Soc Psychiatry Psychiatr Epidemiol (2005) 40 : 209–213 DOI 10.1007/s00127-005-0876-3

ORIGINAL PAPER

Yvonne Forsell

The Major Depression Inventory versus Schedules for Clinical Assessment in Neuropsychiatry in a population sample

Accepted: 9 September 2004

Abstract *Aims* This study examined the association between the Major Depression Inventory (MDI) and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN). *Methodology* A questionnaire including the MDI was sent out to an adult population and was completed by a total of 10448 persons. Psychiatrists used SCAN and interviewed a subsample (n = 1093). *Results* The specificity of the MDI was 0.22, the sensitivity 0.67 and Kappa 0.25 when Major Depression according to SCAN was considered as the index of validity, and with all depressive disorders the specificity was 0.44, the sensitivity 0.51 and Kappa 0.33. Higher educated persons and those with reported disability were less likely to be false negatives. The sensitivity and specificity for different cut-off scores when using the MDI total score were calculated. *Conclusion* The result from this study suggests that, when MDI is used in population-based samples, cut-off scores rather than the algorithm for depression should be used. The optimal cut-off score must be chosen according to the aims of the study.

Key words depression – population – rating scales – sensitivity – specificity

Introduction

DSM-IV symptoms criteria are designed for clinical use, but Reiger et al. have suggested they can be used in population studies if the symptom scales include symptom thresholds, impairment/disability and duration criteria (Reiger et al. 1998a; Reiger 1998b).

The most commonly used self-rating scales in population studies were developed before the introduction of DSM-III, e. g. the Beck Depression Inventory (BDI) and the Zung Self-rating Depression Scale (Zung-SDS) (Beck et al. 1961; Zung 1965). The BDI was modified according to DSM-IV in 1996 (Beck et al. 1996).

The Major Depression Inventory (MDI) was developed by Bech et al. based on the DSM-IV symptoms of Major Depression and ICD-10 moderate to severe depression (Bech et al. 1997; Bech and Wermuth 1998). The MDI includes symptom thresholds as well as duration criteria.

This study presents data from an ongoing population-based study on mental health in Stockholm, Sweden. The study used a two-phase design including questionnaires and interviews by psychiatrists. The aim of this paper was to study the association between DSM-IV depression according to the MDI and DSM-IV depression assessed by psychiatrists using Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (Wing et al. 1990).

Subjects and methods

In the period 1998–2000, 19744 persons aged 20–64 years registered in the county of Stockholm were randomly selected from the Stockholm city council register. The ethical committee, Karolinska Institutet, Stockholm, approved the study. The selected population received a questionnaire by mail and 10443 persons answered, a participation rate of 53 %. Each questionnaire was accompanied by a consent that was signed by the participant. Participants as well as non-participants were linked to the following official registers: income and wealth, sick leave, hospital discharge register and disability pension. Participation was related to female gender, higher age, higher income and education, being born in Sweden, and having no psychiatric diagnoses in the hospital discharge register or in the disability pension register. The odds ratios for associations between gender, income, country of origin, education and having a psychiatric diagnosis previously according to the registers were similar among participants and non-participants (Lundberg et al. manuscript).

Self-reported depression

Self-reported depression was assessed using the Major Depression Inventory (MDI) (Bech et al. 1997; Bech and Wermuth 1998). This scale

Y. Forsell, M.D., Ph.D. (✉)
Inst. of Public Health
Division of Social Medicine
Norrbacka, Karolinska Hospital
17176 Stockholm, Sweden
E-Mail: yvonne.forsell@phs.ki.se

includes duration criteria (the symptom must have been present for more than 2 weeks) and intensity criteria (six-point scale from 'not at all' to 'all the time'). The MDI can be used as a diagnostic instrument in which the items are dichotomised to indicate the presence or absence of each of the symptoms.

Algorithms for DSM-IV major depression were made according to the instructions given by the author of the scale (Bech et al. 2001). The ten items were also summarised to a score of 0–50 with each item giving a score from 0 ('not at all') to 5 ('all the time').

