Standardised Assessment of Personality – Abbreviated Scale (SAPAS): preliminary validation of a brief screen for personality disorder

PAUL MORAN, MORVEN LEESE, TENNYSON LEE, PAUL WALTERS, GRAHAM THORNICROFT and ANTHONY MANN

Background There is a need for a brief and simple screen for personality disorders that can be used in routine psychiatric assessments.

Aims To test the concurrent validity and test–retest reliability of a brief screen for personality disorder.

Method Sixty psychiatric patients were administered a brief screening interview for personality disorder. On the same day, they were interviewed with an established assessment for DSM–IV personality disorder. Three weeks later, the brief screening interview was repeated in order to examine test–retest reliability.

Results A score of 3 on the screening interview correctly identified the presence of DSM–IV personality disorder in 90% of participants. The sensitivity and specificity were 0.94 and 0.85 respectively.

Conclusions The study provides preliminary evidence of the usefulness of the screen in routine clinical settings.

Declaration of interest None. Funding detailed in Acknowledgements.

Personality disorder can significantly affect the management and outcome of associated mental illness (Patience *et al*, 1995; Yonkers *et al*, 2000). An assessment of personality status should therefore ideally form part of the routine assessments conducted by psychiatric teams (Moran *et al*, 2003; Tyrer & Simmonds, 2003). However, too often the assessment of personality disorder remains one of clinical judgement. Unfortunately, clinical diagnoses are unreliable (Mellsop *et al*, 1982), and although reliability can be improved by the use of standardised assessments, these assessments are lengthy and require training. Self-report questionnaires are useful research tools, but they can be tiring for patients because they require the ability to concentrate on written questions. A brief structured interview with the patient would overcome some of these problems provided it could be easily incorporated into a standard psychiatric interview. This paper reports on the preliminary validation of a brief structured interview for personality disorders that is feasible for use in routine clinical assessment.

METHOD

Participants

A non-random sample of 60 adult patients was recruited from out-patient clinics (*n*=24), in-patient units (*n*=24) and day units (*n*=12) within the South London and Maudsley National Health Service (NHS) Trust. No special attempt was made to select patients with known or suspected personality disorder; however, the sample was chosen to represent patients with a range of psychiatric problems. Patients were also chosen on the basis that they were stable and cooperative with being interviewed. None of the patients was acutely unwell at the time of recruitment. Out-patients and day patients were recruited directly at the time of clinic or day hospital attendance, and in-patients were interviewed on the hospital ward. The sample consisted of 34 women and 26 men, with a mean age of 43 years (s.d.=15.9). The clinical diagnoses of the sample were as follows: affective disorder (*n*=25), anxiety disorder (*n*=11), eating disorder (*n*=9), schizophrenia (*n*=9) and drug or alcohol dependence (*n*=6).

Measures

Screening questionnaire

The screening questionnaire consisted of eight dichotomously rated items taken from the opening section of an informant-based interview, the Standardised Assessment of Personality (SAP) (Mann *et al*, 1981; Pilgrim & Mann, 1990; Pilgrim *et al*, 1993). The SAP allows an ICD–10 or DSM–IV diagnosis of personality disorder to be made (World Health Organization, 1992; American Psychiatric Association, 1994). Each of the eight questions from the opening section of the SAP corresponds to a descriptive statement about the person and can be scored 0 or 1 (see Appendix). The scores on the eight items can be added together to produce a total score between 0 and 8.

An exploratory analysis of the SAP ratings of a sample of 303 primary care attenders (Moran *et al*, 2001; Rendu *et al*, 2002) showed that the total score on these eight official probe items satisfactorily predicted the final SAP diagnosis of personality disorder obtained after more detailed questioning of the informant: area under the curve (AUC)=0.79, 95% CI 0.74–0.84. The performance of these eight items suggested that they might also act as a patient-based screen for a diagnosis of personality disorder. However, the SAP is an informant-based interview and it was unclear how well the probe items would perform when given to patients as opposed to informants. The examination of the psychometric properties of the patient-based screen, the Standardised Assessment of Personality – Abbreviated Scale (SAPAS), formed the basis of this study.

