Original Research

Comparison Between the Montgomery-Asberg Depression Rating Scale–Self and the Beck Depression Inventory II in Primary Care

Carl Wikberg, RN; Shabnam Nejati, RN; Maria E. H. Larsson, PhD;
Eva-Lisa Petersson, PhD; Jeanette Westman, PhD; Nashmil Ariai, MS;
Marie Kivi, PSY; Maria Eriksson, PSY; Robert Eggertsen, MD, PhD;
Dominique Hange, MD; Amir Baigi, PhD; and Cecilia Björkelund, MD, PhD

ABSTRACT

Objective: The Montgomery-Asberg Depression Rating Scale–Self (MADRS-S) and the Beck Depression Inventory II (BDI-II) are commonly used self-assessment instruments for screening and diagnosis of depression. The BDI-II has 21 items and the MADRS-S has 9 items. These instruments have been tested with psychiatric inpatients but not in outpatient primary care, where most patients with symptoms of depression initially seek treatment. The purpose of this study was to compare these 2 instruments in the primary care setting.

Method: Data were collected from 2 primary care randomized controlled trials that were performed from 2010 to 2013 in Sweden: the Primary Care Self-Assessment MADRS-S Study and Primary Care Internet-Based Cognitive Behavioral Therapy Study. There were 146 patients (73 patients each from both trials) who had newly diagnosed mild or moderate depression (per *DSM-IV* recommendations) and who had assessment with both the MADRS-S and BDI-II at primary care centers. Comparability and reliability of the instruments were estimated by Pearson product moment correlation and Cronbach α.

Results: A good correlation was observed between the 2 instruments: 0.66 and 0.62 in the 2 study cohorts. The reliability within the 2 study cohorts was good for both MADRS-S (Cronbach α: 0.76 for both cohorts) and BDI-II items (Cronbach α: 0.88 and 0.85).

Conclusions: The 2 instruments showed good comparability and reliability for low, middle, and high total depression scores. The MADRS-S may be used as a rapid, easily administered, and inexpensive tool in primary care and has results comparable to the BDI-II in all domains.

Depression is an increasing health problem worldwide and a common diagnosis in primary health care.[1] The prevalence of depression in Sweden is 5%–10%.[2] Depression is a complex condition that often coexists with physical conditions. Although general practitioners often are skilled at managing depression during consultation, the diagnosis may be missed in many patients.[3]

During consultation, general practitioners sometimes use a self-assessment instrument to verify or exclude the suspected diagnosis of depression. The 2 most commonly used instruments in Scandinavia for self-assessment of depressive symptoms are the Montgomery-Asberg Depression Rating Scale–Self (MADRS-S)[4] and Beck Depression Inventory II (BDI-II).[5,6] The earlier Beck Depression Inventory I (BDI-I) and MADRS-S have good comparability for inpatient psychiatric treatment and patients with severe depression, but literature review showed no studies that have compared BDI-II and MADRS-S.[7] Other studies have shown similar results regarding self-rating scales.[8,9] However, no reports are available from studies performed in the primary care setting, where depression is mostly mild to moderate, and the prevalence of depression is lower in primary care settings than inpatient psychiatry services.[3,8–10] In primary care, patients frequently have milder symptoms, higher mean age, and other somatic symptoms and diagnoses than patients in psychiatric settings.[11]

The properties of self-assessment instruments are dependent on sensitivity, specificity, and positive and negative predictive values of the instrument.[12,13] The positive predictive value is dependent on prevalence of the disease and is important in the use of screening and self-rating instruments, especially in primary care, where prevalence of individual diagnoses often is < 10%–15%.[14] Both the BDI-II and MADRS-S are instruments for self-assessment that frequently are recommended for use in the treatment and monitoring of patients in primary care, but there is a lack of evidence about their applicability for this purpose. Recently, the Swedish Council on Health Technology Assessment published a review about the diagnostic and monitoring properties of the BDI-II and MADRS-S; the BDI was considered to have sufficient sensitivity, but no comparable studies were available about the MADRS-S.[13] The BDI-II was recommended as a targeted screening instrument for Swedish primary care.[13]

In 2 previous studies, the correlation between the BDI-I and the Comprehensive Psychopathological Rating Scale was evaluated at psychiatric clinics.[8,15] However, limited studies are available about how to use self-assessment instruments in primary care. The purpose of the present study was to assess the comparability of the BDI-II and MADRS-S self-assessment instruments in primary care and to evaluate differences between the instruments in the self-assessment of depressed patients in primary care.

