

RESEARCH ARTICLE

Open Access



Patient health questionnaire-9 versus Edinburgh postnatal depression scale in screening for major depressive episodes: a cross-sectional population-based study

Iná S. Santos^{1*}, Beatriz Franck Tavares², Tiago N. Munhoz¹, Patricia Manzolli², Gisele Bartz de Ávila², Eduardo Jannke² and Alicia Matijasevich^{1,3}

Abstract

Background: Major depressive episodes (MDE) are frequent at the population level and are generally associated with severe symptoms that impair performance of activities of daily living of individuals suffering from this condition. The aim of this study was to compare the accuracy of two tests that separately showed suitable properties in screening for MDE: the Patient Health Questionnaire-9 (PHQ-9) and the Edinburgh postnatal depression scale (EPDS).

Methods: In a previous study, the sensitivity and specificity of the PHQ-9 and the EPDS in screening for MDE were compared with a structured diagnostic interview conducted by psychiatrists and psychologists using the Mini International Neuropsychiatric Interview as the gold standard. In a sample of adults living in the community in Pelotas, Brazil, the PHQ-9 and EPDS were applied at the same interview and the gold standard on a median of 17 days later. The interviews were carried out at the participant's home.

Results: 447 individuals (191 men and 256 women) were assessed. The PHQ-9 and the EPDS results were concordant in 87.5 % of the respondents, with a moderate agreement beyond what was expected by chance alone ($\kappa = 0.61$). The areas below the ROC curves were not statistically different (82.1 % for PHQ-9 and 83.5 % for EPDS) ($p = 0.291$), thus indicating that the two tests had similar moderate accuracy.

Conclusions: PHQ-9 and EPDS may be applied with equal confidence in screening for MDE in the community.

Keywords: Patient health questionnaire-9, Edinburgh postnatal depression scale, Major depressive episode, Screening, Accuracy

Background

Major depressive episodes (MDE) are frequent at the population level and are generally associated with severe symptoms that impair performance of activities of daily living of individuals suffering from this condition [1, 2]. Worldwide, population-based surveys that included more than 37,000 individuals living in ten countries on

four continents have recorded lifetime prevalence of MDE ranging from 8 to 12 % [3].

The age of depressive symptoms onset is in the middle 20s, with the peak risk period for onset ranging from mid-late adolescence to early 40s [1, 3]. The approximately twofold increase in risk of depression among the women in comparison to the men is consistent over cultures and most age groups [4]. Conjugal situation (more frequent among individuals who are unmarried and live without a partner) and family genetic factors (parental depression increases the risk of the offspring also developing depressive episodes) are recognized risk factors for MDE [1, 3]. Epidemiological studies using a range of

*Correspondence: inasantos@uol.com.br

¹ Post-graduate Program in Epidemiology, Federal University of Pelotas, Rua Marechal Deodoro, 1160, 3o andar, Pelotas 96020220, Brazil
Full list of author information is available at the end of the article

different methods have documented higher overall rates of depression over time, with increasing rates in the young, associated with a shift forward to younger ages [1]. Although older individuals may be less likely to recognize depression as a mental disorder, and hence are less likely to remember depressive episodes as such, or to report these episodes in interviews on mental health [5], the evidence that rates of depression in the elderly are higher compared with those observed in past studies suggest the presence of a cohort effect.

A number of adverse consequences of major depression have been described. School failure, low probability of ever marrying or higher probability of early marital timing and divorce, teen childbearing, negative parenting behaviors, work absenteeism, lower income-earnings, comorbidity and elevated risk of early death are all associated with major depression [1].

Recognition of depression as a public health problem has led to creation of a variety of screening instruments for use in research and in primary healthcare services, with the aim of identifying individuals at risk of MDE, at an earlier stage [6]. Although two meta-analyses [7, 8] and a quasi-experimental study [9] found no evidence of effectiveness of screening at primary healthcare services, the availability of reliable and valid information is essential for estimating and monitoring depression prevalence and time trends in depression prevalence by means of epidemiological research at the population level.

Because the properties of the screening tests vary as a function of the socio-demographic and cultural characteristics of the population to which the tests are applied, it is recommended that their use should be preceded by studies that evaluate these properties within the context in which they will be used [10]. The sensitivity and specificity of two depression screening instruments, the Patient Health Questionnaire-9 (PHQ-9) [11] and the Edinburgh postnatal depression scale (EPDS) [12], were evaluated on a single sample of adults living in the city of Pelotas, RS, Brazil, and the results have already been published [13, 14]. The objective of the present study was to compare the accuracy of the two tests, which separately showed good sensitivity and specificity in screening for MDE among adults living in the community.