SCID–II

The Structured Clinical Interview for DSM–IV Personality Disorders (SCID–II) (First *et al*, 1997) is a 119-item semi-structured interview with the patient. Each item is scored as 1 (absent), 2 (sub-threshold) or 3 (threshold). Questions may

necessitate further exploration by the interviewer in order to score a particular item. If a threshold is reached on a sufficient number of items, the category of personality disorder is deemed to be present. The SCID–II was designed to generate DSM–III–R (American Psychiatric Association, 1987) diagnoses; however, by eliminating items for passive–aggressive and depressive personality disorders, it can be used to generate DSM–IV personality disorder diagnoses. The instrument demonstrates acceptable test–retest ($\kappa=0.68$) and interrater reliability ($\kappa=0.71$) and takes up to 1 h to administer.

Procedure

A member of the clinical team (either a doctor or a nurse) interviewed the patient with the SAPAS, as part of routine clinical work. Shortly afterwards, the patient was interviewed with the SCID–II by one of the authors (P.M.). The majority (83%, $n=50$) of SCID–II assessments were conducted blind to the results of the screening mini-interview. In the case of 10 patient interviews, no staff member was available to conduct the SAPAS and P.M. therefore conducted both interviews. Approximately 3 weeks later (mean interval 20 days, s.d.=10), each patient was re-interviewed by the same person using the SAPAS.

Analysis

Analyses were performed using STATA version 7 (StataCorp, 1999). The main aim of analysis was to identify an appropriate cut-off score on the SAPAS for predicting a SCID–II (DSM–IV) diagnosis of personality disorder. This was achieved by undertaking an AUC analysis. The performance of the SAPAS at different cut-off scores was assessed by reference to the sensitivity, specificity and predictive values of the screening interview. The internal consistency of the SAPAS was assessed by calculating Cronbach's α on the total score after omitting each item and also overall. The test–retest reliability of each item was estimated by calculating the κ coefficient, and the overall reliability of the total score was estimated using Lin's concordance coefficient (Lin, 1989). Interrater reliability is not a major issue since the questions are largely self-explanatory and no interpretation is placed on responses.

Table I Internal consistency and test–retest reliability of the Standardised Assessment of Personality – Abbreviated Scale items. The alpha coefficient for the total score is 0.68 and Lin's concordance coefficient for the total score is 0.89

Item	Alpha coefficient if item omitted	Kappa coefficient
Difficulty making and keeping friends	0.59	0.81
Usually a loner	0.63	0.83
Trusting others	0.57	0.79
Normally loses temper easily	0.66	0.83
Normally impulsive	0.72	0.61
Normally a worrier	0.62	0.62
Depends on others a lot	0.68	0.82
Generally a perfectionist	0.70	0.73

RESULTS

A total of 33 out of 60 patients received a SCID–II diagnosis of personality disorder, giving an overall prevalence of 55% (95% CI 42–68). The mean number of personality disorder diagnoses among those with any personality disorder was 2.1 (s.d.=1.2). Table 1 shows the α and κ coefficients of each item from the SAPAS and overall reliability coefficients. This shows that there is a moderate degree of overall internal consistency (0.68). 'Normally impulsive' and 'Generally a perfectionist' are the items least consistent with the rest. The test–retest reliability is reasonable and individual κ values are also acceptable, although the values for 'Normally impulsive' and 'Normally a worrier' are less. 'Normally impulsive' would seem to be the least satisfactory item, taking both internal consistency and test–retest reliability into account.

