ADDICTION

SSA SOCIETY FOR THE STUDY OF ADDICTION

REVIEW

doi:10.1111/add.13462

Buprenorphine compared with methadone to treat pregnant women with opioid use disorder: a systematic review and meta-analysis of safety in the mother, fetus and child

Barbara K. Zedler[1], Ashley L. Mann[1], Mimi M. Kim[2], Halle R. Amick[1], Andrew R. Joyce[1], E. Lenn Murrelle[1] & Hendrée E. Jones[3,4]

Venebio Group, LLC, Richmond, Virginia, USA,[1] Center for Biobehavioral Health Disparities Research, Division of Community Health, Duke University, Durham, NC, USA,[2] UNC Horizons, Department of Obstetrics and Gynecology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA[3] and Departments of Psychiatry and Behavioral Sciences and Obstetrics and Gynecology, School of Medicine, Johns Hopkins University, Baltimore, MD, USA[4]

ABSTRACT

Aims To assess the safety of buprenorphine compared with methadone to treat pregnant women with opioid use disorder. **Methods** We searched PubMed, Embase and the Cochrane Library from inception to February 2015 for randomized controlled trials (RCT) and observational cohort studies (OBS) that compared buprenorphine with methadone for treating opioid-dependent pregnant women. Two reviewers assessed independently the titles and abstracts of all search results and full texts of potentially eligible studies reporting original data for maternal/fetal/infant death, preterm birth, fetal growth outcomes, fetal/congenital anomalies, fetal/child neurodevelopment and/or maternal adverse events. We ascertained each study's risk of bias using validated instruments and assessed the strength of evidence for each outcome using established methods. We computed effect sizes using random-effects models for each outcome with two or more studies. **Results** Three RCTs ($n = 223$) and 15 cohort OBSs ($n = 1923$) met inclusion criteria. In meta-analyses using unadjusted data and methadone as comparator, buprenorphine was associated with lower risk of preterm birth [RCT risk ratio (RR) = 0.40, 95% confidence interval (CI) = 0.18, 0.91; OBS RR = 0.67, 95% CI = 0.50, 0.90], greater birth weight [RCT weighted mean difference (WMD) = 277 g, 95% CI = 104, 450; OBS WMD = 265 g, 95% CI = 196, 335] and larger head circumference [RCT WMD = 0.90 cm, 95% CI = 0.14, 1.66; OBS WMD = 0.68 cm, 95% CI = 0.41, 0.94]. No treatment differences were observed for spontaneous fetal death, fetal/congenital anomalies and other fetal growth measures, although the power to detect such differences may be inadequate due to small sample sizes. **Conclusions** Moderately strong evidence indicates lower risk of preterm birth, greater birth weight and larger head circumference with buprenorphine treatment of maternal opioid use disorder during pregnancy compared with methadone treatment, and no greater harms.

Keywords Buprenorphine, dependence, fetus, harm, methadone, opioid use disorder, pregnancy.

Correspondence to: Barbara Zedler, Venebio Group, LLC, 7400 Beaufont Springs Drive, Suite 300, Richmond, VA 23225, USA. E-mail: barb.zedler@venebio.com
Submitted 16 November 2015; initial review completed 16 February 2016; final version accepted 6 May 2016

INTRODUCTION

The incidence of prescription and illicit opioid use during pregnancy has increased substantially in the United States since 2000, paralleling a similar escalation in the general population [1]. Moreover, the prevalence of opioid use disorder (OUD) during pregnancy more than doubled between 1998 and 2011, to four per 1000 deliveries [2].

All pregnancies have a background risk of adverse consequences. Pregnant women with OUD have a higher frequency of additional risk factors for adverse pregnancy outcomes than pregnant women who do not use opioids. These risk factors include chronic viral infections, psychiatric conditions, poor health behaviors, adverse social conditions and inadequate prenatal care [3,4].

