Review article

Prenatal exposure to benzodiazepines and Z-drugs in humans and risk of adverse neurodevelopmental outcomes in offspring: A systematic review

Xinchen Wang, Tianyang Zhang, Isabella Ekheden, Zheng Chang, Clara Hellner, Jan Hasselström, Nitya Jayaram-Lindström, Brian M D’Onofrio, Henrik Larsson, David Mataix-Cols, Anna Sidorchuk

* Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
† Stockholm Health Care Services, Region Stockholm, Stockholm, Sweden
‡ Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
| Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden
| Academic Primary Health Care Centre, Region Stockholm, Stockholm, Sweden
| Department of Psychological and Brain Sciences, Indiana University, Bloomington, USA
| School of Medical Sciences, Örebro University, Örebro, Sweden

ARTICLE INFO

Keywords:
- Benzodiazepines
- Z-drugs
- Pregnancy
- Child neurodevelopmental outcomes
- Prenatal exposure
- Anxiolytics
- Hypnotics
- Antiepileptics

ABSTRACT

When used during pregnancy, benzodiazepines (BZDs) and related z-drugs could pass readily through the placenta and the foetal blood-brain barrier, where they can bind to γ-aminobutyric acid (GABA) receptors in the developing foetal brain. Yet, data on long-term safety of prenatal BZD and z-drug use and its impact on offspring neurodevelopment are inconclusive. In this systematic review, we qualitatively synthesize the existing evidence on maternal exposure to various BZDs and z-drugs during pregnancy and offspring cognitive, emotional, behavioural, and motor skills developmental outcomes. Nineteen studies were included. We used harvest plots to visualize the directions of reported associations. Despite several associations between distinct types of BZDs and z-drugs and an increased risk of outcomes within different neurodevelopmental domains were observed, a remarkable scarcity of overall research on the topic and considerable discrepancies in methodology, particularly towards controlling for confounding by indication, precluded drawing conclusions with a reasonable degree of certainty. We outline various research strategies to mitigate methodological limitations and provide directions for future empirical studies on the topic.

1. Introduction

During pregnancy, benzodiazepines (BZDs) and benzodiazepine-related drugs (z-drugs) are prescribed for alleviating symptoms of anxiety and insomnia, and the treatment of maternal psychiatric disorders and epilepsy (Hendrick, 2006; Crawford, 2009; Qato and Gandhi, 2021), with a reported prevalence of prenatal BZD and z-drug use ranging between 0.2% and 3.9% in different countries (Askaa et al., 2014; Bais et al., 2020a; Hanley and Mintzes, 2013; Reis and Kilåén, 2013; Tinker et al., 2019; Martin et al., 2015). Considerable variations in prevalence measures also exist with respect to timing of exposure with some studies reporting higher rates of BZD and z-drug use in early pregnancy (Askaa et al., 2014; Palmsten et al., 2015; Yonkers et al., 2017; Bais et al., 2020b; Qato and Gandhi, 2021), while others showing an increase in the use of such medications in the third trimester (Bais et al., 2020a; Hanley et al., 2020). BZDs and z-drugs pass readily through placenta and the foetal blood-brain barrier, where they can bind to γ-aminobutyric acid (GABA) receptors in the developing foetal central nervous system (Briggs, 2002; Juric et al., 2009; Gielen et al., 2012; Griffin et al., 2013). Thus, it is biologically plausible for these medications to affect foetal growth and development and have an impact on child’s health in short-term and long-term perspectives; although studies on consequences of prenatal BZD and z-drug use are generally scarce.

Prescribing BZDs and z-drugs during pregnancy requires balancing the potential benefits (i.e., treatment of maternal medical/psychiatric conditions) against the risks of adverse obstetric and neonatal outcomes.
and negative long-term consequences for child’s health and development. The clinical guidelines on the use of psychotropic medication in pregnancy (Larsen et al., 2015; McAllister-Williams et al., 2017; Hardy and Reichenbacker, 2019) suggest an individualized approach to starting and continuing prenatal BZD and z-drug treatment based on severity and chronicity of maternal conditions and previous treatment history. The guidelines emphasize that during pregnancy BZDs and z-drugs should be prescribed for the shortest period necessary with a regular review of patient’s needs, and strongly advocate the use of available non-pharmacological and alternative pharmaceutical options for treating maternal indications (Hardy and Reichenbacker, 2019; Larsen et al., 2015; McAllister-Williams et al., 2017). Such careful approach to prescribing BZDs and z-drugs during pregnancy is mainly driven by the findings on the adverse immediate birth outcomes. For example, prenatal use of BZDs and z-drugs was reported to be associated with the increased risks of spontaneous miscarriage, caesarean delivery, preterm birth, low birth weight, small head circumferences, and low Apgar score, and, if medication was used particularly close to delivery, with the elevated risks of low muscle tone (i.e., floppy infant syndrome), neonatal respiratory distress, and neonatal abstinence syndrome (Freeman et al., 2018; Huitfeldt et al., 2020; Ogawa et al., 2018; Wikner and Kallen, 2011; Vonkers et al., 2017). Risk of congenital malformation has previously been suggested, however, meta-analyses published in the last decade concluded non-teratogenicity of prenatal BZDs and z-drug use in general (Enato et al., 2011; Bellantuono et al., 2013; Grigoriadis et al., 2019).

More recently, there has been a growing concern about the lack of evidence on long-term safety of prenatal BZD and z-drug use, with a particular focus on offspring neurodevelopment (Larsen et al., 2015; McAllister-Williams et al., 2017; Hardy and Reichenbacker, 2019; Hjorth et al., 2019; Shyken et al., 2019). Neurodevelopmental outcomes refer to a wide spectrum of conditions characterized by delays or impediments in the acquisition of skills related to cognitive, behavioural, emotional, social, language, or motor developmental domains (Jeste, 2015), with main diagnostic categories including attention-deficit/hyperactivity disorder (ADHD), autism spectrum disorders, intellectual disability, communication disorders, specific learning disabilities, and motor disorders (Thapar et al., 2017). Being multi-factorial in origin and heterogeneous in clinical characteristics, these conditions have a typical onset in childhood and then persist into later life causing a significant impairment. The first systematic review conducted in 2014 summarized the results of seven observational studies on maternal use of BZD-antianxiety during pregnancy and neurodevelopmental outcomes in children, concluding that the available evidence was scarce and inconsistent (El Marroun et al., 2014). Several studies have since then emerged on the topic, holding a potential for further investigating long-term BZD and z-drug safety by studying the associations between prenatal exposure to such medications and various neurodevelopmental outcomes in offspring.

We conducted a systematic review to identify, appraise, and qualitatively synthesize the existing evidence on long-term neurodevelopmental impacts of prenatal exposure to BZDs and z-drugs, with a particular focus on associations of distinct types of such drugs with each neurodevelopmental domain and outlining methodological strategies for future studies.

2. Methods

2.1. Eligibility criteria

The systematic review was preregistered on PROSPERO (ID: CRD42021265828) and reported in accordance with the Preferred Reporting Items for Systematic Reviews (PRISMA 2020 statement) (Page et al., 2021b). The following PECO criteria were applied: Participants – children and adolescents (aged 0–18 years); Exposure – the use of prescribed BZDs and/or z-drugs (i.e., any medications with the Anatomical Therapeutic Chemical Classification System [ATC] codes related to benzodiazepine derivatives in anxiolytics [N05BA], hypnotics/sedatives [N05CD], antiepileptics [N03AE], and z-drugs [N05CF]) by participants’ mothers at any time during index pregnancy (i.e., prenatal exposure), regardless of the ascertainment method (e.g., register or medical record, self-reported); Comparator – no prenatal exposure to BZDs or z-drugs in question; Outcomes – any neurodevelopmental conditions related to cognitive, behavioral, emotional, and motor skills development domains, if assessed through diagnostic criteria or standardized measurement (e.g., the International Classification of Diseases, Diagnostic and Statistical Manual of mental disorders [DSM], standardized test scores, checklists, scales, or structured interviews). Observational studies of cohort, case-control, or cross-sectional design were considered for inclusion. Studies were excluded if: 1) prenatal exposure to BZDs was combined with other (non BZD)-anxiolytics, hypnotics/sedatives, or antiepileptics; 2) exposure was the illicit BZD use; 3) no control group was included; 4) outcomes were self-reported or caregiver-reported without using standardized diagnostic or measurement criteria; 5) outcomes were congenital malformations (e.g., neural tube defect) or chromosomal disorders (e.g., Down Syndrome); 6) the articles were case series, case reports, or secondary data analyses; 7) only neonatal outcomes (i.e., during the first 4 weeks of child’s life) were assessed; 8) less than 5 outcome events were reported.

