Prevalence and risk factors of birth-related posttraumatic stress among parents: A comparative systematic review and meta-analysis

Clara-Sophie Heyne a, Maria Kazmierczak b, Ronnie Souday c, Danny Horesh c, d, Mijke Lambregts-van den Berg e, Tobias Weigl f, Antje Horsch g, h, Mirjam Oosterman i, Pelin Dikmen-Yildiz j, Susan Garthus-Niegel a, k, l, *

a Institute and Polyclinic of Occupational and Social Medicine, Faculty of Medicine, Technische Universität Dresden, Germany
b Institute of Psychology, University of Gdańsk, Poland
c Department of Psychology, Bar-Ilan University, Israel
d Department of Psychiatry, New York University Grossman School of Medicine, United States of America
e Departments of Psychiatry and Child & Adolescent Psychiatry, Erasmus University Medical Center, the Netherlands
f Psychology School, Fresenius University of Applied Sciences, Düsseldorf, Germany
g Institute of Higher Education and Research in Healthcare, Faculty of Biology and Medicine, University of Lausanne, Switzerland
h Department woman-mother-child, Lausanne University Hospital, Switzerland
i Faculty of Behavioral and Movement sciences, Department of Clinical Child & Family Studies, Vrije Universiteit Amsterdam and Amsterdam Public Health Research Institute, the Netherlands
j Department of Psychology, Kirklareli University, Turkey
k Institute for Systems Medicine (ISM) and Faculty of Human Medicine, Medical School Hamburg, Germany
l Department of Child Health and Development, Norwegian Institute of Public Health, Norway

ARTICLE INFO

Keywords:
- Posttraumatic stress
- Posttraumatic stress disorder
- Parents
- Perinatal mental health
- Prevalence

ABSTRACT

This systematic review and meta-analysis aimed to determine mean estimates of prevalence rates for fulfilling all diagnostic criteria of posttraumatic stress disorder (PTSD) or at least showing significant levels of posttraumatic stress (PTSS) in relation to the traumatic event of childbirth. For the first time, both mothers and fathers were included in the synthesis. Studies were identified through systematic database search and manual searches, irrespective of language. Meta-analyses of 154 studies (N = 54,711) applied a random-effects model to four data sets, resulting in pooled prevalence rates of 4.7% for PTSD and 12.3% for PTSS in mothers. Lower rates of 1.2% for PTSD and 1.3% for PTSS were found among fathers. Subgroup analyses showed elevated rates in targeted samples (those with a potential risk status) most distinctly for maternal PTSS. The significant amount of heterogeneity between studies could not be explained to a satisfactory degree through meta-regression. Given the substantial percentage of affected parents, the adoption of adequate prevention and intervention strategies is needed. As this field of research is evolving, attention should be broadened to the whole family system, which may directly and indirectly be affected by birth-related PTSD. Further studies on paternal PTSD/PTSS are particularly warranted.

1. Introduction

Parents-to-be prepare themselves for a much anticipated but also potentially challenging time ahead with their newborn baby. While fear of the birth itself is quite common (Storksen, Eberhard-Gran, Garthus-Niegel, & Eskild, 2012), the possibility of the birth turning into a traumatic event and causing the development of posttraumatic stress disorder (PTSD) probably does not cross the expectant parents’ minds. This is not surprising because traumatic childbirth and its psychological sequelae have been neglected in public discussion, scientific research,
and clinical practice for a long time, despite its potential adverse effects on infant development (Cook, Ayers, & Horsch, 2018; Garthus-Niegel, Ayers, Martini, von Soest, & Eberhard-Gran, 2017) and future reproductive decisions (Gottvall & Waldenstrom, 2002). Due to the traditional image of childbirth and motherhood as one of the most natural and beautiful elements of human life (Horsch & Garthus-Niegel, 2019), parents struggling to cope with their birth experiences do not always receive the necessary acknowledgement and support (Ayers & Sawyer, 2019). Birth-related posttraumatic stress has traditionally been addressed first and foremost in mothers. However, because the event of childbirth impacts the whole family (Horsch & Stuijtszand, 2019), our aim is to close this gap in previous research syntheses and to examine prevalence rates in both parents.

1.1. Theoretical background

1.1.1. The evolution of diagnostic criteria for PTSD

The definition of a potentially traumatic event has evolved ever since PTSD first found its way into the Diagnostic Statistical Manual of Mental Disorders – DSM-III of the American Psychiatric Association (APA, 1980) 40 years ago. Starting out as a disorder commonly associated with war trauma or other unusual life-threatening experiences, recent research has increasingly shed light on the role of childbirth as the triggering stressor. In DSM-III, an event had to be “generally outside the range of usual human experience” (APA, 1980, p. 236) to qualify as traumatic. This was replaced by a two-part description in DSM-IV, where a traumatic event had to involve “actual, or threatened death or serious injury, or a threat to [own or other person’s] physical integrity” (Criterion A1) to which the person responded with “intense fear, helplessness, or horror” (Criterion A2) (APA, 1994, p. 424). The newest edition of the DSM removed criterion A2 (APA, 2013). Although neither DSM-IV nor DSM-5 explicitly list childbirth as a potential traumatic event, the undertaken changes allow childbirth to be defined as a traumatic stressor (for example in cases of perceived threat to the mother’s and/or infant’s physical integrity). Some subjective judgment remains, leading to an intense debate about whether birth qualifies as a traumatic event (e.g., Vythilingum, 2010). Gradually, this discussion has reached a consensus among researchers and clinicians, acknowledging the experience of childbirth as a traumatic event (Ayers & Sawyer, 2019). The notion that childbirth can represent a traumatic stressor is supported by the fact that every third woman describes giving birth as a traumatic experience (Creedy, Shochet, & Horsfall, 2000). In some cases, this leads to the development of PTSD symptoms, with some women meeting full diagnostic criteria for PTSD.

The diagnostic criteria for PTSD are listed in detail in the most well-known diagnostic tools of the past decades, namely DSM-IV, DSM-5, and the International Classification of Diseases – ICD-10 and ICD-11 (APA, 1994, 2013; World Health Organization, 2016, 2018). The symptoms may develop after experiencing or witnessing a traumatic event. They include symptom clusters of intrusion, avoidance, and arousal lasting at least one month and leading to clinically relevant distress or functionality impairment. A fourth cluster “negative alterations in cognitions and mood” has been added in DSM-5 (APA, 2013). Since the introduction of the DSM-5, PTSD is no longer categorized as part of the anxiety disorders, but a new group devoted to all trauma- and stress-related disorders has been added (APA, 2013).

1.1.2. Terminology and measurement of PTSD in the context of childbirth

The choices of terms in the existing literature on PTSD in the postpartum period are manifold, with labels ranging from postpartum PTSD, PTSD following/after childbirth to birth-induced or birth-related PTSD. The perception of the various terms used in scientific research is illustrated in Fig. 1 and is intended as a proposal for establishing a common understanding and use of terms in this field of research. Although these terms have been used interchangeably in some cases, they seem to imply different underlying rationales. In the understanding of this review, they may be categorized into two different layers. First, postpartum PTSD and PTSD following/after childbirth can refer to PTSD caused by any possible stressor already present prior to birth. The event of childbirth may for instance trigger PTSD related to a previous stressor like sexual violence and result in a delayed onset of PTSD (with an index traumatic event other than the birth experience). Second, these terms can refer to PTSD caused by childbirth as the index trauma itself, which is the core area targeted in this review and labelled birth-related PTSD for this purpose. While birth-related PTSD may potentially be triggered in anyone witnessing birth, including health personnel (Beck & Gable, 2012), this study will focus only on the population of parents, i.e., mothers and fathers or partners present during birth.

Some challenges arise when attempting to adapt the diagnostic criteria for PTSD to parents who have recently experienced childbirth. Apart from the prior debate on childbirth as a traumatic event, difficulties remain in assessing symptoms like sleep disturbances, because they are not necessarily psychopathological in the postpartum population (Stramrood et al., 2010). Further, researchers initially had to rely on assessment measures for general PTSD to examine potential psychological consequences of traumatic birth experiences. Increasingly, researchers have modified these instruments to the specific event of childbirth, for example by adapting the instructions of the Impact of Event Scale (IES, Horowitz, Wilner, & Alvarez, 1979) to refer to “childbirth” rather than any traumatic event. The latest step in this evolution is the development of new instruments tailored to measure birth-related PTSD, like the City Birth Trauma Scale (City BiTS, Ayers, Wright, & Thornton, 2018), which for instance specifically assesses potential sleep disturbances unrelated to the baby’s sleep pattern.