Complete opioid abstinence throughout pregnancy is ideal for both mother and fetus, but acute withdrawal during pregnancy is not recommended [5,6]. Relapse rates are high and repeated cycles of intoxication and withdrawal are associated with significant fetal distress that

can lead to placental insufficiency and consequent pregnancy loss, intrauterine growth restriction (IUGR) and preterm labor and birth [5,7–9]. The accepted treatment for OUD during pregnancy is long-acting opioid agonist medication-assisted treatment (OMAT), such as methadone (MET) or buprenorphine (BUP), within the context of a comprehensive program of obstetric care and psychosocial interventions [5,8,10–14]. Adequate medication treatment maintains stable opioid blood levels that reduce maternal craving for and use of heroin or other opioids and improves prenatal care and fetal/infant outcomes compared with untreated opioid use or opioid withdrawal [11,15,16]. MET maintenance treatment during pregnancy has been used widely since the early 1970s via daily visits to government-regulated clinics [17]. BUP maintenance treatment has been used increasingly since its approval in France in 1996 and the United States in 2002, partly because of its availability in the private practitioner setting and pharmacology that enables less than daily dosing, lower overdose risk and fewer drug interactions [11,18]. Three RCTs have been conducted comparing BUP and MET as OMAT in pregnancy, with a primary focus on multiple measures of neonatal abstinence syndrome (NAS) [19–21]. Previous systematic reviews and meta-analyses of these RCTs [5,22] concluded that BUP and MET have similar efficacy for reducing pregnant women's opioid use but that neither opioid agonist was superior for all maternal, fetal and child outcomes [11]. However, uncertainty was high regarding the conclusions due to the small body of evidence, particularly for outcomes other than NAS, due largely to their infrequency. For NAS, the meta-analyses identified no difference between BUP and MET in the frequency of NAS requiring treatment, the amount of morphine or time required to treat or the length of hospitalization. However, the single, large RCT ($n = 131$) [19] observed significantly less severe NAS, based on 89% less morphine required to treat and 43% shorter hospitalization, compared with no difference in the two small RCTs ($n = 14$ [20] and $n = 21$ [21]). No additional RCTs are available or likely to be performed, but the cumulative body of relevant observational studies has not been reviewed rigorously or synthesized quantitatively for any pregnancy outcomes.

The objectives of this review were to assess systematically all available evidence from clinical studies regarding the safety of buprenorphine compared with methadone treatment of opioid-dependent pregnant women and provide quantitative treatment effect estimates for selected pregnancy outcomes, as feasible.

METHODS

The conduct and reporting of this review conform with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Supporting information, Figure S1) [23].

Search strategy and inclusion criteria

We searched the PubMed and Embase databases and the Cochrane Database of Systematic Reviews from their inception through February 2015 using the search strategy in Supporting information, Table S1 without language restrictions. We searched manually the reference lists of review papers and the included studies to identify additional papers.

We included studies if they: (1) were RCTs or observational (cohort or case–control) studies (OBSs); (2) enrolled opioid-dependent pregnant women; (3) compared buprenorphine or buprenorphine–naloxone with methadone as OMAT; and (4) reported original data on one or more specified pregnancy-related outcomes representing important potential pregnancy-related harms (Table 1). This review focused on outcomes other than NAS for practical reasons. There are several pregnancy outcomes besides NAS which lack systematic review or meta-analysis to date and, due to its numerous measures, NAS would benefit from a separate systematic review that includes eligible OBSs. Two researchers reviewed each title and abstract independently and then assessed the full texts of potentially eligible papers. Disagreements between reviewers regarding eligibility were resolved by consensus.

