study, raters performed the

relevant neurological tests in a pre-determined random order on every participant in separate treatment rooms. Raters were blinded to the participant health status i.e. presence, absence, or extent of DPN, though were aware that all of the participants had diabetes. Raters were also blinded to each other's results as well as to their own results from the first testing session when undertaking the retest. The order of application of the tests was randomised using an online random number generator ([www.randomizer.org](http://www.randomizer.org)). The order of raters was randomised in a manner that was not pre-determined and the order of site application of the monofilament was randomised at the discretion of the individual raters. Participants were blind to all results, though were provided with a plain language summary on request at study completion. The tests were performed only on the right limb in order to satisfy the assumption of independence of data [31], with the right limb chosen rather than a random limb in order to minimise rater confusion. Participants were required to attend the retest after seven days at the same location and were required to close their eyes for each test procedure. In addition, each test was first demonstrated on the dorsal aspect of the participant's hand and in relation to vibration, ‘buzzing’ was differentiated from pressure sensation.

### Statistical analysis

SPSS version 25 was used for statistical analysis. Results for all neurological tests were broken down into dichotomous variables, namely abnormal or normal results, with abnormal being indicative of neuropathy. The intra-rater reliability was calculated using an unweighted Cohen's Kappa (K) statistic [32]. In order to calculate the inter-rater reliability and effect of experience on reliability, Cohen's K was initially determined between the following pairs of raters: R1 and R2; R1 and R3; and R2 and R3 (monofilament and neurothesiometer) and R4 and R5 (tuning fork tests). Fleiss' K was then calculated to determine the overall reliability between raters R1-R3 [33]. Interpretation of the Cohen's and Fleiss' K statistic was performed using the method proposed by Landis and Koch [34] (Values indicating: 0.01–0.20 = slight, 0.21–0.40 = fair, 0.41–0.60 = moderate, 0.61 to 0.80 = substantial, and 0.81–1.0 = almost perfect). Values below 0.4 were interpreted as clinically unacceptable for reliability of a test [35].

### Results

Fifty participants volunteered for testing with monofilament and neurothesiometer, of whom 44 returned for the retest. Six participants were unable to return within the required period of seven days and thus did not take part in the intra-rater reliability component of this study. Twenty-four participants volunteered for tuning fork

testing, all of whom returned for the re-test. Participant characteristics are detailed in Table 1.

### Monofilament

**Intra-rater reliability:** The four-site 10 g monofilament examination demonstrated variable intra-rater reliability ( $n = 50$ ) with Cohen's K ranging from fair ( $K = 0.34$ , 95%CI: 0.06 to 0.63,  $p = 0.02$ ) to substantial ( $K = 0.67$ , 95%CI: 0.45 to 0.89,  $p < 0.01$ ), Table 2. The 10-site monofilament test demonstrated intra-rater reliability ( $n = 50$ ) ranging from moderate ( $K = 0.44$ , 95%CI: 0.09 to 0.79,  $p < 0.01$ ) to substantial ( $K = 0.77$ , 95%CI: 0.55 to 0.99,  $p < 0.01$ ) and was not related to increasing clinical experience, Table 2.

**Inter-rater reliability:** Determined by Fleiss' K (Table 3), the four-site monofilament test ( $n = 44$ ) displayed substantial inter-rater reliability ( $K = 0.61$ , 95%CI: 0.45 to 0.77,  $p < 0.01$ ) compared to moderate inter-rater reliability for the 10 site test ( $K = 0.54$ , 95%CI: 0.38 to 0.70,  $p = 0.02$ ). There was very little discrepancy between reliability when analysing pairs of raters for the 10-site test, however for the four-site test the reliability was relatively higher for the pooling of the more experienced podiatrists ( $K: 0.72$ , 95%CI: 0.53 to 0.91,  $p < 0.01$ ) compared to the pooling of the two Podiatrists with less experience ( $K: 0.55$ , 95%CI: 0.31 to 0.78,  $p < 0.01$ ), Table 3.