METHOD

Study Population

We used data from 2 primary care randomized controlled trials that were performed from 2010 to 2013: the Primary Care Self-Assessment MADRS-S Study (PRI-SMA)[16] and Primary Care Internet-Based Cognitive Behavioral Therapy Study (PRIM-NET).[17] There were 31 different primary health care centers involved (PRIM-NET: 16 centers; PRI-SMA: 15 centers). All included patients had visited a primary care center as a planned or a drop-in visit and had been diagnosed with a new episode of mild or moderate depression during the visit. The diagnosis was made in the PRIM-NET study during a psychologist/psychotherapist consultation according to the Mini-International Neuropsychiatric Interview[18] or in the PRI-SMA study during a general practitioner consultation according to the Primary Care Evaluation of Mental Disorders[19] (PRIME-MD) (Figure 1). Both diagnostic procedures followed the *DSM-IV* recommendations. We analyzed data on patients from the PRI-SMA and PRIM-NET cohorts aged ≥ 18 years who were on antidepressant medication and who had no change in antidepressant medication during the previous 4 weeks. Exclusion criteria were severe depression (patients with severe depression often need referral to specialist care), previous suicide attempt, current suicide risk, known abuse, serious psychiatric disorders, dementia, and cognitive or language difficulties that prevented the patient from completing the forms.

Data (MADRS-S and BDI-II) were collected from the patients from both studies (PRIM-NET and PRI-SMA). In patients from the PRIM-NET study, a psychologist administered the BDI-II and a nurse administered the MADRS-S; in patients from the PRI-SMA study, a general practitioner administered the MADRS-S and a nurse administered the BDI-II (Figure 1). The patients completed both instruments within 1 to 24 hours in the PRI-SMA study and 1 hour to 2 weeks in the PRIM-NET study; most patients completed both instruments within the same day (Figure 1). We included 73 patients from each of the 2 studies (total of 146 patients in the present study).

Instruments

In this study, we compared the BDI-II and MADRS-S.[4,20] The BDI-II is a later version of the original BDI reported by Aaron T. Beck in 1961 and includes 21 multiple-choice items that are each rated on a 4-point scale (severity: 0 = minimum, 3 = maximum).[5,20] Using the self-assessment scale, the patient was asked to assess how he or she had felt during the previous 2 weeks. The total scores for all 21 items were added, and cutoff scores for the BDI-II were defined for level of depression (depression: 0–13 = minimal, 14–19 = mild, 20–28 = moderate, 29–63 = severe).

The MADRS-S is the self-rating version of the original MADRS published in 1979.[4,21] The MADRS-S has 9 items, and each item is scored between 0 (minimum) and 6 (maximum). The patient is asked to assess how he or she has felt during the previous 3 days. The scores for all 9 items were added, and cutoff scores for the MADRS-S were defined for level of depression (depression: 0–12 = minimal, 13–19 = mild, 20–34 = moderate, ≥ 35 = severe). The MADRS-S has been evaluated concerning validity, reliability, and sensitivity to change, showing acceptable psychometric properties.[22]

Ethical Considerations

Both cohort studies were approved by the regional ethical review board of Gothenburg, Sweden. All patients were provided with written and oral patient information and signed a written consent.

Study Power

We used all available data for this study to increase the power of the study. Calculation of power for the comparison between the BDI-II and MADRS-S was based on previous available studies.[7,8] The sample size required for α = .05 and power = 0.80 was 19 cases with assessments on each instrument (BDI-II and MADRS-S). Due to access to a larger sample, we included 73 cases from each study, which was the highest number possible to obtain 2 groups equivalent in size.

Statistical Analysis

Data analysis was performed with statistical software (IBM SPSS Statistics for Windows, Version 21.0, IBM Corp, Armonk, New York). The association between the items from the 2 instruments was evaluated with intraclass correlation and Cronbach α. Cronbach α was used to test reliability of the BDI-II and MADRS-S items. Pearson product moment correlation was used to test associations between BDI-II and MADRS-S total scores. Descriptive statistics were reported as number (%) and were compared with Pearson χ^2 test. The level of significance was set at $P \leq .05$.

