An investigation of the adequacy of psychiatric interviews conducted through an interpreter

Saeed Farooq, Christopher Fear and Femi Oyebode

We evaluated clinical information gained directly from 10 English-speaking and from 10 non-English-speaking subjects both directly and through interpreter-mediated interviews. High levels of agreement between raters, when assessing both cohorts, were found for all data with a non-significant tendency towards better agreement in the Asian than the English-speaking sample for family history data. Analysis of the interview contents showed a number of errors of interpretation which were similar to those noted in previous studies. The addition of quantitative data represents a significant advantage over previous studies, allowing the qualitative results to be placed into perspective. Recommendations are made for optimising and avoiding the pitfalls of interpreter-mediated interviews.

The communication needs of non English-speaking UK populations have never been systematically assessed. Surveys of British Asians attending hospitals suggested that more than half have difficulty in communicating with doctors and show dissatisfaction with existing interpretation services (Leatherdale *et al*, 1978; Stevens & Fletcher, 1989; Madhock *et al*, 1992). In psychiatry it is often difficult to understand a patient's symptoms or psychopathology even when the therapist understands the patient's language. Interviews performed through an interpreter would be expected to introduce greater difficulties. Sabin (1975), reviewing cases of suicide by Spanish-speaking patients who had been assessed through interpreters, concluded that the patients' emotional suffering and despair had been selectively underestimated in the process of interpretation.

In an interview it must be realised that interpreters do not simply translate what they hear from patients; they interpret. Thus the colloquial English expression "It's raining cats and dogs" could be translated verbatim into any Asian language but would be meaningless. An interpreter, familiar with the expression would imbue it with meaning as the phrase "It's raining very heavily". While the former is a translation, the latter is an interpretation. Our study evaluates and compares information gained by a psychiatrist working through an interpreter and by a bilingual psychiatrist fluent in the patient's own language.

The study

Our sample comprised patients aged 18 to 65 with clinical evidence of psychopathology but without cognitive impairment or speech disorder. Ten were selected who spoke Mirpuri or Punjabi, and whose knowledge of English was absent or rudimentary. These were compared with ten English-speaking patients: nine Caucasian and one third-generation British Afro-Caribbean.

A 30-item check-list was derived from the Present State Examination (Wing *et al*, 1974), but with a rating scale of severity from 0 (absent) to 4 (very severe). To this was added a 12-item family history check-list since factual, family history data would be expected to contain less error than the difficult concepts of mental state examination.

An interpreter was selected who had nationally-recognised qualifications and more than six years' experience. She was informed that the study aimed to compare symptoms elicited from patients when interviewed directly and through an interpreter. English-speaking patients were interviewed jointly by the psychiatrists, and ratings compared after each interview to provide a control-group of patients and thus test inter-rater variability under optimal interviewing conditions. Non English-speaking subjects were interviewed separately by the bilingual psychiatrist (SF) alone and by the English-speaking psychiatrist (CFF) through the interpreter, in all cases within 12 hours of one another. In each case, equal numbers of patients were assigned, randomly, to be interviewed first by each psychiatrist. Interviews were audiotaped and, in the case of those conducted first by the bilingual psychiatrist, the questions used were translated, by him, into English and used as a basis for the second interview. In addition, in the reverse

situation, the interviewing psychiatrist recorded his questions (but not the interpreter's rendering) into a dictating machine.

Verbatim transcripts were prepared by the bilingual psychiatrist and used to compare the English-speaking psychiatrist's questions, the interpreter's interpretations and the patient's responses. Inter-rater agreement for continuous variables (mental state items) was calculated using 95% confidence intervals applied to the mean difference scores between rating 1 (English-speaking psychiatrist via interpreter) and rating 2 (by bilingual psychiatrist) (Bland & Altman, 1986). The mean difference is considered to be significant if zero lies outwith the confidence interval. For dichotomous variables (family history items), agreement between raters was compared using an exact calculation of binomial probabilities with associated 95% confidence intervals.

Findings

There were no statistically significant differences on demographic variables of age, gender or duration of contact with psychiatric services. The Asian sample had a mean stay in the UK of 20.1 (s.d.=8.6) years. Seven were primarily Mirpuri-speaking, the other three Punjabi and one had received an elementary education. The ten English-speaking patients had been born, and educated to secondary school level, in Britain.

