Are symptom features of depression during pregnancy, the postpartum period and outside the peripartum period distinct? Results from a nationally representative sample using Item Response Theory (IRT)

Nicolas Hoertel, M.D., M.P.H.[1,2,3], **Saioa López, M.D.**[4], **Hugo Peyre, M.D., M.P.H.**[5], **Melanie M. Wall, Ph.D.**[1,7], **Ana González-Pinto, M.D., Ph.D.**[4], **Frédéric Limosin, M.D., Ph.D.**[2,3], and **Carlos Blanco, M.D., Ph.D.**[1]

[1] Department of Psychiatry, New York State Psychiatric Institute / Columbia University, New York, NY 10032, USA.

[2] Assistance Publique-Hôpitaux de Paris (APHP), Corentin Celton Hospital, Department of Psychiatry, 92130 Issy-les-Moulineaux; Paris Descartes University, PRES Sorbonne Paris Cité, Paris, France.

[3] INSERM UMR 894, Psychiatry and Neurosciences Center; Paris Descartes University, PRES Sorbonne Paris Cité, Paris, France.

[4] Hospital Universitario de Alava (Santiago). Biomedical Research Centre in Mental Health Net (CIBERSAM). Department of Psychiatry. Vitoria, Spain.

[5] Assistance Publique Hôpitaux de Paris (APHP), Robert Debré Hospital, Child and Adolescent Psychiatry Department, Paris, France. 6. Cognitive Sciences and Psycholinguistic Laboratory, Ecole Normale Supérieure, CNRS, EHESS, Paris, France.

[7] Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, USA.

Abstract

Background—Whether there are systematic differences in depression symptom expression during pregnancy, the postpartum period and outside these periods (i.e., outside the peripartum period) remains debated. The aim of this study was to use methods based on item response theory (IRT) to examine, after equating for depression severity, differences in the likelihood of reporting DSM-IV symptoms of major depressive episode (MDE) in women of childbearing age (i.e., aged 18-50) during pregnancy, the postpartum period and outside the peripartum period.

Corresponding author: Nicolas Hoertel, M.D., M.P.H. Department of Psychiatry, Corentin Celton Hospital, Paris Descartes University 4 parvis Corentin Celton; 92130 Issy-les-Moulineaux, France Phone: 33 (0) 1 58 00 44 21 Fax: 33 (0) 1 58 00 44 53 nico.hoertel@yahoo.fr.

Method—We conducted these analyses using a large, nationally representative sample of women of childbearing age from the USA (n = 11,256) who participated in the second wave of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC).

Results—The overall 12-month prevalence of all depressive criteria (except for worthlessness/ guilt) was significantly lower in pregnant women than in women of childbearing age outside the peripartum period, whereas the prevalence of all symptoms (except for "psychomotor symptoms") were not significantly different between the postpartum and the non-peripartum group. There were no clinically significant differences in the endorsement rates of symptoms of MDE by pregnancy status when equating for levels of depression severity.

Conclusions—This study suggests that the clinical presentation of depressive symptoms in women of childbearing age does not differ during pregnancy, the postpartum period and outside the peripartum period. These findings do not provide psychometric support for the inclusion of the peripartum onset specifier for major depressive disorder in the DSM-5.

The Diagnostic and Statistical Manual of psychiatric disorders-fourth edition (DSM-IV) specifier "with postnatal onset" for major depressive episode (MDE) has been expanded to "with peripartum onset" in the DSM-5. In addition to onsets within 4 weeks after delivery, "peripartum onset" includes in the new version of the DSM onsets during the entire pregnancy. This change has been the subject of intense debate and brought to the fore considerations regarding whether peripartum depression is a distinct entity, different from other existing depressive disorders.[1-5] This debate is important because the nosological status of peripartum major depressive disorder (MDD) is a crucial determinant of its lines of treatment models, research and approaches to policy and funding for treatment and research.[1, 6, 7]

A line of research has suggested that peripartum depression is part of a spectrum of emotional disorders related distinctly to pregnancy and childbearing, and therefore a distinct diagnostic category. Proponents of this view argue that the clinical presentation of peripartum depression is unique[8-11] and could be caused by specific neurobiological or psychological mechanisms.[5] An opposite line of research has argued that there is unsufficient evidence to support a distinction between peripartum and non-peripartum depression and that the perinatal period may trigger or exacerbate emotional disorders like many other stressful life events.[12, 13]

The peripartum onset specifier in DSM-5 has opened another debate. This change reflects the finding that a proportion of "postpartum" depressive episodes actually begin during pregnancy and continue and often worsen during the postpartum[14] and emphasizes the importance of managing mood disorders both during and after pregnancy.[5] However, some authors have pointed out that the combination of prenatal and postnatal onset in a single specifier may obscure important differences in the epidemiology,[15] clinical

presentation,[16-18] prognosis[19] and treatment response[5] of major depression in these two periods.

Despite the clinical and research importance of these questions, data comparing the phenomenology of depression during pregnancy, the postpartum period and outside these 2 periods (i.e., the non-peripartum period) are sparse,[20] and whether there are systematic differences in depression symptom expression by pregnancy status (i.e., pregnancy, postpartum and outside the peripartum period) that could guide treatment remains uncertain.[8, 20] Identifying these potential differences could have yet important implications because it may help clinicians identify this often underdiagnosed and undertreated condition[21-24] and researchers explain potentially differential pathophysiology[22, 25-27] and treatment response.[20]

While some studies have suggested that the symptoms of depression during pregnancy or the postpartum period are more similar than different from those of non-peripartum depression,[28-32] others indicate that certain depressive features may be distinct. Specifically, non-peripartum depression may be associated with higher rates of suicidal ideation and attempts,[8-10] depressed mood,[8] anhedonia,[8] feelings of guilt,[9] insomnia,[9] and psychomotor retardation[9] compared to depression during pregnancy or the postpartum period. In contrast, psychomotor symptoms (restlessness/agitation),[8] impaired concentration/decision making[8] and anxiety symptoms[11, 20] may be more prominent in peripartum depression. In addition, depressive episodes with postpartum onset may be more often associated with obsessive-compulsive and psychotic symptoms than those with pregnancy-onset[16] and prior research supports differences in depression symptom levels from pregnancy to postpartum.[33, 34]

However, prior research has often been limited by methodological constraints such as reliance on samples of convenience, incomplete assessment of full range of DSM-IV MDE criteria, or absence of control for overall depression severity. This last point is critical because it is unclear whether any differences in symptom expression are due to true phenomenological differences, or whether such differences are reflective of greater overall symptom severity in one group compared to the others. Only one study, conducted in a clinical sample,[8] controlled for overall depression severity between groups, but its results may not be applicable to women in the general population.[35, 36]

Accordingly, our aim was to examine in a large nationally representative sample of women of childbearing age whether, when equating for levels of depression severity using item response theory (IRT) methods,[37] there were differences in the likelihood of reporting DSM-IV MDE symptoms[9, 10, 20, 29, 30] during pregnancy, the postpartum period and outside the peripartum period. Methods based on IRT,[37] which are rapidly becoming the gold standard for examining differential symptom expression,[38-42] provide significant improvements over previous statistical techniques as IRT approaches allow examining the likelihood that a particular symptom will be reported given a particular level of depression severity.[40-42] Our null hypothesis was that there would be no significant differences in psychometric properties of DSM-IV symptoms of MDE by pregnancy status.