RESULTS

Patients from the PRI-SMA study were more frequently older and had higher frequency of university education than patients from the PRIM-NET study (Table 1). In the PRIM-NET group, mean score decreased with education level; in the PRI-SMA group, mean scores increased with education level (Table 2). Analysis of partial correlation of age showed a correlation between education level and BDI-II total score with 0.185 ($P \leq .03$) in the total sample and 0.338 ($P \leq .004$) in the PRI-SMA study sample.

Comparison of mean scores with the BDI-II and MADRS-S in the 2 study samples showed significant sex

Figure 1. Study Flowchart
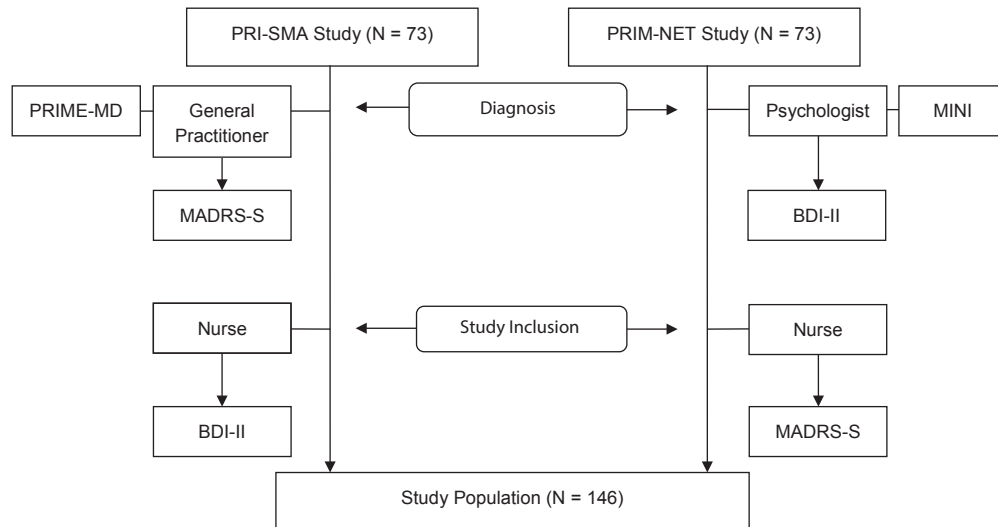
Abbreviations: BDI-II = Beck Depression Inventory II, MADRS-S = Montgomery-Asberg Depression Rating Scale–Self,
MINI = Mini-International Neuropsychiatric Interview, PRIME-MD = Primary Care Evaluation of Mental Disorders,
PRIM-NET = Primary Care Internet-Based Cognitive Behavioral Therapy Study, PRI-SMA = Primary Care Self-
Assessment Montgomery-Asberg Depression Rating Scale–Self Study.

Table 1. Descriptive Characteristics of Patients[a]

Characteristic	Total	PRIM-NET Study	PRI-SMA Study	$P \leq$[b]
No. of patients	146	73	73	
Sex				NS[c]
Male	39 (27)	21 (29)	18 (25)	
Female	107 (73)	52 (71)	55 (75)	
Age (y)				.005
<40	83 (57)	50 (69)	33 (45)	
≥40	63 (43)	23 (31)	40 (55)	
Education level				.02
Grade school (9 y)	13 (9)	3 (4)	10 (14)	
High school	82 (57)	49 (67)	33 (46)	
University	50 (34)	21 (29)	29 (40)	

[a]Data are reported as n (%).
[b]Pearson χ^2 test.
[c]$P > .05$.
Abbreviations: NS = not significant, PRIM-NET = Primary
  Care Internet-Based Cognitive Behavioral Therapy Study,
  PRI-SMA = Primary Care Self-Assessment Montgomery-Asberg
  Depression Rating Scale–Self Study.

