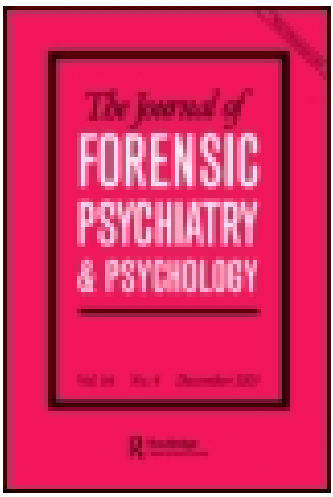

The Journal of Forensic Psychiatry & Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rjfp20

Detecting personality disorders in the prison population of England and Wales: comparing case identification using the SCID-II screen and the SCID-II clinical interview

Simone Ullrich [a], Derek Deasy [b], Jane Smith [c], Ben Johnson [d], Maria Clarke [e], Nick Broughton [f] & Jeremy Coid [a]

[a] Forensic Psychiatry Research Unit, Queen Mary College, University of London, UK

[b] St. Clare's Unit, Children's University Hospital, Dublin, UK

[c] Department of Health Sciences, University of Leicester, UK

[d] Merseycare NHS Trust, Liverpool, UK

[e] Patterson Centre for Mental Health, London, UK

[f] West London Mental Health NHS Trust, UK

Published online: 29 Jul 2008.

The Journal of Forensic Psychiatry & Psychology
Vol. 19, No. 3, September 2008, 301–322

RESEARCH ARTICLE

Detecting personality disorders in the prison population of England and Wales: comparing case identification using the SCID-II screen and the SCID-II clinical interview

Simone Ullrich[a]*, Derek Deasy[b], Jane Smith[c], Ben Johnson[d], Maria Clarke[e], Nick Broughton[f] and Jeremy Coid[a]

[a]*Forensic Psychiatry Research Unit, Queen Mary College, University of London, UK;* [b]*St. Clare's Unit, Children's University Hospital, Dublin, UK;* [c]*Department of Health Sciences, University of Leicester, UK;* [d]*Merseycare NHS Trust, Liverpool, UK;* [e]*Patterson Centre for Mental Health, London, UK;* [f]*West London Mental Health NHS Trust, UK*

(*Received in final form 3 March 2008*)

The cross-instrument validity of the SCID-II screen and the SCID-II structured clinical interview was examined in a sub-sample of 496 male and female prisoners from a national survey in England and Wales. Participants completed the SCID-II screen and were interviewed two weeks later by clinicians using the SCID-II interview. The screen was adjusted on the basis of interview results. Diagnostic categories were compared before and after adjustment of the screen. Diagnostic cut-off adjustments to the screen resulted in moderate agreement between instruments for categorical and dimensional evaluations. The screen retained good internal consistency with improved discrimination between personality disorders and clinical syndromes following adjustment. The SCID-II screen is a valid self-report instrument for diagnosing personality disorders. It may be preferable for prison populations but is not a substitute for a structured clinical interview.

Keywords: personality disorder categories; screen; prisoners

Introduction

It is essential to obtain accurate epidemiological data on personality disorders among prison populations, as prevalences may determine service needs and have important implications for inmate management. The presence of these disorders may also predict the behaviour of some prisoners while in custody and after release (Coid, 1998; Salekin, Rogers, & Sewell, 1996). Previous epidemiological evidence indicates that prisoners in the UK and USA have substantial levels of psychiatric morbidity compared to the

*Corresponding author. Email: s.ullrich@qmul.ac.uk

ISSN 1478-9949 print/ISSN 1478-9957 online

DOI: 10.1080/14789940802045182
http://www.informaworld.com

general population (Coid, 1984; Lamb & Weinberger, 1998), with the most extreme differences demonstrated for substance misuse and antisocial personality disorder. However, data on the full range of Axis II categories are not available from previous prison surveys, where a range of different diagnostic approaches have been applied (Moran, 1999). Most studies have been carried out in a single or limited number of penal institutions where the survey samples may not be representative of prisoners in general.

The validity of case definition procedures still remains highly problematic in epidemiological studies of personality disorders (Zimmerman, 1994). Clinician interviewers, using structured clinical diagnostic interviews, are considered the 'gold standard' in reducing measurement error. However, the costs of employing experienced clinicians, as compared to lay interviewers, is a principle deterrent in large-scale studies. It has been recommended that a two-stage design should be used in general population surveys, whereby clinicians only interview subjects with positive results from a screen administered in the first phase and a random sample of subjects with negative results (Dohrenwend & Shrout, 1981). This is intended to make financial savings, but to be valid the method requires a screen with high sensitivity, whereas the economic gain would partly depend on the screen's specificity. A serious problem with reliance on a two-phase design, however, is loss of Axis II data potentially available in the first phase when using a brief screening instrument designed to measure only the presence or absence of any Axis II disorder. Although the two-phase design may result in greater confidence in diagnostic prevalence, loss of statistical power may then preclude the study of associations between individual Axis II categories and less common variables. Self-report instruments have traditionally been used as a screen in the first phase of two-phase studies. However, there are disagreements over the validity of Axis II diagnostic data derived from self-report instruments (Loranger, 1992).

This study aimed to investigate the cross-instrument validity of a self-report questionnaire and structured clinical interview for DSM-IV personality disorders (SCID II) in a representative sample of prisoners in England and Wales. In a second step, a modified diagnostic algorithm for diagnosis of personality disorder was applied to the screening instrument, to investigate whether better concordance between self-report and interview could be achieved.

Method

National survey of prisoners

The National Survey of Psychiatric Morbidity in Prisoners in England and Wales (Singleton, Meltzer, Gatward, Coid, & Deasy, 1998) provided an opportunity to compare measurements of DSM-IV Axis II personality disorders from a self-report screen administered by lay interviewers and

from structured clinical interviews administered by trained clinicians. This survey was commissioned by the Department of Health for England and Wales, and followed three previous national surveys of general households (Meltzer, Gill, Pettigrew, & Hinds, 1995), homeless people (Gill, Meltzer, Hinds, & Pettigrew, 1996), and persons in institutions (Meltzer, Gill, Hinds, & Pettigrew, 1996). These surveys were designed to provide a national estimate of the prevalence, severity, and duration of mental health problems, and their accompanying disability, to provide a baseline measure when setting the Department's target for the *Health of the Nation* strategy (Department of Health, 1992) to improve the health and social functioning of people with mental illness in the UK. However, the three previous surveys did not include measures of personality disorder.