Mean differences between ratings for the individual items of mental state examination (MSE) are shown in Table 1 and percentage agreement between raters for the family history (FH) items are shown in Table 2 respectively, together with their respective 95% confidence indices raters. No significant differences were found between the two interviewers' ratings for any of the MSE or FH items. There was a non-significant tendency towards poorer agreement between raters on the family history items obtained from the English-speaking as opposed to the Asian group.

Analysis of content

The following were common errors.

Omission A message is completely or partially deleted.

Table 1. Mean ratings by two interviewers and indices of agreement (95% CI) for the mental state items

Item	English-speaking patients		Asian patients	
	Mean difference in ratings (s.d.)	95% CI (d.f.=9)	Mean difference in ratings (s.d.)	95% CI (d.f.=9)
Affective symptoms				
Anxiety	0.70 (1.16)	−0.13–1.53	0.50 (1.27)	−0.40–1.40
Appetite	−0.50 (0.97)	−1.20–0.20	−0.70 (1.16)	−1.53–0.13
Concentration	−0.60 (1.17)	−0.23–1.43	0.90 (1.60)	−0.22–2.03
Depression	0.60 (1.84)	−0.71–1.90	0.30 (1.57)	−0.83–1.43
Diurnal variation	0.50 (1.08)	−0.27–1.27	1.00 (1.41)	−0.02–2.02
Elation	0.20 (0.63)	−0.25–0.65	−0.40 (1.26)	−1.30–0.50
Guilt	0.10 (0.57)	−0.30–0.51	0.50 (1.08)	−0.27–1.27
Hopelessness	0.00 (0.82)	−0.59–0.59	−0.20 (1.62)	−1.35–0.95
Interest	0.00 (0.82)	−0.59–0.59	−0.10 (0.99)	−0.80–0.60
Irritability	0.60 (0.84)	−0.01–1.21	0.40 (1.35)	−0.57–1.37
Panic	−0.30 (0.48)	−0.64–0.04	0.20 (1.55)	−0.90–1.30
Sleep	−0.20 (0.79)	−0.76–0.36	0.00 (0.94)	−0.68–0.68
Suicide	0.60 (1.35)	−0.37–1.57	−0.10 (1.37)	−1.07–0.87
Weight	0.60 (1.17)	−0.23–1.43	0.60 (1.43)	−0.41–1.61
Perceptual symptoms				
Auditory hallucinations	−0.10 (0.57)	−0.51–0.30	0.10 (1.10)	−0.69–0.89
Somatic hallucinations	0.00 (0.00)	0.00–0.00	−0.50 (1.35)	−1.47–0.47
Visual hallucinations	0.00 (0.00)	0.00–0.00	0.00 (0.00)	0.00–0.00
Thought symptoms				
Compulsion	0.30 (0.95)	−0.37–0.97	−0.10 (1.66)	−1.29–1.09
Delusions	0.10 (0.88)	−0.53–0.73	−0.20 (1.32)	−1.15–0.74
Formal disorder	0.00 (0.00)	0.00–0.00	0.10 (0.32)	−0.13–0.33
Ideas of reference	−0.20 (1.03)	−0.94–0.54	1.10 (1.97)	−0.29–2.50
Thought insertion	−0.10 (0.32)	−0.33–0.13	0.70 (1.49)	−0.36–0.76
Thoughts fast	−0.80 (1.75)	−2.04–0.44	0.20 (0.79)	−0.36–0.76
Thoughts slow	−0.20 (1.55)	−1.30–0.90	0.40 (2.12)	−1.11–1.91
Obsession	0.20 (0.63)	−0.25–0.65	−0.20 (1.62)	−1.35–0.95

Table 2. Percentage agreement between two interviewers for the family history data

Item	English-speaking patients		Asian patients	
	% agreement	95% CI (d.f.=9)	% agreement	95% CI (d.f.=9)
Fater				
Alive (yes/no)	60	26.2–87.8	100	69.2–100
Age (exact)	70	34.8–93.3	100	69.2–100
Age (within 5 years)	90	55.5–99.8	100	69.2–100
Well (yes/no)	70	34.8–93.3	90	55.5–99.8
Good relationship (yes/no)	100	69.2–100	100	69.2–100
Mother				
Alive (yes/no)	90	55.5–99.8	90	55.5–99.8
Age (exact)	90	55.5–99.8	90	55.5–99.8
Age (within 5 years)	100	69.2–100	100	69.2–100
Well (yes/no)	100	69.2–100	100	69.2–100
Good relationship (yes/no)	90	55.5–99.8	100	69.2–100
Number of brothers	80	44.4–97.5	100	69.2–100
Number of sisters	80	44.4–97.5	100	69.2–100
Position in family (eldest, youngest etc.)	70	34.8–93.3	100	69.2–100
Family history of mental illness (yes/no)	90	55.5–99.8	90	55.5–99.8

CF: "How many brothers and sisters do you have"?
Interpreter: "How many sisters do you have?"
Patient: "Four"
CF: "Four of each?"
Interpreter: "You said four"
Patient: "Yes".