Table 2. Scores for 2 Depression Instruments With Patients From 2 Study Samples[a]

Patient Characteristics	Total	PRIM-NET Study	PRI-SMA Study
No. of patients	146	73	73
MADRS-S total score			
All patients	22 ± 7	20 ± 7	24 ± 8
Male	21 ± 8	18 ± 7	24 ± 8
Female	22 ± 7	21 ± 6	24 ± 8
Age <40 y	23 ± 8	20 ± 7	27 ± 7
Age ≥40 y	21 ± 6	20 ± 5	21 ± 7
Grade school (9 y)	22 ± 8	24 ± 7	22 ± 8
High school	21 ± 7	20 ± 7	23 ± 8
University	23 ± 7	19 ± 6	27 ± 6
BDI-II total score			
All patients	27 ± 10	26 ± 9	27 ± 10
Male	23 ± 8[b]	22 ± 7[b]	25 ± 9
Female	28 ± 10[b]	28 ± 9[b]	28 ± 10
Age <40 y	28 ± 10	26 ± 9	31 ± 10
Age ≥40 y	25 ± 9	27 ± 10	24 ± 9
Grade school (9 y)	21 ± 11[c]	26 ± 11	20 ± 11[c]
High school	27 ± 9	27 ± 9	27 ± 9
University	29 ± 10	25 ± 10	31 ± 9

[a]Data are reported as mean ± SD.
[b]Sex, t test: $P \leq .02$ in total sample; $P \leq .02$ in PRIM-NET sample.
[c]Education, partial correlation: $P \leq .02$ in total sample; $P \leq .001$ in
 PRI-SMA sample.
Abbreviations: BDI-II = Beck Depression Inventory II,
  MADRS-S = Montgomery-Asberg Depression Rating Scale–Self,
  PRIM-NET = Primary Care Internet-Based Cognitive Behavioral
  Therapy Study, PRI-SMA = Primary Care Self-Assessment
  Montgomery-Asberg Depression Rating Scale–Self Study.

differences, with females having higher mean BDI-II scores than males (all patients combined and patients in the PRIM-NET group) (Table 2). In addition, patients who had a grade school education level had lower mean BDI-II scores than patients who had high school or university education levels (all patients combined and patients in the PRI-SMA study) (Table 2).

The MADRS-S and BDI-II scores had sufficient agreement concerning all depression levels, with slightly less agreement in the lowest level, below the cutoff for mild depression (minimal depression: MADRS-S = 0–12, BDI-II = 0–13) (Figure 2). Significant correlations were observed between total scores of the BDI-II and MADRS-S (correlation coefficient: PRI-SMA = 0.66, PRIM-NET = 0.62).

The reliability was good in both cohorts. In the PRI-SMA sample, the Cronbach α was 0.76 for MADRS-S items

and 0.88 for BDI-II items. In the PRIM-NET sample, the Cronbach α was 0.76 for MADRS-S items and 0.85 for BDI-II items.

DISCUSSION

This study showed good comparability between 2 commonly used self-assessment instruments (MADRS-S and BDI-II) for patients who attended primary care visits

