Clinical validation of methods of diagnosis of neuropathy in a field study of United Kingdom sheep dippers

D Buchanan, G A Jamal, A Pilkington, S Hansen

Objectives: To investigate the reproducibility of measured indices of chronic peripheral neuropathy from a field study of sheep dippers when compared with similar measurements carried out in a clinical setting.

Methods: A stratified random sample of field study subjects was invited to attend a clinic. Neuropathy was measured both in the field and at the clinic with a modified version of a standard symptoms questionnaire and quantitative sensory thresholds for hot, cold, and vibration. These were combined into a classification of the likelihood of neuropathy with a neuropathy scoring system. Indicators of sensory abnormality were based on comparison of sensory thresholds to age dependent reference values derived from an external reference group.

Results: Only 51% of subjects were assigned similar classifications in the field and clinic based on the neuropathy scoring system. Of the component indices, grouped symptom scores, with 65% of subjects showing exact agreement, proved to be more reproducible than quantitative sensory test indicators. There were biases in the comparison of field and clinic measurements of hot and vibration sensations, but no evidence of greater variation between individual people in sensory thresholds in the field relative to at the clinic.

Conclusions: The neuropathy scoring system proved to be of limited reproducibility, due in a large part to the lack of reproducibility of the indicators of sensory test abnormality caused by inadequate temperature control. However, the symptoms score and measured sensory thresholds could be used separately as indices of neuropathy in exposure-response analyses.

The Mayo Clinic method for diagnosing neuropathy has only been validated in a clinical setting with professional staff. The current study allowed the opportunity to compare the method in the clinic and the field.

METHODS

Subject selection
Selection of field study subjects for invitation to the clinic (at the Institute of Neurological Sciences, Glasgow) was by stratified random sample with the strata corresponding to predicted diagnosis of neuropathy based on field survey results as described later. Eighty subjects were invited from each of the none and possible neuropathy categories. Due to small numbers in the group, all 44 sheep farmers from the probable or definite group were invited. In the letter of invitation, subjects were given some indication of their results from the field study. Non-responders were followed up by phone about 2 weeks after the initial invitation. Only farmers or farm workers were invited to attend, although this included those who had never been exposed to sheep dips.

In the field study, subjects were assessed in farm premises, usually the living quarters. Surveys were carried out in the winter months to avoid distortion by acute effects resulting from recent dipping, which normally took place over the summer. Clinical assessment took place at a single clinic (the Institute of Neurological Sciences, Glasgow) between 7 months and 18 months after the field survey, with a mean (SD) gap of 13.4 months (1.9 months). As the field study had investigated indices of chronic peripheral neuropathy, it was

Abbreviations: O Ps, organophosphates; Ip, index of nerve pathology; QST, quantitative sensory tests
considered likely that clinical subjects would not show any major changes in their neurological deficit during the interval between the field and clinical studies.

**Diagnosis of neuropathy**

The method used to diagnose neuropathy in the field was a modified version of the Mayo Clinic method\(^1\) which was originally designed to detect the prevalence and severity of polyneuropathy among subgroups of the general population in a hospital setting. The full battery of clinical tests included in the Mayo Clinic method consisted of a neurological symptoms score, neurological disability score, quantitative sensory tests (QST), and nerve conduction studies. It was considered that only the neurological symptom score and the QST components, both after modification, could be used in the field study by non-clinical personnel.

**Symptoms questionnaire**

A modified version of the Mayo Clinic symptoms questionnaire was devised specifically to detect chronic neurological effects which may be associated with exposure to OP pesticides.\(^1\) Peripheral nerves were assessed for muscle weakness and sensory functions in the upper and lower limbs. Both negative and positive sensory symptoms were investigated. The autonomic nervous system questions considered symptoms occurring within specific organ groups—for example, bladder function or gastrointestinal symptoms. The modification excluded questions involving the cranial nerves and included others to ensure that symptoms were only scored if they occurred bilaterally. Only symptoms that had been present for at least 1 month within the previous year were recorded. Symptoms were summed to give a symptoms score, although autonomic symptoms were down weighted by a factor of 0.5 to increase specificity. The questionnaire was administered in the field by a trained technician and in the clinic by a neurologist.