Several structured clinical interviews were potentially available in the planning stages for the prisoners' survey, to measure DSM-IV Axis II diagnoses. However, the final choice was influenced by two major considerations. First, any measure of personality disorder would have to be made while carrying out a range of additional measures of psychopathology. Diagnostic reliability therefore had to be balanced with the inevitable time constraints of an extended interview. Second, an instrument was required that could be used in future surveys with a two-phase design. An ideal instrument would combine a screen for Axis II disorders that performed efficiently in case identification, but would also provide sufficiently accurate measures of personality disorder for comparison with other measures taken from the larger sample in the first phase of the two-phase survey. The Structured Clinical Interview for DSM-IV Personality Disorders (First, Gibbon, Spitzer, Williams, & Benjamin, 1997) was chosen for this purpose. It consists of 119 items and involves a 60–90-minute interview between clinician and participant, covering each personality disorder category in turn; each criterion is evaluated by specified questions and subsequent probes. A unique feature is the option to precede the interview with a self-report questionnaire covering the same items, although this option is seldom used in a research context. The questionnaire is expected to produce a certain amount of false positives in comparison to interview ratings, but only few false negatives. Although its development was originally intended to save time in routine clinical practice, its format does offer considerable potential for use as a diagnostic screen in the context of a survey, but has not previously been used as such.

The survey facilitated the investigation of the effectiveness of the SCID-II screen in case identification among prisoners, and investigated its potential utility for a future two-stage procedure which might be applied to the general population. The design of the survey did involve two stages, but was not a true two-phase procedure in that selection for clinical interview was random rather than based on the results of the questionnaire administered as a screen in the first stage.

Survey design

It was anticipated that the prevalence rate for certain mental disorders would show considerable variation between different types of prisoner (remand and sentenced, male and female), as in previous studies. The survey was therefore designed to provide separate prevalence estimates for these groups. In addition, some prisons were expected to have a higher proportion of prisoners with mental disorder than others, for example those with special therapeutic units or those with higher security which allows them to hold more dangerous offenders. To avoid the possibility of over- or under-sampling any such clusters of prisoners, all prisons in England and Wales were included in the survey and samples of inmates were taken from all locations within each prison. To ensure that persons in all prisons had an equal chance of being selected, it was also necessary to use a fixed sampling fraction so that the size of the sample in each prison was proportional to numbers in each prison.

The sampling design was based on 61,944 prisoners in 131 penal establishments at the end of July 1997, which included 46,872 male sentenced prisoners, 12,302 male remand prisoners (including civil prisoners), and 2,770 women prisoners. To obtain the necessary number of interviews for each type of prisoner, different sampling fractions were required in each group, and included one in 34 male sentenced prisoners, one in eight male remand prisoners, and one in three women prisoners (whether remand or sentenced). The sampling fraction for the male sentenced prisoner group was changed to one in 50 for the last four weeks of the survey because a larger number of male sentenced prisoners than expected was accessed in the early part of the fieldwork.

Since the prison population constantly changes due to release and transfer between institutions, some selected prisoners were expected to leave before they could be interviewed, while others not on the sampling list will arrive. To ensure the correct number of interviews was obtained, replacement of those leaving with new arrivals was carried out for remand but not sentenced prisoners.

The survey was carried out in two stages. The first stage involved initial interviews by lay interviewers who asked questions and entered the prisoners' responses into laptop computers. Parts of the interview were self-administered by the prisoners, using these computers, unless they were unable to read or the interviewer judged that they would be unable to self-administer reliably. The second stage included every fifth person interviewed in the first stage, who was referred to a clinician for a follow-up interview.

Response

All 131 prison establishments agreed to participate in the survey; a total of 3,563 prisoners were selected to take part. After complete description of the

study to the prisoners, written informed consent was obtained and 3,142 (88%) were interviewed in the first phase. A further 37 agreed to take part but failed to complete the long interview. Only 198 (6%) refused an interview, 53 (1%) were unable to participate (mainly because of language difficulties), and the interviewers were unable to contact 118 (3%) prisoners. Interviewers were advised not to see 15 prisoners, because they were too dangerous or too disturbed at the time. Response rates were good in all prisons, and below 80% in only 12 (in two there was a high proportion of immigration detainees who had language difficulties preventing their participation).

In the second phase, 505 (76%) of the 661 prisoners selected for follow-up were interviewed. Because of the time delay between interviews, 105 (16%) people could not be contacted at the follow-up stage, in most cases because they had left prison. A further 50 (8%) refused the follow-up, mainly due to reluctance to take part in another long interview so soon after the first. The mean delay between first and second phase interviews was two weeks.

Assessment instruments (Stage 1)

Lay interviewers administered a questionnaire using laptop PCs. This questionnaire assessed sociodemography, general health, use of services in prison, service use before current prison term, lifetime experience of services, smoking, alcohol consumption, drug use, difficulties with daily living, intelligence (using the Quick Test; Ammons & Ammons, 1962), whether there was a history of deliberate self-harm, key life events, and post-traumatic stress, whether there were neurotic disorders and symptoms (using the Clinical Interview Schedule-Revised (CIS-R); Lewis, Pelosi, Araya, & Dunn, 1992), and whether psychosis was present (using the Psychosis Screening Questionnaire; Bebbington & Nayani, 1995). Information on criminal charges or convictions was obtained from each prison. A history of convictions was obtained from prisoner self-report.

The original SCID-II questionnaire consists of 119 questions covering 13 DSM-IV personality disorder categories. Three questions were modified to become more understandable to persons in the UK. A further question was omitted as it was culture-specific to the USA and could not be understood in a UK context. An additional 14 questions were added to the questionnaire by one of the authors (JWC), to assess adult antisocial behaviour. Questions on depressive and passive-aggressive personality disorder were omitted since these diagnoses are only listed in the appendix of DSM-IV. The 116-item version was then self-completed on a lap-top computer by 88% of participants. The modified version of the screening questionnaire can be found in Appendix A.

Categorical DSM-IV diagnoses were made according to the number of criteria recommended as a diagnostic cut-off by the screening instrument.

When more than one question referred to a single diagnostic criterion, a positive response to any of these on the questionnaire would result in the criterion being coded as present.