1.1.3. Prevalence of posttraumatic stress after childbirth

Previous quantitative studies using a multitude of different instruments have confirmed that a significant percentage of women experience PTSD after childbirth. The first meta-analysis found that 2.9% of women in community samples and 14.1% in risk samples were affected by birth-related PTSD between one to 18 months postpartum (Grekin & O’Hara, 2014). A second meta-analysis by Dikmen-Yildiz, Ayers, and Phillips (2017) found slightly higher prevalence rates of 4.0% in community and 18.5% in high-risk groups one to 14 months after birth (25 of 28 included studies focused on birth as index event). Dekel, Stuebe, and Dishy (2015) conducted a review focussing on birth-related PTSD in the first six months after full-term births, primarily including community samples (four of 36 studies at-risk groups) and found rates between 4.5% and 6.3%. This suggests that specific groups may be particularly vulnerable to develop PTSD and meeting full diagnostic criteria as a result of their birth experience. These vulnerable groups, in the following referred to as targeted samples, could for example be characterized by pregnancy complications (e.g., Polachek, Dulitzky, Margolis-Dorfman, & Simchen, 2016) or a history of childhood trauma (e.g., Oh et al., 2016). Prior reviews found a relatively wide range of prevalence rates across individual studies, which is commonly
attributed to differences in sample selection, measurement time point, and/or use of self-report instruments vs. clinical diagnostic interviews (Dikmen-Yildiz et al., 2017).

Numerous studies indicate that even without meeting full diagnostic criteria, sub-threshold levels of posttraumatic stress symptoms can negatively affect a family (e.g., child development: Garthus-Niegel et al., 2017; relationship satisfaction: Garthus-Niegel et al., 2018; mother-infant bond: Stuijfzand, Garthus-Niegel, & Horsch, 2020). Similar to the above-described variability regarding terminology of postpartum/birth-related PTSD, various terms are utilized to describe this phenomenon, including but not limited to, clinically relevant or significant symptom levels, partial PTSD, PTSD profile, and subclinical PTSD. Hereinafter, the term posttraumatic stress symptoms (PTSS) will be used to refer to symptoms not fulfilling the whole set of PTSD diagnostic criteria, though clinically noteworthy. So far, only one review has looked at both PTSD and PTSS and found a prevalence of 9.6% for birth-related PTSS in controlled high-quality studies, almost twice as high as the one for birth-related PTSD (Dekel et al., 2017).

The father-to-be or the woman’s partner present during childbirth may also experience substantial distress. However, the mental health of fathers or partners has long been neglected and is still an underrepresented topic in this field of research (Singley & Edwards, 2015). Only a handful of studies has quantitatively addressed the assumption that apart from the woman giving birth, the father’s/partner’s mental health can also be affected by witnessing a traumatic birth (e.g., Ayers, Wright, & Wells, 2007; Bradley, Slade, & Levston, 2008; Schobinger, Stuijfzand, & Horsch, 2020). As it has become more and more common for fathers/partners to be present during birth and expectations rise for them to be equally involved in their children’s lives, the need to explore both parents’ mental health has increased (Fisher et al., 2012; Horsch & Stuijfzand, 2019). To the best of the authors’ knowledge, there have been no prior attempts to quantitatively review existing studies on birth-related posttraumatic stress in fathers/partners.

To summarize, to date, prior research has reviewed maternal postpartum PTSD in various time frames (up to 14, or 18 months postpartum), in some cases without clear definitions of the traumatic event (i.e., any event vs. childbirth) or sample type (i.e., community vs. at-risk). Therefore, this systematic review and meta-analysis aims to close the gaps in previous syntheses and to incorporate the most recent research results in this evolving field. This will be accomplished by for the first time including both, evidence on a symptomatic, as well as a diagnostic level of birth-related posttraumatic stress in mothers and fathers/partners. This approach will be combined with a clear focus on childbirth as a unique traumatic index event (Horesh, Garthus-Niegel, & Horsch, 2021), instead of a general investigation into postpartum PTSD. We will also transparently report the sample characteristics distinguishing targeted from non-targeted samples. Extensive insight into these questions is important for early identification of those at risk and in need of professional support, as well as the development and implementation of appropriate (preventive) interventions.

1.2. Aims and review questions

The study has two objectives. (a) The main research objective is to determine the prevalence rates of birth-related PTSD/PTSS in both parents. To accurately reflect longitudinal primary research and possibly widely spaced-out measurement time points, we will also explore the course of prevalence rates over time (as pursued by Dikmen-Yildiz et al., 2017). Provided that the analyzed studies hold sufficient data, (b) possible moderating effects on prevalence rates will be explored in a second step. Young age, primiparity, emergency cesarean sections (c-sections), and assisted vaginal births have been proposed as risk factors of birth-related PTSD and will therefore be included in the meta-regression (Dekel et al., 2017). Analyzing the effects of methodological variables, i.e., measurement type, year of publication, geographical region, and risk of bias will allow for comparison to the meta-analytic results of Dikmen-Yildiz et al. (2017). An exhaustive investigation into the second research aim goes beyond the scope of this review; consequently the choice of included covariates should be understood as exemplary. Below, the research objectives are elaborated in detail:

(a) To summarize the prevalence rates of birth-related PTSD/PTSS in mothers and fathers/partners in targeted and non-targeted samples, if possible supplemented by examination of the course over time (comparative meta-analysis)
(b) To explore possible moderating effects on prevalence rates by age, parity, mode of birth, type of PTSD/PTSS measure, year of publication, geographical region, or risk of bias (exploratory meta-regression)

2. Methods

This study is part of COST Action CA18211, a research network funded by the European Cooperation in Science and Technology (COST). The COST Action is called “DEVOTION: Perinatal Mental Health and Birth-Related Trauma: Maximizing best practice and optimal outcomes” and unites parties from all over Europe and beyond.

The protocol was registered on PROSPERO a priori and can be accessed at https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020175813. Methods were aligned with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (Moher et al., 2009) and the Joanna Briggs Institute (JBI) Manual for Evidence Synthesis (Aromataris & Munn, 2020).

2.1. Search strategy and screening procedure

A search string was defined combining terms for the postpartum period with terms for the population of parents as well as the disorder or symptoms of posttraumatic stress. The search string can be found in Appendix A. It was applied to six electronic databases (i.e., PsycInfo, PubMed, Web of Science, Scopus, PTSDpubs, and CINHALH), including comprehensive medical, psychological, and scientific databases along with databases specialized in nursing and PTSD. Wherever possible, subject headings were added to the free text terms and a filter for human studies was applied. The search was limited to studies published since 1994, corresponding to the release of the DSM-IV (APA, 1994), in which the trauma definition first made inclusion of childbirth as a traumatic event possible. The database search was not restricted regarding language of the publication and was performed on June 4th, 2020.

Subsequently, further studies were identified by scanning the reference lists of previous reviews, book chapters on the topic, and studies included in this review. Additionally, grey literature was explored via various sources (e.g., databases for dissertations, Google Scholar, ResearchGate). Manual searches were repeated prior to submission up to February 28th, 2021, to allow inclusion of studies published later than June 2020.

The screening process was conducted by Clara Heyne (CH) as first reviewer and an independent second reviewer for each screening step (title-abstract: Maria Kazmierczak [MK], full text: Ronnie Souday [RS]). Agreement between the two raters was calculated by CH using simple kappa statistics and can be considered excellent (kappa = 0.91) for title-abstract and good (kappa = 0.73) for full text screening (Higgins et al., 2008). Discrepancies between the two reviewers were resolved by discussion, with the involvement of a third reviewer (Susan Garthus-Niegel [SGN]) if necessary. Native speakers were consulted to assess eligibility of non-English articles. Authors of articles not providing sufficient detail to make a judgment regarding inclusion criteria (k ≤ 0.7) were contacted and asked for clarification. Almost half of them (k = 39) replied with additional details allowing for an informed decision. If establishing contact was unsuccessful or no response was received within a four-week period, the study was excluded from the review.

When multiple papers reported data from the same or an overlapping
sample or database, we excluded those not published in English. If there was more than one English report, the one with the largest sample size (or if sample size was identical, the most recently published one) was included. Although we aimed to include all articles regardless of language, two Persian articles had to be excluded because it was not possible to recruit a Persian-speaking person for data extraction. The study selection process is visualized in a PRISMA flow diagram (Moher et al., 2009) in Fig. 2.

In order for our work to be as current as possible at the time of publication, an update of the database search was conducted after the review process on December 13th, 2021. New database results were identified and screened in an analogous manner. Manual searches (i.e., checking reference lists of newly included studies and sources of grey literature) were also repeated and were completed by January 21st, 2022. The whole team contributed to the search update.

2.2. Inclusion and exclusion criteria

Inclusion criteria were specified for four different aspects. First, in order for a study to be included in the synthesis, it had to report prevalence data from a quantitative, validated measure of PTSD/PTSS. The review aimed at including data from both, clinical diagnostic interviews and self-report measures. Intervention studies were eligible if they incorporated a control group and/or a pre-intervention measurement and did not solely include subjects already diagnosed with PTSD or subjects reporting clinically significant symptom levels of PTSS. Apart from this restriction, observational (prospective/retrospective longitudinal, cross-sectional) or experimental (randomized/non-randomized controlled trials, quasi-experimental trials) study designs were considered.

