

## EDITORIAL

# Clinical practice, measurement and information technology<sup>1</sup>

Psychiatrists appear to order fewer ancillary tests than other physicians (Andrews & Hadzi-Pavlovic, 1988). Is a clinical diagnosis unsupported by other assessments still good mental health practice? Most branches of medicine make considerable use of ancillary investigations, both to tell clinicians what they cannot know (e.g. cause of the anaemia), or do know, but need better data before acting (e.g. an X-ray to check bone healing). In many countries there is a legal requirement that the basis for the diagnosis should be made explicit before treatment is commenced. In psychiatry these clinical and legal requirements for additional information could be met by the use of structured diagnostic interviews and by questionnaires and rating scales that measure symptom profiles, disabilities and risk factors.

Information does not have to be new to be useful. Forty years ago, Meehl (1954) compared the accuracy of clinical *versus* statistical prediction applied to the same clinical data. He concluded that statistical predictions based on clinical data were at least as accurate as the judgements of trained clinicians. Clinicians were, and frequently still are, indignant at Meehl's conclusions. They rightly considered that as their judgements had been the 'gold standard' against which the tests had been developed, how could a statistical compilation of tests be better than the judgements on which the tests were based? The answer lies in reliability. One clinician's judgement about the same set of clinical data may vary from occasion to occasion, and likewise, such judgements may vary between clinicians. Statistical compilations do not vary in this manner. Reliability is a pre-requisite for validity. If a clinician cannot reach the same diagnosis on two occasions when reviewing the same clinical data then the validity of at least one of the decisions is questionable. If two clinicians given the same clinical data cannot agree, then at best one, and at worst both, will be incorrect. In consequence, the treatment prescribed by one or both of the clinicians may be wrong. Valid judgements depend on reliable judgements, a fact that researchers in psychiatry have learned slowly and painfully.

Reliable judgements are only one part of the story. Clinicians vary considerably in the ways they ask their patients about illness and, consequently, can obtain different clinical information from the same patient. Until the diagnostic criteria were specified, first in DSM-III and now in the ICD-10 Diagnostic Criteria for Research, diagnosis was an art, taught by senior doctors to junior doctors. Under these conditions, while intra-hospital consensus might have been high, inter-hospital and, indeed, inter-country agreement was often low. The uniform diagnostic criteria have helped to change this. Schizophrenia, at least theoretically, now means the same in Beijing as it does in Birmingham, agoraphobia the same in Athens as it does in Atlanta.

## IMPROVING DIAGNOSTIC ACCURACY

A new class of structured diagnostic interviews has been developed to improve how symptoms are elicited and to show how well those symptoms satisfy the diagnostic criteria. Such interviews have usually achieved their goal by modelling a clinical interview, with the gains in reliability coming from standardization of the content of the questions, the form of the interview, and the scoring process. However, no matter how perfect the structured diagnostic interview, validity will be limited by the exactness with which the interview elicits the behaviours, thoughts and feelings described by the diagnostic criteria, and by the inherent validity of the diagnostic criteria themselves.

<sup>1</sup> Address for Correspondence: Dr G. Andrews, Anxiety Disorders Unit, 299 Forbes Street, Darlinghurst, NSW 2010, Australia.

There are a number of such interviews. They differ in emphasis, target populations and methods. The Schedule for Affective Disorders and Schizophrenia (Endicott & Spitzer, 1978) is a semi-structured, relatively flexible interview that evaluates specific symptoms to generate DSM-III-R diagnoses. This schedule has acceptable reliability but, being time consuming, is most appropriate for specific types of clinical research. The Structured Clinical Interview for DSM-III-R (Spitzer & Williams, 1988) is a semi-structured interview that allows a clinician to map symptoms efficiently onto the appropriate diagnostic criteria. SCAN, or Schedule for Clinical Assessment in Neuropsychiatry (Wing *et al.* 1990), is a semi-structured interview developed from the PSE-9 and based on a comprehensive glossary of symptoms and signs. Developed under the aegis of the World Health Organization it is an educational and research instrument to be used by trained clinicians. A shorter form, which omits the psychosis sections, can be used by trained lay interviewers. Scoring is complex because this schedule aims to describe symptoms in a fashion independent of the current classifications, even though ICD and DSM diagnoses can be generated.

