

Repeatability of nerve thickness assessment in the clinical examination for leprosy

C. KOLAPPAN,* R. SELVARAJ,
ABDUL KHUDOOS, B. N. APPE GOWDA,
MANJULA DATTA & R. PRABHAKAR
*Tuberculosis Research Centre, Epidemiology Unit, Tamil Nadu,
India*

Accepted for publication 2 December 1994

Summary The assessment of the thickness of the superficial peripheral nerve trunks to document nerve involvement is an important aspect of clinical examination in case finding for leprosy, and is usually done by trained paramedical workers (PMWs). This assessment is subject to variability and has implications on the outcome of the survey. The present study proposes to quantify this variability. In this study, 242 individuals, consisting of 50 neuritic cases, 143 nonneuritic cases of leprosy and 49 normal controls, selected from the records of the trial of BCG prophylaxis in leprosy in South India, were examined by a doctor and paramedical workers. Repeatability of nerve thickness assessment for ulnar and popliteal nerves between the medical officer (MO) and the PMWs was quantified using Kappa statistics. The Kappa values for repeatability between the MO and the PMWs ranged from 0.45 to 0.54 and 0.52 to 0.69 for ulnar and popliteal nerves, respectively. The implications of the variability in nerve assessment are discussed.

Introduction

Leprosy surveys very often employ trained PMWs for screening the population. The MO examines the individuals diagnosed by the PMWs as having leprosy for confirmation. The clinical examination for the diagnosis of leprosy includes the palpation of superficial peripheral nerve trunks, especially the ulnar and popliteal nerves in order to assess their thickness. The proper assessment of nerve thickness is essential for the diagnosis and classification of disease. The clinical examination by palpation is the only way to diagnose nerve thickening and, in neuritic leprosy, the assessment of nerve thickening is often the only means of diagnosis.

Since the clinical assessment of nerve thickness is not based on objective criteria, variability between 2 assessors or the same assessor at 2 different times could be

* Correspondence: Tuberculosis Research Centre, Epidemiology Unit, Madras 600 031. India.

expected. Neuritic leprosy constitutes about 14% of the prevalent cases of leprosy.¹ But the proportion of leprosy patients with peripheral nerve trunk involvements will be much higher. Therefore the variability in clinical assessment of nerve thickening is likely to have a considerable influence on the estimates of prevalence rates so obtained. This study presents an attempt to actually quantify this variability.

Materials and methods

Since the prevalence of thickened nerves among the general population is very low, the study sample is selected to include a fair proportion of individuals with thickened nerves. We selected 242 subjects, consisting of 36 (21%) patients with neuritic leprosy, 143 (59%) with nonneuritic leprosy and 49 (20%) normal controls, from the records of the South Indian Chingleput trial of BCG prophylaxis in leprosy. The assessment of repeatability was carried out between 8 PMWs and 1 MO who were engaged in the BCG Prophylaxis Trial in leprosy and had at least 10 years experience in the clinical examination for leprosy. Out of the 244 selected, 242 were examined by the MO (2 were not available for examination). Following the examination by the MO of an individual, each of the 8 PMWs examined him/her independently. The interval between the MO and PMW examination for any individual varied between 0 and 7 days. The PMWs recorded their findings independently on a separate preceded sheet without knowing the tidings of the other PMWs and the MO. Each single examination was independent of any other examination.

The findings of the clinical examination were recorded on the prescribed form in a preceded format for each individual. While assessing the nerves, both thickness and consistency were assessed and recorded.

STATISTICAL ANALYSIS

The measurement of interexaminer agreement

Suppose that each of a sample of *n* subjects is rated independently by 2 examiners on a categorical scale consisting of 2 categories:

		First examiner	
		Positive	Negative
Second examiner	Positive	a	b
	Negative	c	d

A measure of agreement is the mean pair agreement index P_o , which is obtained as

$$\begin{aligned}
 P_o &= \text{number of agreements/Total No. of pairs} \\
 &= (a+d) / (a+b+c+d)
 \end{aligned}$$

This varies considerably with the prevalence. Any characteristic with a low prevalence will have a higher value for d and a higher repeatability, i.e. the value of P_0 is inflated. Thus it is clear that except in the most extreme circumstances some degree of agreement is to be expected by chance alone.

Let us consider an index that assumes the value 1 when there is complete agreement. Let P_0 denote the observed value of the index. Let P_e denote the value expected on the basis of chance alone.

The obtained excess beyond chance is $P_0 - P_e$, where the maximum possible excess is $1 - P_e$. The ratio of these 2 differences is denoted by Kappa (K) and is defined by

$$K = (P_0 - P_e) / (1 - P_e).$$

Kappa is a measure of a chance corrected agreement between 2 ratings. A value of '1' for Kappa indicates complete agreement. A value of '0' implies agreement no better than chance. Negative values indicate more disagreement than expected by chance.

In this study, interobserver variations were measured using the Kappa statistic for the chance corrected agreement as described by Fleiss⁴. This approach was preferred because of its ability to quantify the maximum possible agreement on abnormalities.