Second, measurements had to be conducted within four weeks to 14 months postpartum. The minimum was set at four weeks in alignment with the DSM diagnostic criterion of duration for PTSD and to ensure differentiation from acute stress disorder (APA, 1994, 2013). This corresponds to a large number of studies choosing four to eight weeks postpartum as their (first) point of measurement. In accordance with the most recent review (Dikmen-Yildiz et al., 2017) and in order to allow for an overview of roughly one year (from the most commonly used initial measurement time at one to two months postpartum), 14 months was chosen as latest measurement time. The time frame was not set to 12 months postpartum to avoid the risk of excluding studies aiming to...
measure at one year post birth, while exceeding this exact time point due to various practical reasons. Many studies assess populations of a wider range postpartum, rather than (roughly) at one specific time point following birth (Dikmen-Yildiz et al., 2017). Studies in which the complete time range exceeded our pre-defined limit (e.g., population of women between one and 18 months postpartum) were included as long as the range of one standard deviation of the mean lay within the time frame of one to 14 months postpartum. Longitudinal studies had to include at least one measurement point within our range.

Third, study subjects had to be mothers who had experienced childbirth themselves and/or fathers/partners who attended birth. The mention of fathers or partners takes into account that (a) the mother’s partner may be female or male and may not necessarily be the child’s biological father, and (b) a biological father may choose to be present at birth, regardless of his relationship status with the mother.

The fourth and last criterion was the explicit reference to childbirth as the stressor. Childbirth was defined as live birth or stillbirth, the latter referring to fetal death occurring after 20–24 weeks of pregnancy (Diamond & Diamond, 2016). Studies implement the link to childbirth as the traumatic event by using a measure specifically designed for birth-related PTSD/PTSS like the City BITS (Ayers et al., 2018), by defining childbirth as the index event on a general PTSD/PTSS measure such as the IES (Horowitz et al., 1979), or by assessing and reporting PTSD/PTSS prevalence while distinguishing the underlying trauma, for example through assessment with the Structured Clinical Interview for DSM-IV (SCID, First, Spitzer, Gibbon, & Williams, 1996). If specification of the index trauma was not explicitly reported and author confirmation could not be obtained, this criterion was considered fulfilled if the article’s aim was clearly expressed in terms of exploring consequences of traumatic birth experiences.

### 2.3. Data extraction and quality assessment

Having identified the relevant work, data were extracted into a data extraction form in Excel with the help of a coding manual (see Appendix B) by the first author (CH). The coding manual included sections, such as study characteristics, participants and sample, as well as methods and results of the outcome analysis. The results of the main outcome were proofread for all studies to ensure objectivity and prevent transcription errors (Danny Horesh [DH], Mijke Lambregtse-van den Berg [MLB], Tobias Weigl [TW]). Roughly 20% of the studies (k = 29) were double coded in full by a second reviewer (Mirjam Oosterman [MO]). Interrater reliability was calculated by TW for categorial variables (e.g., adequate sample size, country of study) using kappa and for continuous variables (e.g., prevalence, mean age) using intraclass correlation coefficients (ICC). Interrater agreement was good to excellent with kappa >0.60 and ICC > 0.90 (Cicchetti, 1994). Data extraction and risk of bias rating of the French studies was taken on by Antje Horsch (AH). If prevalence rates were unclear, authors were contacted and asked to provide the missing data. Ten of the 30 contacted authors complied with this request, the other 20 studies had to be excluded.

The studies’ risk of bias was assessed by the first author (CH) as well as an independent second reviewer (Pelin Dikmen-Yildiz [PDY]) using the JBI’s critical appraisal checklist (Munn, Moolla, Lisy, Ritiano, & Tufanaru, 2015). This checklist was developed specifically for evaluating studies reporting prevalence data. It includes nine items assessing the appropriateness of target population and sampling method, sample size, description of subjects and setting, conduction of analyses, validity and reliability of measurement, as well as management of response rate. Please refer to pages 7–9 of Appendix B in the Supplementary Materials for a detailed item description. Kappa = 0.71 indicated good interrater agreement according to Higgins et al. (2008). Disagreements were handled in the same manner as throughout the selection process, involving a third reviewer (SGN) when necessary. Of the nine items, three were considered especially important for the estimation of prevalence rates and were defined as major domains: sample size, validity, and reliability of measurement. To be included in the quantitative analyses, a study had to have a minimum sample size of 20 for each subgroup and a minimum of three out of all nine items had to be rated as fulfilled, including at least one of the three major domains. A study was considered as holding a low risk of bias if at least 60% of the JBI checklist’s items were rated as fulfilled.

### 2.4. Data analysis

All statistical analyses were conducted with the Comprehensive Meta-Analysis software (Version 3, Borenstein, Hedges, Higgins, & Rothstein, 2008). The applied level of confidence intervals was 95% and significance was set at p < .05. Following the recommendation by Borenstein, Hedges, Higgins, and Rothstein (2009), a random-effects model was applied because it takes into account that the true prevalence rate may vary across different studies. Heterogeneity was quantified using Q, tau-squared (τ²), and I-squared (I²) statistics. Data transformation was incorporated in terms of logit transformation.

Separate meta-analyses were carried out for PTSD and PTSS as well as for mothers and fathers/partners, yielding four separate data sets (i.e., PTSD in mothers, PTSS in mothers, PTSD in fathers/partners, and PTSS in fathers/partners). Where applicable, subsamples from one study were allocated to different data sets. An instrument was considered for assessment of PTSD if it included the complete set of DSM-IV diagnostic criteria A-F (APA, 1994) or DSM-5 criteria A-G (APA, 2013). Criterion H (which excludes symptoms due to physiological effects of a substance or medical condition) was defined as optional for the purpose of this review. Rates were categorized as PTSS if they only targeted the symptom criteria (i.e., B-D in case of DSM-IV and B-E in case of DSM-5), with no or incomplete assessment of other diagnostic criteria. Subgroups of risk status were introduced to obtain prevalence estimates for non-targeted (i.e., without specific risk characteristics) and targeted samples (i.e., with potential risk due to a variety of characteristics). The original study authors usually proposed whether their sample was at risk for PTSS or PTSD. This appraisal was adopted if it was supported by existing literature on risk factors for birth-related posttraumatic stress (e.g., Dekel et al., 2017).

Whenever studies reported data on both, diagnosis and symptom level, only prevalence for diagnoses was used because it was not always clear whether PTSD cases were included in reported PTSS rates or not. For studies incorporating more than one instrument to assess PTSD or PTSS for the whole sample, only results for the instrument with the strongest psychometric properties (e.g., clearly defined cutoff) were chosen. Studies with subsamples of different risk statuses were separately assigned to the relevant analysis if reported data allowed for this distinction. Moderators besides risk status were tested via random-effects meta-regression to examine whether participants’ age, percentage of primiparous mothers, percentages of assisted vaginal birth and emergency c-section, assessment type, year of publication, geographical region, or risk of bias explain a significant amount of variance between the studies. The method of moments (also known as DerSimonian and Laird method) was used for estimation of τ² and, as recommended by Borenstein et al. (2009), rather than utilizing a Z-distribution, the Knapp-Hartung method was applied for random-effect models.

To explore the potential effect of outliers, leave one out analyses were conducted (Borenstein et al., 2009). Publication bias was assessed visually by inspection of funnel plots and statistically using Egger’s test and Begg and Mazumdar’s rank correlation test.

### 3. Results

#### 3.1. Study selection

The systematic database search brought forward more than 15,000 results (s. Fig. 2). They were downloaded into a reference management program (Zotero) and more than half of them were discarded as
duplicates. Including the records found through manual searches, a total of 7654 records were screened for eligibility based on title and abstract. The vast majority of studies were excluded in this first step, most of them due to lack of primary quantitative data (e.g., reviews, case studies, editorials), because the keyword of trauma was meant in a medical sense only (i.e., physical trauma) instead of referring to psychological trauma, or because they focused on samples other than parents. This left 812 references for full-text review, leading to the final inclusion of 169 studies.

3.2. Characteristics of included studies

The included 169 studies were published journal articles except for 14 dissertations, two theses, and one unpublished open access preprint. Of the 169 studies, six were in French, one in German, and one in Czech, the rest of them were written in English. The reports were published between 1994 and 2021 and originated in 29 different countries—most coming from Europe (k = 97), followed by West Asia (k = 30), and North America (k = 28). The most frequently represented countries were the United Kingdom (k = 34) and the United States (k = 22). Samples were predominantly female (k = 153). Only five studies focused on fathers and 11 studies included both, mothers and fathers. Although this review was open to include partners who were not the biological father, all collected studies explored only mothers and/or biological fathers. Participants were on average 30.95 years old (SD = 2.4). Thirteen studies reported on primiparous women only, with a mean of 58% of primiparous women in all studies. Less than half of the studies specified the type of birth method including assisted vaginal birth and emergency c-section. For those that did, on average 13% of the participants experienced assisted vaginal birth and 20% had an emergency c-section. Eight and 14 studies excluded participants with assisted vaginal birth and emergency c-section, respectively. Time postpartum varied between one and 14 months, 28 studies performed multiple measurements of PTSD/PTSS within this time frame. Samples were labelled as targeted if they had specific characteristics implying risk for PTSD/PTSS development. This comprised traumatic birth experience, fear of childbirth, prematurity, admission to the neonatal intensive care unit (NICU), emergency c-section, stillbirth, pregnancy complications, ethnic minorities, and maternal history of trauma. Appendix C provides an overview of the studies’ characteristics.

