Risk factors of postpartum depression and depressive symptoms: umbrella review of current evidence from systematic reviews and meta-analyses of observational studies

Chiara Gastaldon, Marco Solmi, Christoph U. Correll, Corrado Barbui and Georgios Schoretsanitis

Background
Evidence on risk factors for postpartum depression (PPD) are fragmented and inconsistent.

Aims
To assess the strength and credibility of evidence on risk factors of PPD, ranking them based on the umbrella review methodology.

Method
Databases were searched until 1 December 2020, for systematic reviews and meta-analyses of observational studies. Two reviewers assessed quality, credibility of associations according to umbrella review criteria (URC) and evidence certainty according to Grading of Recommendations-Assessment-Development-Evaluations criteria.

Results
Including 185 observational studies (n = 3,272,093) from 11 systematic reviews, the association between premenstrual syndrome and PPD was the strongest (highly suggestive: odds ratio 2.20, 95%CI 1.81–2.68), followed by violent experiences (highly suggestive: odds ratio (OR) = 2.07, 95%CI 1.70–2.50) and unintended pregnancy (highly suggestive: OR=1.53, 95%CI 1.35–1.75). Following URC, the association was suggestive for Caesarean section (OR = 1.29, 95%CI 1.17–1.43), gestational diabetes (OR = 1.60, 95%CI 1.25–2.06) and 5-HTTLPR polymorphism (OR = 0.70, 95%CI 0.57–0.86); and weak for preterm delivery (OR = 2.12, 95%CI 1.43–3.14), anaemia during pregnancy (OR = 1.47, 95%CI 1.17–1.84), vitamin D deficiency (OR = 3.67, 95%CI 1.72–7.85) and postpartum anaemia (OR = 1.75, 95%CI 1.18–2.60). No significant associations were found for medically assisted conception and intra-labour epidural analgesia. No association was rated as ‘convincing evidence’. According to GRADE, the certainty of the evidence was low for Caesarean section, preterm delivery, 5-HTTLPR polymorphism and anaemia during pregnancy, and ‘very low’ for remaining factors.

Conclusions
The most robust risk factors of PDD were premenstrual syndrome, violent experiences and unintended pregnancy. These results should be integrated in clinical algorithms to assess the risk of PPD.

Keywords
Depressive disorders; perinatal psychiatry; umbrella review; risk factors; postpartum depression.

Copyright and usage
© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Postpartum depression (PPD) is a disorder characterised by symptoms of major depression occurring after delivery. The DSM-5 uses a peripartum specifier for affective symptoms occurring within 4 weeks after childbirth, but longer time frames of up to 1 year postpartum are also used in clinical and research settings. This differentiation in terms of onset timing may account for the considerable phenomenological heterogeneity of PPD, which is increasingly gaining attention. Additionally prevalence rates of PPD vary worldwide with estimates ranging between 10 and 25%. These rates suggest that an epidemiologically relevant proportion of women in the peripartum phase suffer from PPD. Therefore, PPD introduces globally a major public health problem with adverse sequelae for maternal and offspring well-being. Specifically, PPD has been associated with maternal and familiar distress, suicidal risk, impaired child development and behaviour outcomes. PPD has been associated with maternal and familiar distress, suicidal risk, impaired child development and behaviour outcomes. A substantial body of literature has contributed data on various types of PPD risk factors mainly using case-control and cohort design studies, which are the most appropriate type of study to answer such epidemiological questions. History of affective disorders and life stress are among the early identified major risk factors for PPD. Moreover, consistent data suggests higher risk of PPD in women having experienced violence. Different types of violence, such as child and adult physical, sexual and psychological abuse, have been strongly associated with PPD symptoms. Further, lower socioeconomic status and social resources have been also investigated in the context of PPD, implying a clear role for lack of social support when facing obstetrical complications. The effects of complications during pregnancy and delivery have been addressed in an interdisciplinary fashion. Specifically, the effects of conditions such as gestational diabetes (GDM), preeclampsia, vitamin D deficiency and anaemia, as well as exercise and dietary habits, on the prevalence of affective symptoms in the peripartum have been investigated. Parallel to complications during pregnancy, aggravating or protective factors related to delivery, such as preterm birth, use of labour analgesia and Caesarean section, have also been assessed. Although data are increasingly available, methodological controversies introduce limited generalisability, and hamper a global understanding of risk factors and a pragmatic risk quantification. Moreover, as quality indicators directly affect the certainty of risk estimates,
formal syntheses of the credibility of the exponentially emerging fragmented meta-analyses are required. This umbrella review aimed to identify, quantify and measure the degree of credibility of the association of PPD with different risk factors, including peripheral markers, obstetric complications and psychological factors.