Materials and Methods

Sample

The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) is a survey of a nationally representative sample of the civilian and non-institutionalized adult population of the United States.[43] Participants were drawn from the 2004–2005 NESARC, the second Wave following upon the Wave 1 NESARC, conducted in 2001–2002 and described in detail elsewhere.[44, 45] The cumulative response rate at Wave 2 was 70.2%, reflecting 34,653 completed Wave 2 interviews.[43] Weighted Wave 2 data were adjusted to be representative of the civilian population on sociodemographic variables including region, age, race/ethnicity and sex, based on the 2000 Decennial Census. The research protocol, including informed consent procedures, received full human subjects review and approval from the US Census Bureau and the Office of Management and Budget.[43]

Assessment of pregnancy status

All women of childbearing age (i.e., aged between 18 to 50 years[46]) were asked in Wave 2 whether they were pregnant at the time of the Wave 2 interview and whether they had been pregnant at any point in the preceding 12 months. Women who were pregnant at the time of the interview were included in the "pregnancy group", those who had been pregnant at some time during the past-year were included in the "postpartum group", whereas those who answered negatively to both questions were included in the "non-peripartum group".

Assessments of DSM-IV symptoms of MDE

The diagnostic interview was the Alcohol Use Disorders and Associated Disabilities Interview Schedule-DSM-IV version (AUDADIS-IV),[47] a fully structured diagnostic interview designed for experienced interviewers who are not clinicians. Test-retest reliability[47, 48] and validity[49] of the AUDADIS-IV diagnoses of MDE are good (k = 0.64– 0.67), and a clinical reappraisal study[44] of major depression indicated good agreement between AUDADIS-IV and psychiatrist diagnoses (k = 0.64–0.68). Analyses for this study focused on the nine DSM-IV MDE symptoms (Table 1). All NESARC participants were asked whether they had a 2-week episode of depressed mood and/or anhedonia between the 2 Waves interviews. Participants who answered affirmatively were further queried about the other DSM-IV depressive symptoms that occurred between the 2 Waves interviews. Depressive symptoms had to co-occur to be counted.[47] Participants who met the DSM-IV criteria for MDE were queried on whether they had these depressive symptoms during the past year.

Our purpose was to examine the full depressive spectrum in the general population of women of childbearing age. Because we were sure of the occurrence of the depressive symptoms during the year preceding the Wave 2 interview for women with a past-year MDE but not for women with subthreshold depressive symptoms (i.e., without a DSM-IV diagnosis of MDE), we performed two sets of analyses. First, we compared depressive symptoms expression stratifying by pregnancy status in the full sample of women of childbearing age, assuming that all women reporting subthreshold depressive symptoms between the 2 Waves interviews had these symptoms during the year preceding the

interview. Second, we conducted a sensitivity analysis by excluding all women who reported subthreshold depressive symptoms from our analyses to ensure including only depressive symptoms that occurred the same year we assessed pregnancy status (referred to as the "sensitivity analysis subsample").

Analyses

Comparing prevalence of depressive symptoms

Differences in prevalence of each DSM-IV depressive symptom by pregnancy status were tested using chi-square tests in both the full sample of women of childbearing age and the subsample of women with a past-year DSM-IV diagnosis of MDE.

Assessing unidimensionality

Prior to item response theory modeling, including testing for differential item functioning (DIF), it is necessary to confirm the unidimensionality of the set of symptoms. Eigenvalue analysis of the tetrachoric correlation matrices of the nine DSM-IV depressive symptoms was performed separately in the three groups (i.e., pregnancy, postpartum group and non-peripartum group). A scree plot with an elbow after the first eigenvalue and a well- fitting one-factor model according to standard goodness of fit tests (RMSEA<.05, TLI>.95 and CFI>.95) were used to indicate unidimensionality in both sets of analyses.

Item response theory (IRT)

Item response theory (IRT) modeling was conducted to examine whether the DSM-IV depressive symptoms were endorsed similarly in women during pregnancy, the postpartum period and outside the peripartum period[50] after controlling for differences in underlying depression severity.[38] In IRT, a severity and a discrimination parameter are estimated for each item (i.e., symptoms of MDE in this case) indicating where along the latent severity continuum a symptom has a probability of 50% of being endorsed (for the severity parameter), and how rapidly the probability of observing the symptom changes across increasing levels of the latent severity continuum (for the discrimination parameter). Once IRT parameters for each item were estimated, Spearman rank order correlation coefficient was used to evaluate the ordering of the severity and discrimination parameters in the three groups.

All factor analyses were conducted in Mplus Version 7.0[51] which takes into account the NESARC sampling weights and design effects, in the parameter and standard error estimation as well as model fit calculations. The default estimator for the analysis was the variance adjusted weighted least squares (WLSMV), a robust estimator that does not assume normally distributed variables and provides the best option for modeling categorical data.

Testing differential item functioning (DIF)

Multisample confirmatory factor analysis (MCFA) was used[41, 42] to test DIF in the three groups of (a) factor loadings (i.e., item discrimination parameters) and (b) thresholds (i.e., severity parameters). The different models varied in terms of the parameters constrained to be equal across groups.[52] Initially, an unconstrained model was fit allowing all parameters

to differ across groups in both sets of analyses (Tables 2 and 3). Analyses were conducted iteratively to determine which IRT parameters differ across groups.[53] Nested models were compared using the chi-square difference test (implemented using the DIFFTEST option in Mplus). To detect DIF in IRT parameters of each item, the discrimination and severity parameters of the others items were constrained to be equal across groups. For each item, a chi-square difference test DIF comparing models with and without the discrimination parameter of the item constrained to be equal across groups was used to identify DIF in the discrimination parameter. To identify DIF in the severity parameter, a chi-square difference test DIF was used to compare models with and without the severity parameter of the item constrained to be equal across groups.

To minimize the number of false positives that could derive from the multiple comparisons conducted in this study, we set alpha at 0.05 and used the Benjamini–Hochberg procedure to adjust p-values for all 1 df tests.[54] Furthermore, small differences in severity between groups can be statistically significant but may not be clinically meaningful.[55] Thus, it was decided *a priori* that only differences greater than 0.50 in symptom discrimination and severity would be considered as clinically meaningful.[56] Such differences in symptom severity can be interpreted as one "half of the standard unit difference between the values of the (underlying) trait necessary to have a 50–50 chance of responding positively in one group compared to another".[56] This difference may be associated with differences as large as 10% in the probability of response.[56] Because graphical display is helpful in placing the "effect sizes in a practical and theoretical context",[57] we present the item response curves (IRC) for all symptom parameters that reached our *a priori* threshold of both statistical and clinical significance.