**Quantitative sensory tests**

Two thermal sensory tests, for hot and cold sensation, and a vibration sensory test were used in both field and clinic.

The Somedic Vibrameter was used for measurement of vibration threshold.\(^7\) This equipment follows physiological principles by indicating the stimulus strength as the amplitude of vibration and not the voltage input to the vibration transducer.\(^7\) The vibration threshold, which tests the fibre function of large peripheral nerve, was measured over the right metatarsal bone in the foot.

The method for measuring thermal thresholds has been developed at the the Institute of Neurological Sciences, Glasgow\(^8\) and the instruments used were manufactured by Medelec.\(^8\) This assesses the function of the small peripheral nerve fibres, cold threshold for small diameter myelinated A\(^\delta\) fibres, and hot threshold for small diameter unmyelinated (C) fibres. The thresholds were measured over the dorsum of the foot and calculated as the change from basic skin temperature (34°C) under the probe. The foot skin temperature was measured at the start of the assessment, and if lower than 31°C was warmed using warm water.

The QSTs were administered in the field by a trained technician\(^7\) and in the clinic by the scientist who developed the thermal threshold technique. The same equipment models were used in both the field and clinical surveys. In the field study, where ambient temperatures were not easily controlled, it was often necessary to warm the foot during and before thermal sensory testing.

Each QST threshold was scored as normal or abnormal by comparison with clinical reference values. The reference values for the thermal sensory tests were calculated with a sample of 68 healthy male and female volunteers aged between 16 and 76 to estimate the upper 95th percentiles adjusted for age. For the vibration threshold, age dependent reference values were obtained from the handbook for the Somedic Vibrameter based on a study of 100 normal subjects.

**Neuropathy scoring system**

A neuropathy scoring system, based on the Mayo Clinic method, was used to classify subjects into categories of none, possible, probable, and definite neuropathy (table 1). It was originally intended that all three QST results would be used in the scoring system. However, after initial investigation of the field survey results it was found that there was an unexpectedly high number of apparently abnormal cold thresholds in comparison with clinical reference values.\(^1\) This resulted in many subjects falling into the possible neuropathy category, even among the non-exposed groups. This was thought to be due to the adverse effect of cold ambient temperatures on thermal thresholds in the field and therefore the inapplicability of hospital based clinical reference values. Therefore, to ensure a wide selection of true disease states among those attending the clinic, a modified scoring system was used for selection that excluded results from the cold QST.

**Statistical methods**

Agreement between field and clinical classifications of neuropathy was investigated, initially by cross tabulation. More formal comparisons were made using \(\chi^2\) tests of association and \(\kappa\) statistics.\(^11\) The \(\kappa\) statistic is a unitless measure for the level of agreement between two categorical variables where 0 represents no agreement and 1 perfect agreement. With ordered categories, a weighted \(\kappa\) statistic takes account of the level of disagreement in terms of number of categories by which the two classifications differ, and weights accordingly. Altman\(^12\) presents guidelines for interpreting \(\kappa\) statistics which indicate that values less than 0.20 signify poor agreement. The level of agreement between continuous variables was assessed with scatter plots, linear correlation coefficients, and the method of paired differences,\(^7\) after log transformation of thresholds to normality. Statistical analysis was carried out with Genstat for Windows version 5.4.1.\(^7\)

**RESULTS**

**Study group**

Seventy nine subjects attended the clinic, most of whom were men (87%), and all of whom had been exposed to sheep dips. Ages ranged from 20 to 66 years with a mean (SD) of 45 (11.2) years. There was no evidence of a difference in recruitment rates from the two regions surveyed, with 65% of those attending being from Scottish farms. Subjects drank on average 9.5 units of alcohol a week (SD 8.6 units).