Eight siblings were recorded, although the patient had four sisters only.

Substitution A concept or theme was replaced by another. A patient was asked: "Do you ever feel that you would like to go to sleep and not wake up?" This was interpreted in terms of a desire for good sleep rather than a suicidal idea.

Condensation A tendency to simplify and explain a complicated and lengthy response. This is important in considering thought disorder. Thus the response: "When I was born I have left land, land of India, Handsworth and Bengal, Prime Ministers sign, nations kept fighting, Rajah came to me . . ." was condensed into an assertion of involvement with Prime Ministers and the fighting between nations giving the impression of a grandiose idea.

Similar phonetic sounds The use of words borrowed from other languages may result in error where there is a similar-sounding word in the patient's native language. One of our patients was asked if she felt guilty. The interpreter, unable to find a straightforward construction for the word "guilty" or perhaps thinking that the patient might be familiar with the English word, used the word "guilty" in her interpretation. Interestingly, there is a similar-sounding word in Punjabi which means "swelling". The patient's response, in Punjabi, made it obvious that she did not feel a swelling!

Conceptual errors A patient's response clearly described a delusional perception but, unable to make sense of this only the delusional belief was related.

Closed/open questioning Subtle changes in the way a question is asked also led to errors and were more likely where a patient gave predominantly "I don't know" responses".

CF: Do you feel happy or sad in your spirits?"
Patient: "If I am not unhappy or sad . . . (pause) . . . then I am happy"
Interpreter (without interpreting this response): "You feel sad now?"
Patient: "Yes"
Interpreter: "She is unhappy".

Comment

Previous studies which have examined the problems associated with using interpreters in medical settings (Marcos *et al*, 1973; Launer, 1978; Westermeyer, 1989), have concluded that errors occur in the elucidation of clinical information by such methods. These studies have tended to be qualitative, retrospective or both in design. Further, they have often relied upon an evaluation of patients' satisfaction with the service they received, or upon specific examples of communication failure, rather than quantifying areas of discrepancy in the communication process as a whole.

Our findings support the qualitative discrepancies discussed by previous researchers but, by using a quantitative measure of agreement

The language barrier

211

between raters, we have shown that such difficulties produce only minimal errors in eliciting information. The interviews were rendered more comparable through our use of the questions asked by the first interviewer as a basis for the questions of the second interviewer. This allowed for unstructured interviews with considerable freedom in phrasing the questions, while ensuring that each interview covered essentially the same ground. The joint interviewing of the English-speaking patients may be criticised in that it produced identical terms of reference for the two interviewers, whereas interviews of the Asian patients were conducted separately and would therefore expect to give rise to less agreement between raters. In practice, no such discrepancies were discovered. Our use of a transcript of the first interviewer's questions as the basis for the second interview where patients were interviewed separately may have aided this. Additionally, where separate interviews were conducted, they were within the space of the same day so as to exclude errors due to change in phenomenology over time. Also, interviewing the English-speaking sample jointly allowed the two raters to become more aware of one-another's interviewing and rating style so as to ensure a closer match for the separate interviews of the Asian patients.

It should be emphasised that we used a qualified professional interpreter whose experience greatly increased the value of the information obtained. On many occasions she modified questions and answers, without significantly altering their content or the concepts, to clarify them so as to convey the point in such a way as to be acceptable and comprehensible to the recipient. However well-qualified an interpreter is, certain sources of error can be identified. The most often overlooked is that patient and interpreter, especially if the latter is British-born may share a common language but different cultural values.

Clinicians who are not cognisant of the task may be equally, if not more, responsible for distortions in interpretation. Speaking too quickly, introducing long sentences, addressing the patient in the third person are just a few examples of the problems which clinicians can pose. The use of technical language which may not readily be translated and phrasing of questions in an indirect fashion were sources of error in this study for which the clinician was primarily responsible. In many emergency circumstances an experienced interpreter will not be available so that the use of relatives or other patients as interpreters may be considered. There is little justification for this, however, as it raises considerable difficulties including issues of confidentiality and the transgression of family boundaries. Moreover most units, particularly in areas with ethnic minority populations, provide lists of interpreters who can be contacted in an emergency.