The Composite International Diagnostic Interview (CIDI) (Robins *et al.* 1988; Wittchen, 1994) is a fully structured diagnostic interview that can be administered by a trained non-clinical interviewer. Developed under the auspices of the World Health Organization, the CIDI generates both ICD and DSM diagnoses. The CIDI is in modular form and it takes about an hour to cover all the major diagnoses. The fully structured format makes it exceptionally reliable; it has been widely used and is available in 15 languages. Developed for epidemiological surveys, it is being used increasingly for clinical research, and may have some applications for clinical decision making. Validity studies show good to excellent sensitivity but suggest a need for greater specificity (Andrews *et al.* 1995).

The fully structured nature of the CIDI is unique and this has made a self-administered computerized form possible. The CIDI-Auto can be completed by the patient, with the computer taking care of the logic of the complex interview (World Health Organization, 1993). Patients who have never used a computer before find the computerized version acceptable. As the self-administered version takes no staff time the CIDI-Auto is the first structured diagnostic interview with the potential to be a cost effective clinical tool. The CIDI-Auto does not assess cognitive impairment and neither the CIDI nor the CIDI-Auto are suitable for use with regressed psychotic patients. Shortened forms, especially of the substance use modules, would be an advantage for clinical use, and some method of identifying the principal diagnosis from among the current comorbid disorders would also be valuable.

In a clinical setting, these diagnostic instruments function as decision support instruments yet do not lessen the importance of the clinician's opinion. They are of especial value in informing non-psychiatrists about the diagnostic possibilities. The sensitivity and specificity of the CIDI, for instance, is comparable with most routine investigations used in general medicine. What is different is that the CIDI, like the other structured diagnostic interviews, is not measuring a physical phenomenon like haemoglobin concentration, but the presence of a difficult abstract concept like 'obsessive-compulsive disorder'. Because we are dealing with scoring consensus-based diagnostic criteria, it is important that responsibility for the accuracy of such diagnostic interviews remains with some public body such as the World Health Organization. In short, psychiatrists should make greater use of such interviews, provided they are used in a valid fashion that improves patient care.

## MEASURING SYMPTOMS, DISABILITY AND RISK FACTORS

The relevance of measurement to treatment is obvious: if a patient's clinical status does not alter with treatment the clinician needs to know more than whether the patient is 'relieved, unimproved or worse'. The clinician actually needs to know, if he or she is to be properly informed about response to treatment, which symptoms have altered and which have not. The clinician also needs to know whether the patient's level of disability has been lessened, for treatment should not conclude until the patient reports that the disorder no longer interferes with their life and activities. Last, when prognostic or risk factors have been identified, the advisability of assertive follow-up will

depend, in part, on changes in these risk factor measures, as well as on changes in symptoms and disability.

There are now a large number of rating scales and questionnaires for the measurement of such clinical information and there are excellent source books (McDowell & Newell, 1987; Thompson, 1989; Bowling, 1991). Our own units, like many others, have test libraries that contain hundreds of such instruments, but unfortunately too few have demonstrated reliability and validity. This is not surprising when one considers the effort involved in producing an instrument with good psychometric properties (see Andrews *et al.* 1993, for an example of the tedium involved). Nevertheless, there are diagnosis-specific measures of symptoms, including behaviour deficits and cognitions; measures of disablement, quality of life and consumer satisfaction; and measures of risk factors relevant to specific disorders, which do provide reliable and valid data.