**Neuropathy scoring system**

The comparison of field and clinical classifications of neuropathy is shown in table 2. The probable and definite categories have been combined due to the small numbers in the definite category. Forty (51%) of the 79 subjects were classified in the same category in both the field and clinic. A \(\chi^2\) test of association was significant (\(p=0.02\)). However, agreement based on the \(\kappa\) statistic (\(\kappa=0.26\), SE 0.08) would only be described as fair based on the guidelines of Altman.\(^12\) The weighted \(\kappa\) statistic was very similar (\(\kappa=0.27\)), indicating that the level of disagreement was
as likely to be by two categories as one. There was no evidence of bias in the direction of disagreement, 18 (46%) of the 39 off diagonal elements being in the direction of greater likelihood of neuropathy in the clinic compared with the field. However, 11% were classified at opposite ends of the diagnostic scale on the separate occasions.

**Symptoms score**

Table 3 shows the cross tabulation of symptoms score in the field and clinic. Symptoms scores have been grouped into three categories that correspond to how it is used within the overall neuropathy scoring system. Exact agreement of the grouped neuropathy symptom score was found for 51 (65%) of the 79 subjects. Overall agreement was significantly better than chance, based on the $\chi^2$ test of association ($p<0.001$). The unweighted $\kappa$ statistic, $\kappa = 0.37$ (SE 0.10), points to reasonable reproducibility of the questionnaire between field and clinic, whereas the higher value of the weighted $\kappa$ of 0.46 suggested that often disagreement was only by one, rather than two, categories.

**Agreement of QST outcomes**

Among the three QST outcomes, the proportion of subjects for which field and clinical outcomes were concordant, representing exact agreement, ranged from 60%–68% (table 4). This was significantly better than chance for the heat and vibration tests, but not the cold test ($p=0.10$). Relative to chance, overall agreement was highest for the vibration test ($\kappa = 0.30$, SE) than for either the hot ($\kappa = 0.22$) or cold tests ($\kappa = 0.18$), but, in all three, was poorer than for the symptoms score.

**Measured QST thresholds**

A comparison of the actual measured sensory test thresholds is more informative about magnitude of differences between field and clinic than comparing binary test outcomes. Figure 1 shows scatter plots of the three QST thresholds, on the log scale, as they were measured in both field and clinic. These show, for each test, a positive linear correlation between field and clinic. Linear correlation coefficients were 0.71, 0.44, and 0.66 for the log transformed hot, cold, and vibration thresholds respectively. There was evidence of bias in the hot and vibration, but not the cold, measurements (table 5). Hot thresholds tended to be lower by a factor of two, and vibration thresholds higher by a factor of two, in the field relative to the clinic. Adjusted for these biases, differences between field and clinic varied within subjects with a GSD of about 2.8 across all three sensory thresholds. Based on the ratio of GSDs, variation between subjects in the field was comparable with that in the clinic for cold and vibration thresholds (0.99 for cold, 0.96 for vibration) and was only slightly higher in the field for hot thresholds (1.25).

**Other factors**

In the main epidemiological phase of the study it was found that, for the same age, sex, and exposure, farmers in the English regions reported symptoms more often than farmers in Scotland, with an odds ratio (OR) of 2.0. Therefore, the reproducibility of symptoms (categorised as <1 and ≥1 due to generally low prevalence of any symptoms) in the clinic was analysed by region. Reproducibility among the 26 English farmers ($\kappa = 0.55$; SE 0.16) was found to be better than that...
among the 53 Scottish farmers (κ 0.24; SE 0.14), principally because the Scottish farmers tended to underreport symptoms in the field relative to the clinic. The reason for this is unclear but may have reflected greater media exposure in England about the health effects of organophosphates around the time of the epidemiological field survey. Also in the main epidemiological phase, women of the same age and exposure were found to report symptoms more often than men (OR 2.9). However, there was no evidence of markedly different reproducibility of symptoms between men and women, although data were only available for 11 women who attended the clinic.

The biases noted for the hot and vibration thresholds in the field relative to the clinic were found to be consistent when subjects were categorised by region and sex.

**DISCUSSION**

The current study allowed the opportunity to compare a modified version of the Mayo Clinic method for diagnosis of neuropathy in both a field and clinic setting. The response rates, representativeness, and the selection of the subgroup of subjects for the clinical study from the epidemiological survey has been described elsewhere. The question of response rates and representativeness is less important as this paper only compares the findings from the field and clinic measurements in the same subjects.