Method

We used the Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines to guide this umbrella review (Supplementary Table 1, available at https://doi.org/10.1192/bjp.2021.222).29

The review protocol was registered with International Prospective Register of Systematic Reviews (PROSPERO; registration number CRD42020168468).

Study design

We summarised the evidence from multiple research syntheses, performing an umbrella review,26 which is a form of review of previously conducted systematic reviews and meta-analyses. It consists in repeating the meta-analyses adopting a uniform approach for available factors to enable their comparison.30 Considering the growing number of systematic reviews and meta-analyses available, this approach allows us to compare and contrast the findings of separate reviews related to a topic of interest.31 The conduct of an umbrella review also provides a comprehensive overview of healthcare areas, to highlight whether the evidence is consistent or contradictory, and to explore possible heterogeneity sources for existing evidence.31,32 This type of review is considered among the highest level of evidence,33 and is particularly useful to inform clinical practice and policies.34 Further details are described in the Supplementary Methods.

Two reviewers (C.G. and G.S.) independently conducted the literature search; the screening of the titles, abstracts and full-text papers; and the data extraction. Details of the process are described in the Supplementary Methods.

Eligible reviews were exclusively systematic reviews with a meta-analysis. We considered systematic reviews of observational studies (prospective or retrospective cohort and case-control studies) that investigated the association between exposure to any risk factor and the risk of developing PPD. We excluded systematic reviews that did not present study-level data, such as odds ratios or relative risks with 95% confidence intervals. When more meta-analyses on the same research topic were available, the meta-analysis with the largest number of included studies, providing effect sizes at a study level, was considered for inclusion, as previously described.35–37 From each included systematic review, two investigators (C.G. and G.S.) independently extracted information on first author, publication year, outcomes of interest, number of studies per meta-analysis and summary meta-analytic estimates. Primary studies included in all systematic reviews were retrieved and inspected by two reviewers (C.G. and G.S.). Details of the selection and extraction process are described in the Supplementary Methods.

Reporting quality of included systematic reviews and meta-analyses

Two reviewers (C.G. and G.S.) independently assessed the quality of each systematic review, using A Measurement Tool to Assess Systematic Reviews (AMSTAR-2), a 16-point assessment tool of the methodological quality of systematic reviews. AMSTAR-2 has good interrater agreement, test–retest reliability and content validity (details in the Supplementary Methods).38

It assesses the methodological strength of reviews, through 16 domains, which include the research question and design, literature search, data extraction, explicit reports of each step and choice made by reviewers to allow transparency (e.g. presence of a list of excluded studies, with reason for exclusion), quality and statistical assessments. Each item allows for the following response options: yes, partial yes or no. Of these 16 domains, seven are considered critical domains, as they can particularly affect the validity of the review and its conclusion. The AMSTAR-2 is a qualitative tool, not designed to be scored. AMSTAR-2 offers a scheme for interpreting weaknesses identified in critical and non-critical items: reviews with no or only one non-critical weakness are considered 'high quality'; reviews with more than one non-critical weakness but no critical flaws are considered 'moderate quality'; reviews with only one critical flaw with or without non-critical weaknesses are considered 'low quality'; finally, reviews with more than one critical flaw with or without non-critical weaknesses are considered 'critically low quality' (see Supplementary Methods and Supplementary Box 1 for further details).38,39

Statistical analysis and umbrella review criteria

We extracted effect sizes of individual studies included in each meta-analysis for every association, and afterwards we re-estimated the summary effect sizes with 95% confidence intervals, using random-effects models as we expected large heterogeneity.40 Additionally, we calculated the 95% prediction intervals for the summary random effect sizes, which account for heterogeneity between studies and specify the uncertainty for the effect that would be expected in a new study examining that same research question.41 We evaluated heterogeneity with Cochran’s Q-statistic (statistical significance set at P-value <0.10) and estimated with the I² metric.42 The I² ranges from 0 to 100%, and it is considered very large, large, moderate and low for values ≥75%, 50–74%, 25–49% and <25%, respectively.43 Potential publication and small-study effects biases were evaluated with Egger’s test.44,45 Specifically, small-study effects bias was considered to be present when a more conservative effect in the largest study and a P-value ≤0.10 in the regression asymmetry test were found. Further, we assessed the excess significance, which is a test that examines if the observed number of studies (O) with statistically significant results (i.e. with P < 0.05) in the meta-analysis is higher than the expected number (E).46 For each meta-analysis, E was calculated as the sum of the statistical power estimates for every study in the meta-analysis. The power of each study was calculated assuming a non-central t distribution.47 The estimated power depends on the plausible effect size. As the true effect size for any meta-analysis is unknown, this approach assumes that the most plausible effect is the one provided by the largest study included. Excess significance for all meta-analyses was set at P-value ≤0.10.48,49 Statistical analyses were performed with RSTUDIO version 1.3.1056 for Windows (RStudio, Boston MA, US; see https://www.rstudio.com/products/rstudio/download/).