To investigate the use of alternative cut-off scores on the SAPAS, a logistic regression was employed with the SAPAS total score as predictor and SCID–II diagnosis as dependent variable. This analysis produced an AUC of 0.94 (95% CI 0.88–0.99). To assess the sensitivity and specificity of the SAPAS for various cut-off scores, a sensitivity–specificity plot was obtained (Fig. 1). This indicates that a probability cut-off of 0.65 for a positive SCID diagnosis (equivalent to a total SAPAS score of between 3 and 4) has approximately equal sensitivity and specificity, with both around 0.8. The performance of the SAPAS at a range of cut-off scores is displayed in Table 2; this shows that a cut-off score of 3 or 4 correctly classified over 80% of the patients. Although both thresholds
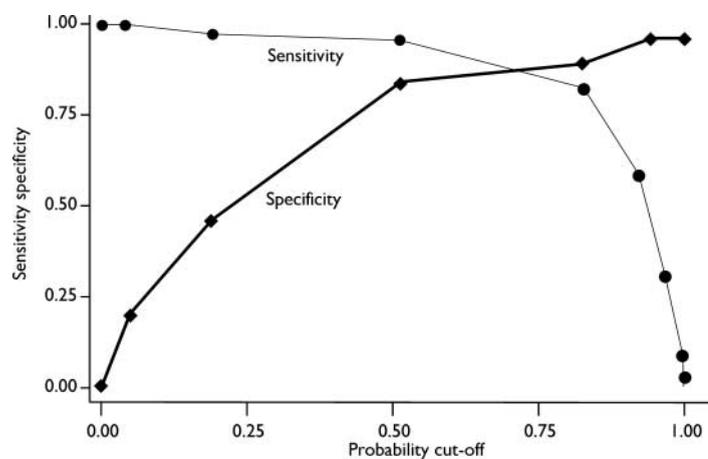

Fig. I Sensitivity–specificity plot relating Structured Clinical Interview for DSM–IV Personality Disorders positive diagnosis to total score on the Standardised Assessment of Personality – Abbreviated Scale.

Table 2 Sensitivity, specificity and power to predict personality disorder at different cut-off scores of the Standardised Assessment of Personality – Abbreviated Scale

Cut-off score	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Correctly classified (%)
2 or more	0.97	0.44	0.68	0.92	73
3 or more	0.94	0.85	0.89	0.92	90
4 or more	0.82	0.89	0.90	0.80	85
5 or more	0.58	1.0	1.0	0.66	77
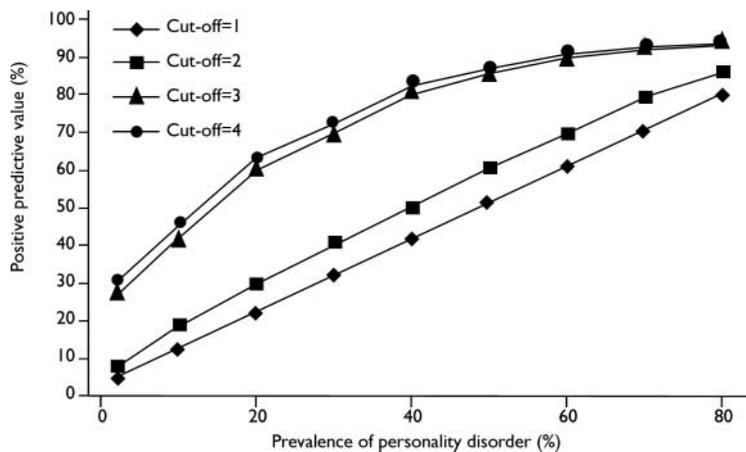
Fig. 2 Scatter plot showing the effect of prevalence of personality disorder on the positive predictive value of the Standardised Assessment of Personality – Abbreviated Scale.

performed similarly, arguably the cut-off score of 3 offers the best balance of sensitivity (0.94) and specificity (0.85) and gives the maximum total of these two measures. When the ten non-blind assessments were excluded the AUC was 0.92 (95% CI 0.85–0.99), and at a cut-off of 3 the sensitivity was 92% and the specificity was 84%, indicating that the full sample had not been biased by the inclusion of these cases.

A scatter plot showing the positive predictive value of the screen at different cut-off scores of the SAPAS (Fig. 2) allows the effect of assuming various levels of population prevalence to be judged.

