Perinatal Depression

A Systematic Review of Prevalence and Incidence

Norma I. Gavin, PhD, Bradley N. Gaynes, MD, MPH, Kathleen N. Lohr, PhD, Samantha Meltzer-Brody, MD, MPH, Gerald Gartlehner, MD, MPH, and Tammeka Swinson

OBJECTIVE: We systematically review evidence on the prevalence and incidence of perinatal depression and compare these rates with those of depression in women at nonchildbearing times.

DATA SOURCES: We searched MEDLINE, CINAHL, PsycINFO, and Sociofile for English-language articles published from 1980 through March 2004, conducted hand searches of bibliographies, and consulted with experts.

METHODS OF STUDY SELECTION: We included cross-sectional, cohort, and case-control studies from developed countries that assessed women for depression during pregnancy or the first year postpartum with a structured clinical interview.

TABULATION, INTEGRATION, AND RESULTS: Of the 109 articles reviewed, 28 met our inclusion criteria. For major and minor depression (major depression alone), the combined point prevalence estimates from meta-analyses ranged from 6.5% to 12.9% (1.0–5.6%) at different trimesters of pregnancy and months in the first postpartum year. The combined period prevalence shows that as many as 19.2% (7.1%) of women have a depressive episode (major depressive episode) during the first 3 months postpartum; most of these episodes have onset following delivery. All estimates have wide 95% confidence intervals, showing significant uncertainty in their true levels. No conclusions could be made regarding the relative incidence of depression among pregnant and postpartum women compared with women at nonchildbearing times.

CONCLUSION: To better delineate periods of peak prevalence and incidence for perinatal depression and identify high risk subpopulations, we need studies with larger and more representative samples.

(Obstet Gynecol 2005;106:1071–83)

From RTI International, Research Triangle Park, North Carolina; Department of Psychiatry, University of North Carolina School of Medicine, Chapel Hill; University of North Carolina School of Public Health, Chapel Hill; and Cecil G. Sheps Center for Health Services Research, University of North Carolina School of Medicine, Chapel Hill, North Carolina.

This research was funded through a contract from the Agency for Healthcare Research and Quality and the Safe Motherhood Group to the RTI International-University of North Carolina Evidence-Based Practice Center (contract no. 290–02–0016). Dr. Gaynes was supported in part by a National Institute of Mental Health K23 Career Development Award (MH01951–03).

The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services of a particular drug, device, test, treatment, or other clinical service.

Corresponding author: Norma I. Gavin, PhD, RTI International, 3040 Cornwallis Road, P.O. Box 12194, Research Triangle Park, NC 27709-2194; e-mail: gavin@rti.org.

ISSN: 0029-7844/05

Perinatal depression encompasses major and minor depressive episodes that occur either during pregnancy or within the first 12 months after delivery. This disorder can have devastating consequences, not only for the women experiencing it but also for the woman's child and family. Perinatal depression has been associated with lower quality interactions between mothers and their children,[1] missed pediatric appointments and greater use of emergency department services,[2] higher levels of psychiatric disturbances among children,[3] and greater child insecurity in attachment relationships.[4]

Estimates of the prevalence of the condition vary widely–from 5% to more than 25% of pregnant women and new mothers.[5–7] To estimate disease burden more accurately and thereby better target and prioritize health care expenditures, we need more precise estimates of the prevalence of perinatal depression.

To address these gaps, the federal Agency for Healthcare Research and Quality (AHRQ) in association with the Safe Motherhood Group, a collaboration of 8 federal agencies and offices within the U.S. Department of Health and Human Services led by the Office on Women's Health, commissioned the RTI International–University of North Carolina (RTI-UNC) Evidence-based Practice Center (EPC) to systematically review evidence about the prevalence and incidence of perinatal depression. Our focus was on

major depression alone and in combination with minor depression during pregnancy and the first year postpartum, specifically among the general U.S. population of women of childbearing age. We also compared these rates with the prevalence of depression among women during nonchildbearing periods.

SOURCES

We used 3 strategies to identify studies: systematic searches of electronic databases, hand searches of reference lists of included articles, and consultations with a Technical Expert Panel. The Technical Expert Panel comprised 4 individuals, including a psychologist, a psychiatrist, and 2 obstetricians with both clinical and research experience in perinatal depression.

The electronic databases searched are MEDLINE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), PsycINFO, and Sociofile. We initially searched the databases in April 2003 for a feasibility study.[8] That study included 6 key questions on perinatal depression. Besides prevalence and incidence, the key questions addressed the natural history, risk factors, screening accuracy, screening effectiveness, and treatment effectiveness for the disease. We found relevant articles for the prevalence and incidence of perinatal depression under the literature searches for all these questions. We updated all searches for this review in March 2004.

We searched for English language articles only. The key MeSH terms were disease identifiers, including "puerperal disorders" with "depression" or "depressive disorder," or "postpartum depression," or "perinatal depression"; study type identifiers, including "natural history," "cohort studies," or "longitudinal studies"; and study topic identifiers, including "incidence," "prevalence," "mass screening," "treatment," "therapeutics," "treatment failure," "treatment outcomes," "treatment duration," "treatment errors," "treatment delay," or "treatment complications."

STUDY SELECTION

Inclusion Criteria

To identify relevant studies, we generated a priori a list of inclusion and exclusion criteria. Studies had to report on original data, be conducted in a developed country, and be published from January 1980 through March 2004. We excluded studies from less developed countries to ensure that the study results could be generalized to all or a major subset of pregnant and postpartum women in the United States. The time frame ensured that the applied reference standards were consistent with *Diagnostic and Statistical Manual of Mental Disorders, Third Edition* (DSM-III) or later criteria for the diagnosis of depression. In addition, studies had to assess women for major depression, either alone or together with minor depression, during pregnancy or the first year postpartum by means of a clinical assessment or structured clinical interview. We excluded studies of the prevalence and incidence of perinatal depression that relied solely on self-report screens to identify depression. We also excluded studies of women with major or minor depression in which the outcomes of interest were not distinguishable from those for women with bipolar disorder, primary psychotic disorders, or maternity blues (a mild mood disturbance experienced by approximately half of childbearing women within 3–6 days after delivery that resolves within a few hours to a few days).

We included cross-sectional, cohort, and case-control studies. We also included clinical studies that were conducted for purposes other than determining the prevalence and incidence of perinatal depression (eg, randomized clinical trials of treatment efficacy) but nevertheless estimated a population-based prevalence or incidence for perinatal depression that met the other inclusion criteria.