As previously stated, a study had to fulfill minimum quality criteria to be included in statistical analyses (i.e., n ≥ 20, a minimum of three out of all nine JBI checklist items fulfilled, including at least one of the three major domains). Applying these criteria, 15 studies were excluded because risk of bias was deemed too high for statistical analyses, leaving 154 studies. Of the remaining studies, 68% were categorized as high risk of bias (< 60% of JBI items rated “yes”) and 32% as low risk of bias (≥ 60% of JBI items rated “yes”). Fig. 3 shows the distribution of ratings on each of the nine JBI checklist items (Munn et al., 2015), please refer to Appendix D for further details. Although 83% of included studies fulfilled the validity of measurement item, it should be noted that validity of measurement was based on general validation only. Thus, prior psychometric validation of the instruments in postpartum populations or for the specific translated or adapted version used is not guaranteed. A particularly high proportion of studies was rated “unclear” in the domains sample, sampling method, data-analysis, and response rate. Almost two thirds of the studies did not fulfill the JBI checklist’s sample size criterion. This item was checked by comparing the reported sample size to the one necessary according to power analyses. If studies did not report power analyses, the sample size considered necessary to estimate a prevalence rate with good precision was calculated (Daniel, 1999; Naing, Winn, & Rusli, 2006).

Sample sizes spanned from n = 21 to n = 4438, providing a total N = 54,711 for quantitative synthesis. Instruments to measure posttraumatic stress (PTSD or PTSS) were manifold and included self-report measures (k = 121) almost 8 times more often than clinical interviews (k = 16). In total, this meta-analysis reports on the results of 29 different measures of posttraumatic stress. PTSD rates were derived from the Clinician-Administered PTSD scale (CAPS, Blake et al., 1995), Clinician-Administered PTSD scale for DSM-5 (CAPS-5, Weathers et al., 2018), Composite International Diagnostic Interview for women (CIDI-V, Steinert et al., 2007), City BiTS (Ayers et al., 2018), Childbirth-Related PTSD Questionnaire (CR-PTSD-Q, Zeníšková, 2019), Mini-International Neuropsychiatric Interview (MINI, Sheehan et al., 1998), National Women’s Study PTSD module (NWS-PTSD, Kilpatrick, Resnick, Saunders, & Best, 1989), Posttraumatic Diagnostic Scale (PDS, Foa, Cashman, Jaycox, & Perry, 1997), Posttraumatic Stress Disorder Symptom Scale Interview for DSM-5 (PSSI-5, Foa et al., 2016), SCID (First et al., 1996), and Traumatic Event Scale (TES, Wijma, Soderquist, & Wijma, 1997). Scoring rules were based on DSM-IV or DSM-5 (APA, 1994, 2013). Prevalence rates of PTSS were obtained via the Davidson Trauma Scale (DTS, Davidson et al., 1997), IES (Horowitz et al., 1979)
(in one study in combination with items from the General Health Questionnaire [GHQ, Goldberg, 1988]), IES-R (Weiss & Marmar, 1997), Los Angeles Symptom Checklist (LASC, King, King, Leskin, & Foy, 1995), Mississippi Scale for PTSD (M-PTSD, Keane, Caddell, & Taylor, 1988), PTSD-Checklist (PCL, Weathers, Litz, Herman, Huska, & Keane, 1993) (also in specific PCL-S and civilian PCL-C version), PTSD-Checklist for DSM-5 (PCL-5 Weathers et al., 2013), Primary Care PTSD Screen (PC-PTSD-IV, Cameron & Gusman, 2003), Perinatal Posttraumatic Stress Disorder Questionnaire (PPW, DeMier, Hynan, Harris, & Manniello, 1996), modified PPQ (mPPQ, Callahan, Borja, & Hynan, 2006), PTSD Symptom Scale (Foà, Riggs, Dancu, & Rothbaum, 1993) (in interview PSS-I and self-report PSS-SR version), Posttraumatic Stress Disorder Symptom Scale (Evren et al., 2016), and Trauma Screening Questionnaire (TSQ, Brewin et al., 2002). There was a mixture of applied scoring rules as well as cutoff scores to determine clinically relevant symptom levels, in some cases varying for the same instrument (e.g., IES and IES-R).

3.3. Prevalence of birth-related PTSD/PTSS in mothers

An overview of prevalence rates for all populations is provided in Table 1 and forest plots are displayed in Figs. 4, 5, 6, and 7, respectively. All results are reported in more detail in the following paragraphs.

3.3.1. Posttraumatic Stress Disorder

Prevalence of birth-related PTSD among mothers was summarized based on 46 studies, including 51 subsamples. PTSD was assessed through clinical interviews in 14 and via self-report measures in 32 studies. Prevalence rates ranged from 0.0% to 30.0%. The expected high level of heterogeneity was statistically confirmed with $Q = 721.383$ and $I^2 = 93.069%$ ($p < .001$), indicating substantial variance between studies. The overall estimated prevalence rate was $4.7\%$ (95\% CI [0.036, 0.061]), with slightly lower rates in non-targeted ($k = 40$, 4.4\%, 95\% CI [0.033, 0.059]) than in targeted samples ($k = 11$, 6.8\%, 95\% CI [0.033, 0.125]). The difference between targeted and non-targeted samples was not statistically significant ($Q = 1.160$, $p = .282$).

### Table 1

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean prevalence</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>Mothers</td>
<td>4.7%</td>
</tr>
<tr>
<td>Non-targeted samples</td>
<td>4.4%</td>
<td>[0.033, 0.059]</td>
</tr>
<tr>
<td>Targeted samples</td>
<td>6.8%</td>
<td>[0.033, 0.135]</td>
</tr>
<tr>
<td>Overall</td>
<td>12.3%</td>
<td>[0.175, 0.252]</td>
</tr>
<tr>
<td>Non-targeted samples</td>
<td>6.7%</td>
<td>[0.054, 0.084]</td>
</tr>
<tr>
<td>Targeted samples</td>
<td>21.1%</td>
<td>[0.175, 0.252]</td>
</tr>
</tbody>
</table>

3.3.2. Posttraumatic Stress Symptoms

Ninety-four studies with 114 subsamples were incorporated in the calculation of the overall prevalence of birth-related PTSS in mothers, composing the largest of the four data sets. Except for three studies with interview assessment, all studies utilized self-report measures. Reported prevalence ranged from 0.0\% (non-targeted samples) to 90.0\% (targeted sample due to traumatic experience of stillbirth). High $Q$ (2430.489) and $I^2$ (95.351\%) values confirmed the presence of notable heterogeneity ($p < .001$). The pooled prevalence was 12.3\%, 95\% CI [0.107, 0.252]. Again, non-targeted samples ($k = 80$, 6.7\%, 95\% CI [0.054, 0.084]) showed lower rates than targeted ones ($k = 34$, 21.1\%, 95\% CI [0.175, 0.252]), in this case reaching statistical significance ($Q = 61.260$, $p < .001$).

3.4. Prevalence of birth-related PTSD/PTSS in fathers

3.4.1. Posttraumatic Stress Disorder

Four studies with six subsamples reported prevalence of birth-related PTSD in fathers, making this the smallest of the four data sets. Interviews were conducted in two studies, whereas the other two studies assessed participants via self-report measures. As may be expected given the small number of studies (including four subgroups in which no participants filled cutoffs for PTSS), studies in this data set were relatively homogenous ($Q = 12.498$, $I^2 = 59.955\%$, $p = .029$) compared to the ones with female participants. Reported prevalence varied from 0.0\% to 7.2\%, yielding a mean rate of 1.2\%, 95\% CI [0.004, 0.035]. The pattern of higher rates for targeted compared to non-targeted samples did not emerge in this group. On the contrary, overall prevalence was slightly higher for non-targeted ($k = 3$, 2.1\%, 95\% CI [0.003, 0.135]) than targeted groups ($k = 3$, 0.8\%, 95\% CI [0.002, 0.033]), but this difference was not statistically significant ($Q = 0.596$, $p = .440$). Meta-analytic results for this data set rely on a very limited sample size, therefore possessing explorative relevance at the most.