DISCUSSION

Performance of the SAPAS

A score of 3 or 4 on the SAPAS correctly identified the presence of personality disorder in over 80% of participants. The study therefore provides preliminary evidence of the usefulness of the SAPAS as a screen for personality disorder in routine clinical settings. The findings should, however, be treated with caution, taking into account a number of limitations.

First, the study relied on a small, non-random sample of stable and cooperative patients with a high prevalence of personality disorder. Although the screen performed acceptably in this population, if it were to be applied to a population with a lower prevalence of personality disorder, its predictive power would diminish (Fig. 2). Consequently, the screen is probably not suitable for use in general community or primary care settings, where the prevalence of personality disorder is in the range 10–20%. Samuels et al (2002) estimated that the prevalence of DSM–IV personality disorders in a community sample was 9%. Thus, from Fig. 2, based on this prevalence, the positive predictive power of the SAPAS in a community sample would be between 40% and 50%. In addition, although sensitivity and specificity are independent of the prevalence of a disorder in a population, measures may be more or less applicable to different populations. The

findings therefore require replication in larger and more diverse populations of psychiatric patients.

Second, our choice of the SCID–II as the criterion for validation of the SAPAS may be questioned. However, the validity of the assessment measures for personality disorder has yet to be firmly established and none has been proved superior to any other (Zimmerman, 1994). The SCID–II was chosen as the gold standard because it has been widely used and its psychometric properties are well established (Zimmerman, 1994).

Third, we did not examine the ability of the SAPAS to discriminate between either sub-categories or clusters of personality disorder. In clinical practice, patients with personality disorders usually fulfil diagnostic criteria for more than one sub-category of disorder (McGlashan et al, 2000) and it therefore makes little sense to screen for individual categories of personality disorder. In addition, the identification of sub-categories and clusters of personality disorder requires a more sophisticated diagnostic approach than that afforded by the SAPAS.

Comparison with existing screening methods for personality disorder

A number of self-report questionnaires are available for the purpose of screening for personality disorder. These include the International Personality Disorder Examination Screen (Lenzenweger et al, 1997), the Personality Diagnostic Questionnaire – Revised (Hyler et al, 1992) and the SCID–II Screen (Ekselius et al, 1994). Although these instruments are of some value to researchers interested in identifying 'high-risk' populations, when compared with a structured interview their specificity is invariably poor. In addition, they require the ability of the respondent to concentrate on a long set of questions.

To the best of our knowledge, only two other interviewer-administered screens for personality disorder have been published. Langbehn et al (1999) have developed the Iowa Personality Disorder Screen (IPDS) to provide a mini-structured interview that the authors estimate can be completed in 5 min. The IPDS consists of 11 questions that address general personality disorder criteria as well as specific criteria. The instrument has been validated against the Structured Interview for DSM–IV Personality Disorders (SIDP–IV) (Pfohl et al, 1997).

The authors reported excellent sensitivity (92%) and good specificity (79%), although the validation was a somewhat circular exercise, as the IPDS items were derived from the DSM–III–R version of the SIDP. Van Horn *et al* (2000) have developed a structured patient interview for personality disorders, the Rapid Personality Assessment Schedule (PAS–R). However, the PAS–R requires staff training and performs moderately well as a screen for personality disorder when compared with the full version of the PAS (sensitivity 64%, specificity 82%).

In this preliminary validation exercise, the SAPAS showed superior psychometric performance compared with both the IPDS and the PAS–R. In addition, the SAPAS is short (no interview took longer than 2 min to complete), does not require training, is simple to use, and was acceptable to the respondents in this study. It therefore fulfils many of the criteria for a desirable screening test (Brewin *et al*, 2002).

Potential applications of the SAPAS

The SAPAS could be used to identify individuals who are at potentially high risk of having any type of personality disorder in a general adult psychiatric setting. The screen itself should not be used to make a diagnosis of personality disorder or cluster of personality disorders, and we would advise that a person scoring more than 3 on the SAPAS should be interviewed with a detailed structured assessment of personality. Clinicians and investigators might wish to adopt higher or lower thresholds, depending on the nature of the sample and the relative importance to them of sensitivity and specificity.