Methods

A cross-sectional population-based study was conducted in the urban zone of the municipality of Pelotas between February and June 2012 to evaluate the health of adolescents, adults and elderly people. A sampling design of two-stage conglomerates with probability proportional to size was used. According to the 2010 Population Census there were 495 census tracts, the primary sampling units. The secondary sampling units were households. All

private households with permanent resident as of December 2011 in the 130 census tracts randomly selected were listed. In each census tract drawn, around 12 households were randomly selected for the survey. All the people living in the households drawn who were 10 years of age or over were eligible. The participants were interviewed at home, by trained interviewers, through applying a structured questionnaire that included questions about their economic condition, schooling, marital status, skin color, occupation, health, and behavior. The adults (≥ 20 years of age) answered the PHQ-9 and EPDS questionnaires, and these were applied by general interviewers. Individuals who had cognitive or mental disabilities confirmed by the fieldwork supervisor, as well as those institutionalized (hospitals, elderly homes, among others), were excluded.

Validation studies on PHQ-9 [13] and EPDS [14] were conducted on a subsample of adults (≥ 20 years of age). The sampling process for the validation studies was conducted weekly, starting from the interviews that were conducted for the main study. Through simple random draws, one-third of the households included in the main study were selected for the validation studies. The person in charge of the draw was unaware of the results from the PHQ-9 or EPDS tests that were applied in the main study. In each household thus selected, all the people living there who were 20 years of age or over, independently of the PHQ-9 or EPDS scores, were invited to receive a second visit for a supplementary interview. This second interview was conducted by a mental health professional (psychiatrist, psychologist or medical resident in psychiatry), who had previously been trained to apply and interpret the gold-standard instrument and was blind to the scores achieved by the participant in the PHQ-9 and EPDS questionnaires. The participants were unaware of the professional training of these interviewers, so that this would not influence the responses.

The PHQ-9 consists of nine questions that assess the presence of each of the symptoms of MDE, as described in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) [15] (depressed mood; loss of interest or pleasure in doing things; problems relating to sleep, tiredness or lack of energy; changes in appetite or weight; feelings of guilt or uselessness; problems of concentration; feelings of being slow or restless; and suicidal thoughts). The frequency of each symptom over the preceding 2 weeks was evaluated on a Likert scale from 0 to 3.

The EPDS was originally constructed to identify postpartum depression, but it can be applied to screen for depression in the community, including among men [16]. The EPDS consists of a scale of 10 items, each with four possible responses from 0 to 3, which express the intensity of depressive symptoms over the 7 days preceding the interview.

The questionnaire for the first interview (main study) was set up in sections. The PHQ-9 was applied after the participants had answered the questions in the sections relating to socio-demographic factors, behavioral factors, chronic diseases and use of medications. Following the PHQ-9 application, there were questions on the subjects' use of and access to healthcare services and their dietary habits, and then the EPDS was applied. Further details on the methodology of the validation studies for the PHQ-9 and EPDS, along with the Portuguese-language versions used, can be obtained in other published papers [13, 14].

To calculate the sample size, the following parameters were used: sensitivity and specificity of 80 %, acceptable error of 10 % points upwards or downwards and significance level of 95 %. Thus it was necessary to include around 200 subjects with MDE and 200 without MDE. Given that the point-prevalence of depressive symptoms among the adult population of Pelotas had been found to be around 30 % [17], it was estimated that with a sample of around 600 individuals, it would be possible to locate around 200 with MDE.

The gold standard used was the Mini International Neuropsychiatric Interview (MINI) [18], which has been validated for use in Brazil [19]. This structured diagnostic questionnaire assesses the presence of mental disorders, in accordance with DSM-IV and ICD-10. For depressive disorders, it has sensitivity and specificity of 92 % [19]. In the present study, the gold standard was used to diagnose the presence of MDE. All individuals who were considered to be positive for MDE gave responses to an additional group of questions that investigated other possible causes for the symptoms, such as direct effects of substances, organ disorders, medical illness or presence of psychotic symptoms, or whether the symptoms would be better explained as reactions to grief, for which the diagnosis of MDE would be rejected.