Based on the above calculations, we assessed the umbrella review criteria (URC) to classify the strength of associations as ‘convincing’ (class I), ‘highly suggestive’ (class II), ‘suggestive’ (class III) or ‘weak’ (class IV) (details in Supplementary Box 2).28,50–55 Precisely, meta-analyses were considered without bias (class I) if they met the following criteria: ≥1000 cases, random-effects P-value ≤10⁻⁵ of the meta-analysis, low or moderate between-study heterogeneity (I² ≤ 50%), 95% prediction intervals that excluded the null value, and absence of both small-study effects and excess significance. Associations were considered highly suggestive (class II) when the following criteria were met: ≥1000 cases, highly significant summary associations (P-value ≤10⁻⁵ in the random-effects estimate) and 95% prediction intervals not including the null value.
value. Suggestive evidence (class III) criteria required ≥1000 cases and P-value ≤0.001 in the random-effects model. Weak association (class IV) criteria required only P-value ≤0.05. Associations were not considered significant if the P-value in the random-effects model was ≥0.05. All P-values were two tailed.

Grading of Recommendations, Assessment, Development and Evaluations

Additionally, the overall certainty in the estimates was qualitatively assessed by two reviewers (G.S. and C.G.) – with one author (C.B.) adjudicating the decision in case of discrepancies – using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) method.46

GRADE allows to rate the certainty of estimate for each outcome and give an overview of findings easily understandable for patients, clinicians, researchers, guideline developers and policy makers.57 The following factors were considered for each outcome of interest, according to the GRADE method: study design, risk of bias, inconsistency, imprecision, indirectness, presence of large effect, dose–response gradient and publication bias.45

Based on GRADE assessments, the certainty of estimates was classified as high, moderate, low or very low (further details in Supplementary Box 3). In the case of observational studies, the certainty of evidence is low when there are no reasons to downgrade the certainty of evidence, and very low when at least one reason to downgrade the certainty of evidence is found. The only case in which evidence from observational studies could be ‘moderate’ is when some reason to upgrade the certainty of evidence is found (e.g. strong association), with no downgrades on the other items.

Tables of summary of findings were developed with the GRADEproGDT app (GRADEpro Guideline Development Tool [Software], McMaster University and Evidence Prime, Canada; see gradepro.org).

Overall evaluation

Risk factors were quantified together with a formal assessment of the certainty of estimates, using quantitative URC and GRADE. We employed both methods because they are complementary. The URC quantitatively evaluate the strength of the associations, and GRADE qualitatively assesses the certainty of evidence. We ranked all risk factors based first on the strength of each association (URC), then on the certainty of evidence (GRADE), and finally, on the quality of the systematic review (AMSTAR-2).

Sensitivity analyses

We performed three sensitivity analyses to assess whether the credibility of the evidence and the strength of the association varied when the following studies were retained in the analysis: cohort studies; studies in which the PPD diagnosis/symptoms were assessed with standardised criteria (e.g. ICD or DSM diagnosis, The Mini-International Neuropsychiatric Interview (M.I.N.I.), Structured Clinical Interview for DSM Disorders (SCID) or Edinburgh Postnatal Depression Scale (EPDS) score with a cut-off score of ≥13), excluding non-validated or self-assessed screening tools; and studies assessing mood symptoms at least 7 days after delivery (further details in Supplementary Methods).

Results

Description of studies included in the meta-analyses

The systematic search yielded 703 records. After duplicate removal, title and abstract screening, 73 full-text articles were retrieved and checked for inclusion. Eleven systematic reviews were included, including 12 meta-analyses, with 185 primary studies and 3 272 093 participants (Fig. 1 and Supplementary Table 2). The excluded articles and the reasons for their exclusion are provided in the Supplementary Material.