Data Assessment

Two clinicians with expertise in perinatal depression reviewed each abstract of the identified articles against the inclusion criteria and resolved any differences in inclusion by consensus. From each article meeting the inclusion criteria, we abstracted study features and all estimates of the prevalence and incidence of major and minor depression together and of major depression alone. The primary author (N.G.) read and abstracted each article; a second member checked the table entries for accuracy against the original article.

We grouped estimates by whether they were point prevalence, period prevalence, or incidence estimates. *Point prevalence* is the percentage of the population with depression at a given point in time (eg, at 24 weeks gestational age or 9 weeks postpartum); *period prevalence* is the percentage of the population with depression over a period of time (eg, during pregnancy or from delivery to the end of the first 3 months postpartum); and *incidence* is the percentage of the population with depressive episodes that begin within a given period of time.

We further subdivided the estimates by the reference time period. Time periods for point prevalence estimates were trimesters during pregnancy and months during the first postpartum year. Estimates for

different weeks of gestation within the same trimester of pregnancy were considered as being conducted in the same time period (eg, estimates for weeks 14 through 27 of gestation were considered the second trimester). Similarly, estimates for different weeks postpartum but within the same month postpartum were considered within the same time period (eg, estimates for weeks 1 through 4 postpartum were considered month 1; weeks 5 through 9 postpartum, month 2). Relevant time periods for period prevalence and incidence estimates were either single trimesters and months or multiple trimesters and months.

During abstraction, we graded the quality of the study based on selected study features.[9] We developed a quality rating form by modifying the Downs and Black[10] instrument for randomized controlled trials and observational studies. Our form rated the reporting completeness and clarity of the article, the external and internal validity of the study design, and the precision of the study prevalence and incidence estimates. The ratings refer to the usefulness or quality of the article for our purposes and not necessarily for the original purpose of the research or article. N.G. completed the quality rating form for each article; another project team member reviewed the completed form for accuracy and completeness.

Data Analysis

When we found 2 or more point prevalence, period prevalence, or incidence estimates with the same diagnosis, estimate type, and time period, we combined them in meta-analyses using Stata 8.2 software (StataCorp, College Station, TX). Because the inverse-variance weighting method assumes a normally distributed variable, we first transformed the prevalence estimates into log odds estimates. The procedure produces forest plots that show the individual study and combined estimates and their 95% confidence intervals (CIs).[11] It also produces a Q test of the homogeneity of the estimates.[11] This statistic tests the null hypothesis that the estimates come from the same distribution; a $P < .05$ suggests that they do not.

We reviewed the forest plots of the studies in each summary estimate to determine whether we could identify the source of any heterogeneity between studies. We then reran the meta-analyses excluding studies that were obvious outliers (ie, studies for which the graphed CIs consistently did not overlap those of estimates from the other studies on visual examination). We used random effects models to account for any remaining heterogeneity in the estimates across studies.

To analyze associations between the prevalence of depression and study characteristics, we conducted cumulative meta-analysis and a series of meta-regressions. In the cumulative meta-analysis, we added studies one by one, based on publication year, to produce a new combined estimate with the cumulative evidence for each year.[11] This procedure allowed us to see trends in the estimate over time. We conducted cumulative meta-analysis on the 2-month point prevalence estimates for both major and minor depression together and major depression alone.

We then conducted a series of meta-regressions to estimate the effect of several different population and study design factors on the point prevalence estimates.[12] In every model, we included 1) the time point at which depression was assessed, 2) whether the population was generally of low risk, 3) whether the population was restricted to those with low socioeconomic status, and 4) one of the following factors: publication year, study country, interview type, diagnostic criteria, whether depression was assessed only for women who were designated as at risk based on a screening instrument, and the quality rating score.

To investigate whether the prevalence and incidence of depression are higher during pregnancy and the first year postpartum than during nonchildbearing periods, we computed odds ratios (ORs) for studies with a comparison group of women of similar age during nonchildbearing times. Because the types and timing of prevalence and incidence estimates did not overlap in these studies, except for one time point, we did not conduct meta-analyses of the ORs.

RESULTS

We found a total of 837 unduplicated citations in the electronic searches and picked up an additional 9 citations through hand searches and discussion with the Technical Expert Panel. Of these 846 articles, 737 did not meet the inclusion criteria and were excluded. The remaining 109 articles were pulled for full review. Of these, 28 prospective studies met our inclusion criteria for analysis of the prevalence and incidence of perinatal depression. Only 3 of the 28 studies included a comparison group of women in nonchildbearing years.

Table 1 summarizes the major characteristics of the included studies by study type and alphabetically within type. The 25 studies without a comparison group are shown first,[13–37] followed by the 3 studies with a comparison group.[38–40] Study sample sizes ranged from 54 to 4,964 women (median, 202 women). Although all studies had an adequate sample

Table 1. Major Characteristics of Studies of the Prevalence and Incidence of Perinatal Depression