Results

In leprosy surveys and control programmes, PMWs are used for screening purposes only. Case diagnoses are made by MOs whose clinical examination is considered as standard. In this study the variability between the standard examiner (MO) and test examiner (PMW) was assessed for each of the 8 PMWs independently. The prevalences of thickened ulnar popliteal nerves of the study population were 23% and 26%, respectively. The prevalences of altered consistency of the same nerves were 55% and 9.2%, respectively. Since the prevalences of altered consistency were low in the study population, the findings on consistency are not reported here. Table 1 shows the distribution of various categories of study population as seen by the MO and PMWs. The study population consisted of a reasonable mix of all types of leprosy in which nerve thickening could be expected. Table 1 shows the number of individuals examined by each PMW. The MO had examined 242 cases, out of which at least 212 had been examined by

Table I. Distribution of the study population according to disease status

Case category	MO	Number examined by the PMWs							
		1	2	3	4	5	6	7	8
Maculo anaesthetic	35	30	30	31	31	29	29	31	30
Tuberculoid	17	16	16	16	15	16	15	15	16
Neuritic	50	44	49	50	48	49	49	49	47
Lepromatous	14	10	11	11	11	10	10	11	10
Borderline	49	39	43	43	41	41	41	43	40
suspect	28	28	28	28	24	28	26	28	26
Normals	49	45	49	49	48	48	49	49	49
Total	242	212	226	228	218	221	219	226	218

Table 2. Interobserver agreement in the assessment of nerve thickening

PMW	No. of nerves assessed	Kappa statistics			
		Ulnar nerve Kappa	(95% CI)	Popliteal nerve Kappa	(95% CI)
1	424	0.50		0.69	(0.60, 0.79)
2	452	0.51	(0.40, 0.60)	0.57	(0.48, 0.66)
3	456	0.54	(0.45, 0.63)	0.59	(0.49, 0.63)
4	436	0.53	(0.44, 0.62)	0.56	(0.47, 0.65)
5	442	0.49	(0.40, 0.58)	0.54	(0.45, 0.63)
6	438	0.45	(0.36, 0.55)	0.55	(0.46, 0.64)
7	452	0.47	(0.38, 0.56)	0.58	(0.49, 0.67)
8	436	0.49	(0.40, 0.58)	0.52	(0.43, 0.61)

each PMW. Table 2 shows the Kappa statistics for ulnar and popliteal nerve thickness assessment between the MO and each PMW. It is seen that the Kappa values for the 8 PMWs for ulnar thickening ranged from 0.45 to 0.54. The corresponding ranges for popliteal nerve thickening was 0.52-0.65.

Discussion

The sample is selected in such a way that a satisfactory level of prevalence of study characteristics can be ensured in the study population. The study was also confined to nerve thickening, since this is the sign that will be used for case detection in the field, and therefore other signs such as nerve tenderness and sensory deficit were not looked for.

The repeatability of a qualitative measurement like the thickening of a nerve is dependent upon the 2 components of variability, namely the biological variability and measurement variability. Dr Noordeen¹ has reported spontaneous regression of thickened nerves in leprosy patients (biological variability). In the present study, the biological variability is minimized by having the paired examinations for each observer within 7 days, It is therefore reasonable to assume that the interobserver variation in this study is almost entirely due to the measurement variability. There is a certain inherent lack of precision in the method adopted for the assessment of nerve status. Obesity, occupation and the size of the corresponding nerve in the contralateral limb are the factors that influence the decision on nerve thickness status. Since the above-mentioned criteria are subjective in nature rather than objective, the repeatability as expressed by the Kappa statistic is not very good. Earlier studies have studied the repeatability for the diagnosis of leprosy between MOs and between the MO and senior PMWs, respectively.

Neelan *et al.*³ have studied the repeatability of diagnosis and classification of early lesion of leprosy among medical officers. They have not,-however, studied the repeatability of nerve thickening assessment in that study. Gupte *et al.*² reported Kappa values (K=0.78) for agreement between 3 pairs of examinations for assessment of nerve thickness with sensory deficit. They had studied nerve thickening not in isolation but as a part of total clinical examination. However, they did not study ulnar and popliteal nerve

thickening separately as was done in this study. The objective of the present study was to view nerve thickening as a clinical sign by itself. Therefore a complete examination of the patients was not carried out and only their nerves were examined. This was expected to eliminate expectation bias due to knowledge of the case status of the individual. In the present study the Kappa values for ulnar thickening range from 0.45 to 0.54 and the popliteal thickening from 0.52 to 0.69. According to the classification of Landis & Koch⁵ Kappa values of between 0.4 and 0.6 could be taken as moderate agreement. We note that agreement is better for popliteal thickening than for the ulnar thickening, but for both nerves it is still only moderate.

In view of the above findings, it would appear that one cannot expect more than a moderate level of reliability for a clinical examination undertaken by PMWs with respect to nerve thickening. However, standardizing PMWs and selecting those with high levels of interobserver agreement would minimize measurement errors, especially in rapid prevalence surveys.

Conclusion

The assessment of thickness of ulnar and popliteal nerves is an important aspect of routine clinical examination. Leprosy screening examination is usually carried out by trained PMWs. Each PMW acts as an independent screening test and the MO acts as the confirmatory standard test. In this study, the repeatability of nerve thickening assessment between the MO and the PMWs was studied. The study shows that the Kappa statistic for interexaminer repeatability varies between 0.45 and 0.69. The present study underlines the fact that this clinical sign (nerve thickening) assessed in the clinical examination for leprosy is a soft parameter and even experienced workers show considerable variability in its assessment. This may be kept in mind while assessing prevalence in large surveys.

References

- ¹ Noordeen SK. Epidemiology of (Poly) neuritic type of leprosy. *Lepr Ind.* 1972; **44**: 90-6.
- ² Gupte *et al.* Inter observer agreement and clinical diagnosis of leprosy for prophylaxis studies. *Ind J Lepr.* 1990; **62**: 281-95.
- ³ Neelan PN, Noorden S.K., Ramu.G., Desikan K.V., Prabu KPM., Christian M. Interobserver variation diagnosis and classification of early lesions of leprosy. *Lepr Ind*; 1982, **54**: 485.
- ⁴ Fliess. *Statistical methods for rates and proportions.* 1984:
- ⁵ Landis, Koch. Measurement of inter observer variation in a categorical data. *Biometrics* 1977; **33**: 159-74.