2.2. Literature search and study selection

We developed a sensitive search strategy in collaboration with a librarian at Karolinska Institutet and implemented to the following databases: Medline (OVID), Embase (embase.com), Cochrane Library (Wiley), Web of Science (Clarivate), and PsycInfo (OVID). Grey literature was searched in Dissertation and Theses (Proquest), Open Grey (opengrey.eu) and DART-Europe E-thesis Portal (dart-europe.org). All databases were searched from inception to June 20th, 2021 without language restrictions. Animal studies were excluded. The search strategies are reported in detail in Supplementary materials. After de-duplicating the records found in electronic databases, two authors (XW and AS or IE) independently screened the titles and abstracts for relevance. Next, two authors (XW and IE or TZ) examined the full texts of the relevant papers to identify studies that fulfilled the above-mentioned criteria and were eligible for inclusion. Additionally, the reference lists of relevant reviews and included studies were scrutinized manually. Any disagreement during selection was handled through discussions with the senior author (AS).

2.3. Data extraction and quality assessment

Two authors (XW and AS) independently extracted the following items from each study: first author, country of origin, setting, publication year, study design, exposure definition (the exact drugs), exposure ascertainment method, timing (trimester) when medication were used, daily or cumulative dose, duration of use, outcome definition, outcome ascertainment method, age of offspring at the time of outcome ascertainment, sample size, list of confounders and whether and how the confounding by indication was addressed, and measures of association, if reported. The authors were contacted if further clarification was needed. If the study assessed several outcomes or if the exposure was reported separately for different BZDs or z-drugs, we extracted data on each type of exposure-outcome separately. Also, if the same study used several definitions of exposure (e.g., a broad definition as ‘at least one dispensation’ and a more stringent definition as ‘two or more dispensations’), we extracted information that corresponded to more valid exposure data (i.e., if collected under more stringent definition, e.g., from the study by Bloitreire (Bloitreire et al., 2020)). From each study, we extracted the effect sizes (ES) or description of results if no ES was reported from the model with most comprehensive control for confounders (e.g., from maximally-adjusted models instead of...
minimally-adjusted; from sibling analysis controlling for unmeasured familial confounders instead of analysis of the whole cohort; and from the analysis with control for confounding by indication either by adjustment, restriction, or using the active-comparator design instead of the analysis not controlled for confounding by indication) (Supplementary Table 1).

The quality of studies was assessed by two authors (XW and TZ) using the Newcastle-Ottawa assessment scale on the following items: selection of study population, comparability of groups (where control for confounding by indication was considered as ‘control for the most important factor’), and ascertainment of exposure or outcome (Wells et al., 2000). The study quality was defined as high if 7–9 points were assigned or, otherwise, low. Discrepancies were resolved by discussions with the senior author (AS).

### 2.4. Qualitative data synthesis and harvest plots

Following the guidelines for data synthesis and given the nature of selected studies, a qualitative synthesis was chosen (Borenstein et al., 2009; Campbell et al., 2020; McKenzie and Brennan, 2021). Considerable diversity in definitions of outcomes (i.e., the exact neuro-developmental conditions) and exposures (i.e., the type of BZDs and z-drugs) resulted in a large number of exposure-outcome combinations; each with a small number of included studies. It is not recommended to estimate a pooled ES if the existing evidence is only reported in 2–3 studies, as this precludes drawing meaningful conclusions. Also, the studies greatly varied on how the ES were measured (z-score difference, β-coefficient, odds ratio, hazard ratio, mean score, p-values only, etc.) (Supplementary Table 1). To ease the interpretation of findings from studies where the ESs were not reported at all or for exposure groups other than our exposure of interest (Laegreid et al., 1992; Viggledal et al., 1993; Reebye et al., 2002, 2012), we re-calculated the measure of associations. For study by Misri (Misri et al., 2006), where outcomes were assessed separately by parents and caregivers, we pooled the reported measures together to obtain a single ES.

We used harvest plots (Ogilvie et al., 2008; Crowther et al., 2011) to visualize associations between different types of BZDs and z-drugs and outcomes of interest. For qualitative synthesis, Cochrane guidelines recommend to graphically display the directions of association across multiple variables (McKenzie and Brennan, 2021). We grouped the exposures by type of drugs according to the exposure definitions and the ATC codes provided by the original studies: 1) any BZDs and z-drugs; 2) BZDs only; 3) z-drugs only; 4) BZD-anxiolytics only; 5) BZD-antiepileptics only; 6) BZD-anxiolytics/sedatives/hypnotics and z-drugs (if BZD-antiepileptics were not included); and 7) BZDs unspecified (if the exact drugs were not specified). In line with prior research (Hjorth et al., 2019), we categorized outcomes under four developmental domains based on the outcome definitions provided in the original studies: 1) cognitive development (including communication skills, language competence, overall mental development, intelligence quotient [IQ], overall cognitive development, school performance, intellectual disability with and without delays in developmental milestones, pervasive developmental disorder, and communication-related disorder); 2) emotional development (including internalizing problems and anxiety); 3) behaviour development (including externalizing problems, ADHD diagnosis, ADHD traits, behavioural problems at school, aggressive behaviour, and oppositional defiant disorder); 4) motor skills development (including gross, fine, and unspecified motor skills). For each domain, we created a separate harvest plot with multiple panels (one panel for each type of exposure drugs). Measures of associations from each study were denoted by separate bars (with a bar’s height representing the age of offspring when the outcome was measured) and positioned on the horizontal axis according to the direction of association (based on the reported or recalculated ES) as either ‘decreased risk of outcome’, ‘no association’, and ‘increased risk of outcome’.

Next, to assess the robustness of our results, within each domain we performed a grouping by potential moderators. The moderators included study period (i.e., follow-up time categorized as ending before 1997 and 1997-onwards under the assumption that diagnostic practices have improved over time), study quality (i.e., high and low), and study size (i.e., study population of <100 and ≥100 participants). The latter refers to assessing a ‘small-study effect’, which, if detected, is indicative of publication bias. Small-study effect is considered to be present if there is a tendency in associations to appear significant and stronger, if originated from smaller, less-powered studies, compared to findings from larger studies (Page et al., 2021a).

### 3. Results

Fig. 1 displays the study selection flowchart. The initial search returned 6113 records, including 864 records from grey literature databases. After de-duplication, we examined 5043 titles and abstracts, of which 114 were relevant for full-text examination. We further identified 14 potentially relevant studies from reference lists. The search finally yielded 19 studies eligible for inclusion.