We believe that this is the first empirical study to evaluate the effects of interpreter-mediated interviews in a clinical setting. There are obvious methodological limitations, including small sample size. A further limitation may be our use of an English-speaking sample for comparison and it may have appeared more valid to compare a bilingual Asian sample interviewed in English since these would, at least, have similar cultural traditions. This would, however, have introduced other problems. Previous studies have noted that bilinguals give a different account of symptomatology when interviewed in their favoured and second languages (Marcos *et al*, 1973).

The experience of the English speaking psychiatrist in this study suggests some simple measures which could help to improve effective communication. These include addressing points and questions to the patient directly instead of to the interpreter, thus allowing the clinician to take control of the interview and establish a rapport with the patient. Asking short questions and avoiding excessive jargon are also important, whereas impersonal modes of enquiry ("tell the patient that") can significantly impair the quality of interview. Writing notes during the breaks provided by interpretation can lead to a failure to acquire useful clinical data from non-verbal behaviours, which might not be available in the interpretation of patients' verbal responses. A statement inconsistent with a patient's verbal or non-verbal responses should be clarified by changing its wording, breaking it down into simpler parts or by asking about a related issue. Useful clinical information can be gained by exploring seemingly unconnected issues. In the case of important interview (e.g. for legal purposes), or in case of doubt (e.g. formal thought disorder), it is helpful to tape record the interview so as to ask a bilingual colleague to translate or comment upon it later. Finally, it is helpful to have a pre-interview meeting with the interpreter, clarifying the reason for the interview and the main areas to be covered.

Our findings suggest that interviews conducted through an experienced interpreter, while giving rise to some qualitative distortions, provide a reliable method of collecting information and, together with observations made by the clinician involved, set the basis of a reliable diagnosis. We wish to highlight the importance of using a qualified interpreter. Future studies should concentrate on effectiveness of interpreter-mediated interviews in larger samples, comparing interviews using trained, experienced interpreters with untrained inexperienced interpreters. Finally, it should be stressed that lack of

effective communication with patients should not be regarded merely as an inconvenience to medical practice. Many studies have commented upon higher rates of mental illness in immigrant populations (Cochrane, 1971). In considering these findings, however, the contribution of poor communication has not been raised seriously, neither has this area been studied systematically.

Acknowledgements

References

BLAND, I. M. & ALTMAN, D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, **1**, 307–310.

COCHRANE, R. (1971) Mental illness in immigrants to England and Wales: An analysis of mental hospital admissions. *Social Psychiatry*, **12**, 25.

LAUNER, J. (1978) Taking medical histories through interpreters: Practice in a Nigerian out-patient department. *British Medical Journal*, **2**, 934–938.

LEATHERDALE, B. A., JHITTAY, P. & GILL, A. (1978) Communication with Asian patients. *British Medical Journal*, **2**, 1197–1198.

MADHOCK, R., BHOPAL, R. S. & RAMAJAH, R. S. (1992) Quality of hospital service: A study comparing 'Asian' and 'Non-Asian' patients in Middlesborough. *Journal of Public Health Medicine*, **14**, 271–279.

MARCOS, L. R., ALPET, M., URCUYO, L., et al (1973) The effect of interview language on the evaluation of psychopathology in Spanish-American schizophrenic patients. *American Journal of Psychiatry*, **130**, 549–553.

SABIN, J. E. (1975) Translating despair. *American Journal of Psychiatry*, **132**, 197–199.

STEVENS, K. A. & FLETCHER, R. F. (1989) Communicating with Asian patients. *British Medical Journal*, **299**, 905–906.

WESTERMEYER, J. (1989) *Psychiatric Care of Immigrants in Washington DC.* Washington, DC: American Psychiatric Press.

WING, J. K., COOPER, J. E. & SARTORIUS, N. (1974) *The Measurement and Classification of Psychiatric Symptoms.* London: Cambridge University Press.

Saeed Farooq, *Senior Registrar, Department of Psychiatry, Postgraduate Medical Institute, Lady Reading Hospital, Peshawar, Pakistan;* *Christopher F. Fear, *Consultant in General Adult Psychiatry, Severn NHS Trust, Wotton Lawn Hospital, Horton Road, Gloucester G1 3PX;* Femi Oyebode, *Consultant in General Adult Psychiatry; Queen Elizabeth Psychiatric Hospital, Mindelsohn Way, Off Vincent Drive, Edgbaston, Birmingham B15 2QZ*

*Correspondence