Data extraction and risk of bias assessment

Two researchers extracted data independently from each included paper into standardized tables and resolved discrepancies by consensus. A senior researcher confirmed the accuracy of entries. We contacted authors as feasible if additional information was needed. We categorized studies as RCT or OBS based on elements as reported. Two researchers assessed the risk of bias (ROB) independently for each outcome as high, medium or low, and a senior researcher resolved any conflicts. For RCTs, we assessed randomization adequacy, allocation concealment, missing outcome data, selective outcome reporting and blinding of participants, study personnel and assessors according to standards of the US Agency for Healthcare Research and Quality (AHRQ) [24]. For OBSs, we evaluated the selection of participants, comparability of cohorts, exposure and outcome assessment and follow-up adequacy using the Newcastle–Ottawa Scale as expanded by Guyatt [25,26].

Data synthesis and statistical analyses

The unit of analysis was pregnancies or live births, depending on the outcome. We conducted meta-analyses of the unadjusted study data using random effects models

Table 1 Inclusion criteria and outcome definitions.

Study component	*Criterion/definition*
Population	Opioid-dependent pregnant women
Intervention	Buprenorphine prescribed as opioid agonist medication-assisted treatment for opioid use disorder[a]
Comparator	Methadone prescribed as opioid agonist medication-assisted treatment for opioid use disorder
Outcomes	
Spontaneous fetal death	Miscarriage (death of a fetus or embryo at or before 20 completed weeks gestation) Stillbirth (death of a fetus after 20 completed weeks gestation)
All fetal death	Spontaneous fetal death plus induced fetal death (induced abortion)
Preterm birth	Live birth before 37 completed weeks gestation[b]
Fetal growth outcomes	
*Birth weight (g)	Converted to grams as necessary
*Low birth weight (LBW)	< 2500 g regardless of gestational age
*Small for gestational age (SGA)	Birth weight below an established sex- and gestational week-specific mean value[c]
*Intrauterine growth restriction (IUGR)	Diminished growth velocity documented in two or more intrauterine growth assessments
*Head circumference at birth (cm)	Converted to centimeters as necessary
Fetal/congenital anomalies	An abnormality of structure (malformation), function or metabolism present at birth or identified at fetal death; birth defects
Sudden infant death syndrome (SIDS)	Unanticipated and unexplained death of a live-born infant before age 1 year
Fetal/child neurodevelopment	Cognitive, behavioral, sensory, motor or functional development. Abnormal is a delay or impairment
Maternal adverse events during pregnancy	Categorized by each study as serious (e.g. death) or non-serious[d]
Study designs	Randomized controlled trials, observational (cohort or case–control) studies

EGA = estimated gestational age. [a]We included one study that treated women with an abuse-deterrent combination buprenorphine (BUP)–naloxone formulation [44], but excluded it from quantitative analyses. [b]Preterm birth was defined as < 36 completed weeks gestation in Colombini 2008 [37]. [c]SGA was defined as birth weight below: (a) 2 standard deviations from the sex- and gestational-week specific mean value (Jones 2010 [19]; Kakko 2008 [39]); (b) 10th percentile of the sex- and gestational-week specific mean value (Siedentopf 2004 [47]); or (c) the 5th percentile of the sex- and gestational-week specific mean value (Brulet 2007 [36]; Meyer 2015 [43]). [d]Maternal adverse events were defined as: (a) medical events (Lacroix 2011 [41]); (b) complications (Prasad 2013 [44]); or (c) any untoward medical occurrence (Jones 2010 [19]).

(DerSimonian & Laird method) [27] to account for heterogeneity among the studies and estimated unadjusted treatment effects as weighted mean differences (WMDs) for continuous outcomes and risk ratios (RRs) for binary outcomes. Statistical significance was defined as a 95% confidence interval (CI) for the pooled effect that did not include zero for WMDs or 1.0 for RRs.