Additionally, the following information was used from the questionnaire:

- Age, gender
- Education: divided in ≤ 9 years and ≥ 10 years.
- Somatic disorders: a list of the 25 most common somatic areas that can be affected by disorders was included. Only disorders that were treated by a physician or included hospitalisation were considered.
- Self-reported social disability: social disabilities due to psychological symptoms were assessed using The World Health Organisation Short Disability Assessment Schedule (WHO DAS-S) (Janca et al. 1996). The questions included functions in five areas, leisure activities, daily routines, motivation in work/studies, effectiveness and personal relations. The answers were recorded in a four-point scale: 'not at all', 'sometimes', 'often' or 'a lot'. Social disability was regarded as present if the person answered 'often' or a lot in one or more areas of functioning.

Psychiatric assessment of depressive disorders

1367 persons were randomly selected for interviews, 1104 persons reporting many psychiatric symptoms and 263 persons reporting no or few symptoms. Of the selected persons, 1093 (75 %) completed the interview and non-participation was mostly due to lack of time. The interview was conducted within 2 weeks of receiving the questionnaire. Of the interviewees, 884 had reported many psychiatric symptoms and 208 had reported no or few symptoms in the questionnaire. There were no differences between participants and non-participants in terms of gender, welfare allowance, unemployment benefits, country of origin, sick leave or income. Clinically experienced psychiatrists performed the interview using SCAN (Wing et al. 1990). The interviewers received a 1-week introductory course from one of the WHO-designated trainers and regular training after this. The psychiatrists were not aware of the person's answers in the questionnaire. DSM-IV diagnostic criteria for Axis I disorders were strictly followed. Additionally, diagnoses according to appendix B, criteria sets and axes provided for further study were included (American Psychiatric Association 1994). All depressive disorders included Major Depression, Dysthymia, Bipolar disorders, Mixed Anxiety Depression and Minor Depression.

Statistical analyses

In order to study differences between participants and non-participants, two-tailed Chi-square and t-tests were used. Cronbach coefficient alpha was calculated in order to study the internal consistency of the MDI. As a measure of agreement, Cohen's Kappa index was used. Odds ratio with 95 % CI for having depression according to SCAN was calculated. The Receiver Operating Characteristic (ROC) curve was used to establish the most optimal cut-off point for the MDI score with reference to SCAN. The use of ROC analyses has been described qualitatively and quantitatively in a number of studies (e. g. Hanley and McNeil 1982; Hsiao et al. 1989). ROC analysis is a plot of a measure's true positive rate (sensitivity) and false positive rate (one minus specificity) across the full range of cut-off scores. The line reflects the characteristics of a scale with no discriminating validity. The area under the curve (AUC) is an index of the goodness of the diagnostic scale and the perfect scale has an AUC of 1.0. Analyses were performed using Major Depression and all depressive disorders according to SCAN as the outcome measures.

Results

Cronbach alpha for the MDI was 0.90.

Of the 1093 persons interviewed, 638 were female and 455 male with a mean age of 42.2 (SD 12.5) years. DSM-IV criteria for a depressive disorder according to the MDI were present in 247 persons and according to SCAN in 213 persons. The SCAN depressive disorder diagnoses were as follows: Bipolar disorders 2, Dysthymia 30, Major Depression 81, Mixed Anxiety Depression 29 and Minor Depression 71.

The results when the MDI algorithm for Major Depression was used as a measure are presented in Table 1. The Cohen's Kappa index for Major Depression SCAN/Major depression MDI was 0.25, the sensitivity was 0.67 and the specificity 0.22. For all depressive disorders SCAN/Major Depression MDI, the Kappa index was 0.33, the sensitivity 0.51 and the specificity 0.44.

In persons with all depressive disorders, the false negatives were analysed vs. the true positives. The odds ratios for different characteristics are presented in Table 2. Persons having higher education and those who had reported disability due to psychological symptoms in the questionnaire were less likely to be false negatives.

Table 1 A comparison of Major Depression according to MDI and SCAN, Major Depression and all depressive disorders

	MDI	
SCAN	Major Depression	Non-depressed
Major Depression, n = 81	54	27
No Major Depression, n = 1012	193	819
All depressive disorders, n = 213	108	105
No depressive disorder, n = 880	139	741

Table 2 Predictors of false-negative depressive disorder according to the MDI using SCAN all depressive disorders as the index of validity (108 true positives compared with 105 false negatives)

Variable	Odds ratio (95 % CI)
Male gender	1.0
Female gender	1.0 (0.6–1.8)
Born in Sweden	1.0
Born in other countries	0.5 (0.2–1.1)
18–29 years	1.0
30–44 years	1.5 (0.6–3.5)
45–64 years	1.1 (0.5–2.5)
No reported disability	1.0
Reported disability	0.2 (0.1–0.3)
No somatic disorder	1.0
Somatic disorder	0.6 (0.4–1.1)
Education less than 9 years	1.0
Education more than 9 years	0.5 (0.3–0.9)
5–7 DSM-IV symptoms in SCAN	1.0
7–9 DSM-IV symptoms in SCAN	0.4 (0.3–0.5)

Additionally, persons with more DSM-IV symptoms of Depression in the SCAN more often had reported depression in the MDI. Similar analyses were performed using Major Depression according to SCAN as the index of validity. The tendencies were similar, but the numbers were too small to reach significance.