Results

Sample characteristics

There were 11,342 women of childbearing age in the Wave 2 NESARC sample. Of these, 86 did not report their past-year pregnancy status and were removed from our analyses. Of the 11,256 women included in our analyses, 346 were pregnant at the time of the interview (referred to as the "pregnancy group"), 736 were not currently pregnant but had been pregnant at some time during the year prior to the interview (referred to as the "postpartum group"), while 10,174 women were not (referred to as the "non-peripartum group").

Symptoms of MDE endorsement rates

Table 1 shows the prevalence of symptoms of MDE in women by pregnancy status in the full sample of women of childbearing age and in the subsample of women of childbearing age with a past-year DSM-IV diagnosis of MDE. In the full sample, the prevalence rates of all symptoms of MDE (except for worthlessness/guilt) were significantly higher in the nonperipartum group than in the pregnancy group, whereas the prevalence of all symptoms (except for "psychomotor disturbance" which was less common in the postpartum group) were not significantly different between the postpartum group and the non-peripartum group. The prevalence of MDE was not significantly different across the three groups. In the

subsample of women with a past-year DSM-IV diagnosis of MDE, there were no significant differences in prevalence rates of symptoms of depression across the three groups.

Unidimensionality of symptoms of MDE

In the three groups, fit indices indicated an excellent fit to the data in the full sample (pregnancy group: first factor eigenvalue = 8.5, second factor eigenvalue = 0.2, CFI = 1.00, TLI = 1.00, RMSEA = 0.015; postpartum group: first factor eigenvalue = 8.3, second factor eigenvalue = 0.2, CFI = 1.00, TLI = 1.00, RMSEA = 0.015; non-peripartum group: first factor eigenvalue = 8.3, second factor eigenvalue = 0.3, CFI = 1.00, TLI = 1.00, RMSEA = 0.020) and in the sensitivity analysis subsample (pregnancy group: first factor eigenvalue = 8.8, second factor eigenvalue = 0.1, CFI = 1.00, TLI = 1.00, RMSEA = 0.010; postpartum group: first factor eigenvalue = 8.8, second factor eigenvalue = 0.2, CFI = 1.00, TLI = 1.00, RMSEA = 0.020; non-peripartum group: first factor eigenvalue = 8.8, second factor eigenvalue = 0.1, CFI = 1.00, TLI = 1.00, RMSEA = 0.020). We determined that these fit statistics were adequate to proceed to fitting IRT models.

IRT item parameters

The ranking of IRT parameters was similar in the pregnancy group and in the nonperipartum group both in the main analysis (Spearman's correlation coefficients were respectively 0.97 [differences of rank < 2] for severity parameters and 1.00 [differences of rank = 0] for discrimination parameters) and in the sensitivity analysis (Spearman's correlation coefficients were respectively 0.93 [differences of rank < 3] for severity parameters and 0.93 [differences of rank < 3] for discrimination parameters) (Tables 2 and 3). The ranking of IRT parameters was also similar in the postpartum group and in the nonperipartum group both in the main analysis (Spearman's correlation coefficients were respectively 0.90 [differences of rank < 4] for severity parameters and 0.87 [differences of rank < 4] for discrimination parameters) and in the sensitivity analysis (Spearman's correlation coefficients were respectively 0.95 [differences of rank < 2] for severity parameters and 0.88 [differences of rank < 4] for discrimination parameters).

Differential item functioning (DIF)

In the full sample, after adjusting for multiple testing using the Benjamini-Hochberg procedure, no item differed significantly in symptom severity or discrimination between the postpartum group and the non-peripartum group (Table 2). No item exhibited statistically significant differences in symptom severity between the pregnancy group and the nonperipartum group, while one item (i.e., suicidality) displayed statistically significant difference in symptom discrimination and reached our *a priori* threshold for clinical significance. However, the difference of rank between these groups was null and visual inspection of the item response curve (IRC) for this criterion revealed a small non-uniform DIF (Figure 1), indicating that this difference was not clinically meaningful.

In the sensitivity analysis subsample, similar results were found except for the criterion suicidality which reached our *a priori* threshold for both statistical and clinical significance in symptom discrimination between the postpartum group and the nonperipartum group (Table 3). However, the difference of rank between these two groups was null and visual

inspection of the item response curve (IRC) for this criterion revealed also a small non-uniform DIF (Figure 2), supporting that this difference was not clinically meaningful.

There were no significant differences in symptom severity or discrimination between the pregnancy group and the postpartum group (data available on request) in any of the analyses.

Discussion

This is, to our knowledge, the first study to examine differences in the likelihood of reporting symptoms of MDE during pregnancy, the postpartum period and outside the peripartum period in a nationally representative sample of women of childbearing age using an IRT-based methodology. The benefit of the IRT-based approach is that it accounts for the potential confounding effect of depression severity in evaluating potential group differences in symptom expression. Additional strengths of the present study include (i) the use of a large representative community sample of women of childbearing age, (ii) the incorporation in our analyses of information from the sampling methods of the survey design which has been suggested to substantially improve standard error estimates and tests of model fit,[58] and (iii) the use of an *a priori* defined threshold of clinical significance[56] so that emphasis was placed on differences that are substantively meaningful rather than only on statistical significance.

Several novel findings emerged from the present study.

Our results extend prior findings indicating that, although the overall 12-month prevalence of all depressive criteria in women during pregnancy and the postpartum period was high,[59-63] it was significantly lower in pregnant women (except for the criterion worthlessness/guilt) than in women of childbearing age outside the peripartum period. In addition, we did not find significant differences in the prevalence of depressive symptoms (except for psychomotor disturbance) between women during the postpartum period and women of childbearing age outside the peripartum period. This finding is consistent with several,[64, 65] although not all[66, 67] studies derived from clinical samples. Discrepancies with previous studies results may be due to differences in the diagnostic criteria, the timing of the assessments, sample sizes, or the use of convenience samples rather than population-based samples.[46]

We found that pregnant women were significantly less likely to endorse all depressive symptoms (except for the criterion worthlessness/guilt) than non-peripartum women of childbearing age. Although women without a history of depressive symptoms may be more likely to become pregnant than those with depressive symptoms,[46] which would lead to lower rates of depressive symptoms among pregnant women, no significant differences were found between the postpartum group and the non-peripartum group. Pregnancy may thus have a mood protective effect, either through biological (e.g., hormonal) effects or the often positive psychological effects of upcoming motherhood, that may decrease after delivery. Nevertheless, although the prevalence of depressive symptoms was lower in pregnant than in non-peripartum women, the prevalence of depressive symptoms in peripartum women

was high. In addition, consistent with prior research,[28-31] the prevalence rate of suicidality was similar during pregnancy, the postpartum period and outside the peripartum period. These findings stress the need for continued work to identify the causes of peripartum depression and develop effective treatments among pregnant and postpartum women.[5]

Although substantial differences in depressive symptoms had been previously reported in clinical samples of pregnant and postpartum women,[8-10, 20] the specific contribution of pregnancy and postpartum to the depression symptom expression in the general population of women of childbearing age had not been previously examined. We found that the endorsement rates of the symptoms of MDE were not significantly different during pregnancy, the postpartum period and outside the peripartum period after equating for depression severity. Our results are robust, as suggested by the similarity of the findings of the main and the sensitivity analysis. The only exception was the criterion "suicidality" which reached our *a priori* threshold for both statistical and clinical significance in symptom discrimination in the main and the sensitivity analyses. However, several lines of evidence argue against the clinical impact of this difference. First, the ranking of IRT parameters was similar between groups for both analyses. Second, there were no differences in rank of the discrimination and severity parameters of suicidality between groups in both analyses. Last, the visual examination of the item response curve for suicidality in both analyses, often considered the most accurate method to detect DIF between groups,[57] revealed that the trace lines were not substantially steeper between peripartum groups, indicating that responses to this item are not more discriminating in one group compared to the others in a clinically meaningful way. These results are important because they extend prior findings derived from clinical samples[28-31] to the general population of women of childbearing age.