We anticipated a substantial amount of missing outcome data from attrition based upon the challenges of research with opioid-dependent people, especially during pregnancy [19,28,29]. We decided a priori to include only unadjusted outcome data as available from studies with low or medium ROB in our main analyses. To examine the stability of the main estimates, we conducted sensitivity analyses by including high ROB studies or imputing missing binary data under best- and worst-case scenarios [28,29]. We combined OBSs with similar study methods and clinical variability [30] and calculated summary treatment effect estimates separately by study design [28,29]. We estimated inconsistency (heterogeneity) across studies using the I^2 statistic [28] and investigated sources of clinical and/or methodological variation when we suspected heterogeneity that might affect the results [31]. We synthesized outcome data qualitatively when studies were too heterogeneous to pool quantitatively or when only a single study reported an outcome. For comparisons with 10 or more studies we inspected funnel plots to assess potential publication bias [32]. Analyses were conducted using Comprehensive Meta-Analysis, version 3.2 (Biostat; Englewood, NJ, USA).

Rating strength of evidence (SOE)

SOE is a summary of confidence in our findings. We evaluated the SOE for each outcome based on guidance established by AHRQ using five domains: study limitations, directness, consistency, precision and reporting bias [33]. The assigned grade (high, moderate, low, insufficient) represents the degree of confidence in the effect estimates for an outcome. We graded SOE separately for the bodies of evidence from RCTs and OBSs.

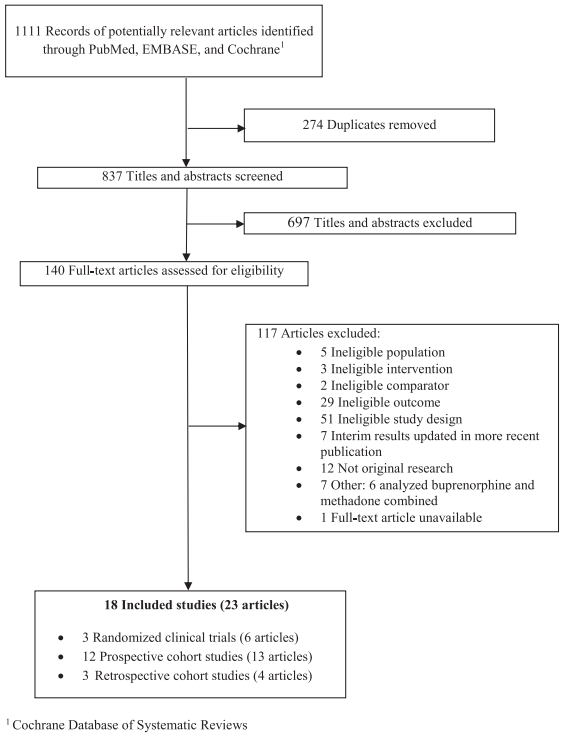


Figure 1 Flow of paper disposition and study selection

RESULTS

Of 1111 citations identified, 140 full-text papers were assessed for eligibility, and 18 studies [18–21,34–49] (reported in 23 papers [18–21,34–52]) satisfied our inclusion criteria (Fig. 1). The only three RCTs [19–21] (six papers [19–21,50–52]) that have been conducted to date were included (223 participants; published 2005–10) (Supporting information, Table S2). Fifteen OBSs [18,34–39,41–48] (17 papers) [18,34–49] enrolled a total of 1923 participants in prospective [18,34–39,41,42,45–47] ($n = 12$) or retrospective [43,44,48] ($n = 3$) cohort studies published 2001–15. No case–control studies were identified. The abuse-deterrent combination BUP–naloxone formulation is prescribed increasingly as OMAT, particularly in the United States and Australia [53,54]. We included one study that treated women with BUP–naloxone [44], but considered it too clinically different from the other studies to include in quantitative analyses. However, sensitivity analyses indicated that effect estimates did not differ for any outcome, whether the BUP–naloxone study was included or excluded. Sample sizes were 15–609. Attrition ranged from 22% [20] to 33% [19] overall in the RCTs and was unbalanced between treatment groups in two RCTs [19,20]. The range of estimated gestational age at study enrollment was 6–37 weeks. The mother's average daily dose at delivery ranged from 5.1 to 18.7 mg for BUP and 35–99.4 mg for MET. Outcome definitions among the studies were generally consistent, except for maternal adverse events (AEs) and small for gestational age (SGA) (Table 1). Statistical heterogeneity was low among the main analyses except for growth outcomes [birth weight (two RCTs, $I^2 = 50\%$); LBW (two OBSs, $I^2 = 62\%$); and SGA (two OBSs, $I^2 = 50\%$)]. ROB was rated medium for each RCT ($n = 3$, Supporting information, Table S3) and medium ($n = 9$) [18,38–43,48,49] or high ($n = 8$) [34–37,44–47] for all OBSs (Supporting information, Table S4). Twelve OBSs assessed the balance between treatment groups for various confounding factors at study enrollment (Supporting information, Table S2), but only two OBSs adjusted for confounding factors in estimating treatment effects [39,48]. If not provided through an integrated, comprehensive prenatal addiction treatment program, MET was generally administered by standalone clinics or pharmacies while BUP was provided by office-based practitioners. Visual inspection of the funnel plot for preterm birth revealed no evidence of publication bias (Supporting information, Figure S2); no other outcomes had enough studies to yield a reliable funnel plot.