Subsequently, the algorithms used to categorise each Axis II category were adjusted, with the aim of increasing concordance between the scores obtained using the screening instrument and the 'gold standard' obtained using the clinical interview in Stage 2. The strategy employed was to manipulate the cut-off points used on the screen questionnaire in order to increase levels of agreement, as measured by the kappa coefficient, firstly between individual criteria, and secondly the screen questionnaire and the Axis II diagnoses derived from the structured clinical interview. A detailed description of the algorithms can be found in Appendix B.

Assessment instruments (Stage 2)

The subsample of one in five participants were interviewed by clinicians (six psychiatrists at specialist registrar level, members of the Royal College of Psychiatrists, and two clinical psychologists) at a mean interval of two weeks following Stage 1. Prisoners were administered the Schedules for Clinical Assessment in Neuropsychiatry (SCAN; World Health Organisation, 1992) and the SCID-II (First et al., 1997). The clinical interviewers were trained in the administration of the SCID-II by the last author. The SCID-II questionnaire was not included in this second stage.

Results

Table 1 shows the prevalence of Axis II personality disorder categories according to the three different systems – the clinician-administered SCID-II and the SCID-II screen, both before and after adjustment – indicating high prevalences in this prisoner sample. Prior to adjusting the algorithms for diagnostic categories, the unadjusted screen diagnosed an overwhelming majority of subjects with a personality disorder (92%), indicating a high level of false positives for a personality disorder diagnosis. The screening instrument appeared to perform most poorly for the categories obsessive-compulsive, schizotypal, schizoid, histrionic, and narcissistic personality disorder – categories with the lowest prevalences derived from the SCID-II administered by clinicians. Adjusting the screen using the new algorithms resulted in a substantial improvement and a substantial reduction in the number of false positives for all categories except schizoid personality disorder.

Table 2 shows the internal consistency, as measured by Cronbach's alpha, of the SCID-II screen for the entire sample of prisoners. The values of the alpha coefficient, based on the relevant SCID-II screen items, ranged from .53 for the adult criteria for antisocial personality disorder to .85 for

Table 1. Comparison of Axis II categorical diagnoses using SCID-II interview, SCID-II screen, and the adjusted version of the SCID-II screen.

	Clinical interview ($n = 496$)		Screen ($n = 3142$)		Adjusted screen ($n = 3142$)	
	n	(%)	*n*	(%)	*n*	(%)
Avoidant	52	(10)	1122	(36)	398	(13)
Dependent	16	(3)	406	(13)	228	(7)
Obsessive-compulsive	46	(9)	1649	(53)	412	(13)
Paranoid	112	(23)	1751	(56)	1190	(38)
Schizotypal	11	(2)	1456	(46)	539	(17)
Schizoid	31	(6)	905	(29)	905	(29)
Histrionic	10	(2)	572	(18)	57	(2)
Narcissistic	35	(7)	1535	(49)	102	(3)
Borderline	91	(18)	1714	(55)	545	(17)
Antisocial	218	(44)	2039	(65)	1712	(55)
No personality disorder	169	(34)	236	(8)	731	(23)

Table 2. Cronbach's alpha reliability of SCID-II screen scores for adjusted and non-adjusted criteria for individual personality disorders and the three clusters of personality disorders ($n = 3142$).

Category	Unadjusted	Adjusted
Adult antisocial	.53	.60
Conduct	.85	.85
Borderline	.81	.78
Narcissistic	.64	.64
Histrionic	.64	.64
Schizoid	.50	.50
Schizotypal	.64	.59
Paranoid	.73	.72
Obsessive-compulsive	.57	.53
Dependent	.68	.68
Avoidant	.71	.71
Cluster A (Schizoid, schizotypal, and paranoid)	.81	.80
Cluster B (Histrionic, borderline, narcissistic, and antisocial)	.79	.78
Cluster C (Avoidant, dependent, and obsessive-compulsive)	.80	.80

conduct disorder, and from .79 to .81 for Cluster A, B, and C personality disorders using the unadjusted algorithms for determining Axis II categories. Following adjustment of the algorithms, little change was observed in the internal consistency of these categories.

Axis I characteristics of second stage sample

Axis I mental disorders in the last 12 months prior to interview, derived from the SCAN, are given in Table 3. The most prevalent diagnoses were of substance misuse, which were highly comorbid with the presence of any Axis II disorder. Before adjustment, the SCID-II screen performed poorly, in that there appeared to be little discrimination between the Axis II disorders and a positive finding of no personality disorder. Following adjustment of the algorithms to determine categorical diagnoses, there was an improvement in the screening instrument's ability to discriminate between Axis I and Axis II disorder.

Comparison of SCID-II screen and structured clinical interview

Table 4 demonstrates that the dimensional scores (derived from the number of items in each category) of the two measures were significantly correlated for all categories of personality disorder, and that each PD score in the screen correlated most strongly with the same diagnosis measured by the interview. However, concordance for individual diagnostic categories was generally poor prior to adjustment (Table 5). Following adjustment, there was improved concordance. However, only four personality disorder categories and the 'any personality disorder' category demonstrated kappas above .40, with only one, antisocial personality disorder, above .50.

Tables 5 and 6 show the kappas, sensitivity, specificity, and positive and negative predictive power of the SCID-II screen diagnoses before and after adjustment, according to diagnoses obtained using clinician interviews. The sensitivities of the screen were more than 80% for all but schizoid and histrionic personality disorder prior to adjustment, but fell for most categories after adjustment. The mean specificity was 62% (41.8–90.4%) but rose to a mean of 83% (61.6–98.3%) after adjustment. Negative predictive values were all above 90% prior to adjustment and remained above 90% subsequently, despite a fall for most categories, except antisocial personality disorder. The predictive positive values of the unadjusted screen were moderate to poor for most categories, except antisocial and borderline personality disorder, but demonstrated improvement following adjustment. Nevertheless, the positive predictive values for schizotypal, schizoid, and histrionic personality disorder still remained poor.

Discussion

We compared the prevalence of personality disorder categories obtained through the SCID-II self-report screen with those derived from the research diagnostic interview administered by clinicians, and found that the screen over-diagnosed personality disorder categories, as had been expected. Prior

Table 3. SCAN Axis I diagnosis (12 months) and those screening positive and negative for personality disorders using SCID-II interview, SCID-II screen, and adjusted version of the SCID-II screen.