Author, Year	Country	Sample Size	Who Interviewed	When Interviewed	Interview Type	Diagnostic Criteria	Quality Rating
Studies without comparison groups							
Affonso et al, 1990[13]	US	202	All	Pregnancy & PP	SADS-PPG	RDC	8
Areias et al, 1996[14]	Portugal	54	All	Pregnancy & PP	SADS	RDC	12
Berle et al, 2003[15]	Norway	411	All EPDS ≥ 8 & some < 8	PP	MINI-V4.4/ MADRS	DSM-IV	9
Campbell and Cohn, 1991[16]	US	1,033	All	PP	SADS	RDC	12
Cooper et al, 1996[17]	England	4,964	EPDS ≥ 8	PP	SCID	DSM-III-R	12
Cox et al, 1982[18]	Scotland	105	All	PP	SPI	Pitt's	11
Garcia-Esteve et al, 2003[19]	Spain	1,123	All EPDS ≥ 9 & some < 9	PP	SCID-NP	DSM-IV	13
Gotlib et al, 1989[20]	Canada	295	All BDI ≥ 10 & some < 10	Pregnancy & PP	SADS	RDC	11
Hobfoll et al, 1995[21]	US	192	All	Pregnancy & PP	SADS	RDC	12
Kent et al, 1999[22]	Australia	710	GHQ28 > 4	PP	CIDI-A	DSM-III-R	12
Kitamura et al, 1993[23]	Japan	120	All	Pregnancy	SADS/ SADS-C	RDC	13
Kitamura et al, 1999[24]	Japan	111	All	Pregnancy & PP	SADS	RDC	10
Kumar and Robson, 1984[25]	England	196	All	Pregnancy & PP	SPI	RDC	11
Lee et al, 2001[26]	Hong Kong	781	All GHQ > 4 & some ≤ 4	PP	Modified SCID	Modified DSM-III-R	12
Lee et al, 2001[27]	Hong Kong	145	All	PP	Modified SCID	Modified DSM-III-R	8
Lucas et al, 2001[28]	Spain	641	BDI > 21	PP	Not specified	DSM-III-R	9
Matthey et al, 2003[29]	Australia	408	All	PP	DIS	DSM-IV	11
Murray and Cox, 1990[30]	England	100	All	Pregnancy	SPI	RDC	10
O'Hara et al, 1984[31]	US	99	All	Pregnancy & PP	SADS	RDC	10
Pop et al, 1993[32]	Netherlands	293	All	Pregnancy & PP	Not specified	RDC	13
Watson et al, 1984[33]	England	128	All	Pregnancy & PP	SPI	ICD-9	13
Whiffen, 1988[34]	Canada	115	All	PP	SADS	RDC	10
Yamashita et al, 2000[35]	Japan	88	All	PP	SADS	RDC	10
Yonkers et al, 2001[36]	US	802	All IDS ≥ 18 or EPDS ≥ 12 & some < 12	PP	SCID	DSM-IV	14
Yoshida et al, 1997[37]	England	98	All	PP	SADS	RDC	11
Studies with comparison groups							
Cooper et al, 1988[38]	England	483 cases 313 controls	All GHQ ≥ 12 & some < 12	PP	PSE/ MADRS	PSE ID/ Catego Class	10
Cox et al, 1993[39]	England	232 cases 232 controls	All EPDS ≥ 9 & some < 9	PP	SPI	RDC	12
O'Hara et al, 1990[40]	US	182 cases 179 controls	All	Pregnancy & PP	SADS	RDC	13

PP, postpartum; SADS-PPG, Schedule for Affective Disorders and Schizophrenia–Pregnancy and Postpartum Guidelines; RDC, Research Diagnostic Criteria; EPDS, Edinburgh Postnatal Depression Scale; MINI-V4.4, Mini International Neuropsychiatric Interview, Version 4.4; MADRS, Montgomery-Asburg Depression Rating Scale; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; SCID, Structured Clinical Interview for DSM-III-R; DSM-III-R, Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised; SPI, Standardized Psychiatric Interview; SCID-NP, Structured Clinical Interview for DSM-III-R, Non-Patient Version; BDI, Beck Depression Inventory; GHQ, General Health Questionnaire; CIDI-A, Composite International Diagnostic Interview; SADS-C, SADS Change Version; DIS, Diagnostic Inventory Schedule; ICD-9, International Classification of Diseases, Ninth Edition; IDS, Inventory of Depressive Symptomatology; PSE, Present State Examination; PSE ID, PSE Index of Definition.

size to provide a reliable prevalence estimate, most were not large enough to allow subgroup analyses.

Included studies represented a wide array of developed nations but generally had a limited racial/ethnic mix. Fourteen did not specify the racial and ethnic composition of the study subjects. Among the other 14 studies, 5 included only white non-Hispanic women,[16,22,28,31,32] 2 studies included only Chinese women,[26,27] and 2 others included only Japanese women.[35,37] The remaining 5 studies noted a racially mixed population, but all had a predominant race or ethnicity. In 4 of these studies, 73–90% of the women were white non-Hispanic,[13,20,21,33] and in the fifth, 75% were Hispanic.[36]

We rated included studies as generally good on reporting, poor on external validity, and fair on internal validity and precision. For prospective studies without comparison groups, the average overall quality score was 11.1 points (of a possible 20 points); for prospective studies with comparison groups, the average overall quality score was 11.7 points (of a possible 25 points).

Depression Assessment

The studies differed in who received a clinical interview, the interview instrument, the diagnostic criteria used to identify a depressive episode from the interview responses, and when the interview was conducted.

Of the 28 prospective studies, 18 conducted a clinical interview with all study women. The remaining 10 studies first had study subjects complete a self-report depression screening instrument, such as the Edinburgh Postnatal Depression Scale (EPDS), the Beck Depression Inventory (BDI), or the General Health Questionnaire (GHQ). These studies then administered a clinical interview to women scoring over a predetermined cutoff on the screening instrument. Seven of the 10 studies also interviewed a small sample (eg, 10%) of the women scoring below the cutoff, but few of the studies used the results from these interviews to adjust their final prevalence estimates for false negatives. Most studies used low enough cutoff scores that the resulting downward bias in the estimates was minimal.

The most frequently used instrument among our studies was the Schedule for Affective Disorders and Schizophrenia (SADS). This semistructured interview is widely used in clinical research and has well-established reliability and validity.[41] O'Hara et al[31] adapted the Schedule for Affective Disorders and Schizophrenia for use with pregnant and postpartum women. Twelve of the 28 prospective studies used this interview instrument. Five of the studies used the section of the Structured Clinical Interview for DSM Diagnoses (SCID) that covers depressive disorders.[42,43] Five other studies used the Standardized Psychiatric Interview (SPI) of Goldberg et al.[44] Other interview instruments used are the Composite International Diagnostic Interview (CIDI-A),[45] the Diagnostic Interview Schedule,[46] the Mini International Neuropsychiatric Interview (MINI-V4.4),[47] the Present State Examination (PSE),[48] and the Montgomery-Asberg Depression Rating Scale (MADRS).[49]

All studies that used the Schedule for Affective Disorders and Schizophrenia and 3 of the studies that used the Standardized Psychiatric Interview based depression diagnosis on the Research Diagnostic Criteria (RDC).[50] To be diagnosed with major or minor depression, women had to have reported that they felt sad, tearful, or blue for at least 2 weeks. The 2-week criterion rules out women who were experiencing postpartum blues only. In addition, for a diagnosis of major depression, the women had to have reported a depressed mood and at least 3 or 4 additional symptoms, such as sleeping disturbances, loss of appetite, fatigue, loss of interest in usual activities or of the ability to concentrate, psychomotor retardation, and suicidal thoughts. Depressed women with fewer symptoms were classified as having minor depression.