The data analysis included calculation of the sensitivity and specificity for each score on a continuous scale for each of the tests. For each PHQ-9 and EPDS cutoff point, the sensitivity (proportion of individuals with MDE according to MINI criteria that were correctly identified by the test), specificity (proportion of individuals without MDE according to the gold standard correctly identified as such by the test), positive predictive value (proportion of true positives among all positives identified by the test), accuracy (proportion of true positives and true negatives identified by the test), and positive likelihood ratio (the odds that the given cutoff point would be expected in an individual with in opposed to one without MDE according to the gold standard) with 95 % confidence intervals were calculated.

To compare the accuracy of the tests for identifying individuals at risk of MDE, the sensitivity and

1-specificity values of each of the cutoff points for the PHQ-9 and EPDS were plotted on a single receiver operating characteristic (ROC) curve. The cutoff point with greatest sensitivity and specificity on the ROC curve was defined as the lowest value for the equation $\{(1 - \text{sensitivity})^2 + (1 - \text{specificity})^2\}$. The accuracies of the PHQ-9 and EPDS were compared by means of the areas under the respective ROC curves. The concordance between the two tests, i.e. beyond what would be expected by chance, was calculated by means of the kappa statistic.

The main study and the validation studies were approved by the Research Ethics Committee of the School of Medicine, Federal University of Pelotas, in accordance with protocols 77/2011 and 14/2012, respectively. A free and informed consent statement was signed by each participant in the main study before information was gathered. The individuals who were diagnosed as positive through the gold-standard assessment were attended at home and/or were referred to the healthcare services.

Results

A total of 533 individuals were identified as candidates for the gold-standard interview. Of these, 447 (84 %) were assessed: there were 29 refusals; 51 were deemed to be losses after three attempts to find them; and six were found to have moved away from this municipality. The interviews with mental health professionals (gold standard) were held on average 24 days after application of the two tests (median of 17 days).

The individuals evaluated comprised 191 men and 256 women. The majority (83.2 %) were under the age of 60 years and 76.5 % self-declared as having white skin color. With regard to socioeconomic variables, more than one-third (39.7 %) were living in families with mean monthly incomes ≤ 3 minimum wages and 15.3 % had only attended school until the 4th year of elementary school. More than half of the participants were doing paid work at the time when the PHQ-9 and EPDS were applied (58.8 %) and were living with partners (65.5 %).

The individuals who were lost from the gold-standard interview were similar to those who were interviewed by the general interviewers, with regard to all the characteristics investigated, except in relation to paid work. The frequency of unemployment among individuals evaluated by means of the gold standard was greater than when the PHQ-9 and EPDS were applied (respectively, 60.4 and 41.2 %).

The gold-standard interview identified 40 individuals (32 women and eight men) presenting MDE (8.9 %; 6.3–11.6 %). For the PHQ-9, values ≥ 9 were more accurate for identifying individuals at greater risk of presenting MDE (Fig. 1). At this point, the sensitivity was 77.5 %

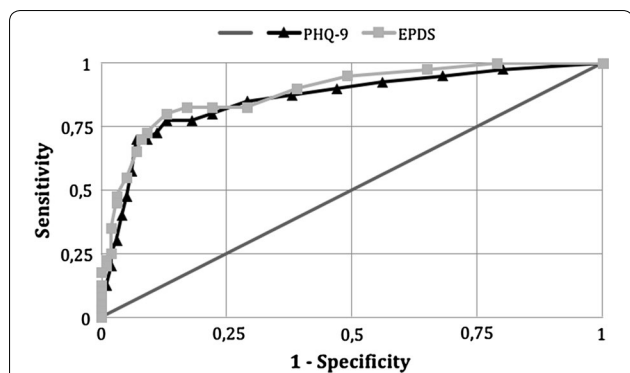


Fig. 1 Receiver operating characteristic (ROC) curves of the Patient Health Questionnaire (PHQ-9) and Edinburgh Postnatal Depression Scale (EPDS) for screening for major depressive episodes among adults living in the community. Areas under the ROC curve: PHQ-9 = 0.821; EPDS = 0.835

(61.5–89.2 %), specificity 86.7 % (83.0–89.9 %), positive predictive value 36.5 % (26.3–47.6 %), and positive likelihood ratio 5.8 (4.3–7.9) (Table 1). A total of 85 individuals (19.0 %) scored ≥ 9 . The area under the ROC curve indicated a general test accuracy of 82.1 %.