Overall, our findings suggest that the clinical presentation of depression during pregnancy and the postpartum period does not substantially differ from non-peripartum depression. Despite this lack of differences, depression during pregnancy and the postpartum constitutes a major public health concern because it is often underdiagnosed and undertreated,[21-24] has disabling consequences for both the mother[68-71] and the newborn,[72-76] and is associated with poorer treatment outcomes than other depressive disorders.[20] At present, there are no established biomarkers for psychiatric disorders. An important avenue for future research that could lend support for the peripartum specifier would be to identify the existence of biosignatures that could reliably discriminate the biology of depression during versus outside the peripartum period or suggest differential treatment.[7, 77]

This study has several limitations.

A first limitation involves the retrospective nature of the assessment of depressive symptoms that could have led to recall bias. Our findings should be further confirmed in studies comparing prospectively depressive symptoms in women during pregnancy, the postpartum period and outside these periods. In addition, although the NESARC was not specifically collected to study depression in pregnant and postpartum women, it represents to date the largest US nationally representative data including information on depressive symptoms and pregnancy status. Second, the postpartum period encompassed in our study up to 12 months after delivery, which may have led to underestimate true differences. Third, women had to

endorse a 2-week period of depressed mood and/or anhedonia to be queried about other DSM-IV depressive symptoms. Therefore, we may have missed information on other depressive symptoms for some participants with lower levels of depressed mood that should be considered to be part of the depressive spectrum. However, about one third of the full NESARC sample endorsed depressed mood or anhedonia sometime in their life and our results are consistent with those of previous studies.[28-30] Fourth, depressive symptoms were rated dichotomously rather than by severity. This may have led the data to be less sensitive to true group differences. Fifth, information on whether women experienced miscarriage or stillbirth and on the stages of pregnancy and postpartum was not available. This last point may be important because women at late stages of pregnancy or peripartum may be more likely to minimize or mis-interpret symptoms of depression as being merely a pregnancy or postpartum-related problem and not indicative of depression. Last, the assessment of 12-month symptoms of MDE may have included women who were early in their pregnancy, and therefore reporting symptoms predating that pregnancy.

Conclusion

Despite these limitations, by addressing several constraints of prior research and using Item Response Theory (IRT), the present study suggests that the clinical presentation of depressive symptoms in women during pregnancy and the postpartum period does not differ from that observed outside the peripartum period. We hope that this information can be helpful in the development of new psychiatric classifications and improve the identification and treatment of depressed women during the peripartum period.

Acknowledgments

References

1. Sharma V, Mazmanian D. The DSM-5 peripartum specifier: prospects and pitfalls. Arch Women Ment Hlth. 2014; 17:171–173.

2. Austin MP. Classification of mental health disorders in the perinatal period: future directions for DSM-V and ICD-11. Arch Women Ment Hlth. 2010; 13:41–44.

3. Jones I, Cantwell R. Nosology Working Group RCoPPS. The classification of perinatal mood disorders--suggestions for DSMV and ICD11. Arch Women Ment Hlth. 2010; 13:33–36.

4. Kornstein SG. Gender issues and DSM-V. Arch Women Ment Hlth. 2010; 13:11–13.

5. Uher R, Payne JL, Pavlova B, Perlis RH. Major depressive disorder in DSM-5: implications for clinical practice and research of changes from DSM-IV. Depress Anxiety. 2014; 31:459–471. [PubMed: 24272961]

6. Wisner KL, Moses-Kolko EL, Sit DK. Postpartum depression: a disorder in search of a definition. Arch Women Ment Hlth. 2010; 13:37–40.

7. Wisner KL, Parry BL, Piontek CM. Clinical practice. Postpartum depression. N Engl J Med. 2002; 347:194–199. [PubMed: 12124409]

8. Bernstein IH, Rush AJ, Yonkers K, et al. Symptom features of postpartum depression: are they distinct? Depress Anxiety. 2008; 25:20–26. [PubMed: 17187349]

9. Manber R, Blasey C, Allen JJ. Depression symptoms during pregnancy. Arch Women Ment Hlth. 2008; 11:43–48.

10. Pitt B. "Atypical" depression following childbirth. Br J Psychiatry. 1968; 114:1325–1335. [PubMed: 5750402]

11. Dalton, K.; Horton, WM. Depression after childbirth: How to recognize, treat, and prevent postnatal depression. 3rd ed.. Oxford University Press; Oxford; New York: 1996.

12. Cooper PJ, Murray L. Course and recurrence of postnatal depression. Evidence for the specificity of the diagnostic concept. Br J Psychiatry. 1995; 166:191–195. [PubMed: 7728362]

13. Riecher-Rossler A, Hofecker Fallahpour M. Postpartum depression: do we still need this diagnostic term? Acta Psychiatr Scand. 2003; (Suppl. 418):51–56. [PubMed: 12807377]

14. Di Florio A, Forty L, Gordon-Smith K, et al. Perinatal episodes across the mood disorder spectrum. JAMA psychiatry. 2013; 70:168–175. [PubMed: 23247604]

15. Viguera AC, Nonacs R, Cohen LS, et al. Risk of recurrence of bipolar disorder in pregnant and nonpregnant women after discontinuing lithium maintenance. Am J Psychiatry. 2000; 157:179–184. [PubMed: 10671384]

16. Altemus M, Neeb CC, Davis A, et al. Phenotypic differences between pregnancy-onset and postpartum-onset major depressive disorder. J Clin Psychiatry. 2012; 73:e1485–1491. [PubMed: 23290332]

17. Banti S, Mauri M, Oppo A, et al. From the third month of pregnancy to 1 year postpartum. Prevalence, incidence, recurrence, and new onset of depression. Results from the perinatal depression-research & screening unit study. Compr Psychiat. 2011; 52:343–351. [PubMed: 21683171]

18. Gavin NI, Gaynes BN, Lohr KN, et al. Perinatal depression: a systematic review of prevalence and incidence. Obstet Gynecol. 2005; 106:1071–1083. [PubMed: 16260528]