Figure 2. Correlation Between 2 Depression Rating Scales in Primary Care: The MADRS-S and BDI-II Scores Were Compared With MADRS-S (I, III) and BDI-II (II, IV) Cutoff Scores as Reference for the 73 Participants in the PRI-SMA Study and the 73 Participants in the PRIM-NET Study; the Table Below the Graphs Shows Mean Score Within Each Classification (MADRS-S and BDI-II)
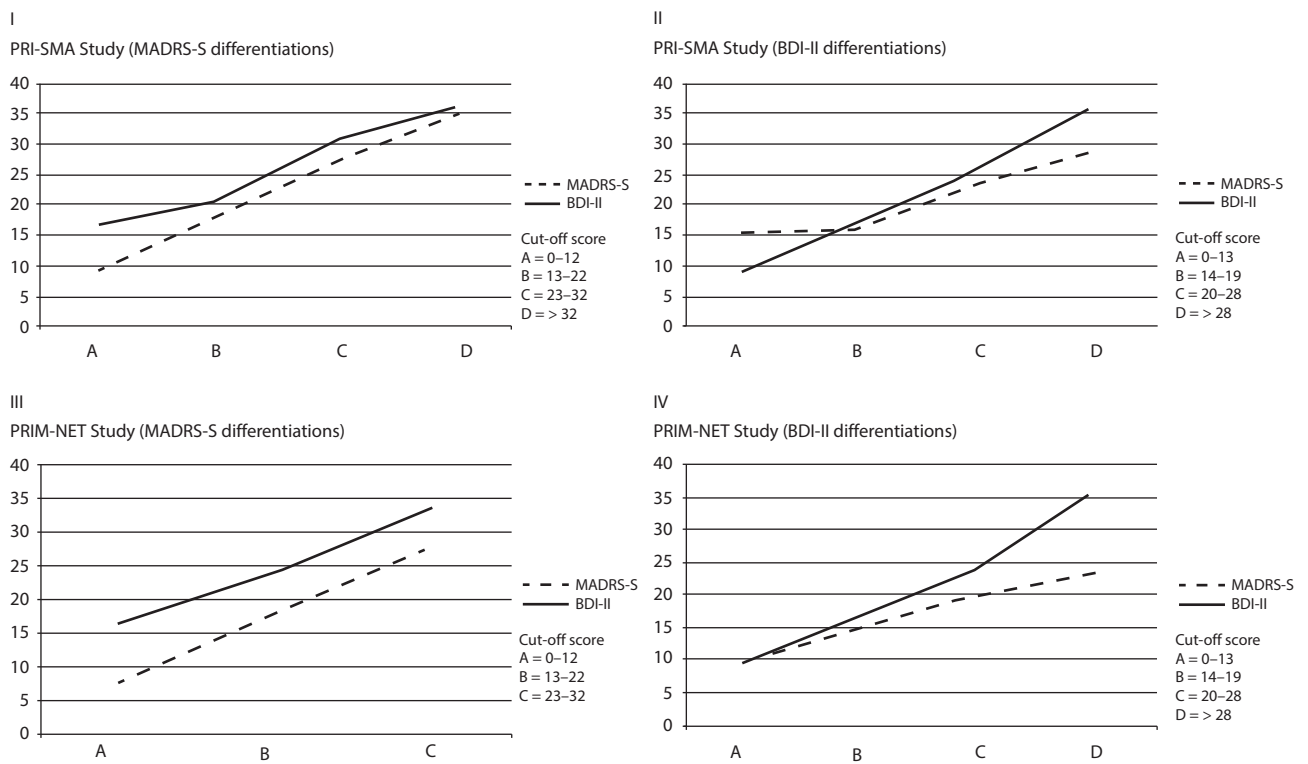
I	PRI-SMA (MADRS-S differentiations)					II	PRI-SMA (BDI-II differentiations)			
MADRS-S	9	18.05	27.17	34.78		MADRS-S	15.29	16.1	23.09	28.58
BDI-II	17	20.95	30.72	36.11		BDI-II	9.14	16.9	25.3	35.73
	A	B	C	D			A	B	C	D
III	PRIM-NET (MADRS-S differentiations)					IV	PRIM-NET (BDI-II differentiations)			
MADRS-S	7.57	18.13	27.95			MADRS-S	9.6	15.21	20.19	23.57
BDI-II	16.71	24.15	33.65			BDI-II	9.6	16.71	24.19	35.36
	A	B	C				A	B	C	D

Abbreviations: BDI-II = Beck Depression Inventory II, MADRS-S = Montgomery-Asberg Depression Rating Scale–Self, PRIM-NET = Primary Care Internet-Based Cognitive Behavioral Therapy Study, PRI-SMA = Primary Care Self-Assessment Montgomery-Asberg Depression Rating Scale–Self Study.

and had symptoms of depression. This result suggests that it would be equally reliable to use the MADRS-S assessment scale instead of the BDI-II in consultations with patients who have mild or moderate depression. The MADRS-S requires less time than the BDI-II, and the MADRS-S is easily administered and free of charge. Although a previous report suggested that the BDI-II was the self-assessment instrument with best evidence for primary care use,[13] a literature search showed no previous studies that had evaluated the use of the MADRS-S in primary care. In addition, the MADRS-S previously was judged not to have enough evidence for diagnostic and screening use.[13] However, the present study suggests that a physician may use (1) the longer self-assessment BDI-II instrument for diagnostic assessment and the more rapidly administered MADRS-S in clinical follow-up consultations or (2) the MADRS-S for both diagnostic assessment and follow-up. The present study showed good comparability between the 2 instruments. There is satisfactory correspondence between the instruments, regardless of which self-assessment instrument (MADRS-S and BDI-II) the physician administers to the patient during the consultation. The MADRS-S is particularly useful as a rapid assessment tool to rate the current state of depression.