For the EPDS, the best cutoff score was ≥ 8 (Fig. 1), and this value was reached by 85 individuals (19.0 %). The crude concordance between the two tests was 87.5 %, with a moderate kappa value (0.61). The sensitivity of EPDS ≥ 8 was 80.0 % (64.4–90.9 %), specificity 87.0 % (83.3–90.1 %), positive predictive value 37.6 % (27.4–48.8 %), and positive likelihood ratio 6.1 (4.6–8.3) (Table 1). The area under the ROC curve showed a general accuracy of 83.5 %, i.e. similar to that of the PHQ-9 ($p = 0.291$).

Discussion

The areas under the curves, and also their formats, indicated that the PHQ-9 and EPDS presented similar and moderate accuracy with regard to identifying adults living in the community who were at greater risk of presenting MDE. Such a finding is in line with the results of a

Table 1 Properties and 95 % confidence intervals of the Patient Health Questionnaire-9 (PHQ-9) and Edinburgh postnatal depression scale (EPDS), at the cutoff points of maximum sensitivity and specificity for screening for MDE among adults in the community

Instrument	Sensitivity	Specificity	PPV	PLR
PHQ-9 ≥ 9	77.5 % (61.5–89.2)	86.7 % (83.0–89.9)	36.5 % (26.3–47.6)	5.8 (4.3–7.9)
EPDS ≥ 8	80.0 % (64.4–90.9)	87.0 % (83.3–90.1)	37.6 % (27.4–48.8)	6.1 (4.6–8.3)

MDE major depressive episode, PPV positive predictive value, PLR positive likelihood ratio

systematic review planned to examine the accuracy of depression screening instruments (including PHQ-9 and EPDS), which concluded that no single instrument was superior to another [20].

In the current study, the two tests were concordant in 391 (87.5 %) of the 447 respondents. Good concordance between the tests was seen even with symptom recall times for the two scales differing by 1 week. Similar level of agreement (83 %) was reported by Yawn et al. in a study specifically planned to compare the PHQ-9 and EPDS as screening tools for postpartum depression [21]. The natural history of the depression may have contributed towards the comparability of the two tests, given that once manifested, the depressive symptoms tend to persist for weeks (with a median duration of 3 months) and, in 20 % of the cases, they remain chronic for 2 years or more [22].

Out of the 85 individuals who were screened positive through the PHQ-9, 57 (67.1 %) were also positive according to the EPDS. Among the 56 individuals for whom the two tests had discordant results, the gold standard indicated that in half of the cases, the PHQ-9 result was correct and in the other half, the EPDS result. Further work is required to identify reasons for disagreement.

The positive likelihood ratios, both for the PHQ-9 and for the EPDS, were around six, thus indicating that results from these tests that are ≥ 9 and ≥ 8 , respectively, are six times more likely to occur among individuals with MDE, in consultations with mental health professionals, than among individuals without MDE. The positive predictive values for the two tests were also very similar (36.5 %; 26.3–47.6 % for PHQ-9 at the cutoff ≥ 9 ; and 37.6 %; 27.4–48.8 % for EPDS ≥ 8). Thus, if these two tests are applied for population screening, they will be equally efficient: two in every five individuals with positive screening through either of the tests will present MDE.

Among the limitations of this study, 16 % of individuals could not undergo the gold standard interview and, though, were not included in the validation sample. They were similar to those included in the sample regarding all socioeconomic, demographic and behavioural characteristics investigated, except to be employed. The prevalence of PHQ-9 ≥ 9 and EPDS ≥ 8 among people that failed to be included in the validation sample was similar to those included in the sample (22.1 versus 19.0 %, $p = 0.510$ for PHQ-9; and 24.4 % versus 19.0 %, $p = 0.250$, for EPDS, respectively). It looks like the loss of these individuals may not have impaired the sensitivity estimation in the present study. Additionally, in regard to ethical aspects, neither the lost individuals nor those that refused to attend the gold standard interview had replied positively to the questions on risk of suicide of both the PHQ-9 and EPDS questionnaires.

Both the PHQ-9 and the EPDS were applied in the middle of the interview. There were questions on the subjects' use of and access to healthcare services, as well as on their dietary habits between the PHQ-9 and the EPDS application. However, the fact that both tests were applied to the same sample, in the same interview and in sequence may have jogged the interviewees' memories with regard to answering the questions of the EPDS, in comparison with the PHQ-9, given that the PHQ-9 was applied first. It is possible that this may have introduced some information bias.