Five studies based diagnoses of depression on DSM-III-R criteria, and 4 based diagnoses on DSM-IV criteria. A diagnosis of major depression based on DSM-III-R criteria is comparable to one using the Research Diagnostic Criteria for definite major depression.[51] However, the Research Diagnostic Criteria includes criteria for minor depression, which received its first DSM mention in the fourth edition (DSM-IV)[52] as a proposed category for further study. Other criteria used for diagnoses of depression included Pitt's criteria,[53] the International Classification of Diseases, Ninth Edition (ICD-9), and Present State Examination Index of Definition (PSE ID) and Catego Class.[48]

Finally, most of the studies we reviewed administered the clinical interview at multiple points in time throughout pregnancy and the first postpartum year, allowing for multiple estimates of prevalence and incidence. The 28 prospective studies provided 80 estimates of the prevalence and incidence of major and minor depression and 70 estimates of the prevalence and incidence of major depression alone. Clinical assessments of depression were taken at different points in time throughout pregnancy and the first postpartum year.

Outliers

In a review of the forest plots from the meta-analyses of the prevalence and incidence estimates, we found estimates from several studies consistently to be outliers for all time periods at which they assessed the women's mood. Two studies included only women at low risk of depression.[13,16] Affonso et al[13] included only primigravida women with a viable fetus who were married or living with the infant's father and who had no recent depression episodes. Campbell and Cohn[16] included only primiparous women who delivered full-term, single infants without major complications and who were Caucasian, married, more than 17 years of age, and had at least a high school education. The estimates from these studies were consistently lower than the estimates from the other studies.

Two studies included only women of lower socioeconomic status.[21,36] They generally provided higher estimates of depression prevalence and incidence than the other studies. Lucas et al[28] included only women who screened positive for depression on the Beck Depression Inventory. The cutoff used (> 21) was so high that the bias from false negatives produced consistently lower prevalence estimates than did other studies.

Finally, because of its size, the Cooper et al[17] study dominated the combined 2-month point prevalence estimate for major depression alone. However, the 15.3% estimated point prevalence from this study is outside the 95% confidence interval of the combined estimate for major and minor depression. The purpose of the study was not to produce a prevalence estimate but rather to develop a predictive index for postpartum depression. Furthermore, many of the clinical interviews were conducted by telephone, and the article did not state whether a clinician or lay person conducted the interview. Thus, the procedures for assessing depression in this study may have introduced significant bias in the prevalence estimate.

We reran the meta-analyses excluding these 6 studies to produce "best estimates" of the prevalence of perinatal depression. The final best estimates are shown in Appendix Table 1 for major and minor depression together and in Appendix Table 2 for major depression alone.

Point Prevalence

We show the best estimates for the point prevalence of major and minor depression graphically in Figure 1. It gives the mean estimate and corresponding 95% CI for each trimester of pregnancy and month postpartum in the first year after delivery. The number of studies that we used to compute the estimate and the *P* value for the Q test of homogeneity among the studies are shown above each estimate. For points in time for which no numbers are shown, we found only a single estimate.

As shown in Figure 1, prevalence in the first trimester is 11.0% but drops to 8.5% in the second and third trimesters. After delivery, prevalence of major and minor depression begins to rise and is highest in the third month at 12.9%. In the fourth through seventh months postpartum, prevalence declines slightly, staying in the range of 9.9–10.6%, after which it declines to 6.5%. However, all these estimates have broad 95% CIs.

The best estimates for the point prevalence of major depression alone (Fig. 2) range from 1.0 to 5.6 at different times in the perinatal period and are no more precise than those for major and minor depression together. Episodes of major depression constitute less than half of all cases of depression in the perinatal period, except during 3 seemingly peak times–in the second trimester (4.9%), at 2 months postpartum (5.7%), and 6 months postpartum (5.6%). However, the 95% CIs for these estimates are very wide and overlap those at other times (Fig. 2). Furthermore, the tests for homogeneity show that considerable heterogeneity persists among studies in the combined estimates.

Period Prevalence

We found many fewer estimates of period prevalence. The best estimates suggest that as many as 18.4% of pregnant women are depressed during their pregnancy (ie, from conception to birth), with as many as 12.7% having an episode of major depression. Furthermore, as many as 19.2% of new mothers may have major or minor depression in the first 3 months after delivery, with as many as 7.1% having major depression.

However, all estimates have wide 95% CIs (Appendix Tables 1 and 2). Moreover, the best estimates of different durations are not consistent over longer periods of time. We would expect the period prevalence for major and minor depression from birth to 2 months postpartum to be higher than the period prevalence from birth to 1 month postpartum, and the period prevalence for major depression from birth to 3 months postpartum to be higher than the period prevalence from birth to 2 months postpartum, but we do not see these patterns.

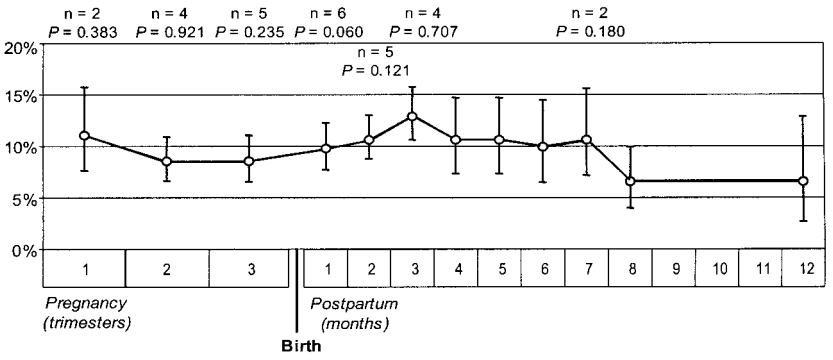

Fig. 1. Point prevalence of major and minor depression. For times with an estimate from a single study, no n or *P* value is shown. n = number of studies on which the combined estimate is based. *P* = the *P* value for the Q test of homogeneity.

Gavin. Perinatal Depression Prevalence and Incidence. Obstet Gynecol 2005.

Incidence

We also found few estimates of the incidence of depression–the percentage of women with depressive episodes that begin during pregnancy or the first year postpartum (Appendix Tables 1 and 2). The studies we found suggest that as many as 14.5% of pregnant women have a new episode of major or minor depression during pregnancy, and 14.5% have a new episode during the first 3 months postpartum. Considering major depression alone, 7.5% of women may have a new episode during pregnancy and 6.5% during the first 3 months after delivery. Although the incidence estimates for major and minor depression in the first 3 months postpartum follow the expected upward trend, the incidence estimates of major depression alone do not.