19. Munk-Olsen T, Laursen TM, Meltzer-Brody S, et al. Psychiatric disorders with postpartum onset: possible early manifestations of bipolar affective disorders. Arch Gen Psychiatry. 2012; 69:428–434. [PubMed: 22147807]

20. Hendrick V, Altshuler L, Strouse T, Grosser S. Postpartum and nonpostpartum depression: differences in presentation and response to pharmacologic treatment. Depress Anxiety. 2000; 11:66–72. [PubMed: 10812531]

21. Ramsay IN, Torbet TE. Incidence of abnormal voiding parameters in the immediate postpartum period. Neurourol Urodyn. 1993; 12:179–183. [PubMed: 7920675]

22. DelRosario GA, Chang AC, Lee ED. Postpartum depression: symptoms, diagnosis, and treatment approaches. JAAPA. 2013; 26:50–54. [PubMed: 23409386]

23. Hoertel N, Le Strat Y, Angst J, Dubertret C. Subthreshold bipolar disorder in a U.S. national representative sample: prevalence, correlates and perspectives for psychiatric nosography. J Affect Disord. 2013; 146:338–347. [PubMed: 23040874]

24. Manetti A, Hoertel N, Le Strat Y, et al. Comorbidity of Late-Life Depression in the United States: A Population-based Study. Am J Geriatr Psychiatry. Aug 26.2013 pii: S1064-7481(13)00237-6. doi: 10.1016/j.jagp.2013.05.001. [Epub ahead of print].

25. Mahon PB, Payne JL, MacKinnon DF, et al. Genome-wide linkage and follow-up association study of postpartum mood symptoms. Am J Psychiatry. 2009; 166:1229–1237. [PubMed: 19755578]

26. Meltzer-Brody S. New insights into perinatal depression: pathogenesis and treatment during pregnancy and postpartum. Dialogues Clin Neurosci. 2011; 13:89–100. [PubMed: 21485749]

27. Yim IS, Glynn LM, Dunkel-Schetter C, et al. Risk of postpartum depressive symptoms with elevated corticotropin-releasing hormone in human pregnancy. Arch Gen Psychiatry. 2009; 66:162–169. [PubMed: 19188538]

28. Cooper PJ, Campbell EA, Day A, et al. Non-psychotic psychiatric disorder after childbirth. A prospective study of prevalence, incidence, course and nature. The Br J Psychiatry. 1988; 152:799–806.

29. Whiffen VE. The comparison of postpartum with non-postpartum depression: a rose by any other name. J Psychiatry Neurosci. 1991; 16:160–165. [PubMed: 1958650]

30. Whiffen VE, Gotlib IH. Comparison of postpartum and nonpostpartum depression: clinical presentation, psychiatric history, and psychosocial functioning. J Consult Clin Psychol. 1993; 61:485–494. [PubMed: 8326051]

31. O'Hara MW, McCabe JE. Postpartum depression: current status and future directions. Annu Rev Clin Psychol. 2013; 9:379–407. [PubMed: 23394227]

32. Cunningham NK, Brown PM, Brooks J, Page AC. The structure of emotional symptoms in the postpartum period: is it unique? J Affect Disord. 2013; 151:686–694. [PubMed: 23978685]

33. De Wilde KS, Trommelmans LC, Laevens HH, et al. Smoking patterns, depression, and sociodemographic variables among Flemish women during pregnancy and the postpartum period. Nurs Res. 2013; 62:394–404. [PubMed: 24165215]

34. Josefsson A, Berg G, Nordin C, Sydsjo G. Prevalence of depressive symptoms in late pregnancy and postpartum. Acta Obstet Gynecol Scand. 2001; 80:251–255. [PubMed: 11207491]

35. Hoertel N, Le Strat Y, Limosin F, et al. Prevalence of subthreshold hypomania and impact on internal validity of RCTs for major depressive disorder: results from a national epidemiological sample. PloS one. 2013; 8:e55448. [PubMed: 23405152]

36. Hoertel N, Le Strat Y, Lavaud P, et al. Generalizability of clinical trial results for bipolar disorder to community samples: findings from the National Epidemiologic Survey on Alcohol and Related Conditions. J Clin Psychiatry. 2013; 74:265–270. [PubMed: 23561233]

37. Lord, F. Applications of Item Response Theory to Practical Testing Problems. Erlbaum; Hillsdale, NJ.: 1980.

38. Weinstock LM, Strong D, Uebelacker LA, Miller IW. Differential item functioning of DSM-IV depressive symptoms in individuals with a history of mania versus those without: an item response theory analysis. Bipolar Disord. 2009; 11:289–297. [PubMed: 19419386]

39. Weinstock LM, Strong D, Uebelacker LA, Miller IW. DSM-IV depressive symptom expression among individuals with a history of hypomania: a comparison to those with or without a history of mania. J Psychiatr Res. 2010; 44:979–85. [PubMed: 20398907]

40. Uebelacker LA, Strong D, Weinstock LM, Miller IW. Use of item response theory to understand differential functioning of DSM-IV major depression symptoms by race, ethnicity and gender. Psychol Med. 2009; 39:591–601. [PubMed: 18588740]

41. Peyre H, Hoertel N, Cortese S, et al. ADHD Symptoms Expression: a Comparison of Individual Age of Onset using Item Response Theory. J Clin Psychiatry. 2014; 75:386–392. [PubMed: 24813406]

42. Hoertel N, Peyre H, Wall M, et al. Examining gender differences in DSM-IV borderline personality disorder symptom expression using Item Response Theory (IRT). J Psychiatr Res. 2014 pii: S0022-3956(14)00262-3. doi: 10.1016/j.jpsychires.2014.08.019. [Epub ahead of print].

43. Grant BF, Goldstein RB, Chou SP, et al. Sociodemographic and psychopathologic predictors of first incidence of DSM-IV substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. Mol Psychiatry. 2009; 14:1051–1066. [PubMed: 18427559]

44. Grant BF, Stinson FS, Dawson DA, et al. Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry. 2004; 61:807–816. [PubMed: 15289279]

45. Grant BF, Stinson FS, Dawson DA, et al. Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry. 2004; 61:361–368. [PubMed: 15066894]

46. Vesga-Lopez O, Blanco C, Keyes K, et al. Psychiatric disorders in pregnant and postpartum women in the United States. Arch Gen Psychiatry. 2008; 65:805–815. [PubMed: 18606953]

47. Grant BF, Dawson DA, Stinson FS, et al. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. Drug Alcohol Depend. 2003; 71:7–16. [PubMed: 12821201]

48. Hasin DS, Goodwin RD, Stinson FS, Grant BF. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiatry. 2005; 62:1097–1106. [PubMed: 16203955]

49. Canino G, Bravo M, Ramirez R, et al. The Spanish Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability and concordance with clinical diagnoses in a Hispanic population. J Stud Alcohol. 1999; 60:790–799. [PubMed: 10606491]

50. Harford TC, Chen CM, Saha TD, et al. An item response theory analysis of DSM-IV diagnostic criteria for personality disorders: findings from the national epidemiologic survey on alcohol and related conditions. Personal Disord. 2013; 4:43–54. [PubMed: 22449066]

51. Muthén, LK.; Muthén, BO. User's guide. Los Angeles: 2007. Mplus: Statistical analyses with latent variables..

52. Millsap RE, Yun-Tein J. Assessing factorial invariance in ordered-categorical measures. Multivar Behav Res. 2004; 39:479–515.