We think that the screen could have both clinical and epidemiological applications. It is feasible for use in busy clinical settings and could therefore be used to identify individuals in need of a more detailed personality assessment. Although the assessment of personality soon after presentation might result in inflated estimates of personality disorder, this is often the time when treatment decisions are made, and if personality assessments are to have useful treatment implications, arguably they should be made at an early stage (Zimmerman, 1994). From an epidemiological perspective, the SAPAS could be used as a first-stage screen as part of a two-stage procedure for

CLINICAL IMPLICATIONS

■ Personality disorder can be predicted effectively with an eight-item screening interview, the Standardised Assessment of Personality − Abbreviated Scale (SAPAS).

■ A score of 3 on the screening interview correctly identified the presence of DSM−IV personality disorder in 90% of participants.

■ The SAPAS is feasible for use in routine clinical settings, where it could be used to identify individuals at risk of personality disorder. It could also be used in epidemiological research as part of a two-stage procedure for case identification.

LIMITATIONS

■ The findings need to be replicated on larger samples of psychiatric patients.

■ The findings need to be replicated in different samples of psychiatric patients.

■ The SAPAS may not be suitable for use in populations where the prevalence of personality disorder is lower.

PAUL MORAN, MD, MORVEN LEESE, PhD, TENNYSON LEE, MRCPsych, PAUL WALTERS, MRCPsych, GRAHAM THORNICROFT, PhD, Health Services Research Department, Institute of Psychiatry; ANTHONY MANN, MD, Section of Epidemiology, Institute of Psychiatry, London, UK

Correspondence: Dr Paul Moran, Health Services Research Department, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK. Tel: 020 7 848 0568; fax: 020 7 848 0333; e-mail: paul.moran@iop.kcl.ac.uk

(First received 20 January 2003, final revision 8 May 2003, accepted 8 May 2003)

case identification (Lenzenweger *et al*, 1997; Mann *et al*, 1999).

ACKNOWLEDGEMENTS

APPENDIX

Standardised Assessment of Personality − Abbreviated Scale

Only circle Y (yes) (or N (no) in the case of question 3) if the patient thinks that the description applies *most of the time* and *in most situations*.

1. In general, do you have difficulty making and keeping friends? . . . . .Y/N
   (yes=1, no=0)

2. Would you normally describe yourself as a loner? . . . . .Y/N
   (yes=1, no=0)

3. In general, do you trust other people? . . . . .Y/N
   (yes=0, no=1)

4. Do you normally lose your temper easily? . .Y/N
   (yes=1, no=0)

5. Are you normally an impulsive sort of person? . . . . .Y/N
   (yes=1, no=0)

6. Are you normally a worrier? . . . . .Y/N
   (yes=1, no=0)

7. In general, do you depend on others a lot? . .Y/N
   (yes=1, no=0)

8. In general, are you a perfectionist? . . . . .Y/N
   (yes=1, no=0)

REFERENCES

American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn, revised) (DSM−III−R). Washington, DC: APA.

— **(1994)** *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM−IV). Washington, DC: APA.

Brewin, C. R., Rose, S., Andrews, B., et al (2002) Brief screening instrument for post-traumatic stress disorder. *British Journal of Psychiatry*, **181**, 158–162.

Ekselius, L., Lindstrom, E., Von Knorring, L., et al (1994) SCID II interviews and the SCID Screen questionnaire as diagnostic tools for personality

disorders in DSM–III–R. *Acta Psychiatrica Scandinavica*, **90**, 120–123.

First, M. B., Gibbon, M., Spitzer, R. L., et al (1997) *The Structured Clinical Interview for DSM–IV Axis II Personality Disorders (SCID–II).* Washington, DC: American Psychiatric Press.

Hyler, S. E., Skodol, A. E., Oldham, J. M., et al (1992) Validity of the Personality Diagnostic Questionnaire – Revised: a replication in an outpatient sample. *Comprehensive Psychiatry*, **33**, 73–77.