Analysis of Confounders

The results of the cumulative meta-analysis (Fig. 3) clearly show the impact of the more precise diagnostic criteria in more recent studies. For both major and minor depression together (Panel A) and major depression alone (Panel B), the cumulative combined 2-month point prevalence estimate drifts downward as more recent studies are added. Thus, the more precise criteria in the more recent studies identify fewer women as depressed. However, we did not find a statistically significant effect of the year of publication in our meta-regression ($P = .17$ for major and minor and $P = .60$ for major alone).

The meta-regression results for major and minor depression show large, positive coefficients for the 2-month postpartum and 3-month postpartum time periods compared with the 4- to 12-month postpartum period. These findings suggest a higher prevalence of depression during these 2 months. However, both coefficients are significant only in the equation that includes diagnostic criteria ($P = .024$). The coefficient for the 2-month postpartum time period is also large and positive for major depression alone, but significant only in the equations including diagnostic criteria ($P = .026$) and whether only women who screened positive for depression were interviewed ($P = .028$). None of the coefficients for the trimesters of pregnancy is statistically significant, suggesting that the prevalence of depression during pregnancy is similar to that during the last 3 quarters of the first postpartum year.

The low-risk indicator has a statistically significant, negative coefficient for both sets of diagnoses, as expected ($P = .000$ to $.036$). Low socioeconomic status has a statistically significant, positive coefficient only for major and minor depression together ($P = .000$ to $.003$).

The meta-regression results also suggest that

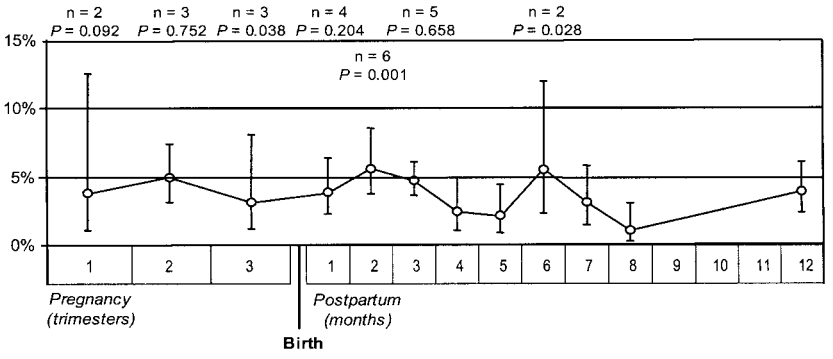

Fig. 2. Point prevalence of major depression. For times with an estimate from a single study, no n or *P* value is shown. n = number of studies on which the combined estimate is based. *P* = the *P* value for the Q test of homogeneity.

Gavin. Perinatal Depression Prevalence and Incidence. Obstet Gynecol 2005.

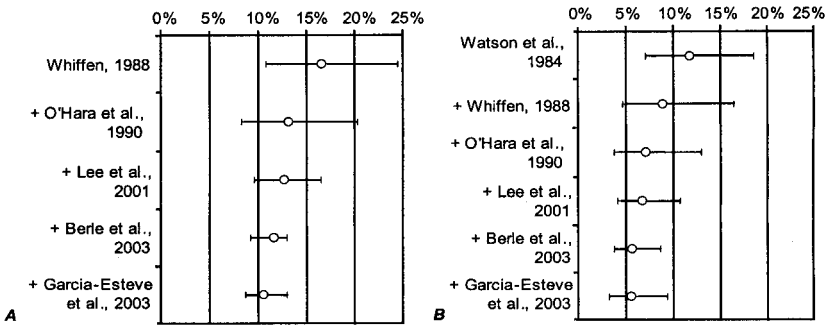

Fig. 3. Cumulative meta-analysis for point prevalence of depression at 2 months postpartum. **A.** Major and minor depression. **B.** Major depression. The prevalence estimates and their 95% confidence intervals were computed by sequentially adding in the studies from the top of the graph (oldest studies) to the bottom (most recent studies). For example, the Garcia-Esteve et al 2003 line for major and minor depression shows the estimate computed by combining the Whiffen 1988, O'Hara et al 1990, Lee et al 2001, Berle et al 2003, and Garcia-Esteve et al 2003 estimates.

Gavin. Perinatal Depression Prevalence and Incidence. Obstet Gynecol 2005.

prevalence can vary by the clinical instrument and diagnostic criteria used to assess depression. The Structured Clinical Interview for DSM Diagnoses instrument defined fewer women with major and minor depression than did the Schedule for Affective Disorders and Schizophrenia interview ($P = .015$), but the coefficient for this variable is not significant in the equation for major depression alone ($P = .736$). The DSM-IV and other diagnostic criteria (eg, Pitt, ICD-9) defined fewer women as depressed than did the Research Diagnostic Criteria in the equation for major and minor depression ($P = .045$ and $P = .000$, respectively), and DSM-III-R and other criteria defined significantly more women as suffering from major depression than did the Research Diagnostic Criteria ($P = .001$).

Finally, studies with higher quality rating scores have lower log odds, but the coefficient of this variable is only marginally significant ($P = .072$) in the equation for major depression alone and is not significant ($P = .783$) in the equation for major and minor depression together. No statistically significant results were found for study country or whether the study interviewed only women who screened positive for depression, although the signs of the coefficients for these variables are as predicted.

Comparison with Other Women

The 3 prospective studies with comparison groups of women of similar age in nonchildbearing periods had adequate data to compute 13 estimates of the relative prevalence and incidence of depression. The estimated ORs and corresponding 95% CIs are shown in Table 2. None of the ORs for prevalence, which covered different time periods in the first postpartum year, indicated a statistically significant difference.

The single study of the incidence of major and minor depression (Table 2) shows a significant 3-fold difference in the odds of having a new episode of major or minor depression among women in their first 5 weeks postpartum compared with women who were not pregnant and had not recently given birth.[39] However, by 6 months postpartum, the difference in the incidence had narrowed and was no longer significant (Table 2).

CONCLUSIONS

The available research suggests that depression is one of the more common complications of the prenatal and postpartum periods. However, a considerable amount of uncertainty remains in the level of its prevalence and incidence. After a careful review of the evidence, we cannot say with certainty that perinatal depression is higher at any particular trimester during pregnancy or month in the first postpartum year, although some evidence points to the second and third months postpartum having slightly higher prevalence.