### Neurothesiometer

**Intra-rater reliability:** The neurothesiometer ( $n = 50$ ) demonstrated a range of intra-rater reliability from moderate ( $K = 0.52$ , 95%CI: 0.21 to 0.82,  $p = 0.01$ ) to substantial ( $K = 0.78$ , 95%CI: 0.58–0.98,  $p = 0.02$ ), Table 2.

**Inter-rater reliability:** Determined by Fleiss' K, the neurothesiometer ( $n = 44$ ) demonstrated substantial inter-rater reliability ( $K: 0.61$ , 95%CI: 0.45 to 0.77,  $p < 0.01$ ). The most experienced pair of raters in this instance produced a substantially lower reliability ( $K: 0.48$ , 95%CI: 0.21 to 0.75,  $p < 0.01$ ) compared with the least experienced pair ( $K: 0.78$ , 95%CI: 0.59 to 0.96,  $p < 0.01$ ), Table 3.

**Table 1** Participant characteristics

Monofilament and VPT ( $n = 50$ )				
	N	Mean Age (SD)	T2-DM	Existing DPN diagnosis
Males	33	72 (11)	33	16
Females	17	73 (7)	17	0
Tuning fork ( $n = 24$ )				
	N	Mean Age (SD)	T2-DM	Existing DPN diagnosis
Males	15	74 (7)	15	1
Females	9	70 (9)	9	0

SD standard deviation, T2 DM type two diabetes, DPN diabetic peripheral neuropathy

### Tuning fork

**Intra-rater reliability:** The conventional and graduated methods outperformed the dampened method for both the testers R4 and R5 ( $n = 24$ ), though all methods demonstrated moderate intra-rater reliability ( $K: 0.41$  to 0.57), Table 2.

**Inter-rater reliability:** The graduated tuning fork ( $k: 0.68$ , 95%CI: 0.41–0.95,  $p < 0.01$ ) demonstrated slightly higher inter-rater reliability ( $n = 24$ ) than the dampened method ( $K: 0.66$ , 95% CI: 0.37–0.95,  $p < 0.01$ ) and conventional method ( $K: 0.63$ , 95% CI: 0.30–0.96,  $p < 0.01$ ), though all demonstrated substantial reliability, Table 3.

### Discussion

The results from our study indicate that monofilament, neurothesiometer and the tuning fork are acceptably reliable methods of testing protective sensation and vibration perception respectively, with some variability demonstrated between inter- and intra-tester reliability as well as with level of clinical experience. Use of a graduated tuning fork or the on/off method using a conventional, non-graduated tuning fork, demonstrated higher reliability than the dampened method and are therefore more appropriate for clinical use. Overall, greater clinician experience resulted in marginally increased reliability of the graduated and conventional (on/off) tuning fork method and substantially increased reliability of the neurothesiometer. Monofilament tests overall, appear to be reliable with clinical experience possibly increasing the reliability of the four-site test. Despite the acceptable levels of reliability demonstrated by these tests, caution must be used in relying on any one test in isolation. Moderate reliability for example still indicates a marked margin of error in test interpretation and it is axiomatic that clinical tests that have the potential to change clinical practice and drive treatment strategies should strive for higher reliability. When considering using these tests for diagnosis and monitoring of DPN we support the current recommendations of using more than one test (e.g. monofilament and tuning fork) as part of a larger screening examination. In addition, we suggest that testing should be performed regularly and repetitively. Of note, our results relate specifically to the reliability of the tests used, i.e. that the results can be replicated, not that they reflect a correct diagnosis of DPN. While use of tests with high reliability is essential for effective clinical management, so too is the need for the tests to be able to diagnose the target condition. It has been stated that two-test combinations have > 87% sensitivity in detecting DPN [36], though further work to determine the combination test with highest reliability that is most diagnostically accurate for identifying presence of DPN is required.