Spontaneous fetal death

Among the seven studies [19–21,35,38,39,41] that assessed spontaneous fetal death, five medium ROB studies reported at least one (Supporting information, Table S5). The difference in treatment effect between BUP and MET was not significant from two RCTs [19,20] and three OBSs [38,39,41] (Table 2 and Fig. 2a), but the direction differed by study design. The risk estimate did not change in sensitivity analyses that included all fetal deaths (spontaneous deaths and elective terminations) or imputed missing pregnancy outcomes.

Fetal/congenital anomalies

Nine studies [19,21,34,38,39,41,43,46,48] that evaluated malformations or other defects at birth or pregnancy loss identified chromosomal defects and cardiovascular, central nervous system, craniofacial and musculoskeletal malformations (Supporting information, Table S7). Only two studies treated essentially all women with MET from preconception to end of pregnancy [41,48]. Among the medium ROB RCT [19] and four OBSs [38,41,43,48] that reported at least one defect, the treatment effect was not significantly different in magnitude or direction between BUP and MET (Table 2 and Fig. 2b).

Preterm birth

Seventeen studies [18–21,34–39,41–45,47,48] reported preterm births. The effect estimate from three RCTs [19–21] indicated lower risk of preterm birth for BUP compared with MET. Similarly, the treatment effect among