53. McBride O, Strong DR, Kahler CW. Exploring the role of a nicotine quantity-frequency use criterion in the classification of nicotine dependence and the stability of a nicotine dependence continuum over time. Nicotine Tob Res. 2010; 12:207–216. [PubMed: 20081041]

54. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Series B. 1995; 57:289–300.

55. Strong DR, Kahler CW, Colby SM, Griesler PC, Kandel D. Linking measures of adolescent nicotine dependence to a common latent continuum. Drug Alcohol Depend. 2009; 99:296–308. [PubMed: 18938047]

56. Steinberg L, Thissen D. Using effect sizes for research reporting: examples using item response theory to analyze differential item functioning. Psychol Methods. 2006; 11:402–415. [PubMed: 17154754]

57. Wilkinson L. Task Force on Statistical Inference, APA Board of Scientific Affairs. Statistical methods in psychology journals: Guidelines and explanations. Am Psychologist. 1999; 54:595–604.

58. Asparouhov T. Sampling weights in latent variable modeling. Struct. Equ. Modeling. 2005; 12:411–434.

59. Andersson L, Sundstrom-Poromaa I, Wulff M, et al. Depression and anxiety during pregnancy and six months postpartum: a follow-up study. Acta Obstet Gynecol Scand. 2006; 85:937–944. [PubMed: 16862471]

60. Adewuya AO, Ola BA, Aloba OO, et al. Prevalence and correlates of depression in late pregnancy among Nigerian women. Depress Anxiety. 2007; 24:15–21. [PubMed: 16845663]

61. Bennett HA, Einarson A, Taddio A, et al. Prevalence of depression during pregnancy: systematic review. Obstet Gynecol. 2004; 103:698–709. [PubMed: 15051562]

62. Gaillard A, Le Strat Y, Mandelbrot L, et al. Predictors of postpartum depression: prospective study of 264 women followed during pregnancy and postpartum. Psychiatry Res. 2014; 215:341–346. [PubMed: 24370337]

63. Le Strat Y, Dubertret C, Le Foll B. Prevalence and correlates of major depressive episode in pregnant and postpartum women in the United States. J Affect Disord. 2011; 135:128–138. [PubMed: 21802737]

64. O'Hara MW, Zekoski EM, Philipps LH, Wright EJ. Controlled prospective study of postpartum mood disorders: comparison of childbearing and nonchildbearing women. J Abnorm Psychol. 1990; 99:3–15. [PubMed: 2307763]

65. Regmi S, Sligl W, Carter D, et al. A controlled study of postpartum depression among Nepalese women: validation of the Edinburgh Postpartum Depression Scale in Kathmandu. Trop Med Int Health. 2002; 7:378–382. [PubMed: 11952955]

66. Augusto A, Kumar R, Calheiros JM, Matos E, Figueiredo E. Post-natal depression in an urban area of Portugal: comparison of childbearing women and matched controls. Psychol Med. 1996; 26:135–141. [PubMed: 8643752]

67. Eberhard-Gran M, Tambs K, Opjordsmoen S, et al. A comparison of anxiety and depressive symptomatology in postpartum and non-postpartum mothers. Soc Psychiatry Psychiatr Epidemiol. 2003; 38:551–556. [PubMed: 14564383]

68. Davidson J, Robertson E. A follow-up study of post partum illness, 1946-1978. Acta Psychiatr Scand. 1985; 71:451–457. [PubMed: 4013805]

69. Barr HM, Bookstein FL, O'Malley KD, et al. Binge drinking during pregnancy as a predictor of psychiatric disorders on the Structured Clinical Interview for DSM-IV in young adult offspring. Am J Psychiatry. 2006; 163:1061–1065. [PubMed: 16741207]

70. Flick LH, Cook CA, Homan SM, et al. Persistent tobacco use during pregnancy and the likelihood of psychiatric disorders. Am J Public Health. 2006; 96:1799–1807. [PubMed: 17008576]

71. Kurki T, Hiilesmaa V, Raitasalo R, et al. Depression and anxiety in early pregnancy and risk for preeclampsia. Obstet Gynecol. 2000; 95:487–490. [PubMed: 10725477]

72. Weinberg MK, Tronick EZ. Emotional characteristics of infants associated with maternal depression and anxiety. Pediatrics. 1998; 102:1298–1304. [PubMed: 9794973]

73. Repokari L, Punamaki RL, Poikkeus P, et al. Ante- and perinatal factors and child characteristics predicting parenting experience among formerly infertile couples during the child's first year: a controlled study. J Fam Psychol. 2006; 20:670–679. [PubMed: 17176203]

74. Huot RL, Brennan PA, Stowe ZN, et al. Negative affect in offspring of depressed mothers is predicted by infant cortisol levels at 6 months and maternal depression during pregnancy, but not postpartum. Ann N Y Acad Sci. 2004; 1032:234–236. [PubMed: 15677418]

75. Rahman A, Bunn J, Lovel H, Creed F. Maternal depression increases infant risk of diarrhoeal illness: --a cohort study. Arch Dis Child. 2007; 92:24–28. [PubMed: 16966339]

76. Rahman A, Harrington R, Bunn J. Can maternal depression increase infant risk of illness and growth impairment in developing countries? Child Care Health Dev. 2002; 28:51–56. [PubMed: 11856187]

77. Hoertel N, de Maricourt P, Gorwood P. Novel routes to bipolar disorder drug discovery. Expert Opin Drug Discov. Aug.2013 8:907–918. [PubMed: 23706065]