There is current enthusiasm (see The Health of the Nation, 1993; First National Mental Health Policy, 1993) that outcome measures should form a routine part of clinical care, applied on admission, at clinical review and on discharge. As most patient consultations take only 30 to 45 min, and because any clinician assessment must detract from clinician–patient interaction time, it is clear that such assessment must take less than ten percent of this interview (i.e. less than 3–4 min), insufficient for the administration of a detailed test battery. Health services do not pay for patient time, hence the attraction of supplementing brief clinician ratings with more detailed instruments to be completed by the patient. Clinicians should make greater use of such measures, provided they are valid, and provided their use improves clinical care.

## THE ROLE OF THE COMPUTERIZED CLINICAL RECORD

There is little disagreement about the need for better assessment of diagnosis, symptom change, disability and risk factors – clinicians simply defer making the effort that such assessment requires. However, two things may change this. First, as noted previously, the administrators are beginning to make funding contingent on standardized assessments, and lists of outcome measures have been proposed (Dean & Foster, 1993). Secondly, the arrival of computerized clinical information systems in mental health is changing the ease of administration of tests. During the last 10 years management information systems that systematize the recording of patient information and which provide information for administration and research have become available. Individual systems are to be found in many clinics, hospitals and even throughout health regions. For example, one of our units supports a system (CRISP, 1993) that is installed in 125 sites throughout Australia and New Zealand.

Computerized management information systems in current use have rarely provided instruments to measure diagnosis, symptoms, disability or risk factors. Systems are being developed to do this and to provide additional clinical support in the form of specimen treatment protocols (cf. CLIMATE Help, 1993). As both the assessment and the treatment protocols will be instantly and easily available to the clinician during the clinical interview, such tools will be used and presumably will enhance patient care. Clinical supervisors will value the information because the computerized clinical information system will be able routinely to draw attention to mismatches between diagnosis and treatment, and to identify patients in whom specific symptoms, disability and risk factors are not responding to treatment. Furthermore, many administrators see that standardized measures of diagnosis and disability will allow the burden on individual service units to be assessed more accurately than do case-mix techniques, with obvious implications for funding by service load. Others see the possibility of using the rate of change in symptom and disability scores as proxy measures for service efficacy. In short, while clinicians will welcome the clinical support that such computerized tests and advice will provide, the information will be useful for many other purposes and there will be a pressure for assessments needed for administrative purposes to further erode the time the clinician has for each patient. Assessments needed for such administrative purposes should be funded separately.

It is necessary to have standards in place to ensure the privacy of such computerised mental health

records given that portions are needed for important administrative purposes. Patients expect the existence and content of their mental health medical record to remain confidential, and concern has been expressed as to whether this could be achieved once the record was computerised (Andrews & Wilkins, 1992). Fortunately, draft standards that will meet these concerns are now appearing (e.g. Standards Australia, 1993). However, despite there being standards for tests (American Psychological Association, 1985) they are rarely applied to the structured diagnostic interviews, questionnaires and rating scales used in psychiatry. Treatment protocols will form part of these computerized medical records and reliability and validity standards could, and should, be developed before such protocols are accepted. Ideally, such protocols should be controlled, developed and maintained by some impartial committee of experts, with further work funded by the contributions of the users or health agencies. A recent WHO sponsored meeting on computers in psychiatry began the task of establishing a mechanism whereby such computer programs could be assessed and registered (World Health Organization, 1994). The US Institute of Medicine Committee on Clinical Practice Guidelines has recommended a standardized method for the assessment and dissemination of treatment guidelines (Lohr & Schyve, 1992).

The matter is too important to be left to the commercial vendors. Clinicians need to be able to recognise which systems are unbiased, reliable and valid – at present there is no mechanism to facilitate this. Perhaps the World Health Organization could take a steering role, along with the national psychiatric organizations, to consider a simple mechanism for identifying valid systems. If standards could be ensured, then clinical diagnosis and treatment will be facilitated by the advent of information technology.