The reproducibility of the overall neuropathy scoring system was only marginally better than chance, with a significant minority of subjects being classified at opposite ends of the scale on the two separate occasions. Therefore, to elucidate this finding, the reproducibility of symptoms score and QST score components of the classification system were assessed separately.

The symptom scores showed an agreement that was significantly better than chance and there was no evidence of bias between the scores obtained in the field and at the clinic. Furthermore, disagreements were often by one category rather than two.

The QST scores reflected an inconsistency between the QST field measurements, especially cold, and the hospital based clinical reference values from which abnormality was defined. There were an unexpectedly high number of abnormal cold thresholds in the field, even among the non-exposed control group of ceramics factory workers (48%), suggesting a methodological problem of this measurement in the field.

Clinical experience with QST measurements show these methods to be very valuable in detecting early nerve damage because of their high sensitivity. For this reason, and because they can be applied by trained field officers, the measurement of QST thresholds was included in the epidemiological study. The repeatability of the thermal sensory tests in the clinic has been evaluated with a group of 320 normal people. The method is used in many centres and the high sensitivity of the technique is supported by a study of 143 patients with neuropathies of diverse aetiologies. A consensus report of the International Panel of the Peripheral Neuropathy Association has endorsed the principle of the technique. About 40 original publications by our group at the Institute of Neurological Sciences, Glasgow and others confirmed the reproducibility of the results and the usefulness of the technique.

Biases in the field QST thresholds were thought to be due to lack of control of the limb temperatures of farmers who were mostly surveyed during winter months, and often just after working out of doors. It was necessary to carry out this study during winter to have a time window in which dipping did not take place and when farmers are able to give up time to take part in a study.

The function of peripheral nerve is highly temperature dependent and the temperature dependency of QST thresholds is well known. In the clinic, room temperature is kept at a constant level, limb temperature measured often, and heating applied if necessary. Steps were taken to rectify this problem in the field study. Digital thermometers for measuring skin temperature were used by the field survey teams. Although heating was applied before the test, it was impossible to maintain the temperature of the foot at a reasonable level during the whole investigation. Skin temperatures measured in the field before sensory testing ranged from 20°C to 35°C, with a median of 31°C, significantly below the optimum 34°C normally used in the hospital setting. Body and core temperatures of limbs may also have been lower than normal.

Receptors in the skin for heat and cold sensation are different and have different temperature dependencies (Kenshalo 1970). In cold temperatures increased sensitivity to heat might reduce hot thresholds, thus biasing field measurements downwards compared with the clinic. Cold temperatures would similarly reduce sensitivity to cold and vibration sensations, both of which were higher on average in the field.

It therefore seems likely that the cold temperatures experienced during the field study is a major factor in the poorer agreement between field and hospital clinic measurements although other causes cannot be excluded.

How can QST measurements be applied in future field studies? It would be difficult to establish a more appropriate reference population. Allowing room and skin temperatures to vary across a larger range would lead to higher variability among thresholds and thus decrease the sensitivity of the method. No data are available to allow correction of threshold values for different temperatures.

The deliberate exclusion of subjects from the clinic, however, with an unreliably high cold threshold, is likely to have reduced the observed bias of clinical measurements relative to those in the field, and this may explain why the observed bias is lower in magnitude than for the vibration threshold. It is therefore likely that the reproducibility of the scoring system, across the study group as a whole, has been underestimated. The time lapse between field and clinic surveys was not thought to be a significant factor when looking at chronic health effects.