Axis I diagnosis (12 months)	Clinical interview				Screen				Adjusted screen			
	PD (n = 327)		NPD (n =169)		PD (n = 460)		NPD (n = 43)		PD (n = 381)		NPD (n = 122)	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Schizophrenia/ delusional	35[c]	(92)	3	(8)	42[a]	(100)	0	–	40[b]	(95)	2	(5)
Unspecified NOPD	24[b]	(92)	2	(8)	30	(100)	0	–	29[b]	(97)	1	(3)
Manic episodes	6	(100)	0	–	6	(100)	0	–	5	(83)	1	(17)
Depressive episodes	49[a]	(79)	13	(21)	62[a]	(98)	1	(2)	55[a]	(87)	8	(13)
Other mood disorder	18	(86)	3	(14)	21	(96)	1	(5)	20	(91)	2	(9)
Brain dysfunction	11[a]	(100)	0	–	12	(100)	0	–	12[a]	(100)	0	–
Phobia	28[a]	(85)	5	(15)	34	(100)	0	–	33[b]	(97)	1	(3)
Anxiety	50[c]	(91)	5	(9)	55	(98)	1	(2)	52[b]	(93)	4	(7)
Obsessive-compulsive	8[a]	(100)	0	–	8	(100)	0	–	8	(100)	0	–
Alcohol	212[c]	(65)	115	(35)	261[b]	(95)	14	(5)	226[c]	(82)	49	(18)
Opioids	127[c]	(84)	25	(16)	149[b]	(97)	4	(3)	136[c]	(89)	17	(11)
Cannabis	102[c]	(86)	17	(14)	119[b]	(99)	1	(1)	111[c]	(93)	9	(8)
Sedative/hypnotics	73[c]	(85)	13	(15)	86[b]	(100)	0	–	77[b]	(90)	9	(11)
Cocaine	104[c]	(86)	17	(14)	119[b]	(98)	2	(2)	110[c]	(91)	11	(9)
Stimulants	106[c]	(87)	16	(13)	123[c]	(100)	0	–	114[a]	(93)	9	(7)
Multiple drug use	44[b]	(86)	7	(14)	51[a]	(100)	0	–	47[b]	(92)	4	(8)

Note: PD = personality disorder present; NPD = no diagnosis of personality disorder. Total numbers and percentages refer to those with Axis I mental disorder with and without personality disorder. Significance based on Pearson χ^2 test (two-tailed), $df = 1$; [a]$p < .05$; [b]$p < .01$; [c]$p < .001$.

Table 4. Correlation coefficients between scores on the SCID-II interview and on the screen ($n = 496$).

	SCID-II screen										
SCID-II interview	AVOI	DEPT	OBCM	PARA	SZTP	ZOID	HIST	NARC	BORD	COND	ADAS
AVOI	.57	.42	.19	.40	.43	.23	.08^{ns}	.25	.46	.18	.11
DEPT	.42	.57	.21	.31	.31	.19	.14	.33	.39	.18	.12
OBCM	.23	.19	.42	.30	.31	.21	.16	.31	.27	.13	.11
PARA	.39	.30	.24	.58	.43	.32	.16	.43	.48	.36	.27
SZTP	.40	.30	.21	.35	.47	.28	.05^{ns}	.25	.41	.16	.13
ZOID	.19	.07^{ns}	.01^{ns}	.15	.19	.32	−0.9	.05^{ns}	.14	.18	.16
HIST	.04^{ns}	.17	.10	.18	.20	.04^{ns}	.44	.34	.23	.11	.14
NARC	.12	.16	.13	.28	.25	.18	.20	.37	.29	.20	.24
BORD	.46	.48	.19	.53	.45	.26	.22	.40	.65	.36	.31
COND	.22	.22	−.02^{ns}	.35	.22	.17	.20	.30	.38	.83	.52
ADAS	.21	.25	.03^{ns}	.37	.23	.13	.25	.36	.47	.60	.65

Note: AVOI = avoidant; DEPT = dependent; OBCM = obsessive compulsive; PARA = paranoid; SZTP = schizotypal; ZOID = schizoid; HIST = histrionic; NARC = narcissistic; BORD = borderline; COND = conduct disorder; ADAS = adult antisocial; ns = non-significant correlation.

Table 5. Results demonstrating the efficiency of the SCID-II screen in detecting personality disorder prior to adjustments ($n = 495$).

Category	Present both[a]		Absent both[a]		Lay only		Clinical only		Kappa	Sensitivity %	Specificity %	PV+ %	PV− %
	n	(%)	*n*	(%)	*n*	(%)	*n*	(%)					
Avoidant	44	(9)	313	(63)	130	(26)	8	(2)	.27	84.6	70.6	25.2	97.5
Dependent	15	(3)	433	(88)	46	(9)	1	(1)	.36	93.8	90.4	24.6	99.8
Obsessive-compulsive	40	(8)	233	(48)	216	(44)	6	(1)	.13	87.0	51.9	15.6	97.5
Schizotypal	11	(2)	263	(53)	221	(45)	0	–	.05	100	54.3	4.7	100
Schizoid	19	(4)	341	(69)	123	(25)	12	(2)	.13	61.3	73.5	13.4	96.6
Paranoid	103	(21)	206	(42)	177	(36)	9	(2)	.30	92.0	53.8	36.8	95.8
Histrionic	6	(1)	391	(79)	94	(19)	4	(1)	.08	60.0	80.6	6.0	99.0
Narcissistic	31	(6)	246	(50)	214	(45)	4	(1)	.11	88.6	53.5	12.7	98.4
Borderline	87	(18)	212	(43)	192	(39)	4	(1)	.27	95.6	52.5	31.2	98.1
Antisocial	209	(42)	155	(31)	122	(25)	9	(2)	.49	95.9	56.0	63.1	94.5
Conduct disorder	285	(58)	125	(25)	78	(16)	7	(1)	.63	97.6	61.6	78.5	94.7
Adult antisocial	296	(60)	76	(15)	106	(21)	17	(4)	.40	94.6	41.8	73.6	81.7

[a]both = in both the screen and the interview.

Table 6. Results demonstrating the efficiency of the SCID-II screen in detecting personality disorder after cut-off and criterion adjustments ($n = 495$).