In order to study the optimal cut-off point for the MDI score, ROC analyses were performed. Fig. 1 presents the ROC curve when persons assessed as having a Major Depression were considered as the index of validity. The AUC was 0.83 (0.78–0.87). Fig. 2 presents the ROC curve using all depressive disorders as the index. The AUC was 0.80 (0.77–0.83). Sensitivity and specificity for different cut-off points are presented in Table 3. At a cut-off score of 26, the sensitivity was 0.61 and the specificity 0.85 for Major Depression. For all depressive disorders, the sensitivity was 0.43 and the specificity 0.89 with 26 as the cut-off score. Choosing a lower cut-off score (20) for Major Depression increased the sensitivity to 0.78 and decreased the specificity to 0.73. Additionally, a lower cut-off score for all depressive disorders (20) increased the sensitivity to 0.67 and decreased the specificity to 0.79.

Discussion

This study evaluated self-reported DSM-IV depression according to the MDI vs. clinician-assessed depression using SCAN in a population sample.

The dropout rate in this population study was fairly high, which might have affected the results. No prevalence rates were calculated since it is likely that persons affected by severe depressive disorders were non-responders. However, the odds ratios for associations between gender, income, country of origin, education and previous psychiatric diagnosis were similar among participants and non-participants.

The use of standardised diagnostic criteria, such as

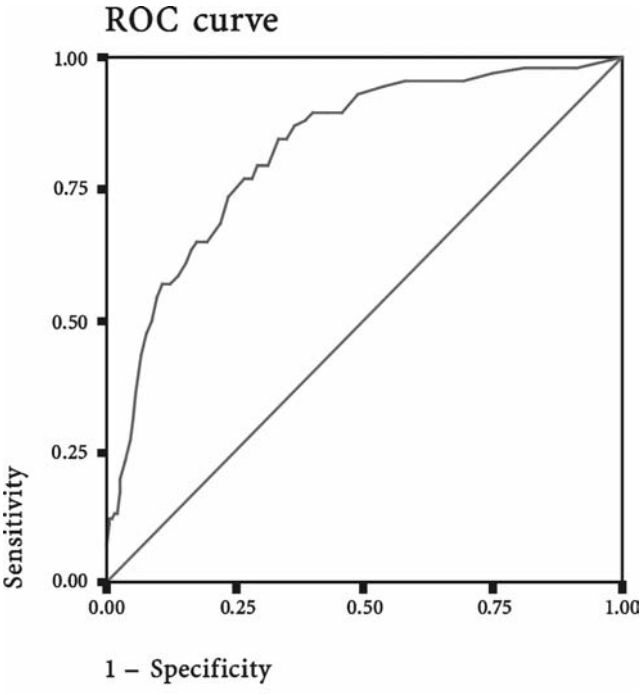

Fig. 1 The Receiver Operating Characteristic curve for the Major Depression Inventory total score with the SCAN diagnosis of Major Depression as the index of validity. The area under the curve is 0.83