The 12 risk factors reported in the 11 systematic reviews were anaemia (during pregnancy and postpartum), GDM, Caesarean section, preterm delivery, intra-labour epidural analgesia, medically assisted conception, violent experiences, premenstrual syndrome (PMS), vitamin D deficiency and unintended pregnancy, whereas one review provided data on a protective factor, the presence of the serotonin-transporter-linked polymorphic region (5-HTTLPR) polymorphism. Supplementary Table 2 summarises the main review and individual study characteristics. The number of studies per meta-analysis ranged between 4 and 33. Seven meta-analyses included ≥1000 cases (range 1074–15 758), and five had <1000 cases. Of the 185 studies, 123 (66.5%) were cohort studies and 62 (33.5%) adopted a case–control or cross-sectional design. Study participants were pregnant females, exposed to one or multiple risk factors. PPD was identified with an EPDS score of ≥13 or with ICD-9 or 10 or DSM-IV criteria in 81 studies; the remaining studies used self-reported or other scales or operational criteria. Assessments of mood symptoms were conducted within the first postpartum week in 31 studies.

Quality assessment of the systematic reviews

The PMS review was of moderate quality according to the AMSTAR-2 scoring system.56 The systematic review on medically assisted conception was of low quality,60 and the remaining nine were of critically low quality (Table 1).17,21,25,27,61–64 The most common weakness was that all the systematic reviews did not contain an explicit statement that the review methods were established before the conduct of the review (Table 1, question 2), with the exception of the review on preterm delivery.25 The nine reviews rated as critically low also did not provide a list of excluded studies with reasons to justify the exclusion (Table 1, question 7).17,21,25–27,61–64 Other critical flaws were that authors did not consider the risk of bias when interpreting results (Table 1, question 13), and the absence of adequate investigation of publication bias or small study bias and their impact on the results (Table 1, question 15).

Summary of associations and URC

Ten meta-analyses showed significant summary random-effects estimates; exposure to PMS, violent experiences, unintended pregnancy, Caesarean section, preterm delivery, GDM, anaemia during pregnancy, vitamin D deficiency and postpartum anaemia increased the risk of PPD (Table 2 and Fig. 2).63

Conversely, the 5-HTTLPR polymorphism was associated with a lower PPD risk.64 Associations between PPD and medically assisted conception or intra-labour analgesia were not significant.27,60

According to the URC, three associations (between PPD and PMS, violent experiences and unintended pregnancy) were highly suggestive (class II); three associations (between PPD and Caesarean section, GDM and 5-HTTLPR polymorphism) were suggestive (class III); four associations (between PPD and preterm delivery, anaemia during pregnancy, vitamin D deficiency and postpartum anaemia) were weak (Table 2); and non-significant associations were reported for medically assisted conception and labour epidural analgesia.

Certainty of evidence according to GRADE

For Caesarean section, preterm delivery, anaemia during pregnancy and 5-HTTLPR polymorphism, the certainty was rated as low

https://doi.org/10.1192/bjp.2021.222 Published online by Cambridge University Press
Table 3), based on the *a priori* GRADE baseline assumption of low certainty for observational studies. We found no reasons to upgrade this baseline evaluation. For postpartum anaemia, PMS, violent experiences, GDM, labour epidural analgesia and unintended pregnancy, the certainty was very low, mainly because of inconsistency.

**Overall ranking**

Figure 2 presents a ranking of associations based on URC, GRADE and AMSTAR-2. The association between PPD and PMS was the most reliable, followed by associations with violent experiences and unintended pregnancy. The association with Caesarean section, GDM and 5-HTTLPR was suggestive, and the association with preterm delivery, anaemia during pregnancy, vitamin D deficiency and postpartum anaemia was weak. No association was found for medically assisted conception or intra-labour epidural analgesia.

**Sensitivity analyses**

In cohort studies, the associations between PPD and violent experiences (class II), unintended pregnancy (class II), GDM (class III), preterm delivery (weak), vitamin D deficiency (weak), postpartum anaemia (weak), medically assisted conception (no association) and intra-labour analgesia (no association) remained at the same strength. Conversely, the association with PMS was downgraded from highly suggestive (class II) to weak because the criterion of ≥1000 cases was not met. Moreover, for PMS, the GRADE certainty assessed in the sensitivity analysis was in contrast with the URC, as it was upgraded from very low to low, as the risk of bias was rated as ‘not serious’ in this case. In the case of Caesarean section, the association was downgraded from suggestive (class III) to non-significant, and the GRADE certainty was downgraded from low to very low. In the remaining cases, the GRADE certainty did not change between the main and the sensitivity analysis (Supplementary Table 3). This sensitivity was not performed for two risk factors: in the case of anaemia during pregnancy, all included studies were cohort, and for the 5-HTTLPR, none of the studies investigated cohorts.