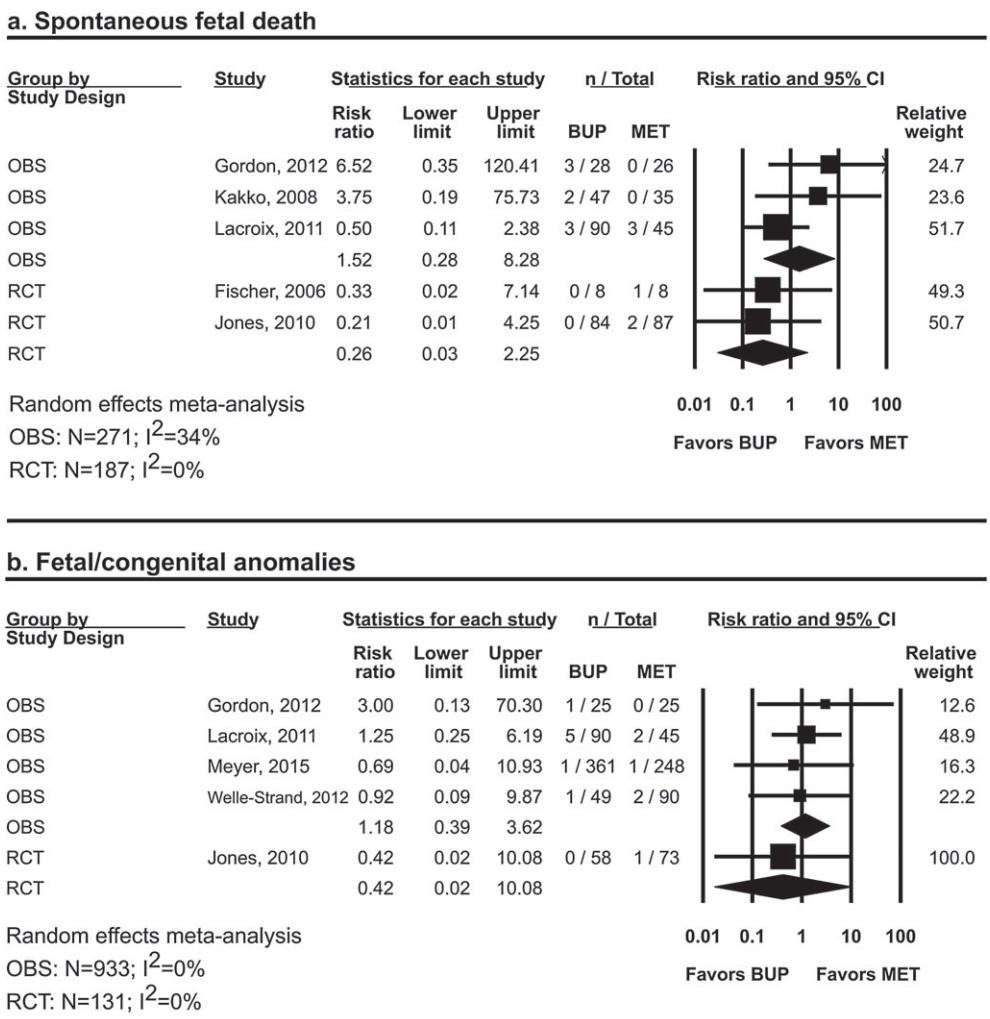


Figure 2 (a) Spontaneous fetal death; (b) fetal/congenital anomalies associated with buprenorphine compared with methadone

seven medium ROB OBSs [18,38,39,41–43,48] showed that BUP was associated with a decreased risk of preterm birth compared with MET (Table 2 and Fig. 3). The treatment effect was similar in sensitivity analyses that included six OBSs with high ROB [34–37,45,47] (including one that defined preterm birth as before 36 weeks gestation) [37] or imputed missing data.

Infant growth outcomes

Eight of 14 studies [19,21,39,41–43,48,49] reporting birth weight had medium ROB. In two RCTs [19,21] BUP-exposed neonates averaged 324 g heavier than MET-exposed neonates. In six OBSs [39,41–43,48,49] the mean difference was 265 g (Table 2 and Fig. 4a). Results were similar in sensitivity analyses that included four high ROB studies [35,37,45,47]. The treatment effect attenuated and was non-significant in two OBS that adjusted for gestational age at birth [39] or maternal age, cigarette smoking, polysubstance use, OMAT dose and duration of dependence [48]. Preterm births in the studies included in the birth weight meta-analysis ranged from 0 to 19% in the two RCTs (0 and 7% for BUP; 9 and 19% for MET) and 4–19% among the five OBSs (4–19% for BUP; 8–17% for MET). One OBS excluded preterm births from the birth weight analysis [49].

Head circumference was similarly significantly larger in infants born to BUP-treated than MET-treated women among seven medium ROB studies [19,21,39,42,43,48,49]. In two RCTs [19,21], BUP-exposed newborns' heads averaged 0.90 cm larger than MET-exposed newborns. In five OBSs [39,42,43,48,49], mean head circumference was 0.68 cm larger in BUP- than MET-exposed infants (Table 2 and Fig. 4b). The treatment effect did not differ after adjustment for a number of factors (excluding gestational age) in one OBS [48].

Table 2 Summary of findings[a] and strength of evidence for buprenorphine compared with methadone treatment of opioid use disorder during pregnancy.