**Table 2** Intra-rater reliability reported as Cohen's K and SE with 95%CI and *p*-values

Monofilament and VPT (n = 44)												
Test	Rater 1				Rater 2				Rater 3			
	K	SE	95% CI	P	K	SE	95% CI	P	K	SE	95% CI	P
VPT	0.52	0.15	0.21 to 0.82	0.01	0.78	0.10	0.58 to 0.98	0.02	0.72	0.13	0.46 to 0.98	< 0.01
Monofilament [4]	0.44	0.14	0.16 to 0.73	0.02	0.67	0.11	0.45 to 0.89	< 0.01	0.34	0.15	0.06 to 0.63	0.02
Monofilament [10]	0.49	0.18	0.14 to 0.83	< 0.01	0.77	0.11	0.55 to 0.99	< 0.01	0.44	0.18	0.09 to 0.79	< 0.01
Tuning fork (n = 24)												
Test	Rater 4				Rater 5							
	K	SE	95% CI	P	K	SE	95% CI	P				
Graduated	0.57	0.17	0.24 to 0.9	< 0.01	0.50	0.17	0.17 to 0.83	< 0.02				
Conventional (on/off)	0.57	0.16	0.26 to 0.88	< 0.01	0.52	0.19	0.15 to 0.89	< 0.02				
Conventional (dampened)	0.41	0.19	0.04 to 0.78	< 0.05	0.49	0.18	0.14 to 0.84	< 0.02				

VPT vibration perception threshold

Previous investigation into the 10 g monofilament has shown mixed reliability. A nine-site monofilament test has been shown to have excellent intra- and inter-reliability [20]. Meijer et al., described moderate to good intra-rater and good inter-rater reliability, respectively, for a two-site test [21] while a three-site test has demonstrated fair to moderate inter- and intra-reliability [37]. Lastly, level of agreement between the four- and 10-site test in 1915 people with diabetes was recently shown to be high (K: 0.87) [26] indicating that these tests may be similarly reliable. Our study supports the relatively high inter-rater reliability of the four- And 10-site 10 g monofilament tests previously reported. The inter-rater reliability of four- and 10-site tests from this present study demonstrated similar levels of reliability overall,

although experience improved reliability for the four-site test. The excellent intra-rater reliability previously described in the nine-site monofilament test [20] was not replicated in the four or 10 site tests used in our study. The large range of intra-rater reliability of the monofilament (fair to substantial) was not associated with greater clinical experience. As these tests rely on subjective responses from a patient, it is possible that these tests will demonstrate variability regardless of the level of experience of the clinician.

The reliability of a variety of methods of assessing vibration perception was determined in this study including an on/off and a dampening method of a conventional, non-graduated tuning fork, a graduated tuning fork and the neurothesiometer. Of these, the

**Table 3** Inter-rater reliability of neurological tests reported as Cohen's or Fleiss' K and SE with 95%CI and *p*-values

Inter-rater reliability between pairs of raters (n = 50)												
Test	Rater 1 & 2				Rater 1 & 3				Rater 2 & 3			
	K	SE	95% CI	P	K	SE	95% CI	P	K	SE	95% CI	P
VPT	0.78	0.9	0.59 to 0.96	< 0.01	0.59	0.13	0.33 to 0.84	< 0.01	0.48	0.14	0.21 to 0.75	< 0.01
Monofilament [4]	0.55	0.12	0.31 to 0.78	< 0.01	0.57	0.12	0.33 to 0.80	< 0.01	0.72	0.10	0.53 to 0.91	< 0.01
Monofilament [10]	0.57	0.14	0.30 to 0.84	< 0.01	0.50	0.14	0.18 to 0.82	< 0.01	0.57	0.16	0.17 to 0.83	< 0.01
Inter-rater reliability between 3 raters (n = 50)												
Test	K	SE	95% CI	P								
VPT	0.61	0.08	0.45 to 0.77	< 0.01								
Monofilament [4]	0.61	0.08	0.45 to 0.77	0.01								
Monofilament [10]	0.54	0.08	0.38 to 0.70	0.02								
Inter-rater reliability for tuning fork between 2 raters (n = 24)												
	T1 & T2 (K)	SE	95% CI	P								
Conventional (on/off)	0.63	0.17	0.30 to 0.96	< 0.01								
Conventional (dampened)	0.66	0.15	0.37 to 0.95	< 0.01								
Graduated	0.68	0.14	0.41 to 0.95	< 0.01								