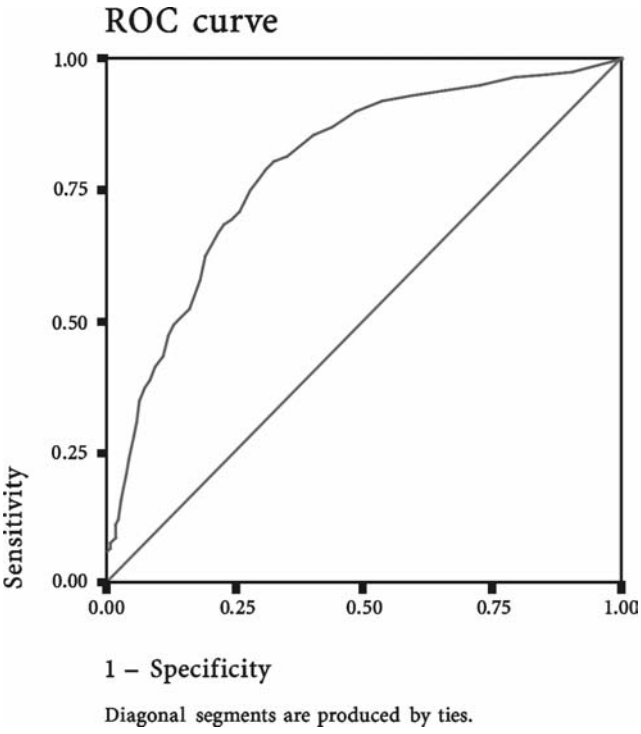

Fig. 2 The Receiver Operating Characteristic curve for the Major Depression Inventory total score when the index of validity was SCAN diagnoses of all depressive disorders. The area under the curve is 0.80

Table 3 Sensitivity and specificity at different cut-off scores for the ten items in the MDI when Major Depression and all depressive disorders were the index of validity

	Major Depression		All depressive disorders	
MDI score	Sensitivity	Specificity	Sensitivity	Specificity
15/16	0.85	0.66	0.76	0.71
16/17	0.80	0.69	0.71	0.74
17/18	0.80	0.70	0.70	0.75
18/19	0.78	0.71	0.68	0.78
19/20	0.78	0.73	0.67	0.79
20/21	0.74	0.76	0.63	0.81
21/22	0.69	0.78	0.59	0.82
22/23	0.65	0.80	0.53	0.84
23/24	0.65	0.82	0.51	0.85
24/25	0.64	0.83	0.50	0.87
25/26	0.61	0.85	0.48	0.88
26/27	0.59	0.86	0.43	0.89
27/28	0.58	0.87	0.42	0.90
28/29	0.58	0.89	0.39	0.91

DSM, is one reason for the advance of psychiatric epidemiology. Reiger et al. have stated that symptom thresholds, impairment/disability and duration criteria need to be included in self-rating scales used in population studies (Reiger et al. 1998a; Reiger 1998b). The MDI is a DSM-based scale for the detection of depression that includes duration and intensity criteria. The internal and external validity has been reported to be higher than for Zung-SDS (Bech and Wermuth 1998). In order to evaluate screening scales, there is a need for a "golden standard" and several studies have used SCAN (Leentjens et al. 2000; Eaton et al. 2000). SCAN incorporates the tenth edition of the Present State Examination and reliability has been reported to be good (Tomov et al. 1990).

The internal consistency of the MDI was high. When Major Depression according to SCAN was the index of validity, the sensitivity of the MDI was 0.67 and the specificity 0.22. Cohen's Kappa was 0.25. When the index was all depressive disorders (SCAN), the sensitivity was 0.51, the specificity was 0.44 and Cohen's Kappa was 0.33. The authors of the scale have reported the sensitivity to be 0.90 and the specificity 0.82 when validating the MDI vs. SCAN in a clinical setting (Bech et al. 2001). This difference in results between clinical settings and population-based studies has previously been demonstrated [validation of Diagnostic Interview Schedule (DIS)] (Robins et al. 1981, Helzer et al. 1985, Anthony et al. 1985). Persons with low education, no reported disability and a less severe depression were more likely to be false negatives. Thus, the MDI seemed to be more accurate for more severely depressed subjects, which explains the discrepancies between the previous study in a clinical setting and this population-based study. This has also been seen in studies evaluating DIS (Eaton et al. 2000).

The authors of the MDI suggest that the optimal cut-off score for DSM-IV diagnosis of Major Depression is 26 (Bech et al. 2001), based on a clinical setting. The results from this study showed that, if a high sensitivity is preferred, a cut-off at 19/20 would be more accurate in a population-based sample (sensitivity 0.78, specificity 0.78).

A way of increasing the specificity might be to add an impairment/disability criterion defined according to DSM-IV. In this study, the odds ratio for being a false negative was 0.2 (0.1–0.3) if the person had reported disability according to WHO-DAS.

Conclusion

The MDI had a high internal consistency. The sensitivity and specificity was relatively low compared to a previous study in a clinical setting. However, when the MDI score was used, a cut-off at 19/20 gave a relatively high sensitivity and specificity. The results from this study suggest that the MDI score rather than the MDI definition of Major Depression should be used in population-based samples. Additionally, an impairment/disability criterion would most likely improve the specificity.