In studies with standardised criteria for the PPD diagnosis/symptoms, the association with PPD was upgraded from suggestive to highly suggestive for GDM, remained suggestive for Caesarean section, and was non-significant for labour epidural analgesia and medically assisted conception. The associations with PMS and unintended pregnancy were downgraded from highly suggestive to weak, violent experiences was downgraded from highly suggestive to suggestive, 5-HTTLPR polymorphism was downgraded from suggestive to weak, and preterm delivery and postpartum anaemia were downgraded from weak to non-significant. The GRADE certainty was upgraded only for anaemia and PMS, from very low to low, whereas the association with 5-HTTLPR polymorphism remained low. In all other cases, GRADE certainty was very low, as in the main analysis (Supplementary Table 4). For anaemia during pregnancy and vitamin D deficiency, sensitivity analysis was not performed, as not enough studies using standardised criteria were available (one and zero studies, respectively).

Excluding studies that assessed mood symptoms within the first postpartum week, the strength of association remained for violent experiences, Caesarean section, vitamin D deficiency, preterm...
Table 1  AMSTAR-2 assessment of the included meta-analyses

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Did the research questions and inclusion criteria for the review include the components of PICO?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>2 Did the report of the review contain an explicit statement that the review methods were established before the conduct of the review and did the report justify any significant deviations from the protocol?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>3 Did the review authors explain their selection of the study designs for inclusion in the review?</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>4 Did the review authors use a comprehensive literature search strategy?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>5 Did the review authors perform study selection in duplicate?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>6 Did the review authors perform data extraction in duplicate?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>7 Did the review authors provide a list of excluded studies and justify the exclusions?</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>8 Did the review authors describe the included studies in adequate detail?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>9 Did the review authors use a satisfactory technique for assessing the risk of bias in individual studies that were included in the review?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>10 Did the review authors report on the sources of funding for the studies included in the review?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>11 If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12 If meta-analysis was performed, did the review authors assess the potential impact of risk of bias in individual studies on the results of the meta-analysis or other evidence synthesis?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>13 Did the review authors account for risk of bias in individual studies when interpreting/discussing the results of the review?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>14 Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>15 If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>16 Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Number of critical weaknesses</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Number of non-critical weaknesses</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Reporting quality of each meta-analysis</td>
<td>Moderate</td>
<td>Critically low</td>
<td>Critically low</td>
<td>Critically low</td>
<td>Critically low</td>
<td>Critically low</td>
<td>Critically low</td>
<td>Critically low</td>
<td>Critically low</td>
<td>Critically low</td>
<td>Critically low</td>
</tr>
</tbody>
</table>

Bolding indicates critical questions. Further details can be found in the Supplementary Material. AMSTAR-2, A Measurement To Assess Systematic Reviews; PMS, premenstrual syndrome; 5-HTTLPR, serotonin-transporter-linked polymorphic region; PICO: population, intervention, comparator, outcome.
### Table 2

<table>
<thead>
<tr>
<th>Risk factors</th>
<th>Number of cases</th>
<th>Meta-analysis effect size</th>
<th>Random-effects effect size of the largest study</th>
<th>Significant studies</th>
<th>Class of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaemia postpartum (Azami et al, 2019)</td>
<td>61</td>
<td>0.006</td>
<td>0.86 (0.06-2.37)</td>
<td>2</td>
<td>NSA</td>
</tr>
<tr>
<td>Anaemia during pregnancy (Azami et al, 2019)</td>
<td>61</td>
<td>0.006</td>
<td>0.86 (0.06-2.37)</td>
<td>2</td>
<td>NSA</td>
</tr>
<tr>
<td>Gestational diabetes (Azami et al, 2019)</td>
<td>62</td>
<td>0.006</td>
<td>0.86 (0.06-2.37)</td>
<td>2</td>
<td>NSA</td>
</tr>
<tr>
<td>Unintended pregnancy (Qiu et al, 2020)</td>
<td>63</td>
<td>0.006</td>
<td>0.86 (0.06-2.37)</td>
<td>2</td>
<td>NSA</td>
</tr>
</tbody>
</table>

#### Discussion

The most reliable association with PPD was found for women suffering from PMS, followed by violent experiences and unintended pregnancy. The risk of PPD was more than doubled in women with PMS or a violent experience, and was 50% higher in women with unintended pregnancy. Women suffering from PMS could have an affective vulnerability underpinned by hormonal fluctuations, which occur both during the premenstrual period and on a much larger scale during postpartum. Experience of violence may be a less specific risk factor, mainly because of increased stress levels, which can activate the hypothalamic-pituitary-adrenal axis, resulting in a release of glucocorticoids influencing psychological functions.

Unintended pregnancy has been previously suggested as a PPD risk factor, mainly because of increased stress levels, which can activate the hypothalamic-pituitary-adrenal axis, resulting in a release of glucocorticoids influencing psychological functions. Specifically, women who did not plan their pregnancy may be unprepared/worried about the health of the foetus, feel a potential conflict between continuing and terminating the pregnancy, and start the prenatal care later than women who planned the pregnancy.