Furthermore, the available evidence from case-control studies of perinatal women and women of similar age at nonchildbearing times does not support the hypothesis that the prevalence of depression is higher during pregnancy or in the first year postpartum compared with nonchildbearing times. These findings are consistent with a comparison of our best combined estimates and those from a national survey of women in childbearing years. The National Comorbidity Survey, fielded in 1990–1992, found that the current 30-day prevalence of major depressive episodes was 5.9% among women aged 15–54 years using the Composite International Diagnostic Interview instrument and DSM-III-R criteria.[54] This find-

Table 2. Odds Ratios for Studies With Comparison Groups of Women During Nonchildbearing Periods

Diagnosis Estimate Type Author, Year	Time Period	Odds Ratio	95% Confidence Interval
Major and minor depression			
Point prevalence			
O'Hara et al, 1990	2nd trimester	1.41	0.61–3.26
O'Hara et al, 1990	9 weeks PP	1.37	0.67–2.83
Cox et al, 1993	6 months PP	1.00	0.54–1.84
Period prevalence			
Cox et al, 1993	Birth to 6 months PP	1.04	0.61–1.76
Incidence			
Cox et al, 1993	Birth to 5 weeks PP	3.26*	1.17–9.06
Cox et al, 1993	Birth to 6 months PP	1.48	0.77–2.82
Major depression			
Point prevalence			
O'Hara et al, 1990	2nd trimester	1.28	0.47–3.51
O'Hara et al, 1990	9 weeks PP	1.33	0.45–3.90
Cooper et al, 1988	3 months PP	0.85	0.33–2.17
Cox et al, 1993	6 months PP	1.00	0.37–2.71
Cooper et al, 1990	6 months PP	1.53	0.65–3.58
Cooper et al, 1990	12 months PP	0.50	0.17–1.46
Period			
Cox et al, 1993	6 months PP	1.16	0.54–2.51

PP, postpartum.
* Statistically significant at $P < .05$.

ing is approximately equivalent to our best 1-month postpartum period prevalence of 5.7% and to the point prevalence at 2 months and 6 months postpartum (5.7% and 5.6%, respectively).

Our estimates of prevalence and incidence were somewhat lower than those found in prior systematic reviews.[5,7] This arises because, in contrast to these reviews, we excluded studies that assessed depression based on self-report screens alone, which tend to overestimate prevalence. Because we required diagnostic confirmation, our estimates are based on only the highest quality evidence. Furthermore, we separated out estimates of major and minor depression from estimates of major depression alone and estimates of point prevalence from estimates of period prevalence, which was not done in these prior reviews, and we included recently published studies that use more precise criteria to identify major depression.

We found in meta-regressions that population characteristics, such as socioeconomic status, and study design characteristics, such as the clinical instrument used to assess depression, can influence the prevalence and incidence estimates. Our results suggest that the prevalence of major depression is similar among socioeconomic status groups but that minor depression may be more prevalent among lower socioeconomic status groups. Furthermore, the different clinical instruments appear to have a different sensitivity for identifying women with minor depression. Studies that we rated as better designed studies tended to have somewhat lower prevalence of depression.

The inconsistent trends found in the period prevalence and incidence estimates and the wide CIs in our best estimates suggest that our results should be interpreted with caution. Furthermore, although we show estimates for sequential time periods, each time period includes a different set of studies. Therefore, our estimates should not be interpreted as describing the actual course of illness.

We conclude that the level of research on perinatal depression warrants both improvement and expansion. The studies on the prevalence and incidence of the disorder were generally of moderate size–too small for reliable subgroup analyses. Furthermore, the study populations were typically restricted to a local community or geographic region served by one provider or a small number of providers of obstetric services and were not representative of the racial and ethnic mix of the countries in which the studies were conducted. Thus, larger studies are needed that better account for the racial and ethnic mix of the U.S. population of pregnant women and new mothers and that better delineate periods of peak prevalence and incidence. Furthermore, researchers need to clarify whether the incidence of perinatal depression is greater than the incidence of depression

in nonchildbearing women of similar ages. A better understanding of differences across population groups and time periods would help clinicians target screening and treatment programs.

Pregnancy and the early postpartum period provide opportunities through regular prenatal and postpartum physician contacts to screen for depression. Because the poor outcomes of suffering from depression during the perinatal period can be farther reaching–affecting not only the woman but her newborn child and other family members–it behooves us to better understand the epidemiology of perinatal depression, as well as the efficacy of screening and treatment programs for these women.