Figure 1.
Differences in the probability of endorsing suicidality across levels of depression severity in the general population of women of childbearing age during pregnancy, the postpartum period and outside the peripartum period in the main analysis.
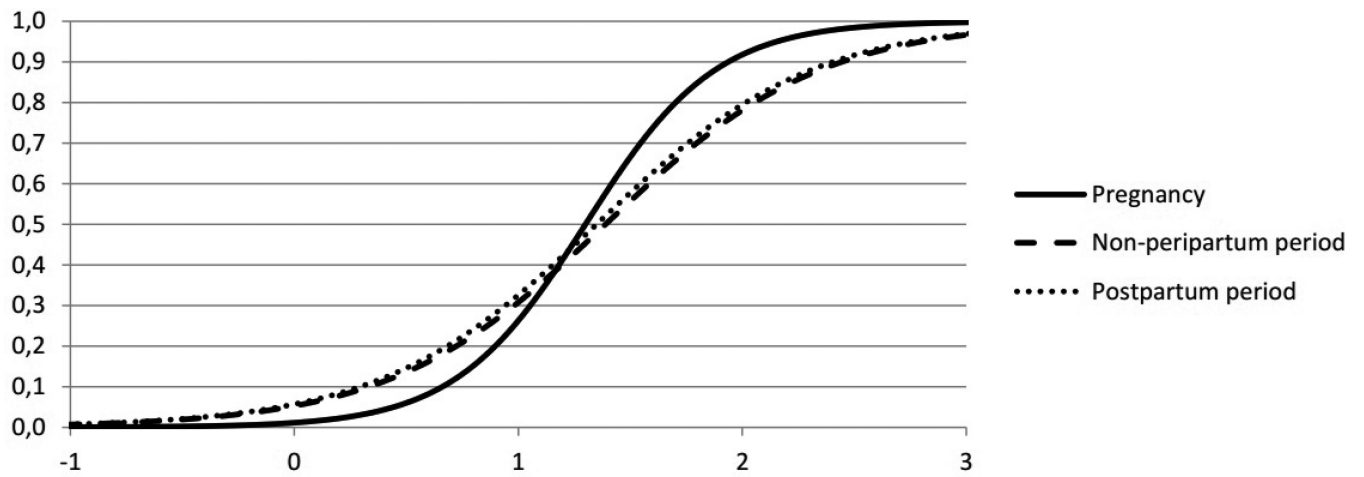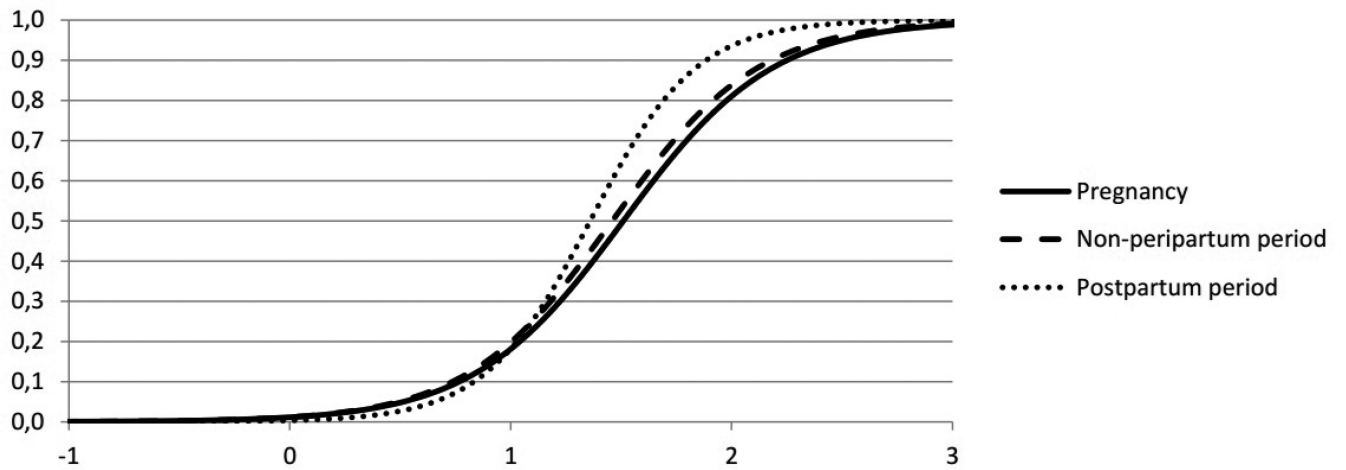
Figure 2.
Differences in the probability of endorsing suicidality across levels of depression severity in the general population of women of childbearing age during pregnancy, the postpartum period and outside the peripartum period in the sensitivity analysis.

Table 1

Prevalence of DSM-IV depressive symptoms in women of childbearing age during pregnancy, the postpartum period and outside the peripartum period in the NESARC sample.

DSM-IV symptoms	Full Sample of women of childbearing age (N = 11,256)					Women of childbearing age with DSM-IV MDE (n = 1,453)				
	Pregnancy (N = 346)	Postpartum period (N = 736)	Non-peripartum period (N = 10174)	Pregnancy vs. Non-peripartum period	Postpartum period vs. Non-peripartum period	Pregnancy (N = 36)	Postpartum period (N = 81)	Non-peripartum period (N = 1336)	Pregnancy vs. Non-peripartum period	Postpartum period vs. Non-peripartum period
	%	%	%	Wald F $(df=1)$ [p-value]	Wald F $(df=1)$ [p-value]	%	%	%	Wald F $(df=1)$ [p-value]	Wald F $(df=1)$ [p-value]
Sadness	15.6	24.3	25.3	**11.6 (0.001)**	0.3 (0.596)	96.3	95.2	85.9	0.0 (0.885)	0.1 (0.791)
Anhedonia	11.3	18.8	20.2	**11.9 (0.001)**	0.6 (0.434)	75.8	85.1	87.2	2.2 (0.143)	0.2 (0.653)
Appetite disturbance	9.2	12.7	17.0	**12.6 (<0.001)**	7.0 (0.010)	74.2	76.5	79.2	0.3 (0.568)	0.3 (0.607)
Sleep disturbance	12.2	16.7	20.4	**10.1 (0.002)**	4.1 (0.048)	95.2	90.1	93.3	0.2 (0.643)	1.0 (0.335)
Fatigue	10.9	16.9	19.6	**11.0 (0.002)**	2.31 (0.133)	93.3	90.5	91.8	0.1 (0.732)	0.1 (0.712)
Psychomotor symptoms	7.9	9.4	13.5	**5.3 (0.024)**	**8.8 (0.004)**	65.4	56.6	70.9	0.3 (0.582)	4.3 (0.045)
Worthlessness / guilt	9.3	12.2	14.1	4.0 (0.051)	1.5 (0.229)	70.9	78.7	74.3	0.1 (0.730)	0.6 (0.438)
Concentration difficulties	12.3	17.4	19.6	**7.2 (0.009)**	1.6 (0.206)	98.4	87.1	93.3	2.0 (0.164)	4.1 (0.050)
Suicidality	5.7	9.5	10.6	**7.6 (0.008)**	0.6 (0.459)	44.1	58.3	54.6	1.2 (0.281)	0.3 (0.571)
DSM-IV diagnosis of MDE	9.1	10.2	13.1	3.3 (0.073)	3.3 (0.074)	-	-	-	-	-

Percentages are weighted values.

P-values in bold are statistically significant (p-value adjusted for multiple testing using the Benjamini-Hochberg procedure < .05).

Table 2

Differential item functioning (DIF) of DSM-IV depressive symptoms in women of childbearing age during pregnancy, the postpartum period and outside the peripartum period in the NESARC sample.