For these risk factors, the effect sizes were generally small (odds ratio < 3.5), despite the URC showing class II associations. Therefore, evidence regarding such associations might only play a partial role with respect to depression onset, and confounding could not be ruled out because of the observational nature of studies, which was reflected by the GRADE certainty (very low).

At an intermediate credibility level, we found associations of PPD with 5-HTTLPR polymorphism, Caesarean section and GDM. The latter increased the risk of PPD by 60%. Conversely, women who underwent Caesarean section had a slightly increased risk of PPD. However, it is noteworthy that in the included meta-analysis, Caesarean section was a significant risk factor for PPD when performed in an emergency situation, but not when elective, suggesting a central role of unexpected complications for PPD. The allelic model of 5-HTTLPR polymorphism was the only protective factor, which slightly reduced the PPD risk. The 5-HTT gene is a key factor that affects risk for depression and other mental disorders.

At an intermediate credibility level, we found associations of PPD with 5-HTTLPR polymorphism, Caesarean section and GDM. The latter increased the risk of PPD by 60%. Conversely, women who underwent Caesarean section had a slightly increased risk of PPD. However, it is noteworthy that in the included meta-analysis, Caesarean section was a significant risk factor for PPD when performed in an emergency situation, but not when elective, suggesting a central role of unexpected complications for PPD. The allelic model of 5-HTTLPR polymorphism was the only protective factor, which slightly reduced the PPD risk. The 5-HTT gene is a key factor that affects risk for depression and other mental disorders.

The transcriptional activity of the 5-HTT gene may be regulated by glucocorticoids influencing psychological functions.

For these risk factors, the effect sizes were generally small (odds ratio < 3.5), despite the URC showing class II associations. Therefore, evidence regarding such associations might only play a partial role with respect to depression onset, and confounding could not be ruled out because of the observational nature of studies, which was reflected by the GRADE certainty (very low).

At an intermediate credibility level, we found associations of PPD with 5-HTTLPR polymorphism, Caesarean section and GDM. The latter increased the risk of PPD by 60%. Conversely, women who underwent Caesarean section had a slightly increased risk of PPD. However, it is noteworthy that in the included meta-analysis, Caesarean section was a significant risk factor for PPD when performed in an emergency situation, but not when elective, suggesting a central role of unexpected complications for PPD. The allelic model of 5-HTTLPR polymorphism was the only protective factor, which slightly reduced the PPD risk. The 5-HTT gene is a key factor that affects risk for depression and other mental disorders.

The transcriptional activity of the 5-HTT gene may be regulated by glucocorticoids influencing psychological functions.
polymorphism was downgraded to weak, suggesting that these factors might be associated with subclinical depression symptoms.

Restricting the analyses to studies with an assessment after the first postpartum week, we found lower certainty for almost all associations, with the exception of the association of PPD with violent experiences, Caesarean section and unintended pregnancy. These associations remained at the same level of credibility indicating that apart from violent experiences, Caesarean section and unintended pregnancy, the other risk (or protective) factors might be associated more with the so-called ‘maternity blues’ or adjustments within the first postpartum week, rather than with depression diagnosed later. This result is notable if we consider that authors previously suggested that maternity blues could be a risk factor for developing depression later on in the postpartum period.13,79,80

Moreover, the credibility of the association with violent experiences was maintained across all analyses. The increased risk was slightly lower in the two sensitivity analyses based on clinical criteria, suggesting that some bias could have inflated the main results. Nevertheless, violent experiences were found as one of the most reliable risk factors for developing PPD, in line also with literature on major depression.66

Finally, the sensitivity analyses confirmed the non-differential association between risk factors and PPD. Moreover, this type of umbrella review, the risk factors listed in this tool might need to be revised and updated along with guidelines. Specifically, considering that currently PMS, Caesarean section and GDM are not included in the tool, we suggest that these risk factors might be worth being considered in future studies.

A second general cautionary note is the observational nature of the primary studies, which does not allow us to establish a causal association between risk factors and PPD. Moreover, this type of studies is vulnerable to bias because of unmeasured confounding and lower internal validity compared with randomised controlled trials.63,64 In fact, it is unclear if some of the investigated factors were proxies for other factors or shared background risk. It should be noted, however, that this risk of bias is taken into consideration by GRADE, which suggests rating the certainty of evidence based on the level of quality of the available evidence.82

A number of limitations need to be considered when interpreting these results. First, although broadly employed in mental health and medicine,50,51,53–55,66 and corroborated by standard statistical tests,42,44–46 the URC classification to classify strengths has been just recently introduced. The criteria of the ≥1000 cases might not be fully applicable for studies that target very specific samples, especially those with low incidence.

