Original Article

Interobserver agreement in eliciting signs of peripheral neuropathy in patients with diabetes mellitus type 2

Abdul Rehman Arshad, Khalid Hussain, Jovaria Masood

Mountain Medical Battalion, Bagh, AJK, Pakistan

Objective: To determine the extent to which two different doctors agree on signs of peripheral neuropathy in patients with diabetes mellitus type 2.

Methodology: In this observational study, patients with diabetes mellitus type 2 were selected by convenience sampling. Those with amputated feet, extensive skin ulceration not permitting clinical examination as well as unwilling patients were excluded from the study. Two doctors elicited three signs of peripheral neuropathy independently on the same day. Vibration sensation was checked over the tips of both great toes using a 128 Hz tuning fork. Light touch sensation was checked over plantar surfaces of both great toes with 10g Semmes-Weinstein monofilament. It was applied perpendicular to the skin surface for a second, using an amount of force that just buckled the monofilament. Both ankle jerks were also checked in a lying position with a reflex hammer. Responses to all the three tests were classified as being present bilaterally, present unilaterally or absent bilaterally.

Results: A total of 105 patients were examined. Kappa values for vibration sensation, light touch sensation and ankle jerks were 0.330, 0.452 and 0.581, respectively.

Conclusion: Suboptimal agreement between the two examiners calls for training prior to large scale studies, so as to reduce the variability in results. (Rawal Med J 201;41:171-174).

Key words: Observer bias, observer variation, peripheral neuropathy, neurologic manifestations.

INTRODUCTION

Distal symmetric polyneuropathy is fairly common in patients with type 2 diabetes mellitus (DM), prevalence being proportional to the duration of diabetes.[1] It is caused by different metabolic abnormalities such as accumulation of advanced glycosylation end products/sorbitol, increased oxidative stress as well as nerve ischemia and impaired repair mechanisms.[2] Presence of neuropathy is associated with an increased risk of foot ulcers, gangrene and amputations.[3] In most cases, it can be diagnosed by simple clinical examination. Certainly, a simple clinical examination has been shown to be a good predictor of future foot ulcer risk.[4] Sensory testing involves using a tuning fork to check vibration sensation and monofilament to look for light touch sensation. Ankle jerks can also be checked easily with a reflex hammer. All these signs are easy to elicit on the bedside. A single test is often not enough to accurately predict neuropathy.[5] American Diabetes Association therefore recommends a careful examination with more than one test.[6]

Nevertheless, it is not essential for two or more examiners to record similar findings for these clinical signs in a given patient. This assumes a greater importance when multiple researchers are carrying out studies involving such a clinical examination. In this case, it is important to first document an acceptable level of agreement between the observers most of the times. To the best of our knowledge, no study on this subject has previously been published in regional literature, at least during the last fifteen years. This study was therefore, planned to determine how two different doctors elicit and grade these signs of neuropathy and how well their findings match with each other.

METHODOLOGY

This observational study was carried out at 1 Mountain Medical Battalion (Bagh, Azad Kashmir) from May to June 2015. Prior approval was obtained from Ethics Review Committee of the institute. Patients with DM type 2 were enrolled by convenience sampling. Informed verbal consent was taken from all of them. Those with amputated feet, extensive skin ulceration not permitting clinical examination as well as unwilling patients

were excluded from the study. Assuming that the two observers would agree at least 50% of the times, there is no possibility of chance agreement, and a relative error of 20%, a sample size of 100 patients was calculated using the following formula:

$$n = \frac{n^*}{1+n^*/N}, \qquad \text{where } n^* = \frac{1}{r^2(P_a - P_e)^2}$$

In this equation, r represents relative error, P_a is the overall agreement probability and P_a is the chance-agreement probability.

Demographic data was collected first. Clinical examination was then done by two doctors the same day. One of them was a medical specialist having five years post fellowship experience and the other a surgeon who had just recently cleared fellowship exam. They elicited the clinical signs independently in separate clinics using identical instruments and were blinded to each other's findings. Vibration sensation was checked over the tips of both great toes using a 128 Hz tuning fork. For this purpose, base of the tuning fork was placed over the distal halux until the patient was unable to feel the vibration. If examiner could feel the vibration for five seconds beyond this, vibration sense of the patient was considered to be impaired. Light touch sensation was tested over plantar surfaces of both great toes with 10g Semmes-Weinstein monofilament. Using an amount of force that just buckled the monofilament, it was applied perpendicular to the skin surface for one second. This sensation was checked at only one point on each foot. Both ankle jerks were also elicited in a lying position. Reinforcement was not used. Responses to all the three tests were classified as being present bilaterally, present unilaterally or absent bilaterally.