Based on these results, in the field study it was decided to use the four neuropathy indices as separate response variables in analyses of exposure-response relations, rather than the neuropathy scoring system based on the Mayo method. In a linear regression framework, a bias in the sensory thresholds that applied independently of exposure, would not affect the

### Table 5: Comparison of field and clinic QST thresholds (log scale)

<table>
<thead>
<tr>
<th>QST</th>
<th>Field</th>
<th>Clinic</th>
<th>Ratio (field/clinic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot</td>
<td>0.77</td>
<td>3.88</td>
<td>1.73</td>
</tr>
<tr>
<td>Cold</td>
<td>0.68</td>
<td>2.76</td>
<td>0.54</td>
</tr>
<tr>
<td>Vibration</td>
<td>3.64</td>
<td>3.35</td>
<td>1.81</td>
</tr>
</tbody>
</table>

G.M., geometric mean; G.S.D., geometric SD; QST, quantitative sensory testing.
Main messages

- The overall neuropathy scoring system showed an agreement that was better than chance with no evidence of a systematic bias between the scores obtained in field and in the clinic.
- The neurology symptom questionnaire in isolation was more reproducible and this method could be used in the field by non-neurologists with basic training.
- There was a significant positive linear correlation between the threshold measurement in the field and the clinic for each QST, but there were biases among the field measurements of sensory tests relative to the clinical measurements.

Policy implications

- The QST measurements should still be considered in epidemiological field studies, but it is essential to enforce much tighter control of conditions, in particular avoidance of low temperatures.

ACKNOWLEDGEMENTS

We gratefully acknowledge the help of all farmers, their families, and farm workers who took part in this study. We also acknowledge the work and diligence of the nursing staff (Miss M MacKinnon and Miss M J Dolan) and clerical staff (Ms E Jackson) who assisted throughout the study at the Institute of Neurological Sciences, Glasgow. We also thank Ms Mags Parker, Mrs Marion Brebner, and Mrs Margaret Burnett of the Institute of Occupational Medicine for their administrative support during the project. Finally we acknowledge, with thanks, the funding from the Health and Safety Executive, the Department of Health, and the Ministry of Agriculture, Fisheries and Food. This work was supported by the Health and Safety Executive, the Ministry of Agriculture, Fisheries and Food, and the Department of Health, UK.

REFERENCES

Sex ratio at birth in Masjidi-Sulaiman (Khozestan province, Iran)

Little is known about the factors that affect the sex ratio in humans. Many animal and human studies have indicated an association between environmental toxins and altered sex ratios. Masjidi-Sulaiman (MIS) is located in the Khozestan province, southwest of Iran. The first oil well in the Middle East was located in MIS (excavated in 1908 by William Darcy). Petroleum—whether gas, oil, or liquid asphalt—that exudes in the form of springs and seepages may reach the surface. Active seepages of oil and gas overlie the MIS oilfield. Seepages of oil and gas may reach the surface. The gas dissolved in the oil of the MIS oilfield contains 40% hydrogen sulphide. It has been reported that the gas dissolved in the oil of the MIS oilfield contains 40% hydrogen sulphide. We determined the sex ratio at birth of families resident in the contaminated area of MIS, and compared this with the sex ratio of the general population of MIS.

We identified 339 offspring within 51 families resident in the above mentioned area, born between 1987 and 2001. Data concerning live births in MIS are routinely compiled, using reports from the Statistical Center of Khozestan province from 1987 to 2001. To test the null hypothesis that the probability of a male live birth in the contaminated area is equal to that in the general population of MIS, the analysis was conducted. A probability of p < 0.05 was considered statistically significant. The sex ratio was expressed as the proportion of total live births that were male. The overall sex ratio in general population of MIS from 1987 to 2001 was 0.504 (total number of live births 44 040). In the contaminated area, the sex ratio was increased (0.563). This difference is statistically significant (χ² = 4.91, df = 1, 0.025 < p < 0.030). We are aware of no other similar reports. Yang et al reported that the association between exposure to petroleum air pollution, near a petroleum refinery plant in Taiwan, and sex ratio at birth was not significant. However, to clarify the effect(s) of subsurface leakage of gas on sex ratio and hormone concentrations in parents, further investigations need to be carried out.

References
another stays healthy have shown the considerable influence of personal factors in the avoidance and causation of disease. It is reasonable to expect that health at work will be subject to the same range of individual factors and that they will be very important in determining the onset of work related diseases due to mechanical or physical stress, to allergens, and to chemicals.