A second general cautionary note is the observational nature of the primary studies, which does not allow us to establish a causal association between risk factors and PPD. Moreover, this type of studies is vulnerable to bias because of unmeasured confounding and lower internal validity compared with randomised controlled trials.63,64 In fact, it is unclear if some of the investigated factors were proxies for other factors or shared background risk. It should be noted, however, that this risk of bias is taken into consideration by GRADE, which suggests rating the certainty of estimates from observational studies as low quality instead of high quality, to acknowledge issues with internal validity.56 On the other hand, findings of meta-analyses of observational studies are more generalisable and pragmatic, as they have larger sample size and are more weighted. Finally, some risk factors already considered may be less strong than commonly considered, such as preterm delivery.

The American College of Obstetricians and Gynecologists recommends that obstetrician-gynaecologists and other obstetric care providers screen patients at least once during the perinatal period for depression and anxiety symptoms, using a standardised, validated tool.83 We suggest that, based on the results of this umbrella review, the risk factors listed in this tool might need to be revised and updated along with guidelines. Specifically, considering that currently PMS, Caesarean section and GDM are not included in the tool, we suggest that these risk factors might be at least considered. Moreover, the assessment of some risk factors, such as violent experience, that are only partially mentioned under the ‘Experiencing stressful life events’ category, could be more weighted. Finally, some risk factors already considered may be less strong than commonly considered, such as preterm delivery.

A number of limitations need to be considered when interpreting these results. First, although broadly employed in mental health and medicine,50,51,53–55,66 and corroborated by standard statistical tests,42,44–46 the URC classification to classify strengths has been just recently introduced. The criteria of the ≥1000 cases might not be fully applicable for studies that target very specific samples, especially those with low incidence.

A second general cautionary note is the observational nature of the primary studies, which does not allow us to establish a causal association between risk factors and PPD. Moreover, this type of studies is vulnerable to bias because of unmeasured confounding and lower internal validity compared with randomised controlled trials.63,64 In fact, it is unclear if some of the investigated factors were proxies for other factors or shared background risk. It should be noted, however, that this risk of bias is taken into consideration by GRADE, which suggests rating the certainty of estimates from observational studies as low quality instead of high quality, to acknowledge issues with internal validity.56 On the other hand, findings of meta-analyses of observational studies are more generalisable and pragmatic, as they have larger sample
## Table 3: Certainty Risk Assessment

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Number of Studies</th>
<th>Study Design</th>
<th>Risk of Bias</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience of violence (Zhang et al., 2019)</td>
<td>17</td>
<td>Observational studies</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Not serious</td>
<td>Strong association</td>
</tr>
<tr>
<td>Unintended pregnancy (Qiu et al., 2020)</td>
<td>63</td>
<td>Observational studies</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Not serious</td>
<td>None</td>
</tr>
<tr>
<td>5-HTTLPR polymorphism (Li et al., 2020)</td>
<td>64</td>
<td>Observational studies</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
</tr>
<tr>
<td>Caesarean section (Xu et al., 2017)</td>
<td>26</td>
<td>Observational studies</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
</tr>
<tr>
<td>Gestational diabetes (Azami et al., 2019)</td>
<td>62</td>
<td>Observational studies</td>
<td>Serious</td>
<td>Very serious</td>
<td>Not serious</td>
<td>None</td>
</tr>
<tr>
<td>Preterm (De Paula Eduardo et al., 2019)</td>
<td>25</td>
<td>Observational studies</td>
<td>Not serious</td>
<td>Serious</td>
<td>Not serious</td>
<td>Strong association</td>
</tr>
<tr>
<td>Anaemia during pregnancy (Azami et al., 2019)</td>
<td>61</td>
<td>Observational studies</td>
<td>Not serious</td>
<td>Not serious</td>
<td>Not serious</td>
<td>None</td>
</tr>
<tr>
<td>Vitamin D deficiency (Wang et al., 2018)</td>
<td>21</td>
<td>Observational studies</td>
<td>Not serious</td>
<td>Very serious</td>
<td>Not serious</td>
<td>Strong association</td>
</tr>
<tr>
<td>Anaemia postpartum (Azami et al., 2019)</td>
<td>61</td>
<td>Observational studies</td>
<td>Not serious</td>
<td>Serious</td>
<td>Not serious</td>
<td>None</td>
</tr>
<tr>
<td>Labour epidural analgesia (Kountanis et al., 2020)</td>
<td>27</td>
<td>Observational studies</td>
<td>Serious</td>
<td>Very serious</td>
<td>Not serious</td>
<td>None</td>
</tr>
</tbody>
</table>

**Certainty:** From low to very high. **Risk of Bias:** From low to high. **Indirectness:** From low to high. **Imprecision:** From low to high. **Other Considerations:** Not assessed.