The result of a self-assessment during a consultation should not be the basis for diagnosis, treatment, or decisions about sick leave. A diagnosis of depression should be based on standardized criteria (*DSM* or *ICD*) for major depression (≥ 5 symptoms present during a 2-week period, representing a change from previous functioning; ≥ 1 symptom should be either depressed mood or loss of interest or pleasure).[23]

However, the use of assessment instruments in the diagnostic process may facilitate professional judgment. It is important to customize the results about the patient when using the instrument, and professionals using the instruments should know how to interpret the results.[24] Repeated testing or self-assessment may be necessary in health care settings because a single assessment seldom is sufficient. A physician or nurse should gather more information about the patient to assemble a good overview about treatment needed. The results should be related to the patient's situation and how the situation has changed over time.

There are various ways to detect and diagnose depression in patients in primary care. Opportunistic screening is inefficient because it produces an excessive number of false-positive outcomes.[25] Directional screening has sufficient positive predictive value.[12] A patient-centered consultation with open questions enables general practitioners to exclude depression with high precision.[12]

In the present study, more women than men were diagnosed as depressed, which is consistent with the results of earlier studies (Table 1).[26] The descriptive data indicated differences in age and education between the 2 study samples. The age differences between the study groups may be attributed to the lower mean age of the PRIM-NET patients because inclusion in that study presupposed access and ability to operate a computer; this also may have caused study personnel to approach younger patients preferentially for inclusion. Differences in educational level between the groups may be attributed to the younger age of the PRIM-NET population, and younger patients may not have commenced postgraduate studies upon entry to the trial. Nevertheless, differences in education between the groups did not affect the results obtained in comparing the 2 self-assessment instruments. Women scored significantly higher than men in the BDI-II, and further studies are required to explore this finding.

The present data were collected from 2 well-functioning studies that had similar inclusion and exclusion criteria. The patients completed the MADRS-S and BDI-II instruments in almost the same way in both studies. This data collection process may have avoided potential bias that might have occurred because of the order in which the assessments were administered (Figure 1). Despite the fact that the order in which the self-assessments were administered differed in the 2 samples (Figure 1), similar results were obtained. Only patients attending primary care were included in the study population. The study was conducted in patients with mild to moderate depression, in contrast with previous studies conducted in psychiatric inpatient units. The patients included in the present study assessed their own symptoms when diagnosed with a new depressive episode, and this illustrated how well the instruments conformed to patients in the early stages of an untreated depression in primary care.

Not all patients were untreated at the time of inclusion, because some patients were undergoing maintenance therapy; these patients did not change medication during the previous 4 weeks but entered into a new period of depression.

This factor did not affect the results because the degree of depression was not measured in evaluating the correlation between the 2 assessment scales.

The study was performed in 1 Swedish region. Although geographic differences may exist that may affect the results and level of depression, a previous study suggests that this may be unlikely.[26] Further studies are needed, especially with patients who have difficulty completing and understanding the forms because of cognitive problems, dementia, or language difficulties. In the present study, validated translations of the instruments were unavailable.

In conclusion, the present results showed that the MADRS-S and BDI-II have good comparability and reliability in the primary care setting, regardless of level of depression (mild or moderate). General practitioners may obtain comparable results with the MADRS-S or BDI-II.

Author affiliations: Department of Primary Health Care, Institute of Medicine (Drs Petersson, Eggertsen, Hange, Baigi, and Björkelund; Mr Wikberg; and Mss Nejati, Ariai, and Eriksson), Department of Clinical Neuroscience and Rehabilitation, Institute of Neuroscience and Physiology (Dr Larsson), and Department of Psychology (Ms Kivi), University of Gothenburg, Gothenburg, Sweden; Närhälsan Research and Development Primary Health Care, Region Västra Götaland, Gothenburg, Sweden (Drs Larsson and Petersson); and Department of Neurobiology, Care Sciences and Society, Center for Family Medicine, Karolinska Institutet, Stockholm, Sweden (Dr Westman).