VPT vibration perception threshold

neurothesiometer ( $n = 50$ ) demonstrated the highest intra-rater reliability and the graduated tuning fork ( $n = 24$ ) the highest inter-rater reliability. The reliability demonstrated may have been affected by the comparatively low participant numbers in the tuning fork cohort. Overall, the inter-rater reliability of vibration tests was substantial. Our findings regarding the neurothesiometer are supported by two smaller studies investigating the neurothesiometer [22], biothesiometer and Maxivibrometer [25], respectively. In our study, intra-tester reliability of the neurothesiometer was affected by experience, with the new graduate demonstrating substantially lower reliability ( $K = 0.52$ ) than the more experienced clinicians ( $K = 0.72-0.78$ ).

While all tuning fork methods demonstrated substantial inter-rater reliability, the intra-rater reliability was moderate for all methods, and bordering on fair for the dampened method. Previous investigation by Meijer et al., reported substantial intra-rater reliability of the conventional (on/off) method ( $K = 0.69$ ) at the hallux interphalangeal joint [21]. Perkins et al., noted acceptable reliability of the conventional (on/off) method at the hallux dorsum, without reporting a Kappa statistic [23]. Our findings of moderate intra-tester reliability of the graduated tuning fork are somewhat supported by Thivolet et al., who simply stated statistical significance between test and retest at  $p < 0.01$  [24]. A slightly smaller study previously reported low, non-significant inter-rater reliability of the graduated tuning fork [22], which contradicts our findings of substantial reliability. However, the site application and methodology was too dissimilar to our present study to draw any meaningful comparisons. Lastly, the graduated and on/off conventional methods were only marginally affected by experience. We therefore suggest using the graduated tuning fork or conventional on/off method of vibration perception as opposed to the dampened method.

### Limitations

Whilst adding to the paucity of research investigating intra- and inter-rater reliability of vibration perception and monofilament testing in people with diabetes, findings of this study need to be considered in light of several limitations. Though 50 participants attended for test and retest of monofilament and neurothesiometer, only 24 were involved in tuning fork testing. As  $n \geq 30$  is required to satisfy the assumption of normal distribution [38], larger sample studies are warranted. Our study is generalisable to people with type 2 diabetes only, however a strength of this study is that it included people with diagnosed DPN making it generalizable to people requiring testing and ongoing monitoring. In addition, more extensive clinician training and clearer instruction to participants may improve reliability. The findings of

this study are also limited to peripheral neurological testing with neurothesiometer, tuning forks and 10 g monofilament. Other neurological tests such as pain perception, proprioception, ankle reflexes, temperature perception, light touch perception and two-point discrimination were not investigated but may be reliable and of clinical value.

### Conclusion

Neurological screening is routinely performed by health professionals on patients with diabetes as they are at risk of developing DPN. Our research suggests that the neurothesiometer, four-site and 10-site 10 g monofilament are all acceptably reliable in screening for DPN. If using a tuning fork to test vibration perception then consider quantification using a graduated tuning fork or the conventional (on/off) method in place of the dampened method. The results of this study also indicate there is a need for regular and repetitive testing and that a combination of tests should be used for screening and monitoring of DPN for Podiatrists regardless of experience level. The reliability of alternate neurological screening methods warrants investigation.

### Abbreviations

DPN: Diabetic peripheral neuropathy; VPT: Vibration perception threshold

### Acknowledgments

We would like to acknowledge the participants of this study for their time.

### Author's contributions

SL and VC conceived the study and developed the methodology; PT, SV and SL recruited participants; SL, MS and SC and SV collected data; SL, VC and MS completed data analysis; SL, VC, SV and MS developed the manuscript, and all authors approved the final manuscript.

### Funding

There is no funding to declare.

### Availability of data and materials

De-identified data is held securely with the senior author.

### Ethics approval and consent to participate

This research was completed under the ethical approval of the University of Newcastle Human Research Ethics Committee H-2012-0141.

### Consent for publication

Not applicable.

### Competing interests

None of the authors have any conflicts of interests to declare.

### Author details

<sup>1</sup>School of Health Sciences, University of Newcastle, Newcastle, Australia.