3.4.2. Prevalence of birth-related PTSD/PTSS in fathers

Ten studies with 12 subsamples were included in the analysis of prevalence of birth-related paternal PTSS, all using self-report measures. Heterogeneity indicators showed lower values than for the results in mothers, but still suggest a high level of variance ($Q = 125.754$, $I^2 = 91.253\%$, $p < .001$), which is in line with prevalence rates ranging from 0.0\% to 34.7\%. Calculated mean prevalence was 1.3\% (95\% CI [0.006, 0.025]), differing slightly between non-targeted ($k = 9$, 1.2\%, 95\% CI [0.006, 0.023]) and targeted samples ($k = 3$, 5.5\%, 95\% CI [0.004, 0.0490]), although this difference was not statistically significant ($Q = 1.175$, $p = .278$).

3.5. Exploration of the course of maternal PTSD/PTSS over time

For an indication of prevalence in mothers over time, combined estimates were obtained for time points where results were available from more than one study. The results including 95\% CI intervals are shown in Fig. 8. The number of studies reporting on single time points varied widely, from only five studies up to 64 studies providing data for a specific time point, resulting in wide confidence intervals for some time points. Interpretations concerning PTSD/PTSS development in the first year postpartum therefore involve considerable statistical uncertainty and the inconsistent collocation of studies across the different time points does not allow for reliable conclusions. However, available data suggest a modest rise in prevalence of PTSS and a decline of PTSD cases from the first ten weeks to one year postpartum (s. Fig. 8).

Studies reporting data on PTSD in fathers were all conducted between four and six weeks postpartum with one exception of assessment at three months after birth, therefore not providing enough data to explore development over time. Data for paternal PTSS were similarly limited and results inconclusive, which is why they are not presented here.
### 3.6. Test of moderating effects on the prevalence of PTSD/PTSS in random-effects meta-regression

Given the limited number of studies reporting on fathers, meta-regression was performed only for maternal data.

Concerning the PTSD data set, age, parity, and assessment type had no impact on prevalence. Of the tested covariates, year of publication (R² = 0.09) and geographical region (R² = 0.07) explained the largest amount of variance between studies. A scatterplot suggested increasing prevalence rates with more recent year of publication (range in this data set was 2002 to 2022). The graph for geographical region indicated lowest prevalence rates in North America and highest rates in West Asia and Australia. This effect was statistically significant (p = .018). Attempting to simplify categorization of regions (as they were not equally represented in this data set), countries were divided into developed and developing countries or countries in transition, as per United Nations’ definition (UN, 2020). The explained variance was R² = 0.06, with slightly lower mean rates in developed countries, but not reaching statistical significance (p = .069). Assisted vaginal birth and emergency c-section each only explained a negligible amount of
<table>
<thead>
<tr>
<th>Author(s), year</th>
<th>Subgroup</th>
<th>Time point</th>
<th>Assessment type</th>
<th>Risk of bias</th>
<th>Event rate</th>
<th>Lower limit</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmadnejad et al., 2021</td>
<td>non-targeted</td>
<td>2.5-4</td>
<td>self-report</td>
<td>low</td>
<td>0.477</td>
<td>0.426</td>
<td>0.528</td>
</tr>
<tr>
<td>Allen, 1996</td>
<td>non-targeted</td>
<td>7-9</td>
<td>self-report</td>
<td>high</td>
<td>0.048</td>
<td>0.023</td>
<td>0.097</td>
</tr>
<tr>
<td>Allen, 1999</td>
<td>non-targeted</td>
<td>1</td>
<td>self-report</td>
<td>high</td>
<td>0.103</td>
<td>0.047</td>
<td>0.212</td>
</tr>
<tr>
<td>Ayers &amp; Pickering, 2001</td>
<td>non-targeted</td>
<td>1.5</td>
<td>self-report</td>
<td>high</td>
<td>0.069</td>
<td>0.042</td>
<td>0.111</td>
</tr>
<tr>
<td>Ayers et al., 2007</td>
<td>non-targeted</td>
<td>2.25</td>
<td>self-report</td>
<td>high</td>
<td>0.047</td>
<td>0.015</td>
<td>0.135</td>
</tr>
<tr>
<td>Bailham, 2001</td>
<td>non-targeted</td>
<td>1.25-2</td>
<td>self-report</td>
<td>high</td>
<td>0.036</td>
<td>0.014</td>
<td>0.093</td>
</tr>
<tr>
<td>Berman et al., 2021</td>
<td>non-targeted</td>
<td>3 (1.8)</td>
<td>self-report</td>
<td>low</td>
<td>0.163</td>
<td>0.134</td>
<td>0.196</td>
</tr>
<tr>
<td>Boudu et al., 2007</td>
<td>non-targeted</td>
<td>1.5</td>
<td>self-report</td>
<td>high</td>
<td>0.011</td>
<td>0.002</td>
<td>0.072</td>
</tr>
<tr>
<td>Brandon et al., 2011</td>
<td>non-targeted</td>
<td>1</td>
<td>self-report</td>
<td>low</td>
<td>0.016</td>
<td>0.001</td>
<td>0.206</td>
</tr>
<tr>
<td>Catala et al., 2020</td>
<td>non-targeted</td>
<td>4</td>
<td>self-report</td>
<td>high</td>
<td>0.078</td>
<td>0.041</td>
<td>0.142</td>
</tr>
<tr>
<td>Chen et al., 2020</td>
<td>non-targeted</td>
<td>1.4</td>
<td>self-report</td>
<td>high</td>
<td>0.089</td>
<td>0.069</td>
<td>0.114</td>
</tr>
<tr>
<td>Cigolì et al., 2006</td>
<td>non-targeted</td>
<td>3-6</td>
<td>self-report</td>
<td>low</td>
<td>0.013</td>
<td>0.003</td>
<td>0.040</td>
</tr>
<tr>
<td>Cohens et al., 2004</td>
<td>non-targeted</td>
<td>2-2.5</td>
<td>self-report</td>
<td>low</td>
<td>0.003</td>
<td>0.000</td>
<td>0.039</td>
</tr>
<tr>
<td>Creedy et al., 2000</td>
<td>non-targeted</td>
<td>1-1.5</td>
<td>interview</td>
<td>high</td>
<td>0.056</td>
<td>0.039</td>
<td>0.080</td>
</tr>
<tr>
<td>Davies et al., 2008</td>
<td>non-targeted</td>
<td>1.5</td>
<td>self-report</td>
<td>low</td>
<td>0.030</td>
<td>0.015</td>
<td>0.059</td>
</tr>
<tr>
<td>Deferes et al., 2021</td>
<td>non-targeted</td>
<td>2</td>
<td>self-report</td>
<td>low</td>
<td>0.019</td>
<td>0.014</td>
<td>0.026</td>
</tr>
<tr>
<td>Denis et al., 2011</td>
<td>non-targeted</td>
<td>1</td>
<td>self-report</td>
<td>low</td>
<td>0.051</td>
<td>0.027</td>
<td>0.094</td>
</tr>
<tr>
<td>Di Blasio et al., 2015a</td>
<td>non-targeted</td>
<td>3</td>
<td>self-report</td>
<td>high</td>
<td>0.262</td>
<td>0.169</td>
<td>0.381</td>
</tr>
<tr>
<td>Di Blasio et al., 2015b</td>
<td>non-targeted</td>
<td>3</td>
<td>self-report</td>
<td>high</td>
<td>0.056</td>
<td>0.024</td>
<td>0.128</td>
</tr>
<tr>
<td>Di Blasio et al., 2018</td>
<td>non-targeted</td>
<td>1</td>
<td>self-report</td>
<td>low</td>
<td>0.029</td>
<td>0.009</td>
<td>0.085</td>
</tr>
<tr>
<td>Doornbos, 2009</td>
<td>non-targeted</td>
<td>1.5</td>
<td>self-report</td>
<td>high</td>
<td>0.031</td>
<td>0.008</td>
<td>0.115</td>
</tr>
<tr>
<td>Edworthy et al., 2008</td>
<td>non-targeted</td>
<td>1.5</td>
<td>self-report</td>
<td>high</td>
<td>0.009</td>
<td>0.001</td>
<td>0.063</td>
</tr>
<tr>
<td>Fioukouli et al., 2015</td>
<td>non-targeted</td>
<td>1.5-6</td>
<td>interview</td>
<td>high</td>
<td>0.320</td>
<td>0.276</td>
<td>0.367</td>
</tr>
<tr>
<td>Furnet et al., 2014</td>
<td>non-targeted</td>
<td>1.5-2</td>
<td>self-report</td>
<td>low</td>
<td>0.035</td>
<td>0.027</td>
<td>0.044</td>
</tr>
<tr>
<td>Gankanda et al., 2021</td>
<td>non-targeted</td>
<td>1</td>
<td>self-report</td>
<td>low</td>
<td>0.027</td>
<td>0.012</td>
<td>0.058</td>
</tr>
<tr>
<td>Ghorbani et al., 2014</td>
<td>non-targeted</td>
<td>2</td>
<td>self-report</td>
<td>high</td>
<td>0.012</td>
<td>0.002</td>
<td>0.080</td>
</tr>
<tr>
<td>Gökçe Isbir et al., 2016a</td>
<td>non-targeted</td>
<td>1.5-2</td>
<td>self-report</td>
<td>low</td>
<td>0.015</td>
<td>0.005</td>
<td>0.039</td>
</tr>
<tr>
<td>Gökçe Isbir et al., 2016b</td>
<td>non-targeted</td>
<td>1.5-2</td>
<td>self-report</td>
<td>low</td>
<td>0.015</td>
<td>0.005</td>
<td>0.038</td>
</tr>
<tr>
<td>Goutard et al., 2012</td>
<td>non-targeted</td>
<td>1.5</td>
<td>self-report</td>
<td>low</td>
<td>0.061</td>
<td>0.028</td>
<td>0.130</td>
</tr>
<tr>
<td>Haagen et al., 2015</td>
<td>non-targeted</td>
<td>3</td>
<td>self-report</td>
<td>low</td>
<td>0.014</td>
<td>0.006</td>
<td>0.030</td>
</tr>
<tr>
<td>Hirston et al., 2018</td>
<td>non-targeted</td>
<td>1-3</td>
<td>self-report</td>
<td>high</td>
<td>0.053</td>
<td>0.024</td>
<td>0.112</td>
</tr>
<tr>
<td>Hajizadeh &amp; Mirghafourvand, 2021</td>
<td>non-targeted</td>
<td>1</td>
<td>interview</td>
<td>low</td>
<td>0.163</td>
<td>0.125</td>
<td>0.210</td>
</tr>
<tr>
<td>Hall et al., 2017</td>
<td>non-targeted</td>
<td>1-9</td>
<td>self-report</td>
<td>low</td>
<td>0.234</td>
<td>0.180</td>
<td>0.298</td>
</tr>
<tr>
<td>Halperin et al., 2015</td>
<td>non-targeted</td>
<td>1.5-2</td>
<td>self-report</td>
<td>high</td>
<td>0.094</td>
<td>0.058</td>
<td>0.147</td>
</tr>
<tr>
<td>Harrison et al., 2021</td>
<td>non-targeted</td>
<td>median 7.75</td>
<td>self-report</td>
<td>low</td>
<td>0.023</td>
<td>0.019</td>
<td>0.028</td>
</tr>
<tr>
<td>Hernández-Martínez et al., 2019</td>
<td>non-targeted</td>
<td>4-6</td>
<td>self-report</td>
<td>low</td>
<td>0.106</td>
<td>0.056</td>
<td>0.118</td>
</tr>
<tr>
<td>Holt et al., 2018</td>
<td>non-targeted</td>
<td>2-6</td>
<td>self-report</td>
<td>low</td>
<td>0.120</td>
<td>0.104</td>
<td>0.138</td>
</tr>
<tr>
<td>Iles et al., 2014</td>
<td>non-targeted</td>
<td>1.5</td>
<td>self-report</td>
<td>low</td>
<td>0.033</td>
<td>0.016</td>
<td>0.068</td>
</tr>
<tr>
<td>Kersting et al., 2009</td>
<td>non-targeted</td>
<td>6</td>
<td>self-report</td>
<td>high</td>
<td>0.009</td>
<td>0.001</td>
<td>0.131</td>
</tr>
<tr>
<td>Kim et al., 2015</td>
<td>non-targeted</td>
<td>1</td>
<td>self-report</td>
<td>low</td>
<td>0.094</td>
<td>0.040</td>
<td>0.207</td>
</tr>
</tbody>
</table>