#### 3.1. Characteristics of included studies

Table 1 illustrates details of included studies. All 19 studies had a cohort design, were conducted between 1958 and 2016, and published 1975–2020. Twelve studies originated from Western Europe and the rest were conducted in North America. Four studies focused on prenatal exposure to any BZDs and z-drugs (Odsbu et al., 2015; Brandlistuen et al., 2017; Lupattelli et al., 2019; Sundbakk et al., 2019); among that, one study additionally assessed the exposure to only BZDs (Lupattelli et al., 2019), and two studies also assessed the use of z-drugs (Brandlistuen et al., 2017; Lupattelli et al., 2019). Six studies assessed the exposure to only BZD-anxiolytics (either to a single drug, e.g., diazepam (Stika et al., 1990) or chlordiazepoxide (Hartz et al., 1975), or to any BZD-anxiolytics (Brandlistuen et al., 2017; Laegreid et al., 1992; Schechter et al., 2017; Viggledal et al., 1993). One study examined the exposure to any BZD-anxiolytics/hypnotics/sedatives (excluding BZD-antiepileptics) and z-drugs (Radojičić et al., 2017), and three studies did not specify the exact BZDs used (Figueroa, 2010; Mortensen et al., 2003; Mattson et al., 2002). Comparison groups in all abovementioned studies included participants with no prenatal exposure to the corresponding drug(s). Another six studies focused on prenatal exposure to only BZD-antiepileptics (clonazepam); among that, in three studies a combined exposure to clonazepam and SSRI was compared to SSRI monotherapy (Reebye et al., 2002; Misri et al., 2006; Reebye et al., 2012), and in other studies the prenatal use of clonazepam was compared to lamotrigine (Blotière et al., 2020; Daugaard et al., 2020) and/or to no clonazepam use in pregnancy (Daugaard et al., 2020; Elkjær et al., 2018). Exposure data were retrieved from maternal self-reports (Brandlistuen et al., 2017; Laegreid et al., 1992; Lupattelli et al., 2019; Sundbakk et al., 2019; Viggledal et al., 1993), clinical or register records (Stika et al., 1990; Mattson et al., 2002; Reebye et al., 2002, 2012; Mortensen et al., 2003; Misri et al., 2006; Figueroa, 2010; Elkjær et al., 2018; Blotière et al., 2020; Daugaard et al., 2020), or self-reports were complimented with prescription records or verified by clinicians (Hartz et al., 1975; Odsbu et al., 2015; Radojičić et al., 2017; Schechter et al., 2017). All studies measured the exposure across the entire pregnancy, apart from studies by Lupattelli (Lupattelli et al., 2019) (exposure measured separately for mid- and late pregnancy) and Stika (Stika et al., 1990) (exposure measured in the second and third trimesters). Studies by Odsbu (Odsbu et al., 2015) and Brandlistuen (Brandlistuen et al., 2017) subdivided the exposure into short-term (if administered in one trimester only) and long-term (if used in ≥2 trimesters).

Studies varied on assessment of confounding by maternal indication for prescribing BZDs and/or z-drugs during pregnancy (Table 1), with 11 out of 19 studies (57.9%) to control for confounding by indication.
through either restriction to sub-groups of mothers with one or more indication (anxiety, depression, epilepsy, sleeping problems, or pain-related conditions), adjustment for history of indicated conditions and/or symptoms during pregnancy, or active comparator design. Other measured confounders also varied between the studies (Supplementary Table S2), with one study additionally controlling for unmeasured familiar confounders through sibling comparison (Brandlistuen et al., 2017), and three studies reporting only crude associations (Laegreid et al., 1992; Schechter et al., 2017; Viggedal et al., 1993). The study quality assessment is reported in Supplementary Table S3 with 14 out of 19 studies (73.7%) being defined as having high methodological quality.

3.2. Cognitive development

Ten studies assessed offspring outcomes related to cognitive development, including overall cognitive abilities rated by a trained researcher with the General Conceptual Ability and Differential Ability Scales (Schechter et al., 2017), communication skills rated by mothers using the Ages and Stages Questionnaires (Lupattelli et al., 2019), clinical records of communication-related disorder and pervasive developmental disorder (Bloti`ere et al., 2020), clinical records of intellectual disability with and without delays in developmental milestones (Daugaard et al., 2020), clinically measured IQ (Hartz et al., 1975; Mattson et al., 2002), language competence rated by mothers using the language grammar rating scale (Odsbu et al., 2015), clinically assessed overall mental development (Hartz et al., 1975; Reebye et al., 2002, 2012), and registered records of school performance based on academic test results in native language and mathematics (Elkjaer et al., 2018). As shown in the harvest plot (Fig. 2), the measures of associations were controlled for exposure to any BZDs and z-drugs in late pregnancy was associated with an increased risk of lower communication skills at age of 5 years, compared to unexposed counterparts (Lupattelli et al., 2019). Also, prenatal use of BZD-antiepileptics (clonazepam) was associated with clinical diagnosis of intellectual disability with and without delays in developmental milestones in children at a median age of 10 years (interquartile range [IQR]: 6.5–14.0 years) (Daugaard et al., 2020). For clonazepam, the associations were noted when compared to no prenatal exposure to BZD-antiepileptics as well as when compared to prenatal use of lamotrigine (Daugaard et al., 2020). Also, the combined exposure to clonazepam and SSRI during pregnancy was associated with an increased risk of lower overall mental development at 2 months of age, when compared to prenatal use of SSRI only (Reebye et al., 2002). Among studies with no control for confounding by indication, an increased risk of lower school performance in native language at the age 12 years (6th grade) and in mathematics at age 9 (3rd grade) was found in children with prenatal exposure to clonazepam, compared to children with no exposure to such medication (Elkjaer et al., 2018). Similarly, an increased risk of lower overall cognitive ability at the age of 3.6 years was reported in children exposed to BZD-antiepileptics during pregnancy compared to unexposed counterparts (Schechte et al., 2017). All abovementioned studies with increased risks of cognitive outcomes were of high quality (Supplementary Table S1), although the reported measures of associations indicated mainly small-to-medium effects, apart from the study by Daugaard et al. (2020) with the ESs yielding medium-to-large effects (Supplementary Table 1). No significant associations of other types of BZDs and/or z-drugs with any other outcomes within cognitive domain were observed.

3.3. Emotional development

Fig. 3 presents a harvest plot for associations retrieved from four original studies on emotional development outcomes, including internalizing behaviour problems (Misri et al., 2006; Brandlistuen et al., 2017; Sundbakk et al., 2019) and anxiety symptoms (Radoji`ci` et al., 2017), all rated by mothers or both parents, and/or teachers using the Child Behaviour Checklist and the Teacher Report Form. All four studies controlled for confounding by indication. Prenatal use of...
### Table 1
Characteristics of included studies.