Category	Present both[a]		Absent both[a]		Lay only		Clinical only		Kappa	Sensitivity %	Sensitivity %	PV+ %	PV− %
	n	(%)	*n*	(%)	*n*	(%)	*n*	(%)					
Avoidant	27	(6)	410	(83)	33	(7)	25	(5)	.42	51.9	92.6	45.0	94.3
Dependent	13	(3)	453	(92)	26	(5)	3	(1)	.45	81.3	94.6	33.3	99.3
Obsessive-compulsive	20	(4)	409	(83)	40	(8)	26	(5)	.30	43.5	91.1	33.3	94.0
Schizotypal	9	(1)	408	(82)	76	(15)	2	(2)	.15	81.8	84.3	10.6	99.5
Schizoid	19	(4)	341	(69)	123	(25)	12	(1)	.13	61.3	73.5	13.4	96.6
Paranoid	91	(18)	272	(55)	111	(22)	21	(4)	.41	81.3	71.0	45.0	92.8
Histrionic	2	(1)	476	(96)	9	(2)	8	(2)	.17	20.0	98.1	18.2	98.3
Narcissistic	7	(1)	452	(91)	8	(2)	28	(6)	.25	20.0	98.3	46.7	94.2
Borderline	51	(10)	368	(74)	36	(7)	40	(8)	.48	56.0	91.1	58.6	90.2
Antisocial	193	(39)	192	(39)	85	(17)	25	(5)	.56	88.5	69.3	69.4	88.5
Conduct	285	(58)	125	(25)	78	(16)	7	(1)	.63	97.6	61.6	78.5	94.7
Adult antisocial	265	(53)	132	(27)	50	(10)	48	(10)	.57	84.7	72.5	84.1	73.3
Any personality disorder	279	(56)	90	(18)	97	(20)	29	(6)	.42	90.6	48.1	74.2	75.6

[a]both = in both the screen and the interview.

to adjustment, the screen showed generally good sensitivities but only moderate specificities. Following adjustment, this pattern was generally reversed. The primary intention of this study was to improve the positive predictive values of the individual diagnostic categories. This was achieved successfully, following adjustment, for all but one personality disorder category. Certain limitations remained, in terms of the concordance of certain screen-derived diagnoses with those from the structural clinical interview.

Correspondence between screen and clinician diagnosis of most categories was consistently poor prior to adjustment, apart from for antisocial personality disorder. Following adjustment, the weakest diagnoses when using the adjusted screen appeared to be schizoid, schizotypal, and histrionic personality disorder. Similar findings were reported by Hyler, Skodol, Kellman, Oldham, and Rosnick (1990), after using the Personality Diagnostic Questionnaire-Revised (PDQ-R) for schizoid and histrionic personality disorder in a smaller sample of subjects presenting at a psychiatric clinic. Overall, the levels of agreement between the adjusted self-report instrument and the clinical interview in this study were moderate, but were good for antisocial, borderline, and the presence of any personality disorder. This would suggest that the SCID-II screen is useful for examining the correlates between other variables and categories of personality disorder, but only following adjustment of the algorithms to determine these categories. The derived categories from the adjusted instrument were at an acceptable level of concordance for antisocial, borderline, narcissistic, paranoid, avoidant, dependent, and obsessive-compulsive, but of limited accuracy for schizotypal, schizoid, and histrionic personality disorders.

Only one previous study has compared diagnostic convergence between the SCID-II and a self-administered questionnaire in a prison sample. Davison, Leese, and Taylor (2001) administered the Personality Diagnostic Questionnaire-4+ (PDQ-4+) to 62 prisoners convicted of sexual and violent offences. Eight Axis II categories can be compared across the studies. The present study demonstrates higher kappas for all but one category (borderline), and a median kappa of .36 compared to .28 for the eight categories in the previous study. Clark, Livesley, and Morey (1997) reviewed the findings of previous studies measuring diagnostic convergence between interviews and questionnaires in psychiatric clinic samples. The results of this study compare favourably with previous studies in the higher range of diagnostic convergence which compared the Personality Disorder Examination (PDE) and SCID-II with the PDQ (Hyler et al., 1990; Hyler, Skodol, Oldham, Kellman, & Doidge, 1992).

However, although these and the present study are within the higher range of diagnostic convergence among studies that have examined this question, it cannot be claimed that the general level of convergence was other than moderate to poor. Compared to the median kappa of .36 for convergence of the 10 categorical diagnoses in this study, median kappas in

11 previous comparisons ranged from .08 to .42, with a grand median of .27. The kappa of .42 for any personality disorder in this study compares with a range of .19–.47 (grand median = .29) in seven previous comparisons. In the 11 previous comparisons to have reported correlations, median *r* values ranged from .19 to .54 (grand median = .39), compared to .52 for this study. These findings suggest that the SCID-II screen compares favourably to the PDQ in identifying personality disorder diagnoses when both instruments are compared to the SCID-II clinical interview.

Closer examination of the findings of Hyler and colleagues suggests that the two screening instruments differ for individual categories, with the PDQ performing better for categories such as schizotypal, avoidant, and dependant (which were more prevalent in their clinic samples) and the SCID-II screen performing better for the antisocial and paranoid personality disorders (which are more prevalent in correctional samples). This would suggest, on the one hand, that different self-report instruments may be suitable for different samples, but, on the other, that both instruments might be improved by modifying their content according to which self-report questions performed best in relation to the criteria they were intended to identify. However, Clark et al. (1997) concluded that, for the most part, diagnoses are not significantly comparable across different methods beyond chance, arguing for further improvement in assessment instruments but also a change in the conceptualisation of current classifications. They argued for greater use of dimensional measures in future studies to offset the loss of reliable and valid information occasioned by converting data to a nominal scale (Widiger & Francis, 1994).

The SCID-II screen is a useful and valid alternative to currently available self-report instruments for diagnosing Axis II disorders, but is not a substitute for the complete SCID-II interview. Nevertheless, previous studies have demonstrated that even structured clinical interviews show equally divergent results when compared to each other, questioning the very notion of a 'gold standard' in the diagnosis of Axis II disorder, where one research diagnostic instrument cannot simply be substituted one for another (Clark et al., 1997). The SCID-II screen was very useful for identifying prisoners with personality disorder in a sample where prevalence was very high. For the category antisocial personality disorder, these results suggest that the self-report instrument may in fact be a valid substitute. However, the performance of the screen remains uncertain for the general population, where its predictive power is likely to be lower due to lower prevalences of personality disorder.