DSM-IV symptoms	Pregnancy (N = 346)		Postpartum period (N = 736)		Non-peripartum period (N = 10174)		Pregnancy vs. Non-peripartum period				Postpartum period vs. Non-peripartum period			
							DIFFTEST a		DIFFTEST b		DIFFTEST a		DIFFTEST b	
	a [rank]	b [rank]	a [rank]	b [rank]	a [rank]	b [rank]	X² (p-value)	DIFa	X² (p-value)	DIFb	X² (p-value)	DIFa	X² (p-value)	DIFb
Sadness	6.56 [9]	0.67 [1]	5.38 [9]	0.60 [1]	6.05 [9]	0.67 [1]	2.20 (0.138)	0.16	0.08 (0.779)	0.00	1.35 (0.246)	-0.67	1.51 (0.220)	-0.07
Anhedonia	4.31 [5]	0.89 [4]	3.97 [8]	0.81 [2]	3.69 [5]	0.86 [3]	3.20 (0.074)	0.73	0.06 (0.811)	0.03	0.12 (0.732)	0.28	1.61 (0.204)	-0.05
Appetite disturbance	3.78 [3]	1.04 [7]	2.64 [2]	1.11 [6]	2.92 [3]	1.01 [6]	2.55 (0.111)	1.09	0.15 (0.697)	0.03	0.85 (0.356)	-0.28	2.53 (0.112)	0.10
Sleep disturbance	6.00 [8]	0.82 [2]	3.81 [7]	0.89 [5]	5.24 [8]	0.84 [2]	3.68 (0.055)	0.51	0.08 (0.782)	-0.02	3.20 (0.074)	-1.43	0.23 (0.632)	0.05
Fatigue	5.84 [6]	0.90 [5]	3.63 [6]	0.88 [4]	4.38 [6]	0.88 [5]	3.26 (0.071)	1.15	1.38 (0.240)	0.02	2.69 (0.101)	-0.75	1.02 (0.314)	0.00
Psychomotor symptoms	3.77 [2]	1.11 [8]	2.65 [3]	1.30 [8]	2.72 [2]	1.18 [8]	2.36 (0.124)	1.12	1.18 (0.316)	-0.06	0.14 (0.708)	-0.07	2.58 (0.108)	0.12
Worthlessness / guilt	4.31 [4]	1.00 [6]	3.00 [4]	1.12 [7]	3.08 [4]	1.13 [7]	2.97 (0.085)	1.24	0.77 (0.381)	-0.13	0.46 (0.496)	-0.08	0.04 (0.837)	-0.01
Concentration difficulties	5.64 [7]	0.83 [3]	3.59 [5]	0.86 [3]	5.00 [7]	0.87 [4]	2.96 (0.085)	0.65	1.25 (0.264)	-0.05	3.88 (0.049)	-1.41	1.16 (0.281)	-0.01
Suicidality	3.44 [1]	1.30 [9]	2.08 [1]	1.35 [9]	2.08 [1]	1.39 [9]	9.48 (0.002)	1.36	0.09 (0.763)	-0.08	1.12 (0.290)	0.00	1.20 (0.274)	-0.04

Note: a = discrimination parameter estimate, b = severity parameter estimate.

Values for chi square tests in bold are statistically (p-value adjusted for multiple testing using the Benjamini-Hochberg procedure < .05) and clinically significant (method described in the article).

Sampling weights and design effects of the NESARC were taken into account.

DIFFTEST a = Comparing models with the factor loading freed and fixed for the corresponding item.

DIFFTEST b = Comparing models with the threshold respectively freed and fixed for the corresponding item.

Table 3

Differential item functioning (DIF) of DSM-IV symptoms of Major Depressive Episode (MDE) in women of childbearing age during pregnancy and outside the peripartum period in the NESARC sample.

| DSM-IV symptoms | Pregnancy (N = 311) | | Postpartum period (N = 609) | | Non-peripartum period (N = 8736) | | Pregnancy vs. Non-peripartum period | | | | Postpartum period vs. Non-peripartum period | | | |
| | | | | | | | DIFFTEST a | | DIFFTEST b | | DIFFTEST a | | DIFFTEST b | |
	a [rank]	b [rank]	a [rank]	b [rank]	a [rank]	b [rank]	X^2 (p-value)	DIFa	X^2 (p-value)	DIFb	X^2 (p-value)	DIFa	X^2 (p-value)	DIFb
Sadness	16.24 [8]	1.05 [2]	19.64 [9]	1.03 [1]	23.24 [9]	1.05 [1]	0.32 (0.573)	−7.00	2.04 (0.154)	0.00	3.87 (0.049)	−3.6	3.22 (0.073)	−0.02
Anhedonia	8.16 [5]	1.21 [6]	11.71 [6]	1.12 [5]	9.30 [5]	1.12 [5]	6.52 (0.011)	−1.14	1.62 (0.204)	0.09	4.70 (0.030)	2.41	0.11 (0.736)	0.00
Appetite disturbance	5.23 [2]	1.20 [5]	5.81 [4]	1.20 [7]	4.82 [4]	1.20 [6]	1.95 (0.162)	0.41	0.15 (0.696)	0.00	2.90 (0.089)	0.99	1.44 (0.231)	0.00
Sleep disturbance	16.18 [7]	1.05 [3]	13.29 [8]	1.08 [2]	18.09 [7]	1.07 [2]	0.28 (0.594)	−1.91	1.18 (0.278)	−0.02	1.17 (0.280)	−4.80	1.53 (0.216)	0.01
Fatigue	15.08 [6]	1.06 [4]	12.17 [7]	1.08 [3]	13.63 [6]	1.08 [4]	2.98 (0.085)	1.46	1.35 (0.246)	−0.02	1.35 (0.245)	−1.46	0.54 (0.464)	0.00
Psychomotor symptoms	5.35 [3]	1.27 [8]	3.87 [1]	1.38 [9]	4.20 [2]	1.27 [8]	5.30 (0.021)	1.15	2.59 (0.108)	0.00	0.23 (0.631)	−0.33	1.00 (0.317)	0.11
Worthlessness / guilt	7.70 [4]	1.23 [7]	5.50 [3]	1.16 [6]	4.42 [3]	1.24 [7]	3.73 (0.053)	3.28	2.21 (0.137)	−0.01	3.12 (0.077)	1.08	1.87 (0.171)	−0.08
Concentration	16.41 [9]	1.03 [1]	6.69 [5]	1.10 [4]	18.71 [8]	1.07 [3]	0.49 (0.486)	−2.31	0.21 (0.648)	−0.04	5.26 (0.022)	−12.02	2.82 (0.093)	0.03
Suicidality	2.96 [1]	1.51 [9]	4.18 [2]	1.36 [8]	3.04 [1]	1.46 [9]	1.06 (0.304)	−0.08	1.39 (0.238)	0.05	**13.92 (<.001)**	**1.14**	6.34 (0.012)	−0.10

Note: a = discrimination parameter estimate, b = severity parameter estimate.

Values for chi square tests in bold are statistically (p-value adjusted for multiple testing using the Benjamini-Hochberg procedure < .05) and clinically significant (method described in the article).

Sampling weights and design effects of the NESARC were taken into account.

DIFFTEST a = Comparing models with the factor loading freed and fixed for the corresponding item.

DIFFTEST b = Comparing models with the threshold respectively freed and fixed for the corresponding item.