| Study, design | Country / setting (study period) | No. exposed / No. non-exposed | Exposure / comparison (ascertainment) | Timing / (daily dose, duration of use, cumulative dose) | Age of offspring | Outcome (ascertainment) | Control for confounding by indication |
|---------------|----------------------------------|------------------------------|---------------------------------------|--------------------------------------------------------|-----------------|--------------------------|---------------------------------------|
| Blotiere et al. (2020) | France/French national healthcare databases (2011–2016) | 1246/2916 | clonazepam/lamotrigine (disposition records) | All trimesters (NA) | 4.8 (4.1–5.3) yr | Child neurodevelopmental disorders, FSD and mental retardation (medical records) | Restricted to mothers with AED use (epilepsy), adjusted for maternal antipsychotics through IPTW, and active comparator (lamotrigine) |
| Brandistuen et al. (2017) | Norway/ Mola and MBRN (1999–2008) | 557/71 419 | BZD and z-drugs/sibling control (self-administered questionnaires) | All trimesters (NA; short/long-term use; NA) | 1.5 yr, 3 yr | Internalizing behaviour problems (rated by mothers and teachers with CBCL) | Adjusted for symptoms of anxiety or depression and SSRI during pregnancy, lifetime history of major depression, sleep problems through PS |
| Daugaard et al. (2020) | Denmark/Danish national registries (1997–2011) | 314/899 941 | clonazepam/no AED clonazepam/lamotrigine (prescription records) | All trimesters (~4 mg/≤4 mg/ 0 mg; NA; NA) | 10.1 (6.5–14.0) yr | Intellectual disability, delayed developmental milestones (medical records) | Adjusted for psychiatric diagnoses status, maternal epilepsy status, and active comparator (lamotrigine) |
| Elkjær et al. (2018) | Denmark/ National registries (1997–2006) | 188/478 839 | clonazepam/no BZD-AED clonazepam/lamotrigine (prescription records) | All trimesters (NA) | 8, 9, 10, 12, 14 yr (2nd, 3rd, 4th, 6th, 8th grades) | School Performance (register records of academic test results) | No |
| Figueiroa (2010) | The US/The MarketScan data (1997–2006) | 311/37 763 | BZD unspecified/no BZD (medication claims) | Any trimesters (NA) | 5 yr | ADHD (medical claims) | Adjusted for mental health problems during pregnancy |
| Hartz et al. (1975) | the US/University affiliated hospitals (1958–1966) | 501/33 073 | chlordiazepoxide/no chlordiazepoxide (self-reported medication history and clinical records) | Any trimesters (NA) | 8 mo, 4 yr | Motor skills unspecified, mental development, intelligence quotient (clinical examination) | No |
| Laegreid et al. (1992) | Sweden/ The general maternity outpatient unit and two delivery departments and psychiatry department in Gothenburg (1984–1986) | 17/29 | BZD-anxiolytics/no BZDs (self-reported) | Any trimesters (oxazepam (15–60 mg daily) and/or diazepam (5–30 mg daily), lorazepam (5–15 mg daily); NA; NA) | 6 mo, 10 mo, 18 mo | Gross and fine motor skills (clinical neurology examination) | No |
| Lupattelli et al. (2019) | Norway/ Mola (1999–2008) | 283/35 803 | Any BZD and z-drugs/no BZD or z-drugs, BZDs only/no BZDs, z-drugs only/no z-drugs (self-administered questionnaires) | Mid-pregnancy 17–28w, late pregnancy 29w or later (NA; NA) | 5.1 (5.0–5.3) yr | Gross and fine motor skills, and communication skills, and ADHD traits (rated by mothers ASQ and CPRS-R) | Restricted to mothers with depression/ anxiety, sleeping problems, pain-related disorders, and adjusted for severity of depressive and anxiety symptoms during pregnancy and lifetime history of major depression through IPTW |
| Mattson et al. (2002) | The US/the California Teratogens Information Service and Clinical Research Program (NA) | 25/248 | BZDs unspecified/no BZDs (clinical records) | Any trimesters (NA) | 6.3 (2.46) yr | Intelligence quotient (clinical examination with full scale IQ) | No |
| Missi et al. (2006) | Canada/ Reproductive Mental Health Program at British Columbia Women’s Hospital in Vancouver and pediatricians’ offices (1997–2004) | 9/13 | clonazepam – SSRI/SSRI (prescription records) | Any trimesters (NA; 13/142 days; 0.7(0.6) mg) | 4–5 yr | Internalizing behaviour problems (rated by mothers CBCL and TRF) | Active comparator (SSRI) and adjusted for maternal depression |
| | Denmark/ County Health Service data | 82/755 | Any trimesters (NA) | 7–10 mo | | | | (continued on next page) |
Abbreviations: ADHD, Attention-deficit/hyperactivity disorder; AED, Antiepileptics; ASQ, Ages and Stages Questionnaire; CBCL, Child Behavior Checklist; CPRS-R, The Revised Conners’ Parent Rating Scale; IPTW, Inverse probability of treatment weighting; IQ, Intelligence quotient; MBRN, Medical Birth Registry of Norway; Moba, the Norwegian Mother and Child Cohort; NA, not available; SSRI, Selective serotonin reuptake inhibitors; ODD, Oppositional defiant disorder; PDD, pervasive developmental disorder; PS, Propensity score; TRF, Teacher Report Form.

Table 1 (continued)

| Study, design | Country / setting (study period) | No. exposed / No. non-exposed | Exposure / comparison (ascertainment) | Timing / (daily dose, duration of use, cumulative dose) | Age of offsprings | Outcome (ascertainment) | Control for confounding by indication |
|---------------|---------------------------------|------------------------------|-------------------------------------|-----------------------------------------------------|-----------------|------------------------|-------------------------------------|
| Mortensen et al. (2003) | and the County Birth Registry (1991–1996) | 395/51 353 | BZDs unspecified / no psychotropics or AED (prescriptions records) | Any trimesters (NA; short/long-term use; NA) | 3 yr | Motor skills unspecified (examined by a trained nurse with Boel test) | Adjusted for maternal depression, anxiety, sleeping problems and SSRI during pregnancy and history of depression before pregnancy |
| Odbu et al. (2015) | Norway/ Moba (1999–2008) | | Any BZD and z-drugs/no BZD or z-drugs (self-administered questionnaires) | | | | |
| Radojičić et al. (2017) | the Netherlands/ The Generation R cohort (2002–2006) | 104/5609 | BZD-antianxiety and sedative/hypnotics (without antiepileptics) and z-drugs /no BZD or z-drugs (self-reported + prescription records) | All trimesters (NA; 1–66days; 1–145DDD) | 6–7 yr | ODD, anxiety symptoms and aggressive behaviour (rated by mothers and teachers with CBCL and TRF) | Adjusted for maternal anxiety and SSRI during pregnancy |
| Reebye et al. (2002) | Canada/ The Department of Reproductive Psychiatry, Children’s and Women’s Health Centre of British Colombia (1997–1999) | 14/24 | clonazepam + SSRI/SSRI (prescription records) | Any trimesters (NA) | 2 mo | Mental development index and motor skills unspecified (clinical assessment with Bayley scales) | Active comparator (SSRI) and restricted to mothers with psychiatric disorders |
| Reebye et al. (2012) | Canada/ The Department of Reproductive Psychiatry, Children’s and Women’s Health Centre of British Colombia (1997–1999) | 15/22 | clonazepam + SSRI/SSRI (prescription records) | Any trimesters (NA) | 8 mo | Mental development index and motor skills unspecified (clinical assessment with Bayley scales) | Active comparator (SSRI) and restricted to mothers with psychiatric disorders |
| Schechter et al. (2017) | The US/ the Emory Women’s Mental Health Program, a tertiary referral center (NA) | 27/135 | BZD-antianxiety/ no BZD (self-reported and verified by clinicians) | Any trimesters (NA) | 3.6 (2.5–5) yr | Cognitive ability (standardized assessment by trained researcher with the Differential Ability Scales) | No |
| Stika et al. (1990) | Prague* / Districts Medical prescriptions (1974–1986) | 15/15 | Diazepam/no diazepam (prescription records) | Second or third trimesters 20w-later (NA; NA; 2–10 mg) | 10 yr | Behaviour at school (rated by teachers with Teacher’s formulary with analogous scale on various features of behaviour) | No |
| Sundbakk et al. (2019) | Norway/Moba (1999–2008) | 273/36 128 | Any BZD and z-drugs/no BZD or z-drugs (self-administered questionnaires) | Any trimesters (NA) | 5 yr | Internalizing and externalizing behaviour problems (rated by parents with CBCL) | Restricted to mothers with mental or sleeping problems, and adjusted for mental health problems through IPTW | No |
| Viggelstad et al. (1993) | Sweden/The antenatal clinics, two delivery departments and psychiatric departments in Gothenburg (1984–1986) | 17/29 | BZD-antianxiety/ no psychoactive substances (self-reported) | Any trimesters (NA; NA; >5DDD) | 10 mo | Motor skills unspecified (clinically examined Griffiths’ mental development scale I) | |

* The study was performed in 1974–1986 in Czechoslovakia

Abbreviations: ADHD, Attention-deficit/hyperactivity disorder; AED, Antiepileptics; ASQ, Ages and Stages Questionnaire; CBCL, Child Behavior Checklist; CPRS-R, The Revised Conners’ Parent Rating Scale; IPTW, Inverse probability of treatment weighting; IQ, Intelligence quotient; MBRRN, Medical Birth Registry of Norway; Moba, the Norwegian Mother and Child Cohort; NA, not available; SSRI, Selective serotonin reuptake inhibitors; ODD, Oppositional defiant disorder; PDD, pervasive developmental disorder; PS, Propensity score; TRF, Teacher Report Form.

* The age of offspring when outcomes were assessed as either the exact age, range, mean (SD), or age of repeated measurements, yr: years, mo: months.