The pleura is frequently involved in systemic disorders in addition to the diseases that originate in the lung or in the pleura itself. Pleural disease is therefore common and likely to become more common with the predicted increase in the incidence of mesothelioma. Common conditions are not always managed well, and pneumothorax and pleural effusions are two conditions where the expertise needed to manage them efficiently and effectively is often underestimated. A book that focuses on pleural disease and practical aspects of management is therefore welcome.

Pleural diseases, a single author book by Richard W Light, is now in its fourth edition. Its 400 pages provide a comprehensive review of all aspects of pleural disease from the anatomy and physiology of the pleura to the range of disorders that cause pleural disease and their investigation and management. It is written by an enthusiast who draws widely from the literature on animals and man, in addition to his considerable clinical experience. The style of writing is conversational and easy to read, although it can be a little repetitive in places. Details of the individual studies are usually given, so the basis for the author’s conclusions is clear. The book provides a comprehensive and up to date reference source.

The chapters on common conditions such as pneumothorax flow well and are easy to follow. The pros and cons of various management options are debated with practical advice on management. The book is not designed to give didactic advice for the busy junior doctor at the sharp end of managing a patient with a pneumothorax or a pleural effusion, for example, but would be invaluable for drawing up guidelines for management of these conditions. The section on thoracoscopy and chest tubes discusses practical aspects of technique and potential complications, although the latter gives less emphasis to trocar tube thoracostomy than would be merited from practice in the United Kingdom.

The book is well presented, although figures and diagrams are somewhat sparse; the radiographs that are present are clear. I would have welcomed more figures, particularly diagrams to show the anatomy and relation of, for example, the thoracic duct and lymphatics. Better use could be made of line drawings—to explain the physiology of fluid collection in the pleural space and where to insert thoracostomy tubes, for example.

This book will clearly be essential for people with a clinical or research interest in pleural disease. It will also be a very valuable reference source for respiratory physicians and doctors in training.

Anne Tattersfield

NOTICES

27th International Congress on Occupational Health: The Challenge of Equity in Safety and Health at Work, Iguassu Falls, Brazil, 23–28 February 2003

The Congress will have about nine keynote conferences, approaching different angles of the Central Theme; those themes will then be discussed in depth by Panels (60), where different opinions will be debated. There will be about 60 mini-symposia organised by the ICOH Scientific Committees and Work Groups; facilities for the presentation of 1000 posters, and about 500 free papers. Interest groups may schedule meetings in Congress areas.

Conference Secretariat
Tel/Fax: (5541) 353 6719
Email: icoh2003@com.br
Website: www.icoh2003.com.br

First World Congress on Work-Related and Environmental Allergy (1st WOREAL), and Fourth International Symposium on Irritant Contact Dermatitis (ICD), Helsinki, Finland, 9–12 July 2003

Congress on Work-Related and Environmental Allergy

• Work related and environmental aspects of respiratory and skin allergy
• Specific issues related to pathophysiology and skin allergy
• Management and prevention of allergy

Irritant Contact Dermatitis Symposium

• Occupational irritant dermatitis
• Prevention of irritant dermatitis
• Alternative methods for the assessment of irritants
• Irritant dermatitis from cosmetics

Satellite events

• Satellite Symposium, 9 July 2003
• Allergy School, 9–10 July 2003
• 7th International NIVA Course on Work-Related Respiratory Hypersensitivity, 11–15 July 2003

Congress Secretariat
Ms Kirs Saarrela, Congress Manager
Pykkö & Saarrela Ltd
Limingantie 9
FIN-00550 Helsinki, Finland
Tel: + 358 9 79 00 80
Fax: + 358 9 757 36 30
Email: secretariat@worreal.org
Website: www.worreal.org

CORRECTION

We apologise for the following errors in the paper by Buchanan et al. (Clinical validation of methods of diagnosis of neuropathy in a field study of United Kingdom sheep dippers. Occup Environ Med 2002;59:442–6):
• D Buchanan and A Pilkington were affiliated with the Institute of Occupational Medicine, Edinburgh, Scotland, UK.
• The email address for the corresponding author is: s.hansen@clinmed.gla.ac.uk.