Langbehn, D. R., Pfohl, B. M., Reynolds, S., et al (1999) The Iowa Personality Disorder Screen: development and preliminary validation of a brief screening interview. *Journal of Personality Disorders*, **13**, 75–89.

Lenzenweger, M. F., Loranger, A. W., Korfine, L., et al (1997) Detecting personality disorders in a nonclinical population. Application of a 2-stage procedure for case identification. *Archives of General Psychiatry*, **54**, 345–351.

Lin, L. I.-K. (1989) A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, **45**, 255–268.

Mann, A. H., Jenkins, R., Cutting, J. C., et al (1981) The development and use of a standardized assessment of abnormal personality. *Psychological Medicine*, **11**, 839–847.

___ , Raven, P., Pilgrim, J., et al (1999) An assessment of the Standardised Assessment of Personality as a screening instrument for the International Personality Disorder Examination: a comparison of informant and patient assessment for personality disorder. *Psychological Medicine*, **29**, 985–989.

McGlashan, T. H., Grilo, C. M., Skodol, A. E., et al (2000) The Collaborative Longitudinal Personality Disorders Study: baseline Axis I/II and II/II diagnostic co-occurrence. *Acta Psychiatrica Scandinavica*, **102**, 256–264.

Mellsop, G., Varghese, F., Joshua, S., et al (1982) The reliability of axis II of DSM–III. *American Journal of Psychiatry*, **139**, 1360–1361.

Moran, P., Rendu, A., Jenkins, R., et al (2001) The impact of personality disorder in UK primary care: a 1-year follow-up of attenders. *Psychological Medicine*, **31**, 1447–1454.

___ , Walsh, E., Tyrer, P., et al (2003) The impact of co-morbid personality disorder on violence in psychosis. Report from the UK700 trial. *British Journal of Psychiatry*, **182**, 129–134.

Patience, D. A., McGuire, R. J., Scott, A. I. F., et al (1995) The Edinburgh Primary Care Depression Study: personality disorder and outcome. *British Journal of Psychiatry*, **167**, 324–330.

Pfohl, B., Blum, N. & Zimmerman, M. (1997) *Structured Interview for DSM–IV Personality (SIDP–IV).* Washington, DC: American Psychiatric Press.

Pilgrim, J. & Mann, A. (1990) Use of the ICD–10 version of the Standardized Assessment of Personality to determine the prevalence of personality disorder in psychiatric in-patients. *Psychological Medicine*, **20**, 985–992.

___ , Mellers, J. D., Boothby, H., et al (1993) Inter-rater and temporal reliability of the Standardised Assessment of Personality and the influence of informant characteristics. *Psychological Medicine*, **23**, 779–786.

Rendu, A., Moran, P., Patel, A., et al (2002) Economic impact of personality disorders in UK primary care attenders. *Psychological Medicine*, **181**, 62–66.

Samuels, J., Eaton, W. W., Bienvenu, O. J., et al (2002) Prevalence and correlates of personality disorders in a community sample. *British Journal of Psychiatry*, **180**, 536–542.

StataCorp (1999) *STATA Statistical Software: Version 7.* College Station, TX: Stata Corporation.

Tyrer, P. & Simmonds, S. (2003) Treatment models for those with severe mental illness and comorbid personality disorder. *British Journal of Psychiatry*, **182** (suppl. 44), 15s–18s.

Van Horn, E., Manley, C., Leddy, D., et al (2000) Problems in developing an instrument for the rapid assessment of personality status. *European Psychiatry*, **15** (suppl. I), 29–33.

World Health Organization (1992) *International Statistical Classification of Diseases and Related Health Problems* (ICD–10). Geneva: WHO.

Yonkers, K. A., Dyck, I. R., Warshaw, M., et al (2000) Factors predicting the clinical course of generalised anxiety disorder. *British Journal of Psychiatry*, **176**, 544–549.

Zimmerman, M. (1994) Diagnosing personality disorders. A review of issues and research methods. *Archives of General Psychiatry*, **51**, 225–245.