BZD-antianxiety was associated with an increased risk of internalizing behaviour problems in children aged 1.5 and 3 years when compared to unexposed counterparts (Brandlistuen et al., 2017), although the reported increase in risk was small at both ages (Supplementary Table 1). These findings came from a high quality study where a wide range of maternal and child’s measured characteristics and unmeasured confounders were accounted for. Other types of BZDs or z-drugs revealed no associations with emotional development outcomes in children.
X. Wang et al.
Fig. 2. Harvest plot for prenatal exposure to different types of benzodiazepines and z-drugs and cognitive development outcomes. Note: Grey bars represent the associations controlled for confounding by indication (by statistical adjustment or restriction). Patterned bars represent associations not controlled for confounding by indication. If a bar represented an association measured in a subgroup (e.g., mid- or late pregnancy exposure, short- or long-term medication use, full-term offspring only, etc.), or if the comparison group was presented by active controls (i.e., drug other than exposure), or if the exposure was combined with selective serotonin reuptake inhibitors (SSRI) versus SSRI only, a mark is added above a corresponding bar. Outcomes listed below the corresponding bars: cognitive = lower overall cognitive development (Schechter et al., 2017), communication = lower communication skills (Lupattelli et al., 2019), CRD = communication-related disorder (Blotières et al., 2020), ID = intellectual disability (Daugaard et al., 2020), ID + delays = intellectual disability and delays in developmental milestones (Daugaard et al., 2020), IQ = lower Intelligence quotient (Hartuz et al., 1975; Mattson et al., 2002), language = lower language competence (Ostdu et al., 2015), mental = lower overall mental development (Hartuz et al., 1975; Reebye et al., 2002, 2012), PDD = pervasive developmental disorder (Blotières et al., 2020), school (language) = lower school performance in native language (Elkjæra et al., 2018), school (math) = lower school performance in mathematics (Elkjæra et al., 2018).

4. Behavioural development

Five studies focused on outcomes related to behaviour development, including ADHD traits rated by mothers using the Conner’s Parent Rating Scale-Revised (Lupattelli et al., 2019), register records of ADHD diagnosis (Figueroa, 2010), oppositional defiant disorder and aggressive behaviour assessed by mothers and teachers using the Child Behaviour Checklist and the Teacher Report Form based on DSM 4th edition (Radjécōk, 2017), behaviour problems at school rated by teachers using a formulary with analogous scale on various behavioural features (Stika et al., 1990), and externalizing behaviour problems rated by parents using the Child Behaviour Checklist (Sundbakk et al., 2019). As presented in Fig. 4, exposure to BZDs only during late pregnancy was associated with an increased risk of ADHD traits at the age of 5 years (Lupattelli et al., 2019), with the reported ES indicating a small increase in outcome risk (Supplementary Table 1). This association was retrieved from a high quality study where confounding by indication was controlled for. However, no association with an increased risk of ADHD traits were seen for the same group of BZD drugs if administered in mid-pregnancy, or if children were exposed to any BZDs and z-drugs or to z-drugs only, regardless of timing of exposure (Lupattelli et al., 2019). No significant associations with other types of BZDs or z-drugs and behaviour development outcomes were reported.

3.5. Motor skills development

Seven studies focused on motor skills development, including gross and fine motor skills examined clinically (Laergriend et al., 1992) or rated by mothers using the Ages and Stages Questionnaires (Lupattelli et al., 2019), and clinically assessed unspecified motor skills (Hartuz et al., 1975; Mortensen et al., 2003; Reebye et al., 2002, 2012; Viggledal et al., 1993). As presented in Fig. 5, exposure during late pregnancy to either any BZDs and z-drugs, BZDs only, or z-drugs only was associated with an increased risk of lower gross motor skills at the age of 5 years. These findings came from a high quality study (Lupattellie et al., 2019) and were controlled for confounding by indication, with the reported ESs all yielding medium increase in outcome risk (Supplementary Table 1). Several inverse associations also appeared. For example, prenatal exposure to any BZDs and z-drugs was associated with a decreased risk of lower fine motor skills at the age of 5 years and exposure to BZDs only in mid pregnancy was associated with a decreased risk of lower gross motor skills at the same age (Lupattelli et al., 2019), with both ESs indicating medium decrease. Studies where confounding by indication was not accounted for, also reported significant associations with different motor skills problems with the ESs yielding mainly medium-to-large effects. Thus, the prenatal exposure to BZD-antixiolytics was associated with an increased risk of lower unspecified motor skills at 10 months of age (Viggvedal et al., 1993), lower fine motor skills at 6 months and lower gross motor skills at 6 months, 10 months, and 1.5 years (Laergriend et al., 1992), although these two studies were rated as having low methodological quality. Also, prenatal exposure to unspecified BZDs was associated with an increased risk of lower unspecified motor skills in 7–10 months old infants in a high quality study (Mortensen et al., 2003), with the risks being evident both for the whole study cohort and a subgroup of infants born full term. No other types of BZDs and/or z-drugs were shown to be associated with motor skills outcomes.

3.6. Potential moderators

Supplementary Table S4 presents all reported measures of association subdivided by potential moderators, separately for each developmental domain. First, stratification by the years of study periods did not affect the interpretation of results reported for cognitive, emotional, and behavioural domains as all associations with an increased risk of such outcomes were found in cohorts followed-up from 1997 or later. For motor skills development, the increased risks of outcomes by prenatal exposure to BZD-antixiolytics (Laergriend et al., 1992; Viggvedal et al., 1993) and unspecified BZDs (Mortensen et al., 2003) were reported in the cohorts with follow-up ending before 1997; however, these associations were not controlled for confounding by indication and thus should be interpreted with caution. Second, subdivision by the study quality did not change any results since the majority of significant associations originated from high quality studies. The only exceptions were associations between prenatal exposure to BZD-antixiolytics and increased risk of lower motor skills, reported by studies with lower quality and no control for confounding by indication (Laergriend et al., 1992; Viggvedal et al., 1993). Third, only six studies had less than 100 participants (Stika et al., 1990; Laergriend et al., 1992; Viggvedal et al., 1993; Reebye et al., 2002, 2012; Misri et al., 2006). Both small and large studies reported significant and non-significant associations of interest, although the majority of significant associations, in particular those where confounding by indication was controlled for, came from large cohorts. Thus, the risk of small study effect was unlikely to be present since most of the significant associations were originated from well-powered statistical analyses.

3.7. A post-hoc overview of evidence from rodent studies

Studies on prenatal drug safety in humans are methodologically challenging due to the risk of bias and unfeasibility of conducting randomized control trials that together preclude inferring causal effect of maternal drug use on offspring outcomes. To complement our review
Fig. 3. Harvest plot for prenatal exposure to different types of benzodiazepines and z-drugs and emotional development outcomes. Note: Grey bars represent the associations controlled for confounding by indication (by statistical adjustment or restriction). Patterned bars represent associations not controlled for confounding by indication. If a bar represented an association measured in a subgroup (e.g., mid- or late pregnancy exposure, short- or long-term medication use, full-term offspring only, etc.), or if the comparison group was presented by active controls (i.e., drug other than exposure), or if the exposure was combined with selective serotonin reuptake inhibitors (SSRI) versus SRRI only, a mark is added above a corresponding bar.

Outcomes listed below the corresponding bars: anxiety = anxiety symptoms (Radojičić et al., 2017), internalizing problems = internalizing behaviour problems (Misri et al., 2006; Brandlistuen et al., 2017; Sundbak et al., 2019).

# - exposure medication used in one trimester only (i.e., short term).
§ - exposure medication used in two or more trimesters (i.e., long term).
* - exposure combined with SSRI, comparison group – SSRI monotherapy.
X. Wang et al.