<sup>2</sup>Priority Research Centre for Gender, Health and Ageing, University of Newcastle, Newcastle, Australia. <sup>3</sup>Priority Research Centre for Physical Activity and Nutrition, University of Newcastle, Newcastle, Australia.

Received: 6 November 2019 Accepted: 13 January 2020

Published online: 16 January 2020

## References

1. Cho N, Shaw J, Karuranga S, Huang Y, da Rocha FJ, Ohlrogge A, et al. IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Res Clin Pract.* 2018;138:271–81.
2. Dros J, Wewerinke A, Bindels PJ, van Weert HC. Accuracy of monofilament testing to diagnose peripheral neuropathy: a systematic review. *Ann Fam Med.* 2009;7(6):555–8.
3. Dyck PJ, Kratz KM, Karnes JL, Litchy WJ, Klein R, Pach JM, et al. The prevalence by staged severity of various types of diabetic neuropathy, retinopathy, and nephropathy in a population based cohort. *Neurology.* 1993;43(4):817.
4. Perkins BA, Orszag A, Ngo M, Ng E, New P, Bril V. Prediction of incident diabetic neuropathy using the monofilament examination. *Diabetes Care.* 2010;33(7):1549–54.
5. Sadosky A, McDermott AM, Brandenburg NA, Strauss M. A review of the epidemiology of painful diabetic peripheral neuropathy, postherpetic neuralgia, and less commonly studied neuropathic pain conditions. *Pain Pract.* 2008;8(1):45–56.
6. Lavery LA, Armstrong DG, Vela SA, Quebedeaux TL, Fleischli JG. Practical criteria for screening patients at high risk for diabetic foot ulceration. *Arch Intern Med.* 1998;158(2):157–62.
7. Caselli A, Pham H, Giurini JM, Armstrong DG, Veves A. The forefoot-to-rearfoot plantar pressure ratio is increased in severe diabetic neuropathy and can predict foot ulceration. *Diabetes Care.* 2002;25(6):1066–71.
8. Armstrong DG, Lavery LA, Vela SA, Quebedeaux TL, Fleischli JG. Choosing a practical screening instrument to identify patients at risk for diabetic foot ulceration. *Arch Intern Med.* 1998;158(3):289–92.
9. Turner R, Holman R, Stratton I, Cull C, Matthews D, Manley S, et al. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). *Lancet.* 1998; 352(9131).
10. Nathan D, Genuth S, Lachin J, Cleary P, Crofford O, Davis M, et al. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med.* 1993;329(14):977–86.
11. Armstrong DG, Harkless LB. Outcomes of preventative care in a diabetic foot specialty clinic. *J Foot Ankle Surg.* 1998;37(6):460–6.
12. Plank J, Haas W, Rakovac I, Görzer E, Sommer R, Siebenhofer A, et al. Evaluation of the impact of chiropodist care in the secondary prevention of foot ulcerations in diabetic subjects. *Diabetes Care.* 2003;26(6):1691–5.
13. Schaper N, van Netten J, Apelqvist J, Bus S, Hintchcliffe R, Lipsky B. IWGDF Practical guidelines on the prevention and management of diabetic foot disease. [www.iwgdfguidelines.org](http://www.iwgdfguidelines.org). 2019.
14. American Diabetes Association. Microvascular complications and foot care: Standards of Medical Care in Diabetes-2018. *Diabetes Care.* 2018;41(Suppl 1):S105.
15. Mayfield JA, Sugarman JR. The use of the Semmes-Weinstein monofilament and other threshold tests for preventing foot ulceration and amputation in persons with diabetes. *J Fam Pract.* 2000;49(11):S17.
16. Mueller MJ, Hastings M, Commean PK, Smith KE, Pilgram TK, Robertson D, et al. Forefoot structural predictors of plantar pressures during walking in people with diabetes and peripheral neuropathy. *J Biomech.* 2003;36(7): 1009–17.
17. Singh N, Armstrong DG, Lipsky BA. Preventing foot ulcers in patients with diabetes. *J Am Med Assoc.* 2005;293(2):217–28.
18. Young MJ, Breddy JL, Veves A, Boulton AJ. The prediction of diabetic neuropathic foot ulceration using vibration perception thresholds: a prospective study. *Diabetes Care.* 1994;17(6):557–60.