### Discussion

- **Risk factors:** Development of an umbrella review to systematically summarise data on the association between PPD and several risk factors. Risk factors include demographic, obstetric, psychological, and social factors.

- **Methodology:** Systematic review and meta-analysis of observational studies. The review included a wide range of risk factors, including demographic, obstetric, psychological, and social factors.

- **Limitations:** The reviewed studies were not homogeneous, and the quality of the evidence was variable. The results are limited by the lack of randomised controlled trials and the small number of studies.

- **Conclusions:** The review identified several risk factors for PPD, including a history of mental illness, a personal history of PPD, a history of depression or anxiety, and a history of substance use. However, the evidence for most of these factors was inconclusive and limited by the quality of the evidence. The lack of high-quality evidence makes it difficult to establish causality and to develop effective prevention strategies.

- **Strengths:** The review included a wide range of risk factors, and it was the first umbrella review to systematically summarise data on the association between PPD and several risk factors.

- **Weaknesses:** The reviewed studies were not homogeneous, and the quality of the evidence was variable. The results are limited by the lack of randomised controlled trials and the small number of studies.

- **Future directions:** Further research is needed to establish the causality of the identified risk factors and to develop effective prevention strategies. The field of perinatal mental and physical health research is still in its infancy, and more research is needed to understand the complex interplay of risk and protective factors.
Furthermore, this work provides methodological directions for future studies aiming to improve our understanding of predictors of PPD. First, there is an urgent need to generate further multidisciplinary evidence to more effectively tackle mental (and physical) health challenges for women during pregnancy, as well as the postpartum phase.9,69 Second, further replication of the evidence regarding biological and psychosocial factors in more sophisticated models is hoped to lay the groundwork for a comprehensive PPD screening tool. Moreover, timing of assessment of depression seems extremely relevant. Therefore, more well-designed, prospective observational studies on risk factors for PPD, collecting data pre-intrapartum and following up on women after the first week postpartum are needed to determine whether risk factors are associated with perinatal depression.95,98 Ultimately, our data may enhance efforts to develop interdisciplinary prevention and care targeting patients at high risk for PPD.

Our results could be of help in updating postpartum screening tools employed to identify and screen women at risk of PPD, such as the one developed by the American College of Obstetricians and Gynecologists.82 The early recognition and management of these patients will improve treatment outcomes improving maternal health and new-born development.

**References**


12 Kuehn BM. Postpartum depression screening needs more consistency. *JAMA* 2020; 323(24): 2454.


**Supplementary material**

To view supplementary material for this article, please visit https://doi.org/10.1192/bjp.2021.222

**Data availability**

The data that support the findings of this study are available from the corresponding author, C.G., upon request.

**Author contributions**

C.G., G.S., CB, M.S. and C.U.C. designed the study. C.G. and G.S. drafted the manuscript. C.G., G.S. and C.B. contributed to the database preparation and double-checked the data. C.G., G.S. and M.S. conducted data analyses. C.G. and G.S. assessed the quality of meta-analyses with AMSTAR-2. All authors critically revised the manuscript and approved the draft and final manuscripts.

**Funding**

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Declaration of interest**

C.B. and C.G. have nothing to disclose. G.S. has served as a consultant for H&S Therapeutics. C.U.C. has been a consultant and/or advisor to, or has received honoraria from, Alkermes, Allergan, Argenti, Boehringer-Ingelheim, Gedeon Richter, Gerson Lehrman Group, Indivior, Intracel Therapies, Janssen/ALI, LB Pharma, Lundbeck, Medfuture-PROPhase, Medicon, Medscape, Merck, Mylan, Neuriconne, Noven, Otsuka, Pfizer, Recordati, Rovi, Seriver, Sumitomo Dainippon, Sunovion, Supersum, Takeda and Teva. He has provided expert testimony for Bristol Myers Squibb, Janssen and Otsuka. He served on a Data Safety Monitoring Board for Boehringer-Ingelheim, Lundbeck, Roxi, Supersum and Teva. He has received royalties from UpToDate and grant support from Janssen and Takeda. He is also a shareholder of LB Pharma. M.S. has been a consultant and/or advisor to, or has received honoraria, from Argenti and Lundbeck.


Committee on Ethics. ACOG Committee Opinion No. 646: ethical consideration for including women as research participants. Obstet Gynecol 2015; 126(5): e100–7.