Neuroscience and Biobehavioral Reviews 137 (2022) 104647

(caption on next page)
and to gain better insight into the potential mechanisms that underlie the associations of interest, we performed a post-hoc (not preregistered in PROSPERO) overview of rodent studies. Rodent experiments eliminate the risk of exposure misclassification and allow randomization to ensure no systematic differences between exposed and unexposed groups. However, the translation of animal experiment results to humans may also be challenging due to interspecies differences in drug metabolism and other factors (Knight, 2007). A comprehensive review of rodent experiments on prenatal and early postnatal exposure (to model late pregnancy exposure in humans) to BZDs and z-drugs and neurodevelopmental outcomes in offspring has recently been published (Zucker, 2017). In our post-hoc overview, we focused on the results reported by Zucker (2017) and collected rodent studies from PubMed by applying the similar search strategy as we used for studies in humans, but including postnatal exposure. We selected a representative sample of papers where various neurodevelopmental outcomes and potential mechanisms of their development in rodents were discussed.

As reported by Zucker (2017), prenatal and early postnatal exposure to BZDs was associated with an increased depression, anxiety, and aggression (for exposure to alprazolam), increased cognitive deficits and decreased social behaviour (alprazolam, oxazepam) and increased multiple learning and memory deficits (diazepam, lorazepam, clonazepam, and chloridiazepoxide) in adult rodents. An increase in locomotor activity was seen in offspring of rodents exposed to diazepam, while a decrease in such activity was reported for exposure to oxazepam (Zucker, 2017). Other studies on rodents with offspring prenatally exposed to BZDs, particularly to diazepam, reported a decreased adaptation (Benesova et al., 1994), hyper-emotional responsiveness and anxiety state (Singh et al., 1996), as well as delayed appearance of neonatal reflexes, which is indicative of brain maturation (Nicosia et al., 2003). In contrast, a perinatal exposure to z-drugs, despite scarcely reported (Zucker, 2017), was not associated with learning, memory, emotionality, or locomotor activities among rodent offspring. Rodent studies also indicated several potential mechanisms through which BZDs and z-drugs may impact brain development. Thus, one study reported that prenatal BZD exposure in rats can result in deficient GABA neurotransmission (Nicosia et al., 2003), and such transmission dysfunction was implicated to underlie the symptoms of autism spectrum disorders in humans (Pizzarello and Cherubini, 2011; Zhao et al., 2021). It has also been suggested that BZDs can trigger widespread cell death in the infant rat brain (apoptotic neurodegeneration) (Olney et al., 2002; Ikonomidou, 2009).

4. Discussion

4.1. Overall summary

The scientific literature in this area appears to be limited, with substantial variability in applied methodology and approaches to control for confounding, including confounding by indication. The overall scarcity of data on associations of interest and considerable clinical heterogeneity of existing studies (i.e., inconsistency in definitions and measures of exposures and outcomes, and in statistical approaches) precluded us from performing quantitative synthesis of the available evidence.

The results of our qualitative synthesis are, however, indicative of several features that might be important for highlighting gaps in current research and directing future studies on long-term safety of BZD and z-drug prescription during pregnancy. First, regarding the types of drugs, BZD derivatives in anxiolytics and in antiepileptics appeared to be studied more often than any other related medications. Prenatal exposure to other drugs, for example z-drugs, were studied to much lower extent (only in 2 studies), while the exposure to BZD derivatives in sedatives/hypnotics were not at all addressed on their own in the literature. Studying the exposure to distinct types of BZDs and z-drugs is important due to the differences in their pharmacological properties, half-life, production of active metabolites, and selectivity in binding to neuronal GABA receptor subtypes (Griffin et al., 2013). Similar conclusions were reached by a review on the teratogenicity of prenatal exposure to BZDs (Bellantuono et al., 2013), which reported differential associations with the risk of malformations by distinct BZD drugs and advised against studying the ‘class effect’ of BZDs, instead suggesting to focus on the safety profile of each specific compound.

Second, the reviewed studies highlight the importance of assessing the timing of exposure. For example, a study by Lupattelli et al. (2019) reported positive associations with some outcomes among women with depression and anxiety if BZDs and z-drugs were administered in late pregnancy, whereas no such associations were observed with exposure in mid-pregnancy. Unlike other psychotropics, BZDs and z-drugs are often taken ‘as needed’ and it has been shown that a substantial proportion of women discontinues their treatment after becoming aware of pregnancy (Raitasalo et al., 2015; Bais et al., 2020a). At the same time, large population-based studies from the US and Canada reported that nearly 70% of all pregnant women with filled BZD and z-drug prescriptions were incident recipients, i.e., with BZD and z-drug treatment being initiated after the beginning of pregnancy (Hanley and Mintzes, 2014; Hanley et al., 2020). If timing of exposure is not specified, the exposed cohort may become heterogeneous and include women who were dispensed medication only once at the beginning of pregnancy (Raitasalo et al., 2015; Bais et al., 2020a). However, if timing of exposure is specified, the exposed cohort may become heterogeneous and include women who were dispensed medication only once at the beginning of pregnancy together with chronic users and the incident users for whom the treatment was initiated due to, e.g., worsening sleep problems during late pregnancy. Exposure to any medication in the first trimester may be harmful due to immaturity of the blood-brain barrier (Thorpe et al., 2013), while exposure in late pregnancy may affect offspring brain development, for which the second and third trimester are the sensitive periods (Ross et al., 2015). Furthermore, due to a small number of studies we were unable to specifically focus on exposure duration or frequency, which precluded us from assessing whether other characteristics related to timing of exposure would make the prenatal BZD and z-drug use differentially associated with the outcomes (e.g., if there is any duration threshold for the change in outcome risk). Thus, the timing of drug administration with regards to adverse neurodevelopmental
outcomes represents a clear knowledge gap.

Third, the reviewed studies mainly focused on outcomes related to cognitive and motor skills development (10 and 7 studies, respectively), with less information available on outcomes within behavioural and emotional development domains. Despite a small number of studies available for each exposure-outcome combination and a high clinical heterogeneity between the original studies, we cannot completely rule out the possibility that prenatal exposure to BZDs and/or z-drugs may be
associated with the increased risk of some neurodevelopmental outcomes. Indeed, several significant associations with outcomes in each developmental domain were observed, although the reported effects were mainly small-to-medium in magnitude. It is worth mentioning that in this review the outcomes in 12 studies were retrieved from medical records or registers (among that, 7 studies reported at least one significant association), while in other 7 studies the outcomes were assessed by mothers and/or teachers (with 2 studies presenting significant associations). A recent commentary and review on prenatal exposure to medication and child neurodevelopment (Hjorth et al., 2019; Vigod and Dennis, 2019) highlighted the importance of using valid and reliable outcome measures (e.g., a detailed clinical neurological assessment by healthcare professional) given the complexity and diversity of neurodevelopmental conditions. Ideally, future studies should focus on blinded clinician-reported rather than non-professional assessor-reported outcomes, in order to reduce the risk of bias and misclassification.