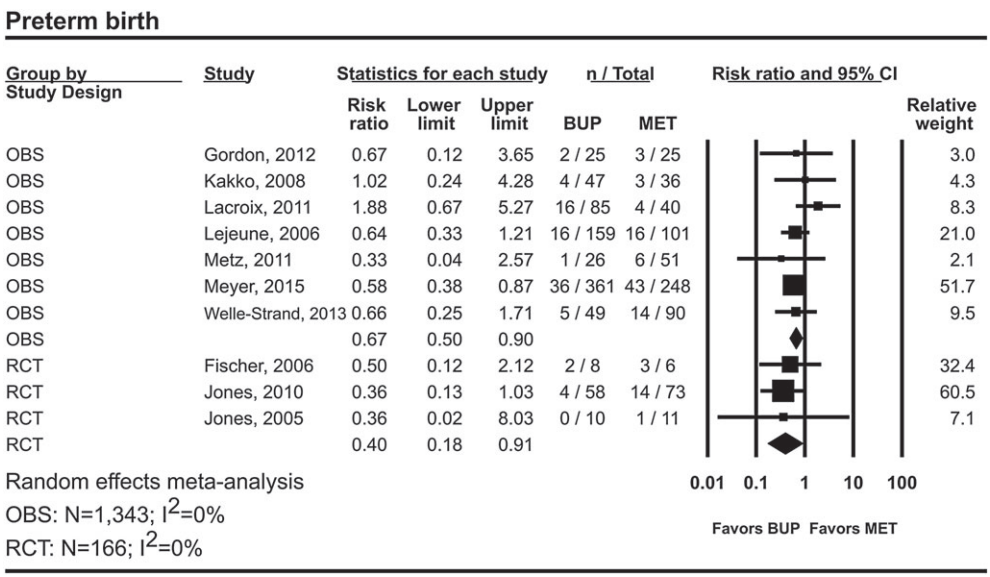
Outcome	*No. of studies*[a] *(n pregnancies or live births)*	*Summary effect size*[a] *(95% CI)*	*Strength of evidence grade*[b]
Spontaneous fetal death			
RCT	2 (187)	RR = 0.26 (0.03–2.31)	Low
Observational	3 (271)	RR = 1.17 (0.32–4.27)	Low
Fetal/congenital anomalies			
RCT	1 (131)	RR = 0.42 (0.02–10.08)	Insufficient
Observational	4 (933)	RR = 1.18 (0.39–3.62)	Low
Preterm birth			
RCT	3 (166)	RR = 0.40 (0.18–0.91)[c]	Low
Observational	7 (1343)	RR = 0.67 (0.50–0.90)[d]	Moderate
Birth weight, g			
RCT	2 (150)	WMD = 324 (32–617)	Low
Observational	6 (1085)	WMD = 265 (196–335)	Moderate
Low birth weight			
Observational	2 (222)	0.51 (0.17–1.59)	Low
Small for gestational age			
RCT	1 (131)	RR = 0.63 (0.06–6.77)	Insufficient
Observational	2 (692)	RR = 0.67 (0.34–1.31)	Low
Intrauterine growth restriction			
Observational	2 (385)	RR = 0.80 (0.57–1.12)	Low
Head circumference, cm			
RCT	2 (150)	WMD = 0.90 (0.14–1.66)	Low
Observational	5 (960)	WMD = 0.68 (0.41–0.94)	Moderate
Sudden infant death syndrome (SIDS)			
Observational	1 (83)	0% BUP versus 6% MET ($P = 0.19$)	Insufficient
Neurodevelopment (fetal and child)			
RCT	1 (175)	*Fetal heart rate and motor activity suppression (third trimester): BUP < MET ($P < 0.05$)[e]	Insufficient
Observational	2 (198)	*Visual selective attention at 4 months of age: no significant difference BUP versus MET[e] *Visual latency at 52 months of age: BUP < MET (prolonged) ($P = 0.02$)[e]	Insufficient
Non-serious maternal adverse events			
RCT	1 (175)	77% BUP versus 93% MET ($P = 0.003$)	Insufficient
Serious maternal adverse events			
RCT	1 (175)	9% BUP versus 16% MET ($P = 0.19$)	Insufficient
Maternal death	0 (0)	NA	Insufficient