Fig. 5. Forest plot of maternal PTSS studies by risk status.

Note. CI = confidence interval; HELLP = HELPP syndrome; PE = preeclampsia; PROM = preterm labor rupture of membranes. Time point in months postpartum (M (SD), time frame, or single time point).

variance ($R^2 = 0.02$). However, only 26 and 29 of 46 studies reported rates of assisted vaginal delivery and emergency c-section, respectively, limiting the informative value. The studies’ risk of bias impacted prevalence only marginally ($R^2 = 0.01, p = .536$).

Differences in the data set of maternal birth-related PTSS were explored for the same covariates, except for assessment type, as only three of the studies did not use self-report instruments. Parity, proportion of assisted vaginal birth and emergency c-section, year of publication, and risk of bias were not associated with prevalence. Again, information on the two modes of birth was missing for a large amount of the studies. Although year of publication did not explain a significant amount of variance, the pattern of a slight increase in prevalence over time reappeared in the scatterplot. Geographical region explained 18% of between-study variance, with the lowest rates in Europe. As only one study was left out.

Removal of any single study did not influence maternal PTSS and PTSD prevalence estimates by more than 0.4% and 0.3%, respectively. Paternal rates changed up to 0.4% for PTSS and 0.6% for PTSD when one study was left out.

Significant results in Begg and Mazumdar’s rank correlation test ($p < .001$ for PTSS, $p < .05$ for PTSD) and Egger’s test ($p < .05$ for PTSS, $p < .01$ for PTSD) also support the presence of publication bias for mothers. Following the recommendation made by Higgins et al. (2008), these analyses were not conducted for fathers because the power of the tests might be “too low to distinguish chance from real asymmetry” (p. 317) due to the limited number of available studies.

4. Discussion

This systematic review and meta-analysis summarizes research from the past 25 years on birth-related PTSD/PTSS among parents. The screening process yielded 169 studies, 154 of which were included in the meta-analytic procedures, representing a total of 54,711 participants. The primary objective of this study was to obtain prevalence rates for
Fig. 5. (continued).
samples of mothers and fathers with and without risk status between one and 14 months postpartum. Moreover, the impact of different variables on prevalence rates was explored via meta-regression.

4.1. Prevalence rates

Mean prevalence of birth-related PTSD was found to be 4.7% in mothers and 1.2% in fathers. For birth-related PTSS, results indicated prevalence rates of 12.3% in mothers and 1.3% in fathers. The overall maternal birth-related PTSD rate of 4.7% found in this review is in line with the range of rates determined by Dikmen-Yildiz et al. (2017), Grekin and O’Hara (2014), and Dekel et al. (2017). Nonetheless, because of differences in inclusion criteria (e.g., conceptually more rigorous focus on childbirth as traumatic stressor and time frame covering the first year postpartum) and the larger number of studies, comparison is difficult. For instance, the present review not only roughly confirms the overall prevalence of 5.4% documented by Dikmen-Yildiz et al. (2017), but also extends generalizability by including more recent and non-English studies. However, our subgroup analyses yielded a much lower rate for targeted samples (6.8%) than previously estimated rates of 18.5% for high-risk (Dikmen-Yildiz et al., 2017) or 14.1% for targeted samples (Grekin & O’Hara, 2014). This difference might be explained by the diverse compilation of participants’ risk statuses across included studies. For example, Dekel et al. (2017) defined the following at-risk groups: preeclampsia, low-income Latinas, traumatic delivery, women with experience of child abuse and neglect,
whereas no clear criteria for the classification of “targeted samples” was provided by Grekin and O’Hara (2014).

Prevalence for birth-related PTSS among mothers was 6.7% in non-targeted samples and 21.1% in targeted samples. To date, the only prior review taking PTSS into consideration was conducted by Dekel et al. (2017). They divided studies into two subgroups according to quality ratings (especially considering whether the studies had controlled for PTSS before childbirth) but did not distinguish between samples’ risk status. The overall maternal birth-related PTSS prevalence of 12.3% in the current meta-analysis lies between the rates for high-quality (9.6%) and low-quality studies (16.8%) reported by Dekel et al. (2017).

As expected, mothers’ PTSS rates were higher than their PTSD rates. However, the difference was small in non-targeted samples (4.4% PTSD vs. 6.7% PTSS) and large in targeted samples (6.8% PTSD vs. 21.1% PTSS). These findings should be interpreted with caution, as long as evidence regarding the characteristics defining targeted groups is inconclusive.

The limited amount of literature pertaining to paternal birth-related PTSD/PTSS makes it difficult to put the results of this review into context. With rates of 1.2% for PTSD and 1.3% for PTSS, prevalence was significantly lower in fathers than in mothers. This is in line with prior

Fig. 9. Funnel plots of maternal PTSD/PTSS studies with observed and imputed studies. Note. Imputed studies calculated by Duval and Tweedie’s Trim and Fill method based on random effects model.
Evidence suggesting that development of PTSD/PTSS is generally more likely after direct exposure to an event as compared to witnessing it vicariously (May & Wisco, 2016), i.e., secondary traumatic stress (e.g., Creedy et al., 2000). Further, a gender difference is also found in general PTSD prevalence, with women consistently showing higher PTSD rates than men (Kilpatrick et al., 2013). The almost identical mean rates for PTSD and PTSS among fathers are more surprising and may be attributed to the lack of research in this population, as no convincing etiological explanations are available. Given the confidence intervals’ proximity to zero, it is possible that fathers will only very rarely develop PTSD/PTSS related to childbirth. However, qualitative research suggests that fathers can experience childbirth as traumatic and that they are especially at risk of not finding recognition and support for their distress (e.g., Daniels, Arden-Close, & Mayers, 2020; Etheridge & Slade, 2017; Harvey & Pattison, 2012). Future studies need to examine this quantitatively in order to gain a better understanding of the impact of traumatic birth experiences on the entire family system.