REFERENCES

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3(11):e442.
2. Katon W, Schulberg H. Epidemiology of depression in primary care. *Gen Hosp Psychiatry*. 1992;14(4):237–247.
3. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet*. 2009;374(9690):609–619.
4. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134(4):382–389.
5. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clin Psychol Rev*. 1988;8(1):77–100.
6. Beck AT, Steer RA, Brown GK. *Manual for the Beck Depression Inventory–II.* San Antonio, TX: Psychological Corporation; 1996.
7. Svanborg P, Asberg M. A comparison between the Beck Depression Inventory (BDI) and the self-rating version of the Montgomery-Asberg Depression Rating Scale (MADRS). *J Affect Disord*. 2001;64(2–3):203–216.
8. Martinsen EW, Friis S, Hoffart A. Assessment of depression: comparison between Beck Depression Inventory and subscales of Comprehensive Psychopathological Rating Scale. *Acta Psychiatr Scand*. 1995;92(6):460–463.
9. Schubert DS, Taylor C, Lee S, et al. Detection of depression in the stroke patient. *Psychosomatics*. 1992;33(3):290–294.
10. World Health Organization. Mental health policies in Europe. *Euro Observer*. 2009;11(3):1–4.
11. Magnil M, Gunnarsson R, Björkstedt K, et al. Prevalence of depressive symptoms and associated factors in elderly primary care patients: a descriptive study. *Prim Care Companion J Clin Psychiatry*. 2008;10(6):462–468.
12. Magnil M, Gunnarsson R, Björkelund C. Using patient-centered consultation when screening for depression in elderly patients: a comparative pilot study. *Scand J Prim Health Care*. 2011;29(1):51–56.

13. Swedish Council for Health Technology Assessment. Diagnosis and follow-up of affective disorders: a systematic review. Report 212/2012. ISBN 978-91-85413-52-2. ISSN 1400–1403 [in Swedish].

14. Månsson J, Nilsson G, Strender LE, et al. Reasons for encounters, investigations, referrals, diagnoses and treatments in general practice in Sweden: a multicenter pilot study using electronic patient records. *Eur J Gen Pract*. 2011;17(2):87–94.

15. Tamaklo W, Shubert DS, Mentari A, et al. Assessing depression in the medical patient using the MADRS, a sensitive screening scale. *J Integr Psychiatry*. 1992;8:264–270.

16. Björkelund C. Structured use of a self rating instrument in the primary care consultation with patients with mild and moderate depression. ClinicalTrials.gov Identifier: NCT01402206.

17. Kivi M, Eriksson MC, Hange D, et al. Internet-based therapy for mild to moderate depression in Swedish primary care: short-term results from the PRIM-NET randomized controlled trial. *Cogn Behav Ther*. 2014;43(4):289–298.

18. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for *DSM-IV* and *ICD-10*. *J Clin Psychiatry*. 1998;59(suppl 20):22–33, quiz 34–57.

19. Spitzer RL, Kroenke K, Williams JB; and the Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of PRIME-MD: the PHQ Primary Care Study. *JAMA*. 1999;282(18):1737–1744.

20. Svanborg P, Asberg M. A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. *Acta Psychiatr Scand*. 1994;89(1):21–28.

21. Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961;4(6):561–571.

22. Fantino B, Moore N. The self-reported Montgomery-Asberg Depression Rating Scale is a useful evaluative tool in major depressive disorder. *BMC Psychiatry*. 2009;9(1):26.

23. American Psychiatric Association. *Diagnostic and Statistical Manual for Mental Disorders,* Fifth Edition. Washington, DC: American Psychiatric Association; 2000.

24. Pettersson A, Björkelund C, Petersson EL. To score or not to score: a qualitative study on GPs views on the use of instruments for depression. *Fam Pract*. 2014;31(2):215–221.

25. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. *CMAJ*. 2008;178(8):997–1003.

26. Mattisson C, Bogren M, Nettelbladt P, et al. First incidence depression in the Lundby Study: a comparison of the two time periods 1947–1972 and 1972–1997. *J Affect Disord*. 2005;87(2–3):151–160.

27. Perris C, Eisemann M, Eriksson U, et al. Transcultural aspects of depressive symptomatology. *Psychiatr Clin (Basel)*. 1981;14(2):69–80.