GAVIN ANDREWS AND HANS-ULRICH WITTCHEN

## REFERENCES

- American Psychological Association. (1985). *Standards for Educational and Psychological Testing*. American Psychological Association: Washington, D.C.
- Andrews, G. & Hadzi-Pavlovic, D. (1988). The work of Australian psychiatrists, circa 1986. *Australian and New Zealand Journal of Psychiatry*, **22**, 153–165.
- Andrews, G. & Wilkins, G. E. J. (1992). Privacy and the computerised medical record. *Medical Journal of Australia*, **157**, 223–225.
- Andrews, G., Singh, M. & Bond, M. (1993). The Defense Style Questionnaire. *Journal of Nervous and Mental Disease* **181**, 246–256.
- Andrews, G., Peters, L., Guzman, A.-M. & Bird, K. (1995). A comparison of two structured diagnostic interviews: CIDI and SCAN. *Australian and New Zealand Journal of Psychiatry* (In the press).
- Bowling, A. (1991). *Measuring Health: A Review of Quality of Life Measurement Scales*. Open University Press: Milton Keynes, PA.
- CLIMATE Help. (1993). *Computer Program*. Training and Reference Centre for WHO CIDI: Sydney.
- CRISP. (1993). *Computer Program*. Training and Reference Centre for WHO CIDI: Sydney.
- Dean, C. & Foster, A. (1993). Measuring outcomes. In *Community Mental Health Care*. (ed. C. Dean and H. Greeman), pp. 125–127. Gaskell: London.
- Endicott, J. & Spitzer, R. L. (1978). A diagnostic interview: the Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry* **35**, 837–844.
- First National Mental Health Report*. (1993). Department of Human Services and Health, Australian Government Publishing Service: Canberra, Australia.
- Lohr, K. N. & Schyve, P. M. (1992). Reasonable expectations: from the Institute of Medicine. *Quality Review Bulletin* **18**, 393–396.
- McDowell, I. & Newell, C. (1987). *Measuring Health: A Guide to Rating Scales and Questionnaires*. Oxford University Press: New York.
- Meehl, P. E. (1954). *Clinical versus Statistical Prediction: A Theoretical Analysis and A Review of the Evidence*. University of Minnesota Press: Minneapolis.
- Robins, L. N., Wing, J., Wittchen, H.-U., Helzer, J. E., Babor, T. F., Burke, J., Farmer, A., Jablenski, A., Pickens, R., Regier, D. A., Sartorius, N. & Towle, L. H. (1988). The Composite International Diagnostic Interview. *Archives of General Psychiatry* **45**, 1069–1077.
- Spitzer, R. L. & Williams, J. B. W. (1988). Revised diagnostic criteria and a new structured interview for diagnosing anxiety disorders. *Journal of Psychiatric Research* **22**, 55–64.
- Standards Australia. (1993). *Draft Australian Standard: Information Security and Personal Privacy Protection in Health Care Information Systems*. Standards Australia: Sydney.
- The Health of the Nation. (1993). *Key Area Handbook: Mental Illness*. Department of Health: Heywood, Lancashire.
- Thompson, C. (1989). *The Instruments of Psychiatric Research*. Wiley: Chichester.
- Wing, J. K., Babor, T., Brugha, T., Burke, J., Cooper, J., Giel, R., Jablensky, A., Regier, D. & Sartorius, N. (1990) SCAN: schedules for clinical assessment in neuropsychiatry. *Archives of General Psychiatry*, **47**, 589–593.
- Wittchen, H. U. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI). A critical review. *Journal of Psychiatric Research*, **28**, 57–84.
- World Health Organization. (1993). *CIDI-Auto: Administrator's Guide and Reference*. Training and Reference Centre for WHO CIDI: Sydney.
- World Health Organization. (1994). *Proceedings of Informal Consultation on Computerized Information Systems for Diagnosis Classification and Management of Mental Disorders*. Lubeck, Germany, 2–3 March.