Fig. 10. Funnel plots of paternal PTSD/PTSS studies with observed and imputed studies. Note. Imputed studies calculated by Duval and Tweedie’s Trim and Fill method based on random effects model.
4.2. Moderating effects

Meta-regression of maternal PTSD and PTSS data only explained a small amount of the substantial heterogeneity between studies. Including assessment type as covariate showed no statistical effect on prevalence, confirming the most recent meta-analytic results (Dikmen-Yildiz et al., 2017). Contradicting effects have been found for other psychological disorders, like antenatal depression, where rates seem to be overestimated when assessed through self-report measures (Yin et al., 2021). Even though self-report measures alone are not recommended as tools for diagnosing PTSD (Friedman, 2010) or psychological disorders in general (Antony & Barlow, 2002), our results support their use in acquiring valid prevalence rates.

Geographical region and a country’s status as developed or developing impacted prevalence rates in this meta-analysis. In contrast, Dikmen-Yildiz et al. (2017) found no apparent statistical effect of geographical regions. However, studies on general PTSD or other specific trauma samples documented evidence for differences in prevalence even across European countries (Burri & Maercker, 2014) and for the influence of cultural beliefs on PTSD rates (Oakley, Kuo, Kowalkowski, & Park, 2021). Reviews on perinatal mental disorders (Fisher et al., 2012) and postnatal depression (Dadi, Miller, & Mwanzi, 2020) in low- and middle-income countries have shown high rates compared to high-income countries, which the authors trace back to socioeconomic status, limited access to maternal care, and gender-based violence. The current review’s findings support the notion that this also applies to birth-related PTSD and PTSS across countries and cultures. Future syntheses may be able to detect specific variables responsible for the cultural impact on prevalence rates.

In contrast to Dikmen-Yildiz et al. (2017), lower risk of bias was not associated with lower prevalence. Given the large number of studies with high or unclear risk of bias in our study, this may be reassuring for the generalizability of results. Despite the studies’ varying quality in terms of risk of bias, it is unlikely that studies with high risk of bias distorted the results in this review. Sensitivity analyses further supported robustness of the results.

Unlike the most recent meta-analysis (Dikmen-Yildiz et al., 2017), our meta-regression regarding the year of publication showed an effect for birth-related PTSD. Earlier studies seemed to show lower PTSD rates than more recent ones, which might be explained by the increase in research and awareness over the last decade. As clinicians become more aware of the clinical presentation, they may be less likely to misdiagnose birth-related PTSD as postpartum depression (Alder, Stadmayer, Tschudin, & Bitzer, 2006). Increasing rates could also be a sign of the stigma attached to psychological disorders in the postpartum period being replaced by psychoeducation, possibly encouraging more parents to acknowledge their symptoms and to seek adequate help. For these reasons, the trend towards higher rates may not represent over-estimation of PTSD prevalence.

4.3. Contribution of this systematic review and meta-analysis

This systematic review and meta-analysis was conducted in line with widely recognized guidelines to ensure high methodological rigor, e.g., JBI Manual (Aromataris & Munn, 2020) and Cochrane Handbook (Higgins et al., 2008). Preregistration of the protocol contributed to transparent scientific practice and integrity of this review. A wide selection of databases was searched and screening decisions as well as risk of bias assessment were conducted by two independent raters. The study stands out from the literature on birth-related PTSD due to its inclusive and broad approach. Unlike previous reviews, the population of interest was widened to include not only mothers, but also fathers. A quantitative synthesis of paternal data on birth-related PTSD has, to the best knowledge of the authors, not been attempted before. This clearly adds to the innovation of this study and marks the way towards equal attention to both parents in perinatal mental health research. The restriction to mothers in previous research falls short of this unique kind of trauma, which affects the whole family rather than just one individual.

Additionally, this review explored PTSD as well as PTSS and thus has created a more complete depiction of distress caused by traumatic birth experiences. Including PTSD and PTSS provided the opportunity to address Dikmen-Yildiz et al.’ (2017) concern that restriction to full diagnostic criteria may lead to an underestimation of prevalence rates. The review also aspired to include research from around the world, published in English or other languages. Precautions were taken against excluding possibly relevant work beginning with an extensive database search, followed by a thorough screening process. All these considerations and measures aimed at covering the largest possible extent of research on parental birth-related PTSD/PTSS. More specifically, with these approaches we successfully integrated more than twice as many studies with six times as many participants as the most recent review (Dikmen-Yildiz et al., 2017).

At the same time, the clinical focus was strengthened by clearly defining the relation to childbirth as the underlying traumatic event. This helps to disentangle nonspecific cases of PTSD in the postpartum period from those clearly caused or triggered by birth itself (Harrison, Ayers, Quigley, Stein, & Alderdice, 2021).

4.4. Limitations

The goal of gathering a large number of studies was successful regarding studies compiled on mothers. However, none of the included studies explored partners and the available data on fathers remained limited, restricting performable statistical analyses and explanatory power concerning this population. This highlights the need for future research on parental distress, not limited to mothers. Ideally, both parents should be recruited to create the best possible basis for comparing and interpreting psychological symptoms.

Because studies only sporadically reported the samples’ sociodemographic and obstetric characteristics, these factors could only be explored as covariates to some extent and drawing clear conclusions about their association with prevalence falls beyond the scope of this review. Further, it was not possible to explore other potentially influential variables, such as ethnicity or socioeconomic status, due to inconsistent and/or incomplete description in primary studies. Statistical and graphical explorations suggest that results have been affected by publication bias, despite all attempts to prevent this. Specifically, the search was not limited to English articles, there were no restrictions regarding a study’s geographical origin, and studies were considered for inclusion regardless of their publication status. Implementation of this plan only failed for two articles published in Persian, due to lack of language expertise. It is however noteworthy that the number of non-English articles and studies exploring samples outside of Europe or North America was small.

As is true for all reviews, this systematic review and meta-analysis can only explore available and published data. The validity of interpretation therefore is reliant on the quantity and quality of compiled studies. Compared to maternal data, the meta-analyses on fathers’ PTSD/PTSS rely on a much smaller number of studies. Therefore, caution is warranted when interpreting these results and findings of this study should be viewed as preliminary regarding paternal data.

Missing or vague information in the articles posed a challenge throughout the screening process, data extraction, and quality assessment. For example, several studies did not specify the underlying traumatic event related to their PTSD or PTSS measure. Whenever information for screening decisions was missing and could not be obtained from the author(s), the study had to be excluded. Although this might have led to the exclusion of suitable studies, the effort put into contacting the respective authors should be emphasized. These difficulties also arose during quality assessment. In many cases, it was impossible to distinguish whether details on a specific quality domain
were lacking or not transparently recorded in the study.

The overall summary of the studies’ risk of bias according to the assessment of the JBI checklist (high risk of bias in 68% of included studies) is cause for concern. It should be stressed that the larger part of studies did not primarily aim to determine a valid prevalence rate. Hence, it is not surprising that no study fulfilled all nine items. The advantage of applying the JBI checklist is the use of a standardized measure, suitable for different quantitative study designs. Alternative instruments including summary scores and scales are generally discouraged due to lack of empirical foundation of seemingly detailed and clear ratings, the tendency towards unreliability, and the unlikelihood of transparency to readers (Higgins et al., 2008). However, the usefulness of dichotomous classification of the nine domains is debatable, as it only allows for a simplistic rating. The authors would like to encourage further research and more detailed guidelines for authors on how to assess study quality, including specific operationalization of single items (e.g., which response rates may be considered acceptable).

The downside of the review’s inclusive approach is the resulting heterogeneity among integrated studies. In particular, definitions of significant levels of posttraumatic stress, labelled PTSS in this review, varied substantially. The division of PTSD and PTSS, though based on transparent criteria, remains arbitrary in some respects, as to date there is “no clear agreement on how partial or subclinical PTSD should be defined” (Ayers & Ford, 2016, p. 186). Most of the applied cutoffs (used for PTSS in this study) are based on validation studies in which the use of the score has proven high sensitivity and specificity compared to interview diagnostics, although the assessment itself does not include all criteria necessary for a clinical determination (e.g., PCL-5, as shown in Bovin et al., 2016).