Fourth, the included studies differed in the approaches to control for confounders. Thus, the effect of confounding by indication (controlled in 11 out of 19 studies) or familial confounding (controlled in 1 out of 19 studies) could still persist in the included studies and their results should therefore be interpreted with caution. In fact, the studies varied in their ability to control for confounding by indication and other factors that could account for the associations. Without adequate control for confounding by indication it is difficult to rule out the impact of maternal health on the observed association (i.e., mothers who were prescribed BZDs and/or z-drugs in pregnancy may be more severely ill and have more comorbid disorders than mothers who were not prescribed such medications). As such, to disentangle with the degree to which the associations are due to causal effect of maternal medication on offspring outcome or due to confounding. Because observational studies lack randomization of the treatment, future studies attempting to eliminate confounding by indication and other factors should implement multiple approaches, as we have suggested when studying the effects of other psychiatric medications during pregnancy (Sujan et al., 2019; Li et al., 2020). These include assessment of measures that allow to balance baseline clinical characteristics between groups with and without exposure to BZDs and/or z-drugs, including restriction, active comparator (i.e., a drug with the same or similar indication), matching in study design and adopt stratified analyses, multivariate regression, or propensity score in statistical analyses (Freemantle et al., 2013; Kyriacou and Lewis, 2016; Meuli and Dick, 2018; VanderWeele, 2019; Yoshida et al., 2015), yet which approach to apply requires considerations of the suitability for study design, available data and research question. Furthermore, unmeasured familial confounders were addressed in only one study (Brandlstuen et al., 2017) where exposed and unexposed siblings were compared. As shown in other studies on prenatal psychotropic use and neurodevelopmental conditions (Brandlstuen et al., 2015; Nulman et al., 2015), quasi-experimental family designs are suitable for controlling genetic and early-life environmental confounders shared by full siblings on such associations (Lahey and D’Onofrio, 2010; D’Onofrio et al., 2013). Future studies would benefit from making use of various genetically sensitive designs, such as comparison of siblings or cousins (whose mothers are discordant for exposure to BZD and z-drugs during pregnancy) in order to minimise the impact of shared familial confounders. Also, none of the reviewed studies used paternal exposure to BZDs and z-drugs as a negative control, which could help to assess the effect of unmeasured confounders shared by parents (i.e., any positive associations of paternal exposure with child outcome would suggest familial confounding) (Lipsitch et al., 2010). Altogether, further high-quality studies with a rigorous approach to controlling for confounding by indication and unmeasured confounding are needed before we can understand the extent to which BZDs and z-drug use during pregnancy contributes to offspring neurodevelopmental conditions.

Our post-hoc overview of rodent experiments indicated several outcomes in offspring to be associated with perinatal exposure to BZDs and/or z-drugs. The associations varied between medications, thus, additionally supporting the importance of studying the effects of distinct drugs rather than the ‘class effect’ of BZDs. Despite some similar associations were also observed in our qualitative synthesis of human data, the general scarcity of evidence from both animal and human studies, and particularly concerns of interspecies differences (Knight, 2007), preclude us from generalizing the effects found in rodents to that in humans.

In line with the methodological approaches discussed in the review on causal inference in clinical psychology (D’Onofrio et al., 2020) and similar to the potential pathways that were suggested to underlie the association between prenatal exposure to antidepressants and offspring neurodevelopment in recently published translational review (Sujan et al., 2019), we hypothesized potential causal and non-causal pathways in the corresponding model for the effect of prenatal BZD and z-drug use (Fig. 6). Examples of causal pathways may include a direct effect of medication on brain development through a deficient GABA neurotransmission (Pizzarelli and Cherubini, 2011; Zhao et al., 2021) and apoptotic neurodegeneration (Olney et al., 2002; Ikonomidou, 2009), or the effect of BZDs and z-drugs on developing adverse birth outcomes as, e.g., preterm birth, low birth weight, and low Apgar score (e.g., (Huitfeldt et al., 2020)), which, in turn, are associated with neurodevelopmental problems (e.g., (Class et al., 2014)). As a potential non-causal pathway, maternal underlying health conditions, as indications for BZD and z-drug use, may influence offspring neurodevelopment (e.g., (Dachew et al., 2021)), and can fully or partially explain the associations. Likewise, the associations of interest could also be explained by environmental factors (e.g., maternal social deprivation (Cooke et al., 2021)) and genetic factors (i.e., biological pleiotropy, e.g., (Eilertsen et al., 2021, Nayar et al., 2021)); with each of these factors influencing as maternal indication for prenatal drug use as child’s outcomes per se. The pathways are complex and not mutually exclusive that again emphasizes the importance of using rigorous methodology in studying the impact of prenatal exposure to medication on offspring outcomes.

To complete the discussion on the effects of maternal medication use
on offspring health, it is worth mentioning the exposure to BZDs and z-drugs during breastfeeding, although this is beyond the scope of our review. Rodent models showed a dose-dependent impaired learning in rat offspring exposed to chlordiazepoxide or diazepam during lactation, although the data were scanty (Zucker, 2018). Studies on humans reported a considerable decrease in BZD use prevalence from pregnancy to postpartum, with the continuous use being rather uncommon (Bais et al., 2020a). Also, human studies suggested BZD use to be compatible with breastfeeding because of low levels of BZDs detected in infants’ plasma during lactation, yet longer-acting BZDs (e.g., diazepam, clonazepam) and chronic use should be avoided due to a risk of sedation and poor suckling (Kronenfeld et al., 2017; Nishimura et al., 2021). We were unable to find the corresponding data on z-drug exposure in either rodent or human studies. In our review, the included studies did not report data on medication use during breastfeeding. However, in light of the abovementioned, the confounding effect of drug use during lactation on the associations observed in our review could be considered small, if any, despite mothers may continue or initiate BZD treatment postpartum and transfer medication to children through breastmilk.

4.2. Limitations

Several limitations of the review should be acknowledged. First, an overall scarcity of data on prenatal exposure to BZDs and/or z-drugs and neurodevelopmental conditions in offspring and, in particular, the limited data on exposure to distinct types of drugs, constituted the major obstacle for assessing long-term safety of such medications for offspring health. Despite our attempt to qualitatively synthesize the existing evidence, a substantial diversity in measures of exposure, outcomes and methodological strategies of the original studies made the results of the review inconclusive and not translatable directly to clinical recommendations. Ultimately, clinical guidelines and translational research require an understanding of the absolute and relative risks associated with BZDs and z-drugs. Second, a substantial diversity in controlling for confounding by indication and other measured and unmeasured confounders precluded us from making a conclusion about whether the observed significant associations reflect a causal effect or simply the effect of residual confounding. Third, we cannot completely rule out the risk of publication bias since we were unable to perform any formal (quantitative) tests for such bias. However, the risk of publication bias should be seen as low, if any, due to the following: i) as described above, the review is unlikely to suffer from a small study effect, ii) literature search was comprehensive with a sensitive search strategy and a variety of grey literature sources used, and iii) no restrictions towards language, geography or publication time were applied. With respect to outcome reporting bias, we consider its risk to be low, if any, given a variety of outcomes reported in each study, the use of standardized measurements, and the fact that the studies report both significant and non-significant findings for the same outcomes (but at different ages), which minimize the risk of selective reporting based on the direction of the results.

5. Conclusions

It is currently not possible to conclude with a reasonable degree of certainty whether prenatal exposure to BZDs and/or z-drugs is associated with neurodevelopmental outcomes in offspring. This uncertainty mainly originates from a remarkable scarcity of overall research on the topic. High-quality empirical research on the safety of BZD and z-drug use during different stages of foetal brain development is urgently needed. We have outlined various research strategies to mitigate the impact of confounding by indication and unmeasured confounding.
Declaration of Competing Interest

H.L. has served as a speaker for Medice, Eolverna Pharma and Shire/Thakada and has received research grants from Shire/Thakada; all outside the submitted work. D.M.C reports receiving personal fees from Elsevier, Wolters Kluwer Health, and UpToDate Inc; all outside the submitted work. Other authors report no competing interests.

Data Availability

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

Acknowledgements

We would like to acknowledge Gun Brit Knutsson at the Karolinska Institutet library for helping us with designing the search strategy.

Funding

This study was supported by the Swedish Research Council (grant number: 2019-01408 to A.S.) and Region Stockholm (ALF Medicine grant number 20190379 to A.S.). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neubiorev.2022.104647.

