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 THORNcROFT 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.

Measure

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.

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

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 (κ=0.68) and interrater reliability (κ=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.

**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). ‘Normatively 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 ‘Normatively impulsive’ and ‘Generally a worrier’ are less. ‘Normatively 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

<table>
<thead>
<tr>
<th>Item</th>
<th>Alpha coefficient if item omitted</th>
<th>Kappa coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty making and keeping friends</td>
<td>0.59</td>
<td>0.81</td>
</tr>
<tr>
<td>Usually a loner</td>
<td>0.63</td>
<td>0.83</td>
</tr>
<tr>
<td>Trusting others</td>
<td>0.57</td>
<td>0.79</td>
</tr>
<tr>
<td>Normally loses temper easily</td>
<td>0.66</td>
<td>0.83</td>
</tr>
<tr>
<td>Normally impulsive</td>
<td>0.72</td>
<td>0.61</td>
</tr>
<tr>
<td>Normally a worrier</td>
<td>0.62</td>
<td>0.62</td>
</tr>
<tr>
<td>Depends on others a lot</td>
<td>0.68</td>
<td>0.82</td>
</tr>
<tr>
<td>Generally a perfectionist</td>
<td>0.70</td>
<td>0.73</td>
</tr>
</tbody>
</table>
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 Disorder Examination – 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 fulfills many of the criteria for a desirable screening test (Brewin et al, 2002).

**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


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

PAUL MORAN, MORVEN LESEE, TENNYSON LEE, PAUL WALTERS, GRAHAM THORNOCROFT and ANTHONY MANN

Access the most recent version at DOI: 10.1192/bjp.183.3.228