Lenzenweger, Loranger, Korfine, and Neff (1997) have developed the first self-report screen for use in a two-stage epidemiological survey based on the International Personality Disorder Examination (IPDE) structured clinical interview. The negative predictive value for a definite or probable diagnosis of any personality disorder in a sample of university students was .98. However, the positive predictive value was only .22 for the second stage. Their study has clearly demonstrated many of the difficulties involved in

developing effective screening instruments for Axis II disorders, in that, despite a reduction in the number of persons that should be interviewed in the second stage, and the low risk of missing cases (false negatives), it would still be necessary to interview between four and five cases to identify one with a personality disorder using the IPDE in the second phase of a general population survey.

Despite the inevitable limitations of self-report instruments, the SCID-II screen was selected for the National Household Survey of England and Scotland, where it has been applied in a true two-stage design to determine the participants to be interviewed in the second phase (Singleton, Bumpstead, O'Brien, Lee, & Meltzer, 2001). However, the cost of the survey resulted in the second phase being used primarily to establish a more accurate prevalence of personality disorder than could be achieved using the screen, while correlates with less common variables must be established using measures from the screen in the first phase. The methodological limitation of the reliance of self-report instead of structured clinical interview is not restricted to the Axis II disorders, however, as the predictive values of most Axis II categories obtained from the National Survey of Prisoners compared favourably with those from screening instruments used for the Axis I disorders (Singleton et al., 1998).

Taking into account the limitations we have demonstrated with these Axis II measures, we nevertheless believe that the SCID-II screen has produced important epidemiological data on our national prisoner population for the determination of future service needs, can demonstrate important criminological correlations, and will be employed in future studies to develop measures that may determine the risk of re-offending.

Acknowledgements

We would like to thank Professor Janet B.W. Williams for her advice on the manuscript. Richard Noon and Caroline Nyami collected data for the study.

References

Ammons, R.B., & Ammons, C.H. (1962). The Quick Test: Provisional manual. *Psychological Reports*, *11*, 111–161.

Bebbington, P., & Nayani, T. (1995). The Psychosis Screening Questionnaire. *International Journal of Methods in Psychiatric Research*, *5*, 11–19.

Clark, L.A., Livesley, W.J., & Morey, L. (1997). Personality disorder assessment: The challenge of construct validity. *Journal of Personality Disorders*, *11*, 205–231.

Coid, J.W. (1984). How many psychiatric patients in prison? *British Journal of Psychiatry*, *145*, 78–86.

Coid, J.W. (1998). The management of dangerous psychopaths in prison. In T. Millon, E. Simonsen, M. Birket-Smith, & R.D. Davis (Eds.), *Psychopathy: Antisocial, criminal and violent behaviour* (pp. 431–457). New York: Guildford Press.

Davison, S., Leese, M., & Taylor, P.J. (2001). Examination of the screening properties of the Personality Diagnostic Questionnaire 4 + (PDQ-4+) in a prison population. *Journal of Personality Disorders*, *15*, 180–194.

Department of Health. (1992). *The health of the nation: A strategy for health in England (cm 1986)*. London: HMSO.

Dohrenwend, B.P., & Shrout, P.E. (1981). Toward the development of a two-stage procedure for case identification and classification in psychiatric epidemiology. *Research into Community Mental Health*, *2*, 295–323.

First, M.B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., & Benjamin, L. (1997). *Structured clinical interview for DSM-IV Axis-II personality disorders*. Washington, DC: American Psychiatric Press.

Gill, B., Meltzer, H., Hinds, K., & Pettigrew, M. (1996). *OPCS surveys of psychiatric morbidity in Great Britain, report 7: Psychiatric morbidity among homeless people*. London: HMSO.

Hyler, S.E., Skodol, A.E., Kellman, D., Oldham, J.M., & Rosnick, L. (1990). Validity of the Personality Diagnostic Questionnaire–Revised: Comparison with two structured interviews. *American Journal of Psychiatry*, *147*, 1043–1048.

Hyler, S.E., Skodol, A.E., Oldham, J.M., Kellman, H.D., & Doidge, N. (1992). Validity of the Personality Diagnostic Questionnaire–Revised: A replication in an outpatient sample. *Comprehensive Psychiatry*, *33*, 73–76.

Lamb, H.R., & Weinberger, L.E. (1998). Persons with severe mental illness in jails and prisons: A review. *Psychiatric Services*, *49*, 483–492.

Lenzenweger, M.F., Loranger, A.W., Korfine, L., & Neff, C. (1997). Detecting personality disorders in a non-clinical population. *Archives of General Psychiatry*, *54*, 345–351.

Lewis, G., Pelosi, A., Araya, R.C., & Dunne, G. (1992). Measuring psychiatric disorder in the community: A standardised assessment for lay interviewers. *Psychological Medicine*, *22*, 465–486.

Loranger, A.W. (1992). Are current self-report and interview measures adequate for epidemiological studies of personality disorders? *Journal of Personality Disorders*, *6*, 313–325.

Meltzer, H., Gill, B., Hinds, K., & Pettigrew, M. (1996). *OPCS surveys of psychiatric morbidity in Great Britain, report 4: The prevalence of psychiatric morbidity among adults living in institutions*. London: HMSO.

Meltzer, H., Gill, B., Pettigrew, M., & Hinds, K. (1995). *OPCS surveys of psychiatric morbidity in Great Britain, report 1: the prevalence of psychiatric morbidity among adults living in private households*. London: HMSO.

Moran, P. (1999). *Antisocial personality disorder: An epidemiological perspective*. London: Royal College of Psychiatrists.

Salekin, R., Rogers, R., & Sewell, K. (1996). A review and meta-analysis of the Psychopathy Checklist–Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice*, *3*, 203–215.

Singleton, N., Bumpstead, R., O'Brien, M., Lee, A., & Meltzer, H. (2001). *Psychiatric morbidity among adults living in private households, 2000*. London: Stationery Office.

Singleton, N., Meltzer, H., Gatward, R., Coid, J.W., & Deasy, D. (1998). *Psychiatric morbidity among prisoners in England and Wales*. London: Stationery Office.

Widiger, T.A., & Francis, A. (1994). Towards a dimensional model for the personality disorders. In P.T. Costas & T. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (pp. 19–20). Washington, DC: American Psychological Association.

World Health Organisation. (1992). *SCAN: Schedules for Clinical Assessment in Neuropsychiatry*. Geneva: Author.