Besides the issue of categorization of the assessment as PTSD or PTSS, some instruments did not have clearly determined cutoffs. Two prominent examples are the IES (Horowitz et al., 1979) and its successor the IES-R (Weiss & Marmar, 1997), neither of which were intended to be used in combination with a cutoff score by the original authors (Weiss, 2004). Our review highlights that this original intention has been widely ignored in research because the IES and IES-R were the most frequently deployed measures in the data set of maternal PTS. Studies included in this meta-analysis in which PTSS was measured with the IES (Horowitz et al., 1979) used cutoff scores ranging from nine to 41, in some cases applied to only one subscale, in others to both. This inconsistency may cause data noise and thus explain the small amount of between-study variance explained through meta-regression.

Another potential cause of data noise lies in the lack of clear definitions of risk status and the subsequent classification as “targeted” or “non-targeted”. To the knowledge of the authors, there is no scientific agreement regarding the specific sample characteristics directly increasing the risk for birth-related PTSD or PTSS. This review offers preliminary findings but does not contribute final or novel definitions of risk status.

4.5. Implications for research and practice

4.5.1. Objectives for future research

To accurately determine prevalence rates of birth-related PTSS in parents, a clear definition of relevant symptom levels is needed. In general, researchers should come to an agreement on terms for birth-related PTSD/PTSS (s. proposal in Fig. 1 and equivalent section in introduction) and clearly report whether an assessment was stressor-specific.

When it comes to assessing PTSD/PTSS, researchers face the difficult decision of the most suitable instrument for their studies out of a wide spectrum of interview or self-report measures. Psychometric properties may guide these decisions, even though validation for birth-related PTSD/PTSS or even the postpartum population in general is rare. Instruments designed specifically for childbirth as the traumatic event possibly facilitate linking symptoms to the specific event. Use of general PTSD/PTSS measures is plausible in longitudinal designs or when implementing control groups, as it allows comparability. Measurements without focus on full diagnostic criteria (i.e., IES, Horowitz et al., 1979) are helpful for long-term comparisons. Additionally, high quality instruments for DSM-IV or DSM-5 (APA, 1994, 2013) should be continued to be adapted for future DSM versions to accurately reflect conceptual changes in PTSD criteria. For prevalence estimation, standardized use of instruments in future studies would be helpful, taking into account the above stated benefits of different measures.

Ideally, samples should be large enough to reliably detect even potentially low prevalence rates, especially in fathers/partners. To address the lacking representation of fathers/partners in perinatal mental health research, specific strategies for reaching this population may be adopted. These strategies could include consulting fathers on the questionnaire phrasing or including male interviewers in the research team (Macfadyen, Swallow, Santacrose, & Lambert, 2011).

Especially in cases where ideal sample sizes cannot be reached, sampling methods should enable representativeness in relation to clearly defined target frames concerning major demographic and obstetric characteristics. This is helpful not only for interpretation of population prevalence, but also for any conclusions readers may want to draw from a study. In practice however, many articles did not discuss whether their samples were representative, or if they did, could only rarely prove representativeness. This led to our decision to mark studies without clear risk status as “non-targeted” rather than community samples. However, data can only be generalized to the wider population with certainty if the most important features are comparable. In terms of demographic characteristics, research should pay specific attention to inclusion of individuals from ethnic minorities and different socioeconomic groups, as ethnicity and economic challenges have been linked to mistreatment during childbirth (Vedam et al., 2019).

In order to establish evidence of point prevalence, studies need to explore symptomatology at specific time points, rather than solely providing a sample mean or a wide time frame for time since childbirth. Additionally, further longitudinal studies might allow for more robust insights into the time course of PTSD/PTSS in the first year postpartum.

4.5.2. Recommendations for perinatal health policies

Taken together, the prevalence data emphasize that a substantial percentage of mothers suffers from birth-related PTSD during the first 14 months postpartum in almost all regions of the world. Even more mothers report significant levels of PTSS, affecting one in five women in vulnerable groups. Projecting the prevalence found in this review on the more than 700 million babies born annually worldwide between 2015 and 2020 (UN, 2019), as many as 91 million families may be affected by birth-related PTSD/PTSS and could be identified and subsequently referred to appropriate treatment each year. Even though the public health burden may be less serious for fathers or partners, judging by the mean prevalence rate found in this review, further quantitative research needs to be conducted before this conclusion may be drawn with certainty.

It has taken many years of research and campaigns to achieve recognition of postpartum depression (PPD) as the most frequent complication of pregnancy (Moran Vozar et al., 2020). Today, obstetric care providers routinely screen for PPD in many countries and researchers justifiably argue that the same routine should be implemented for birth-related PTSD (Moran Vozar et al., 2020). Screening methods may be introduced only with focus on targeted groups, possibly relying on short instruments (e.g., four item screen PC-PTSS, Cameron & Gusman, 2003). As has been demanded repeatedly, policy in perinatal health services needs to sharpen awareness for birth-related PTSD/PTSS and adopt strategies targeting modifiable risk factors, such as social support by staff.

There is no doubt that adequate prevention and intervention could counteract a considerable amount of suffering on the part of the affected family. A conclusive framework for implementation has been
suggested by Ayers and Ford (2016). It begins with primary prevention by “screening women during pregnancy for key vulnerability factors and adapting care to prevent PTSD from occurring” (pp. 193–194). Subsequently, it continues with secondary prevention after birth to “identify those who appraise the birth as traumatic or who have initial PTSD symptoms [and offer them] brief interventions, such as psychoeducation or midwife counseling, to help symptoms resolve” (p. 194). The final component of their proposition is tertiary intervention, making sure that mothers with PTSD are offered psychotherapy, if appropriate in combination with pharmacotherapy (Ayers & Ford, 2016). According to a recent review (de Bruijn, Stramrood, Lambregtse-van den Berg, & Rius Ottenheim, 2020), debriefing, cognitive behavioral therapy (CBT), and eye movement desensitization and reprocessing (EMDR) seem to be particularly beneficial treatment options. Building on previous and future insights on the effect of birth-related posttraumatic stress on the family system, it would be important to design and investigate interventions targeting the couple dyad or the family as a whole.

5. Conclusion

This systematic review and meta-analysis for the first time summarized research from 154 studies on birth-related PTSD/PTSS among parents. Mean prevalence of birth-related PTSD was 4.7% in mothers and 1.2% in fathers. For birth-related PTSD, prevalence rates of 12.3% in mothers and 1.3% in fathers were found. Future studies should include several time points in order to measure the course of prevalence rates over time. Given the relatively small number of available studies on fathers or other co-parents, future research investigating the impact of traumatic birth experiences on the entire family system is needed. Furthermore, we recommend use of the term “birth-related PTSD/PTSS” in future studies to clarify that symptoms are specifically related to childbirth as a traumatic stressor. We also propose the use of instruments tailored to specifically measure birth-related PTSD/PTSS. Raising awareness of birth-related PTSD/PTSS in perinatal health services policy, adopting strategies targeting modifiable risk factors, as well as screening methods with a focus on targeted groups should be prioritized.

Author contribution

CH prepared and conducted the systematic search. CH and RS performed the manual searches. CH and MK carried out the title/abstract screening. The full-text screening was conducted by CH and RS. AH, MK, PDY, CH, and SGN screened the French, Polish, Turkish, German, and Norwegian/Danish studies, respectively. Interrater reliability for the screening steps was calculated by CH. CH extracted the data of included studies; French studies were coded by AH. Extracted data was proof-read by DH, MLB, and TW. MO double coded 20% of the studies. CH and PDY performed the quality assessment; French studies were rated by AH. TW calculated the interrater reliability for quality assessment and double coding. CH analyzed the data and wrote the manuscript. SGN served as third reviewer for screening, double coding, and quality rating decisions and supervised the implementation of this systematic review and meta-analysis and the drafting of the manuscript. All authors contributed to the search update as well as the manuscript revision and approved the submitted version.

Role of funding sources

No funding was provided for this work, but all authors are (management committee) members of COST action CA18211.

Declaration of Competing Interest

All authors declare that they have no conflicts of interest.

Acknowledgements

We are very grateful to the support of the COST Action network, especially all the members and elsewhere recruited native speakers who readily assisted in screening the non-English articles: Ema Hresianová, Adriana Sampaio, Ernesto Gonzalez Mesa, Jihye Jeong, Xiao Zhou, and Leonie Warlo.

We would also like to thank all the authors who kindly provided additional information for our review. Their number is too big to list here, which gives an idea of the extent of helpfulness we experienced. We appreciate every single one!

Thank you to Marie Kopp and Michael Höfler for their helpful support with methodological issues, as well as to Franciska Rebherr for her assistance in the literature search and Freya Thiel for the language editing of our revised manuscript.

We deeply appreciate the reviewers’ dedication and the time they invested to provide feedback and possibilities for improvement in our manuscript, which have been invaluable for further increasing the quality of our work.

Appendix. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2022.102157.

References1


1 Note: References included in the meta-analyses may be found in Appendix E.
C.-S. Heyne et al.

Clinical Psychology Review 94 (2022) 102157