REFERENCES

1. Stein A, Gath DH, Bucher J, Bond A, Day A, Cooper PJ. The relationship between post-natal depression and mother-child interaction. Br J Psychiatry 1991;158:46–52.
2. Flynn HA, Davis M, Marcus SM, Cunningham R, Blow FC. Rates of maternal depression in pediatric emergency department and relationship to child service utilization. Gen Hosp Psychiatry 2004;26:316–22.
3. Murray L, Stein A. The effects of postnatal depression on the infant. Baillieres Clin Obstet Gynaecol 1989;3:921–33.
4. Marmorstein NR, Malone SM, Iacono WG. Psychiatric disorders among offspring of depressed mothers: associations with paternal psychopathology. Am J Psychiatry 2004;161:1588–94.
5. O'Hara MW, Swain AM. Rates and risk of postpartum depression: a meta-analysis. Int Rev Psychiatry 1996;8:37–54.
6. Llewellyn AM, Stowe ZN, Nemeroff CB. Depression during pregnancy and the puerperium. J Clin Psychiatry 1997;58 suppl 15:26–32.
7. Bennett HA, Einarson A, Taddio A, Koren G, Einarson TR. Prevalence of depression during pregnancy: systematic review. Obstet Gynecol 2004;103:698–709.
8. Gaynes B, Gavin N, Meltzer-Brody S, Sleath B, SuttonS. Perinatal depression: feasibility study. Final Report from the RTI-International - University of North Carolina Evidence-Based Practice Center to the Agency for Healthcare Quality and Research (AHRQ) under Contract No. 290-02-0016. Research Triangle Park (NC): RTI-International; 2003.
9. Gaynes BN, Gavin N, Meltzer-Brody S, Lohr KN, Swinson T, Gartlehner G, et al. Perinatal depression: prevalence, screening accuracy, and screening outcomes. Evidence Report/ Technology Assessment No. 119 (Prepared by RTI-University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016). AHRQ Publication 05-E006-1. Rockville (MD): Agency for Healthcare Research and Quality; 2005. Available at: http://www.ahrq.gov/clinic/epcsums/peridepsum.pdf. Retrieved August 31, 2005.
10. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998;52:377–84.
11. Sterne JAC, Bradburn MJ, Egger M. Meta-analysis of Stata. In: Egger, M, Smith G, Altman D, eds. Systematic reviews in health care: meta-analysis in context. 2nd ed. Boston (MA): Blackwell BMJ Books; 2001. p. 347–72.
12. Morton SC, Adams JL, Suttorp MJ, Shekelle PG. Meta-regression approaches: what, why, when, and how? Technical Review 8 (Prepared by Southern California-RAND Evidence-based Practice Center, under contract no. 290-97-0001). AHRQ publication no. 04-0033. Rockville (MD): Agency for Healthcare Research and Quality; 2004.
13. Affonso DD, Lovett S, Paul SM, Sheptak S. A standardized interview that differentiates pregnancy and postpartum symptoms from perinatal clinical depression. Birth 1990;17:121–30.
14. Areias ME, Kumar R, Barros H, Figueiredo E. Comparative incidence of depression in women and men, during pregnancy and after childbirth: validation of the Edinburgh Postnatal Depression Scale in Portuguese mothers. Br J Psychiatry 1996;169:30–5.
15. Berle J, Aarre T, Mykletun A, Dahl A, Holsten F. Screening for postnatal depression. Validation of the Norwegian version of the Edinburgh Postnatal Depression Scale, and assessment of risk factors for postnatal depression. J Affect Disord 2003;76: 151–6.
16. Campbell SB, Cohn JF. Prevalence and correlates of postpartum depression in first-time mothers. J Abnorm Psychol 1991; 100:594–9.
17. Cooper PJ, Murray L, Hooper R, West A. The development and validation of a predictive index for postpartum depression. Psychol Med 1996;26:627–34.
18. Cox JL, Connor Y, Kendell RE. Prospective study of the psychiatric disorders of childbirth. Br J Psychiatry 1982;140: 111–7.
19. Garcia-Esteve L, Ascaso C, Ojuel J, Navarro P. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in Spanish mothers. J Affect Disord 2003;75:71–6.
20. Gotlib IH, Whiffen VE, Mount JH, Milne K, Cordy NI. Prevalence rates and demographic characteristics associated with depression in pregnancy and the postpartum. J Consult Clin Psychol 1989;57:269–74.
21. Hobfoll SE, Ritter C, Lavin J, Hulsizer MR, Cameron RP. Depression prevalence and incidence among inner-city pregnant and postpartum women. J Consult Clin Psychol 1995;63: 445–53.
22. Kent GN, Stuckey BG, Allen JR, Lambert T, Gee V. Postpartum thyroid dysfunction: clinical assessment and relationship to psychiatric affective morbidity. Clin Endocrinol 1999;51: 429–38.
23. Kitamura T, Shima S, Sugawara M, Toda MA. Psychological and social correlates of the onset of affective disorders among pregnant women. Psychol Med 1993;23:967–75.
24. Kitamura T, Sugawara M, Shima S, Toda MA. Temporal variation of validity of self-rating questionnaires: improved validity of repeated use of Zung's Self-Rating Depression Scale among women during the perinatal period. J Psychosom Obstet Gynecol 1999;20:112–7.
25. Kumar R, Robson KM. A prospective study of emotional disorders in childbearing women. Br J Psychiatry 1984;144: 35–47.
26. Lee D, Yip A, Chiu H, Leung T, Chung T. A psychiatric epidemiological study of postpartum Chinese women. Am J Psychiatry 2001;158:220–6.
27. Lee D, Yip A, Chiu H, Leung T, Chung T. Screening for postnatal depression: are specific instruments mandatory? J Affect Disord 2001;63:233–8.
28. Lucas A, Pizarro E, Granada ML, Salinas I, Sanmarti A. Postpartum thyroid dysfunction and postpartum depression: are they two linked disorders? Clin Endocrinol 2001;55:809–14.
29. Matthey S, Barnett B, Howie P, Kavanagh DJ. Diagnosing postpartum depression in mothers and fathers: whatever happened to anxiety? J Affect Disord 2003;74:139–47.

30. Murray D, Cox JL. Screening for depression during pregnancy with the Edinburgh Depression Scale (EPDS). J Reprod Infant Psychol 1990;8:99–107.
31. O'Hara MW, Neunaber DJ, Zekoski EM. Prospective study of postpartum depression: prevalence, course, and predictive factors. J Abnorm Psychol 1984;93:158–71.
32. Pop VJ, Essed GG, de Geus CA, van Son MM, Komproe IH. Prevalence of post partum depression–or is it post-puerperium depression? Acta Obstet Gynecol Scand 1993;72:354–8.
33. Watson JP, Elliott SA, Rugg AJ, Brough DI. Psychiatric disorder in pregnancy and the first postnatal year. Br J Psychiatry 1984;144:453–62.
34. Whiffen V. Vulnerability of postpartum depression: a prospective multivariate study. J Abnorm Psychol 1988;97:467–74.
35. Yamashita H, Yoshida K, Nakano H, Tashiro N. Postnatal depression in Japanese women: detecting the early onset of postnatal depression by closely monitoring the postpartum mood. J Affect Disord 2000;58:145–54.
36. Yonkers KA, Ramin SM, Rush AJ, Navarrete CA, Carmody T, March D, et al. Onset and persistence of postpartum depression in an inner-city maternal health clinic system. Am J Psychiatry 2001;158:1856–63.
37. Yoshida K, Marks M, Kibe N, Kumar R, Nakano H, Tashiro N. Postnatal depression in Japanese women who have given birth in England. J Affect Disord 1997;43:69–77.
38. Cooper PJ, Campbell EA, Day A, Kennerley H, Bond A. Non-psychotic psychiatric disorder after childbirth: a prospective study of prevalence, incidence, course and nature. Br J Psychiatry 1988;152:799–806.
39. Cox JL, Murray D, Chapman G. A controlled study of the onset, duration and prevalence of postnatal depression. Br J Psychiatry 1993;163:27–31.
40. 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.
41. Endicott J, Spitzer RL. A diagnostic interview: the schedule for affective disorders and schizophrenia. Arch Gen Psychiatry 1978;35:837–44.
42. Spitzer RL, Williams JBW, Gibbon M, First MB. Structured clinical interview for DSM-III-R. Washington, DC: American Psychiatric Press; 1990.
43. Frist MB, Spitzer RL, Gibbon M, Williams JBW. Structured clinical interview for DSM-IV Axis I Disorders (SCID): Clinician Version. Washington, DC: American Psychiatric Press; 1996.
44. Goldberg DP, Cooper B, Eastwood MR, Kedward HB, Shepherd M. A standardized psychiatric interview for use in community surveys. Br J Prev Soc Med 1970;24:18–23.
45. Janca A, Ustun TB, Sartorius N. New versions of World Health Organization instruments for the assessment of mental disorders. Acta Psychiatr Scand 1994;90:73–83.
46. Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. Arch Gen Psychiatry 1981;38:381–9.
47. Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Sheehan KH, et al. The Mini International Neuropsychiatric Interview (M.I.N.I.): a short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatry 1997;12:224–31.
48. Wing JK, Cooper JE, Sartorius N. The measurement and classification of psychiatric symptoms. Cambridge (UK): Cambridge University Press; 1974.
49. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979;134:382–9.
50. Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. Arch Gen Psychiatry 1978;35:773–82.
51. American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders. 3rd ed. Washington, DC: American Psychiatric Association; 1987.
52. American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders. 4th ed. Washington, DC: American Psychiatric Association; 1994.
53. Pitt B. "Atypical" depression following childbirth. Br J Psychiatry 1968;114:1325–35.
54. Blazer DG, Kessler RC, McGonagle KA, Swartz MS. The prevalence and distribution of major depression in a national community sample: The National Comorbidity Survey. Am J Psychiatry 1994;151:979–86.