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

Appendix A: modified SCID-II screening questionnaire

The following questions are answered:

(1) Yes
(2) No
(9) Don't understand/does not apply

Q1: Have you avoided jobs or tasks that involved having to deal with a lot of people?
Q2: Do you avoid getting involved with people unless you are certain they will like you?
Q3: Do you find it hard to be 'open' with people you are close to?
Q4: Do you often worry about being criticised or rejected in social situations?
Q5: Are you usually quiet when you meet new people?
Q6: Do you believe that you're not as good, as smart, or as attractive as most other people?
Q7: Are you afraid of trying new things?
Q8: Do you need a lot of advice or reassurance from others before you can make everyday decisions – like what to wear or what to order in a restaurant?
Q9: Do you depend on other people to handle important areas in your life such as finances, child care, or living arrangements?
Q10: Do you find it hard to disagree with people even when you think they are wrong?
Q11: Do you find it hard to start or work on tasks when there is no one to help you?
Q12: Have you often volunteered to do things that are unpleasant?
Q13: Do you usually feel uncomfortable when you are by yourself?
Q14: When a close relationship ends, do you feel you immediately have to find someone else to take care of you?
Q15: Do you worry a lot about being left alone to take care of yourself?
Q16: Are you the kind of person who focuses on details, order and organisation, or likes to make lists and schedules?
Q17: Do you have trouble finishing jobs because you spend so much time trying to get things exactly right?
Q18: Throughout your life have you or other people felt that you have been so devoted to work (or school) that you have had no time left for anyone else or for just having fun?
Q19: Do you have very high standards about what is right and what is wrong?
Q20: Do you have trouble throwing things out because they might come in handy some day?
Q21: Is it hard for you to let other people help you if they don't agree to do things exactly the way you want?
Q22: Is it hard for you to spend money on yourself and other people even when you have enough?
Q23: Are you often so sure you are right that it doesn't matter what other people say?
Q24: Have other people told you that you are stubborn or rigid?
Q25: Do you often have to keep an eye out to stop people from using you or hurting you?
Q26: Do you spend a lot time wondering if you can trust your friends or the people you work with?
Q27: Do you find that it is best not to let other people know much about you because they will use it against you?
Q28: Do you often detect hidden threats or insults in things people say or do?

Q29: Are you the kind of person who holds grudges or takes a long time to forgive people who have insulted or slighted you?
Q30: Are there many people you can't forgive because they did or said something to you a long time ago?
Q31: Do you often get angry or lash out when someone criticises or insults you in some way?
Q32: Have you often suspected that your spouse or partner has been unfaithful?
Q33: When you are out in public and see people talking, do you often feel that they are talking about you?
Q34: Do you often get the feeling that things that have no special meaning to most people are really meant to give you a message?
Q35: When you are around people, do you often get the feeling that you are being watched or stared at?
Q36: Have you ever felt that you could make things happen just by making a wish or thinking about them?
Q37: Have you had personal experiences with the supernatural?
Q38: Do you believe that you have a 'sixth sense' that allows you to know and predict things that others can't?
Q39: Do you often think that objects or shadows are really people or that noises are actually people's voices?
Q40: Have you had the sense that some person or force is around you, even though you cannot see anyone?
Q41: Do you often see auras or energy fields around people?
Q42: Are there very few people that you're really close to outside your immediate family?
Q43: Do you often feel nervous when you are with other people?
Q44: Are you the kind of person who doesn't think it is important to have any close relationships?
Q45: Would you almost always rather do things alone than with other people?
Q46: Could you be content without ever being sexually involved with anyone?
Q47: Are there really very few things in life that give you pleasure?
Q48: Are you the kind of person who doesn't care what people think of you?
Q49: Are you the sort of person who finds that nothing makes you very happy or very sad?
Q50: Do you like to be the centre of attention?
Q51: Do you flirt a lot?
Q52: Do you try to draw attention to yourself by the way you dress or look?
Q53: Do you often make a point of being dramatic and colourful?
Q54: Do you often change your mind about things depending on the people you're with or what you have just read or seen on TV?
Q55: Do you have lots of friends that you are very close to?
Q56: Do people often fail to appreciate your very special talents or accomplishments?
Q57: Have people told you that you have too high an opinion of yourself?
Q58: Do you think a lot about the power, fame or recognition that will be yours someday?
Q59: Do you think a lot about the perfect romance that will be yours someday?
Q60: When you have a problem, do you almost always insist on seeing the top person?
Q61: Do you feel it is important to spend time with people who are special or influential?
Q62: Is it very important to you that people pay attention to you or admire you in some way?

Q63: Do you think that it's not necessary to follow certain rules or social conventions when they get in your way?
Q64: Do you feel that you are the kind of person who deserves special treatment?
Q65: Do you often find it necessary to step on a few toes to get what you want?
Q66: Do you often have to put your needs above other people's?
Q67: Do you often expect other people to do what you ask without question because of who you are?
Q68: Are you the sort of person who is not really interested in other people's problems or feelings?
Q69: Have people complained to you that you don't listen to them or care about their feelings?
Q70: Are you often envious of others?
Q71: Do you feel that others are often envious of you?
Q72: Do you find that there are very few people that are worth your time and attention?
Q73: Have you often become frantic when you thought that someone you really cared about was going to leave you?
Q74: Do your relationships with people you really care about have lots of extreme ups and downs?
Q75: Have you all of a sudden changed your sense of who you are and where you are headed?
Q76: Does your sense of who you are often change dramatically?
Q77: Are you different with different people or in different situations so that you sometimes don't know who you really are?
Q78: Have there been lots of sudden changes in your goals, career plans, religious beliefs, and so on?
Q79: Have you often done things impulsively?
Q80: Have you tried to hurt or kill yourself, or threatened to do so?
Q81: Have you ever cut, burned, or scratched yourself on purpose?
Q82: Do you have a lot of sudden mood changes?
Q83: Do you often feel empty inside?
Q84: Do you often have temper outbursts or get so angry that you lose control?
Q85: Do you hit people or throw things when you get angry?
Q86: Do even little things get you very angry?
Q87: When you are under a lot of stress, do you get suspicious of other people or feel especially 'spaced out' like you were on drugs?
Q88: Before you were 15, would you bully or threaten other kids?
Q89: Before you were 15, would you start fights?
Q90: Before you were 15, did you hurt or threaten someone with a weapon, like a bat, brick, broken bottle, a knife or a gun?
Q91: Before you were 15, did you deliberately torture someone or cause someone physical pain or suffering?
Q92: Before you were 15, did you torture or hurt animals on purpose?
Q93: Before you were 15, did you rob, mug, or forcibly take something from someone by threatening him or her?
Q94: Before you were 15, did you force someone to have sex with you, get undressed, or touch you sexually?
Q95: Before you were 15, did you start fires?
Q96: Before you were 15, did you deliberately destroy things that weren't yours?
Q97: Before you were 15, did you break into houses, other buildings, or cars?
Q98: Before you were 15, did you lie a lot or con other people?