Acknowledgements We would like to thank all persons that participated in collecting the data, especially Tore Hällström, Ingvar Lundberg and Kerstin Damström-Thakker. This study was supported by grants from the Stockholm county and The Swedish Medical Research Council.

References

1. American Psychiatric Association (1994) Diagnostic and Statistical manual of Mental Disorders. American Psychiatric Association (4th ed, rev). Washington, DC, APA
2. Anthony JC, Folstein MF, Romanovski AJ, Von Korff MR, Nestadt GR, Chahal R, Merchant A, Brown CH, Sapiro S, Kramer M, Gruenberg EM (1985) Comparison of the lay Diagnostic Interview Schedule and a standardized psychiatric diagnosis: experience in eastern Baltimore. Arch Gen Psychiatry 42:667–675
3. Bech P, Stage K, Nair NPV, Larsen JK, Kragh-Sorensen P, Gjeriss P (1997) The Major Depression Rating Scale (MDS). Inter-rater reliability and validity across different settings in randomised moclobemide trials. J Affect Disord 42:39–48
4. Bech P, Wermuth, L (1998) Applicability and validity of the Major Depression Inventory in patients with Parkinson's disease. Nord J Psychiatry 52:305–309
5. Bech P, Rasmussen NA, Raabaer K, Olsen L, Noerholm V, Abilgard W (2001) The sensitivity and specificity of the Major Depression Inventory using the Present State Examination as the index of diagnostic validity. J Aff Disord 66:159–164
6. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961) An inventory for measuring depression. Arch Gen Psychiatry 4: 561–571
7. Beck AT, Steer RA, Ball R, Ranieri W (1996) Comparison of the Beck Depression inventories IA and II in psychiatric outpatients. J Pers Assess 132:381–385
8. Eaton WW, Neufeld K, Chen L-S, Cai G (2000) A comparison of self-reports and clinical diagnostic interviews for depression. Arch Gen Psychiatry 57:217–222
9. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36
10. Helzer JE, Robins LN, McEvoy LT, Spitznagel EL, Stoltzman RK, Farmer A, Brockington IF (1985) A comparison of clinical and Diagnostic Interview Schedule diagnoses: re-examination of lay-interviewed cases in the general population. Arch Gen Psychiatry 42:657–666
11. Hsiao JK, Bartko JJ, Potter WZ (1989) Diagnosing diagnoses. Arch Gen Psychiatry 46:664–668
12. Janca A, Kastrup M, Katschnig H, Lopez-Ibor JJ Jr, MezzIich JE, Sartorius N (1996) The World Health Organisation Short Disability Assessment Schedule (WHO DAS-S): a tool for the assessment of difficulties in selected areas of functioning of patients with mental disorders. Soc Psychiatry Psychiatr Epidemiol 31(6):349–354
13. Leentjens AFG, Verhey FRJ, Lousberg R, Spitsbergen H, Wilmink FW (2000) The validity of the Hamilton and Montgomery-Åsberg depression rating scales as screening for depression and diagnostic tools for depression in Parkinson's disease. Int J Ger Psychiatry 15:644–649
14. Regier DA, Kaelber CT, Rae DS, Farmer ME, Kanauper B, Kessler RC, Norquist GS (1998a) Limitations of diagnostic criteria and assessment instruments for mental disorders. Arch Gen Psychiatry 55:109–115
15. Regier DA (1998b) In reply, Letters to the editor. Arch Gen Psychiatry 55:1146–1147
16. Robins LN, Helzer JE, Croughan J, Ratcliff KS (1981) National Institute of Mental Health Diagnostic Interview Schedule: history, characteristics and validity. Arch Gen Psychiatry 38:381–389

17. Tomov T, Nikolov V (1990) Reliability of SCAN categories and scores: results of the field trials. In: Stefanis CN, Rabavilas AD, Soldatos CR (eds) Vol 1: Classification and Psychopathology, Child Psychiatry, Substance Use: Proceedings of the VIII World Congress of Psychiatry, Athens, 12–19 October 1989. Amsterdam, the Netherlands: Excerpta Medica
18. Wing JK, Babor T, Brugha T, Burke J (1990) SCAN. Schedules for Clinical Assessment in Neuropsychiatry. Arch Gen Psychiatry 47(6):589–593
19. Zung WWK (1965) A self-rating depression scale. Arch Gen Psychiatry 12:63–70