Appendix Table 1. Best Estimates of the Prevalence and Incidence of Major and Minor Depression

Start Date	End Date	Studies	Estimate (%)	95% Confidence Interval (%)	*P* (Test of Homogeneity)
Point prevalence					
	1st trimester	(24, 25)	11.0	7.6–15.8	.383
	2nd trimester	(20, 25, 31, 40)	8.5	6.6–10.9	.921
	3rd trimester	(20, 24, 25, 30, 32)	8.5	6.5–11.0	.235
	1 week PP	(24)	5.5	1.8–12.4	
	1 month PP	(20, 24, 26, 32, 35, 36)	9.7	7.7–12.3	.060
	2 months PP	(15, 19, 27, 34, 40)	10.6	8.7–13.0	.121
	3 months PP	(25, 26, 32, 35)	12.9	10.6–15.8	.707
	4 months PP	(32)	10.6	7.3–14.7	
	5 months PP	(32)	10.6	7.3–14.7	
	6 months PP	(39)	9.9	6.4–14.5	
	7 months PP	(25, 32)	10.6	7.1–15.6	.180
	8 months PP	(32)	6.5	4.0–9.9	
	12 months PP	(25)	6.5	2.7–12.9	
Period prevalence					
Conception	2nd trimester	(14)	9.3	3.1–20.3	
Conception	Birth	(14, 23, 25)	18.4	14.3–23.3	.931
2nd trimester	3rd trimester	(20)	10.2	7.0–14.2	
Birth	1 month PP	(35)	13.6	7.3–22.6	
Birth	2 months PP	(29, 31)	9.6	8.0–11.4	.362
Birth	3 months PP	(14, 35, 37)	19.2	10.7–31.9	.016
Birth	5 months PP	(18)	29.1	20.6–38.9	
Birth	6 months PP	(39)	13.8	9.6–18.9	
Birth	8 months PP	(32)	20.8	16.3–25.9	
Birth	12 months PP	(14)	53.7	39.6–67.4	
Incidence					
Conception	1st trimester	(23, 25)	11.3	7.8–16.3	.757
Conception	2nd trimester	(14)	5.8	1.2–16.0	
Conception	Birth	(14, 23)	14.5	8.1–24.4	.192
1st trimester	2nd trimester	(25)	2.7	0.6–7.6	
2nd trimester	3rd trimester	(20, 25)	2.2	1.1–4.1	.627
Birth	1 month PP	(20, 26, 35)	7.8	3.6–16.1	.003
Birth	2 months PP	(31)	10.3	5.1–18.1	
Birth	3 months PP	(14, 25, 26, 35, 37)	14.5	10.9–19.2	.142
Birth	6 months PP	(39)	11.1	7.3–16.0	
Birth	12 months PP	(14)	49.0	34.4–63.7	

PP, postpartum.
Best estimates reflect the single or combined estimate at each point or period of time remaining after estimates with obvious, identifiable biases have been dropped.

Appendix Table 2. Best Estimates of the Prevalence and Incidence of Major Depression

Start Date	End Date	Studies	Estimate (%)	95% Confidence Interval (%)	*P* (Test of Homogeneity)
Point prevalence					
	1st trimester	(24, 25)	3.8	1.0–12.6	.092
	2nd trimester	(31, 33, 40)	4.9	3.1–7.4	.752
	3rd trimester	(24, 30, 32)	3.1	1.1–8.1	.038
	1 week PP	(24)	0.0	0.0–3.2	
	1 month PP	(24, 26, 32, 35)	3.8	2.2–6.4	.204
	2 months PP	(15, 19, 27, 33, 34, 40)	5.7	3.8–8.7	.000
	3 months PP	(25, 26, 32, 35, 38)	4.7	3.6–6.1	.658
	4 months PP	(32)	2.4	1.0–4.9	
	5 months PP	(32)	2.1	0.8–4.4	
	6 months PP	(38, 39)	5.6	2.4–12.1	.028
	7 months PP	(32)	3.1	1.4–5.8	
	8 months PP	(32)	1.0	0.2–3.0	
	12 months PP	(38)	3.9	2.3–6.1	
Period prevalence					
Conception	Birth	(23)	12.7	7.1–20.4	
1st trimester	Birth	(33)	9.4	4.9–15.8	
Birth	1 month PP	(35)	5.7	1.9–12.8	
Birth	2 months PP	(31)	8.1	3.6–15.3	
Birth	3 months PP	(35, 37)	7.1	4.1–11.7	.626
Birth	5 months PP	(18)	12.6	6.9–20.6	
Birth	6 months PP	(39)	6.5	3.7–10.4	
Birth	8 months PP	(32)	6.8	4.2–10.4	
Birth	12 months PP	(33)	21.9	15.1–30.0	
Incidence					
Conception	Birth	(14, 23, 33)	7.5	3.8–14.2	.116
Birth	1 month PP	(26, 35)	5.2	3.1–8.9	.819
Birth	2 months PP	(33)	8.1	4.0–14.4	
Birth	3 months PP	(26, 35, 37)	6.5	4.2–9.6	.767
Birth	12 months PP	(14)	30.6	18.3–45.4	

PP, postpartum.
Best estimates reflect the single or combined estimate at each point or period of time remaining after estimates with obvious, identifiable biases have been dropped.