Q99: Before you were 15, did you sometimes steal or shoplift things or forge someone's signature?
Q100: Before you were 15, did you run away and stay overnight?
Q101: Before you were 13, did you often stay out very late, long after the time you were supposed to be home?
Q102: Before you were 13, did you often skip school?
Q103: Now thinking of the time SINCE you were 15, do you often do things on the spur of the moment without thinking about how it will affect you or other people?
Q104: Since you were 15, has there been a period when you had no regular place to live, for at least one month or so?
Q105: Have you ever hit or thrown things at your spouse or partner?
Q106: Since you were 15, have you ever hit a child, yours or someone else's, so hard that he or she had bruises, or had to stay in bed or see a doctor?
Q107: Since you were 15, have you been in any fights?
Q108: Since you were 15, have you used a weapon, like a stick, knife, or gun in a fight?
Q109: Do you feel guilty or remorseful for previous behaviour such as having hurt, mistreated, or stolen from other people?
Q110: Since you were 15, besides the offence for which you are currently in prison, have you done things that are against the law – even if you weren't caught? For example, have you stolen things?
Q111: Since you were 15, have you used or sold drugs?
Q112: Since you were 15, have you passed bad cheques?
Q113: Since you were 15, have you been paid for sex?
Q114: Since you were 15, have you ever used an alias or pretended to be someone else?
Q115: Since you were 15, have you often 'conned' others to get what you want?
Q116: Since you were 15, did you ever drive a car when you were drunk or high on drugs?

Appendix B: adjusted diagnostic algorithm for DSM-IV personality disorders

Avoidant personality disorder

The number of criteria to test positive for avoidant personality disorder outlined in the manual for the screen is four. The greatest level of agreement was achieved using a cut-off score of six criteria.

Dependent personality disorder

The number of criteria to test positive for dependent personality disorder in the manual for the screen is five. The greatest level of agreement was achieved using a cut-off of six criteria.

Obsessive-compulsive personality disorder

According to the screen manual, criterion 8 for obsessive-compulsive personality disorder can be met by a positive response to question 23 or 24. Greatest agreement was found when this was changed to scoring positive on question 23 only. The number of criteria to test positive for obsessive-compulsive personality disorder in

the manual for the screen is four. The level of agreement was increased using a cut-off of six criteria.

Paranoid personality disorder

According to the manual, criterion 5 for paranoid personality disorder can be met by a positive response to question 29 or 30. The greatest agreement was found when this was changed to scoring positive on question 29 only. The level of agreement was increased using a cut-off of five criteria for diagnosis instead of four.

Schizotypal personality disorder

According to the screen manual, criterion 1 for schizotypal personality disorder can be met by a positive response to question 33, 34, or 35. The greatest agreement was found when this was changed to scoring positive on at least two of these questions. Manual: criterion 2 for schizotypal personality disorder can be met by a positive response to question 36, 37, or 38. Adjustment: the greatest agreement was found when this was changed to scoring positive on all three of these questions. Manual: criterion 3 for schizotypal personality disorder can be met by a positive response to question 39, 40, or 41. Adjustment: the greatest agreement was found when this was changed to scoring positive on at least two of these questions. Manual: criterion 5 for schizotypal personality disorder can be met by a positive response to question 25, 26, 27, 28, or 32. Adjustment: the greatest agreement was found when the cut-off was changed to scoring positive on at least four of these questions. There was no change to the total number of criteria required for a probable diagnosis.

Schizoid personality disorder

No manipulation increased the agreement between the screen and the clinical interview.

Histrionic personality disorder

The greatest agreement was found when the cut-off was changed to scoring positive on at least six criteria.

Narcissistic personality disorder

According to the screening manual, criterion 1 for narcissistic personality disorder can be met by a positive response to question 56 or 57. The greatest agreement was found when this was changed to scoring positive on both of these questions. Manual: criterion 3 for narcissistic personality disorder can be met by a positive response to question 60 or 61. Adjustment: the greatest agreement was found when this was changed to scoring positive on both questions. Manual: criterion 6 for narcissistic personality disorder can be met by a positive response to question 65, 66, or 67. Adjustment: the greatest agreement was found when this was changed to scoring positive on at least two of these questions. The greatest agreement was found when the cut-off for diagnosis was changed to scoring positive on at least seven criteria.

Borderline personality disorder

According to the screening manual, criterion 3 for borderline personality disorder can be met by a positive response to question 75, 76, 77, or 78. The greatest

agreement was found when this was changed to scoring positive on all four of these questions. Manual: criterion 5 for borderline personality disorder can be met by a positive response to question 80 or 81. Adjustment: the greatest agreement was found when this was changed to scoring positive on both of the two questions. Manual: criterion 8 for borderline personality disorder can be met by a positive response to question 84, 85, or 86. Adjustment: the greatest agreement was found when this was changed to scoring positive on at least two of these questions. The greatest agreement was found when the cut-off for diagnosis was changed to scoring positive on at least seven of these criteria.

Antisocial personality disorder

This personality disorder is diagnosed when both conduct disorder and adult antisocial disorder are present.

Conduct disorder

There were no changes made for this category.

Adult antisocial disorder

According to the screening manual, criterion 1 can be met by a positive response to question 110, 111, 112, or 113. The greatest agreement was found when this was changed to scoring positive on at least two of these questions. Manual: criterion 2 can be met by a positive response to question 114 or 115. Adjustment: the greatest agreement was found when this was changed to scoring positive on both of these questions. Manual: criterion 4 can be met by a positive response to question 105, 106, 107, or 108. Adjustment: the greatest agreement was found when this was changed to scoring positive on at least two of these questions. There was no change in the number of criteria used for cut-off.