Another limitation was the gap of about 17 days (median) between the PHQ-9 and EPDS application and the gold standard administration. Because the validation studies were nested within a large epidemiological study with a complex logistics, a delay on the execution of some of the implementation steps was difficult to prevent. It is possible that depressive symptoms may have changed over this period due to two main reasons. First, the PHQ-9 and EPDS were designed to enquire about feelings over the last fifteen and the last 7 days, respectively. Second, the duration of the MDE may vary with age and with the natural history of the disorder [4]. Recurrent depression typically has shorter episode duration. The young have more frequent episodes of shorter duration whereas the elderly has long episodes and chronic depression. However, the reported median duration of MDE in the community is 3 months [23], what may have minimized at least in part the negative effect of the time lag over the observed sensitivity of the PHQ-9 and EPDS.

Despite the standardization procedures undertaken before the study initiation, another flaw is the lack of assessment of the inter-rater reliability of the gold-standard evaluators.

Finally, there are concerns that EPDS is not suitable for men because it detects distress but not necessarily depression [24], and that EPDS has a different factor structure in men [25–28]. The EPDS, as much of the mental screening instruments assess for common mental disorders including anxiety, depression and psychological distress [20]. In the current study the gold standard interview was planned to identify true cases of MDE so that all cases that scored positive at the screening due to anxiety or psychological distress were classified as false positive results. According to the gold-standard, only eight of the 191 men included in the study presented a MDE. Such a small prevalence prevented us from conducting separate analyses according to the sex of the participants.

Conclusion

In localities with socioeconomic, demographic and morbidity profiles similar to those of the city of Pelotas, both the PHQ-9 and the EPDS can be used confidently for

screening for MDE in the community. Both of these tests have the advantage of containing few questions (nine and ten, respectively) and only taking around 5 min for application among adults [13, 14].

Abbreviations

MDE: major depressive episode; PHQ-9: patient Health Questionnaire; EPDS: Edinburgh postnatal depression scale; PPV: positive predictive value; PLR: positive likelihood ratio; ROC: receiver operating characteristic.

Authors' contributions

ISS participated in the conception, data interpretation and writing of the article. BFT participated in the conception, application of the gold-standard interviews and critical review of the article. TNM participated in the conception and coordination of the fieldwork of both phase of the study, and in the critical review of the article. AM participated in the conception, data analysis and interpretation and critical review of the article. PM, GBA and EJ participated in the data gathering and critical review of the article. All authors read and approved the final manuscript.

Author details

¹ Post-graduate Program in Epidemiology, Federal University of Pelotas, Rua Marechal Deodoro, 1160, 3o andar, Pelotas 96020220, Brazil. ² Department of Mental Health, School of Medicine, Federal University of Pelotas, Pelotas, Brazil. ³ Department of Preventive Medicine, School of Medicine, University of São Paulo, São Paulo, Brazil.

Acknowledgements

The first part of this study (main study) was conducted using funding that came from PRONEX-CNPq (Support Program for Centers of Excellence), which was applied to the research consortium of master's program students 2011–2012, of the Post-graduate Program in Epidemiology of the Federal University of Pelotas. IS Santos and A. Matijasevich are CNPq scientific productivity bursary-holders.

Competing interests

The authors declare that they have no competing interests.

Received: 13 August 2015 Accepted: 9 February 2016

Published online: 27 September 2016

References

- Kessler RC, Bromet EJ. The epidemiology of depression across cultures. *Annu Rev Public Health*. 2013;34:119–38.
- Ferrari AJ, Charlson FJ, Norman RE, Flaxman AD, Patten SB, Vos T, Whiteford HA. The epidemiological modelling of major depressive disorder: application for the global burden of disease study 2010. *PLoS ONE*. 2013;8(7):e69637.
- Andrade L, Caraveo-Anduaga JJ, Berglund P, Bijl RV, De Graaf R, Vollebergh W, Dragomirecka E, Kohn RKM, Kessler RC, Kawakami N, et al. The epidemiology of major depressive episodes: results from the international consortium of psychiatric epidemiology (ICPE) surveys. *Int J Methods Psychiatr Res*. 2003;12(1):3–21.
- Wittchen HU, Uhmann S. The timing of depression: an epidemiological perspective. *Medicographia*. 2010;32(2):115–24.
- Hasin D, Link B. Age and recognition of depression: implications for a cohort effect in major depression. *Psychol Med*. 1988;18(3):683–8.
- Williams JW, Noel PH, Cordes JA, et al. Is this patient clinically depressed. *JAMA*. 2002;287:1160–70.
- Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. *Can Med Assoc J*. 2008;178(8):997–1003.
- Gilbody S, House AO, Sheldon TA. Screening and case finding instruments for depression. *Cochrane Database Syst Rev*. 2005;4:CD002792.
- Christensen KS, Sokolowski I, Olesen F. Case-finding and risk-group screening for depression in primary care. *Scand J Prim Health Care*. 2011;29(2):80–4.

10. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology*. 2nd ed. Boston: Little Brown and Company; 1991.
11. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606–13.
12. Cox JL, Holden JM. Detection of postnatal depression. Development of the 10-item Edinburgh postnatal depression scale. *Br J Psychiatry*. 1987;150:782–6.
13. Santos IS, Tavares BF, Munhoz TN, Almeida LS, Silva NT, Tams BD, Patella AM, Matijasevich A. Sensitivity and specificity of the patient health questionnaire-9 (PHQ-9) among adults from the general population. *Cadernos Saúde Pública*. 2013;29:1533–43.
14. Matijasevich A, Munhoz TN, Tavares BF, Barbosa AP, da Silva DM, Abitante MS, Dall'Agnol TA, Santos IS. Validation of the Edinburgh postnatal depression scale (EPDS) for screening of major depressive episode among adults from the general population. *BMC Psychiatry*. 2014;14(1):1.
15. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th ed. Washington: American Psychiatric Association; 2000.
16. Cox JL, Chapman G, Murray D. Validation of the Edinburgh postnatal depression scale (EPDS) in non-postnatal women. *J Affect Disord*. 1996;39(3):185–9.
17. Rombaldi AJ, Silva MC, Gazalle FK, Azevedo MR. Prevalência e fatores associados a sintomas depressivos em adultos do sul do Brasil: estudo transversal de base populacional. *Rev Bras Epidemiol*. 2010;13:620–9.
18. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The mini-international neuropsychiatric interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatr*. 1998;59:22–33.
19. de Azevedo Marques JM. Validity and applicability of the Mini International Neuropsychiatric Interview administered by family medicine residents in primary health care in Brazil. *Gen Hosp Psychiatr*. 2008;30(4):303–10.
20. Akena D, Joska J, Obuku EA, Amos T, Musisi S, Stein DJ. Comparing the accuracy of brief versus long depression screening instruments which have been validated in low and middle income countries: a systematic review. *BMC Psychiatry*. 2012;12:187.
21. Yawn BP, Pace W, Wollan PC, Bertram S, Kurland M, Graham D, Dietrich A. Concordance of Edinburgh postnatal depression scale (EPDS) and patient health questionnaire (PHQ-9) to assess increased risk of depression among postpartum women. *J Am Board Fam Med*. 2009;22(5):483–91.
22. Spijker J, de Graaf R, Bijl RV, Beekman AT, Ormel J, Nolen WA. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Br J Psychiatr*. 2002;181:208–13.
23. Spijker J, de Graaf R, Bijl RV, Beekman AT, Ormel J, Nolen WA. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Br J Psychiatr*. 2002;181:208–13.
24. Loscalzo Y, Giannini M, Contena B, Gori A, Benvenuti P. The Edinburgh postnatal depression scale for fathers: a contribution to the validation for an Italian sample. *Gen Hosp Psychiatr*. 2015;37(3):251–6.
25. Massoudi P, Hwang CP, Wickberg B. How well does the Edinburgh Postnatal Depression Scale identify depression and anxiety in fathers? A validation study in a population based Swedish sample. *J Affect Disord*. 2013;149(1–3):67–74.
26. Matthey S, Fisher J, Rowe H. Using the Edinburgh postnatal depression scale to screen for anxiety disorders: conceptual and methodological considerations. *J Affect Disord*. 2013;146(2):224–30.
27. Matthey S, Souter K, Mortimer K, Stephens C, Sheridan-Magro A. Routine antenatal maternal screening for current mental health: evaluation of a change in the use of the Edinburgh Depression Scale in clinical practice. *Arch Womens Ment Health*. 2015;19(2):362–72.
28. Matthey S, Valenti B, Souter K, Ross-Hamid C. Comparison of four self-report measures and a generic mood question to screen for anxiety during pregnancy in English-speaking women. *J Affect Disord*. 2013;148(2–3):347–51.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