19. Woodbury MG, Sibbald RG, Ostrow B, Persaud R, Lowe JM. Tool for rapid & easy identification of high risk diabetic foot: validation & clinical pilot of the simplified 60 second diabetic foot screening tool. *PLoS One.* 2015;10(6): e0125578.
20. Young D, Schuerman S, Flynn K, Hartig K, Moss D, Altenburger B. Reliability and responsiveness of an 18 site, 10-g monofilament examination for assessment of protective foot sensation. *J Geriatr Phys Ther.* 2011;34(2):95–8.
21. Meijer J-WG, Smit AJ, Lefrandt JD, Van Der Hoeven JH, Hoogenberg K, Links TP. Back to basics in diagnosing diabetic polyneuropathy with the tuning fork! *Diabetes Care.* 2005;28(9):2201–5.
22. O'Neill J, McCann S, Lagan K. Tuning fork (128 Hz) versus neurothesiometer: a comparison of methods of assessing vibration sensation in patients with diabetes mellitus. *Int J Clin Pract.* 2006;60(2):174–8.
23. Perkins BA, Olaleye D, Zinman B, Bril V. Simple screening tests for peripheral neuropathy in the diabetes clinic. *Diabetes Care.* 2001;24(2):250–6.
24. Thivolet C, El Farkh J, Petiot A, Simonet C, Tourniaire J. Measuring vibration sensations with graduated tuning fork: simple and reliable means to detect diabetic patients at risk of neuropathic foot ulceration. *Diabetes Care.* 1990; 13(10):1077–80.
25. Van Deursen RWM, Sanchez M, Derr J, Becker M, Ulbrecht J, Cavanagh PR. Vibration perception threshold testing in patients with diabetic neuropathy: ceiling effects and reliability. *Diabet Med.* 2001;18(6):469–75.
26. Zhang Q, Yi N, Liu S, Zheng H, Qiao X, Xiong Q, et al. Easier operation and similar power of 10 g monofilament test for screening diabetic peripheral neuropathy. *J Int Med Res.* 2018;46(8):3278–84.
27. Boulton AJ, Armstrong DG, Albert SF, Frykberg RG, Hellman R, Kirkman MS, et al. Comprehensive foot examination and risk assessment a report of the task force of the foot care interest Group of the American Diabetes Association, with endorsement by the American Association of Clinical Endocrinologists. *Diabetes Care.* 2008;31(8):1679–85.
28. Lee S, Kim H, Choi S, Park Y, Kim Y, Cho B. Clinical usefulness of the two-site Semmes-Weinstein monofilament test for detecting diabetic peripheral neuropathy. *J Korean Med Sci.* 2003;18(1):103.
29. Booth J, Young MJ. Differences in the performance of commercially available 10-g monofilaments. *Diabetes Care.* 2000;23(7):984–8.
30. Spruce MC. Neurological Assessment. In: Yates B, editor. *Merriman's Assessment of the lower limb.* 3rd ed. Philadelphia: Churchill Livingstone Elsevier; 2009. p. 146.
31. Menz HB. Two feet, or one person? Problems associated with statistical analysis of paired data in foot and ankle medicine. *Foot (Edinburgh, Scotland).* 2004;14(1):2–5.
32. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20(1):37–46.
33. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychological Bull.* 1971;76(5):378–82.
34. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159–74.
35. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther.* 2005;85(3):257–68.
36. Boulton AJ, Vinik AI, Arezzo JC, Bril V, Feldman EL, Freeman R, et al. Diabetic neuropathies: a statement by the American Diabetes Association. *Diabetes Care.* 2005;28(4):956–62.
37. Davies A. Repeatability of clinical tests for diabetes related peripheral neuropathy. 2016. <https://doi.org/10.13140/RG.2.2.5948.39042>.
38. Portney L, Watkins M. *Foundations of clinical research: applications to practice.* 3rd Edition ed. New Jersey: Prentice Hall; 2009.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