Statistical analysis was done with SPSS version 20. Cohen's kappa (k) was calculated for each of the three clinical modalities to determine the agreement between the two examiners in eliciting the clinical signs. Kappa was interpreted as follows: 0.01- 0.20 (slight agreement); 0.21- 0.40 (fair agreement); 0.41- 0.60 (moderate agreement); 0.61- 0.80 (substantial agreement); 0.81- 0.99 (almost perfect agreement).

RESULTS

The study population comprised of 105 patients with mean age of 50.90±11.53 years. Of these, 44 (41.90%) were males. Median duration of diabetes was 3 years (range: 0 to 23 years). Cross tabulation of vibration sensation, touch sensation and ankle jerks elicited by the two examiners are shown in Tables 1- 3, respectively.

Table 1. Vibration sensation findings.

		Examiner 2			Total
		Absent	Present unilaterally	Present bilaterally	
Examiner 1	Absent	2 (1.90%)	4 (3.81%)	1 (0.95%)	7 (6.67%)
	Present unilaterally	2 (1.90%)	2 (1.90%)	6 (5.71%)	10 (9.52%)
	Present bilaterally	4 (3.81%)	3 (2.86%)	81 (77.14%)	88 (83.81%)
Total		8 (7.62%)	9 (8.57%)	88 (83.81%)	105 (100%)

Table 2. Light touch sensation findings.

		Examiner 2			Total
		Absent	Present unilaterally	Present bilaterally	
Examiner 1	Absent	9 (8.57%)	4 (3.81%)	1 (0.95%)	14 (13.33%)
	Present unilaterally	1 (0.95%)	4 (3.81%)	2 (1.90%)	7 (6.67%)
	Present bilaterally	6 (5.71%)	11 (10.48%)	67 (63.81%)	84 (80.00%)
Total		16 (15.24%)	19 (18.10%)	70 (66.67%)	105 (100%)

Table 3. Ankle jerk findings.

		Examiner 2			Total
		Absent	Present unilaterally	Present bilaterally	
Examiner 1	Absent	27 (25.71%)	8 (7.62%)	6 (5.71%)	41 (39.05%)
	Present unilaterally	6 (5.71%)	7 (6.67%)	3 (2.86%)	16 (15.24%)
	Present bilaterally	2 (1.90%)	2 (1.90%)	44 (41.90%)	48 (45.71%)
Total		35 (33.33%)	17 (16.19%)	53 (50.48%)	105 (100%)

Table 4. Kappa values for different physical signs.

Variable	*k* (95% CI)	Interpretation
Vibration sensation	0.330 (0.151, 0.509)	fair agreement
Light touch sensation	0.452 (0.293, 0.611)	moderate agreement
Ankle jerks	0.581 (0.454, 0.709)	moderate agreement

k= Cohen's Kappa; CI= confidence interval

The two examiners agreed the most on ankle jerks, whereas the lowest agreement was in the use of tuning fork. Cohen's kappa values calculated for each of the three modalities and their interpretation is shown in Table 4.

DISCUSSION

Methods to assess diabetic peripheral neuropathy should be reproducible and should have rather low interobserver variation.[7] These bedside methods detect 85% of insensate diabetic foot, whereas specialized tests are usually reserved for selected patients.[8] Many clinical studies revolve around multiple persons collecting data on patients. Due to the naturally occurring human variability, achieving complete agreement between data collectors is next to impossible. However, reducing bias in clinical trials requires limiting the degree of disagreement between these individuals. Of the various statistical methods measuring interobserver agreement, we have used kappa in this study, because it measures agreement beyond that possible by chance alone.[9] Moreover, we have interpreted kappa as historically suggested by many studies.[10]

Our study has demonstrated fair agreement between the two clinicians in eliciting vibration sensation and moderate agreement in checking light touch sensation and ankle jerks in patients with type 2 diabetes. Similar results have been reported by Maerk et al.[11] In this multicentre study, independent observers carried out sensory examination on 304 diabetic patients. There was a fair agreement in vibration testing and moderate agreement in monofilament and ankle jerk between different observers with kappa scores of 0.31 for tuning fork, 0.59 for monofilament and 0.59 for ankle jerk. Based on these results, the authors recommended monofilament testing as a choice to screen for diabetic foot peripheral neuropathy.[11]

A study on 50 patients, two observers, blinded to each other, applied these three clinical methods on 100 limbs and found kappa values for use of tuning fork were almost the same (0.35; fair agreement), however, values for monofilament and ankle jerks were reversed as compared to ours (0.53 and 0.45 respectively).[12] A study by Edelman et al on 147 diabetic patients found moderate agreement between primary care providers and foot care specialists in testing for light touch using monofilament (k= 0.48).[13] These results are same as ours. A study done on 512 dental professionals revealed kappa of 0.16 for monofilament examination in picking up peripheral neuropathy.[14] However, diabetics made up less than 5% of this cohort. None of these studies had results better than ours.