References

Askaa, B., Jimenez-Solem, E., Enghusen Poulsen, H., Traerup Andersen, J., 2014. Maternal characteristics of women exposed to hypnotic benzodiazepine receptor agonist during pregnancy. Obstet. Gynecol. Int. 2014, 945621.
Bais, B., Molenar, N.M., Bijma, H.H., Hoogendijk, W.J.G., Mulder, C.L., Luijk, A.L., Lambregts-de Van der Berg, M.P., Kamperman, A.M., 2020. Prevalence of benzodiazepine and benzodiazepine-related drugs exposure before, during and after pregnancy: a systematic review and meta-analysis. J. Affect. Disord. 269, 18–27.
Bais, B., Munk-Olsen, T., Bergink, V., Liu, X., 2020. Prescription patterns of benzodiazepine and benzodiazepine-related drugs in the peripartum period: a population-based study. Psychiatry Res. 286.
Bellantuono, C., Tofani, S., Di Sciascio, G., Santone, G., 2013. Benzodiazepine exposure in pregnancy and risk of major malformations: a critical overview. Gen. Hosp. Psychiary 35 (1), 3–8.
Benesova, O., Tejkalova, H., Kristofikova, Z., Klaschka, J., Dostal, M., 1994. Early exposure to psychotropic and analgesic medications—A systematic review. PLoS ONE 2014, e94562.
Bergström, O., Tejalova, H., Kristofikova, Z., Klaschka, J., Dostal, M., 1994. Early exposure to psychotropic and analgesic medications—A systematic review. PLoS ONE 2014, e94562.
Brandlistuen, R.E., Ystrom, E., Hernandez-Diaz, S., Skurtveit, S., Selmer, R., Handal, M., 2014. Prescription patterns of benzodiazepine and benzodiazepine-related drugs exposure before, during and after pregnancy: a systematic review and meta-analysis. J. Obstet. Gynecol. Can. 33 (3), 1–4.
Briggs, G.G., 2002. Drug effects on the fetus and breast-fed infant. Clin. Obstet. Gynecol. 45 (1), 6–21.
Campbell, M., McKenzie, J.E., Seidman, A., Kritzirski, S.V., Brennan, S.E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shephard, S., Thomas, J., Welch, V., Thomson, H., 2020. Synthesis without meta-analysis (SWIM) in systematic reviews: reporting guideline. BMJ 368, 16890.
Crawford, P.M., 2009. Managing epilepsy in women of childbearing age. Drug Saf. 32 (4), 293–307.
Crowther, M., Avenell, A., Maclellan, G., Mowatt, G., 2011. A further use for the Harvest plot: a novel method for the presentation of data synthesis. Res Synth. Methods 2 (2), 79–83.
Dachew, B.A., Scott, J.G., Heron, J.E., Ayungu, G., Alati, R., 2021. Association of maternal depressive symptoms during the perinatal period with postnatal depressive disorder in children and adolescents. JAMA Netw. Open 4 (9), e2125585.
Daugaud, C.A., Pedersen, L., Sun, Y., Dreier, J.W., Christensen, J., 2020. Association of prenatal exposure to valproate and other antiepileptic drugs with intellectual disability and delayed childhood milestones. JAMA Netw. OPEN 3, 11.
D’Onofrio, B.M., Lahey, B.B., Turkheimer, E., Lichtenstein, P., 2013. Critical need for family-based, quasi-experimental designs in integrating genetic and social science research. Am. J. Public Health 103 (Suppl 1), S46–52.
D’Onofrio, B.M., Sjolander, A., Lahey, B.B., Lichtenstein, P., Oberg, A.S., 2020. Accounting for confounding in observational studies. Annu Rev. Clin. Psychol. 16, 25–48.
Elertser, E.M., Hannigson, J.A., McAdams, T.A., Rijjsdijk, F.V., Casijkowski, N., Reichborn-Kjennerud, T., Ystrom, E., Gjerde, L.C., 2021. Parental prenatal symptoms of depression and offspring symptoms of ADHD: a genetically informed intergenerational study. J. Atten. Disord. 25 (11), 1554–1563.
El Marroun, H., White, T., Verhult, F.C., Tiemeier, H., 2014. Maternal use of antidepressant or anxiolytic medication during pregnancy and childhood neurodevelopmental outcomes: a systematic review. Eur. Child Adolesc. Psychiatry 23 (10), 975–992.
Elkjaer, L.S., Bech, B.H., Sun, Y., Laursen, T.M., Christensen, J., 2018. Association between prenatal valproate exposure and performance on standardized language and mathematics tests in school-aged children. JAMA Neurol. 75 (6), 663–671.
Enato, E., Moretti, M., Koren, G., 2011. The fetal safety of benzodiazepines: an updated meta-analysis. J. Obstet. Gynecol. Can. 33 (1), 1–4.
Figueras, R., 2010. Use of antiepileptics during pregnancy and risk of attention-deficit/ hyperactivity disorder in the offspring. J. Dev. Behav. Pediatr. 31 (8), 641–648.
Freman, M.P., Goez-Mogollon, L., Mcinerney, K.A., Davies, A.C., Church, T.R., Soinskiy, A.Z., Noe, O.B., Figueras, A.C., Cohen, L.S., 2018. Obstetrical and neonatal outcomes after benzodiazepine exposure during pregnancy: results from a prospective registry of women with psychiatric disorders. Gen. Hosp. Psychiatry 53, 73–79.
Freudenthal, N., Marston, L., Walters, K., Wood, J., Reynolds, M.R., Petersen, L., 2013. Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. BMJ 347, f6409.
Gienen, M.C., Lumb, M.J., Smart, T.G., 2012. Benzodiazepines modulate GABA receptors by regulating the proactivation step after GABA binding. J. Neurosci. 32 (17), 5707–5715.
Griffin 3rd, C.E., Kaye, A.M., Bueno, F.R., Kaye, A.D., 2013. Benzodiazepine pharmacology and central nervous system-mediated effects. Ochsner J. 13 (2), 214–223.
Griparidis, S., Graves, L., Peer, M., Maminashvilli, L., Dennis, C.-L., Vigid, S.N., Steiner, M., Brown, C., Cheung, A., Dawson, H., Rector, N., Grenette, M., Richter, M., 2019. Benzodiazepine use during pregnancy alone or in combination with an antiepileptic and congenital malformations: systematic review and meta-analysis. J. Clin. Psychiatry 80, 4.
Hanley, G., Mintzes, B., 2014. Patterns of psychotropic medicine use in pregnancy in the United States from 2006 to 2011 among women with private insurance. BMC Pregnancy Childbirth 14, 242.
Hanley, G.E., Miller, T., Mintzes, B., 2020. A cohort study of psychotrophic prescription drug use in pregnancy in England, Scotland, and Wales. Paediatr. Perinat. Epidemiol. 34 (3), 254–266.
Harra, S.C., Heinonen, O.P., Shapiro, S., 1975. Antenatal exposure to methohexmat and chloralazine in relation to malformations, mental development, and childhood mortality. New Engl. J. Med. 292 (14), 726–732.
Hendrick, V., 2006. Psychiatric Disorders in Pregnancy and the Postpartum: Principles and Treatment. Humana Press, NJ.
Hjorth, S., Bromley, R., Ystrom, E., Lupattelli, A., Spigset, O., Nordeng, H., 2019. Use and validity of child neurodevelopmental outcome measures in studies on prenatal exposure to psychotrophic and anagisic medications - a systematic review. PLoS ONE 14, 12.
Huitfeldt, A., Sundbakk, L.M., Skurtveit, S., Handal, M., Nordeng, H., 2020. Associations of maternal use of benzodiazepines or benzodiazepine-like hypnotics during pregnancy with immediate pregnancy outcomes in Norway. JAMA Netw. OPEN 3, 6.
Ionomidou, C., 2009. Triggers of apoplos in the immature brain. Brain Dev. 31 (7), 488–492.
Jeste, S.S. (2015). Neurodevelopmental behavioral and cognitive disorders. Continuum (Minneap Minn) 21(3 Behavioral Neurology and Neuropsychiatry): 690–714.
Juric, S., Newport, D.J., Ritzler, J.C., Galanti, M., Larew, S.Z., 2009. Zolpidem (Ambien) in pregnancy: placental passage and outcome. Arch. Women’s Men. Health 12 (6), 441–446.
Knight, A., 2007. Systematic reviews of animal experiments demonstrate poor human clinical and toxicological utility. Alter. Lab Anim. 35 (5), 641–659.
Kronenfeld, N., Berlin, M., Shaniv, D., Berkovich, M., 2017. Use of psychotropic medications in breastfeeding women. Birth Defects Res 109 (12), 957–997.
Kyriacou, D.N., Lewis, R.J., 2016. Confounding by indication in clinical research. JAMA 316 (17), 1818–1819.