It has been proposed that kappa values less than 0.60 indicate inadequate agreement among the observers and little trust should be placed in the results of a given study.[15] The two examiners were unable to meet these targets. This suboptimal agreement calls for appropriate training and coordination between data collectors so as to reduce the variability in their judgment. Without this effort, there would be a lot of observer bias affecting the findings.

Like most other single-center studies, it might not be possible to generalize the results of this study to other health care setups. The two doctors performing the tests were from different specialities and had differing work experience. It was not possible to overcome this situation since our setup is a very small one, with only one specialist in each speciality. Nevertheless, the strength of our study lies in its design. All diabetic patients enrolled during this study period were examined by both the doctors on the same day using identical sets of instruments. Both of them remained blinded to each other's findings.

CONCLUSION

This study has shown moderate level of agreement amongst the two examiners in demonstrating light touch and eliciting ankle jerks in patients with diabetes mellitus type 2. There was a fair agreement in using the tuning fork. Considering the lack of substantial conformity, it is suggested that appropriate training may be imparted to data collectors in eliciting vibration sensation before starting clinical trials so as to reduce the amount of variability in their findings.

Author contributions:
Conception and design: ARA
Collection and assembly of data: ARA, KH
Analysis and interpretation of the data: KH, JM
Drafting of the article: KH, JM
Critical revision of the article for important intellectual content: ARA
Statistical expertise: ARA
Final approval and guarantor of the article: ARA
Corresponding author email: Dr Abdul Rehman Arshad: maj.abdulrehman@gmail.com
Present address: Military Hospital, Rawalpindi
Conflict of Interest: None declared
Rec. Date: Nov 20, 2015 Accept Date: Jan 25, 2016

REFERENCES

1. Rani PK, Raman R, Rachapalli SR, Pal SS, Kulothungan V, Sharma T. Prevalence and risk factors for severity of diabetic neuropathy in type 2 diabetes mellitus. Indian J Med Sci 2010;64:51-7.
2. Yagihashi S, Mizukami H, Sugimoto K. Mechanism of diabetic neuropathy: where are we and where to go? J Diabetes Investig 2011;2:18-32.
3. Kanji JN, Anglin RE, Hunt DL, Panju A. Does this patient with diabetes have large- fiber peripheral neuropathy? JAMA 2010;303:1526-32.
4. Abbott CA, Carrington AL, Ashe H, Bath S, Every LC, Griffiths J, et al. The North-West Diabetes Foot Care Study: incidence of, and risk factors for, new diabetic foot ulceration in a community-based patient cohort. Diabet Med 2002;19:377-84.
5. Craig AB, Strauss MB, Daniller A, Miller SS. Foot sensation testing in the patient with diabetes: introduction of the quick & easy assessment tool. Wounds 2014;26:221-31.
6. American Diabetes Association. Standards of medical care in diabetes-2008. Diabetes Care 2008; 31:S12S54.
7. Dioquino CPC, Dellosa MAA, Reyes JPT, Panganiban LCR. Usefulness of Monofilament Testing for Detecting Peripheral Neuropathy I. Acta Med Philipp 2009;43:4-8.
8. Jawad SA, Jawaid M. Management of diabetic foot ulcers: some bitter facts and harsh realities. Pak J Med Sci 2006;22:97-100.
9. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85:257-68.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
11. Smieja M, Hunt DL, Edelman D, Etchells E, Cornuz J, Simel DL. Clinical examination for the detection of protective sensation in feet of diabetic patients. J Gen Intern Med 1999; 14:418-24.
12. Taksande B, Ansari S, Jaikishan A, Karwasara V. The diagnostic sensitivity, specificity and reproducibility of the clinical physical examination signs in patients of diabetes mellitus for making diagnosis of peripheral neuropathy. J Endocrinol Metab 2011;1:21-26.
13. Edelman D, Sanders LJ, Pogach L. Reproducibility and accuracy among primary care providers of a screening examination for foot ulcer risk among diabetic patients. Prev Med 1998;27:274-8.
14. Wang Y, Goodrich JM, Werner R, Gillespie B, Basu N, Franzblau A. Agreement between clinical screening procedures for neuropathy in the feet. Muscle Nerve 2012;45:653-8.
15. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012; 22:276-82.