BUP = buprenorphine; MET = methadone; NA = not applicable; WMD = weighted mean difference; RCT = randomized controlled trial; CI = confidence interval; RR = relative risk. [a]Includes only studies with low or medium risk of bias and cases with an outcome available. [b]Based on assessment of five domains: study limitations (risk of bias), consistency, directness, precision and reporting bias (Berkman 2013 [33]). See Supporting information, Table S6 for definitions and the full findings. [c]A relative risk of 0.40 would result in 120 fewer premature infants per 1000 births in BUP-treated pregnant women compared with MET-treated women. See Fig. 3 for details. [d]A relative risk of 0.67 would result in 49 fewer premature infants per 1000 births in BUP-treated pregnant women compared with MET-treated women. See Fig. 3 for details. [e]Preliminary evidence: the clinical significance of these findings is unknown.

Conversely, the risk of low birth weight (< 2500 g) did not differ significantly between BUP and MET exposure in two medium ROB OBSs [39,48] (Table 2 and Fig. 5a), nor did the risk of being SGA in two medium ROB OBSs [39,43] (Table 2 and Fig. 5b). The effect estimate was similar when two high ROB studies [36,47] were included. One RCT [19] observed SGA in 2% of BUP-exposed and 3% of MET-exposed infants, a non-significant difference. The RR for IUGR from two medium ROB OBSs [18,41] was 0.80 (Table 2 and Fig. 5c) and did not change appreciably when a high ROB cohort study [35] was included.

Sudden infant death syndrome

One medium ROB OBS ($n = 83$) [39] observed that no infants exposed to BUP *in utero* experienced sudden infant death syndrome (SIDS) within 4 months of birth compared with two of 36 exposed to MET (5.6%), a non-significant difference (Table 2). One death occurred in a 5-week old male born at 38 weeks and treated for NAS. The MET-treated mother was HIV-positive and smoked 10–15 cigarettes daily. The other death was an 8-week old female delivered at 38 weeks via caesarean section for IUGR and



Weights are from random effects meta-analysis. Size of the data markers reflects the study weight.

The interpretation of *relative risk* (RR) depends on the *actual risk* of the outcome in the population. Risk is defined as the number of cases (*C*) over the size of the population at risk (*N*), i.e., *Risk=C/N*. Relative risk for Treatment 1 compared with Treatment 2 is defined as *Risk1/Risk2*. If *Risk1* is lower than *Risk2* for a population of size *N*, the number of cases averted by using Treatment 1 is defined as: *N*Risk2 - N*Risk1, or N*Risk2*(1-RR).*

This formula can project the number of averted preterm births per population at risk if treated with BUP compared with MET. For example, a RR of 0.4 for preterm birth with prenatal BUP treatment compared to MET would translate into the actual number of averted cases as follows.

The prevalence of preterm births in MET-treated pregnancies among the RCTs was 18/90, or 20%. Among 1000 MET-treated pregnancies, the expected number of infants born prematurely would be 200. Assuming a relative risk of 0.4 (RR in the RCTs), among 1000 pregnancies treated with BUP the number of infants born prematurely would be 0.4*200= 80. Thus, the relative risk of 0.4 for BUP compared with MET would result in 120 *fewer* premature infants per 1000 births.

Similarly, among the OBSs, the prevalence of preterm births in MET-treated pregnancies was 15% (89/591). Among 1000 MET-treated pregnancies, the expected number of infants born prematurely would be 150. Assuming a relative risk of 0.67, the expected number of premature births among 1000 BUP-treated pregnancies would be 0.67*150 = 101, resulting in 49 *fewer* premature infants per 1000 births.

Figure 3 Preterm birth associated with buprenorphine compared with methadone

treated for NAS. The effect estimate was similar when a high ROB OBS [46] was included.

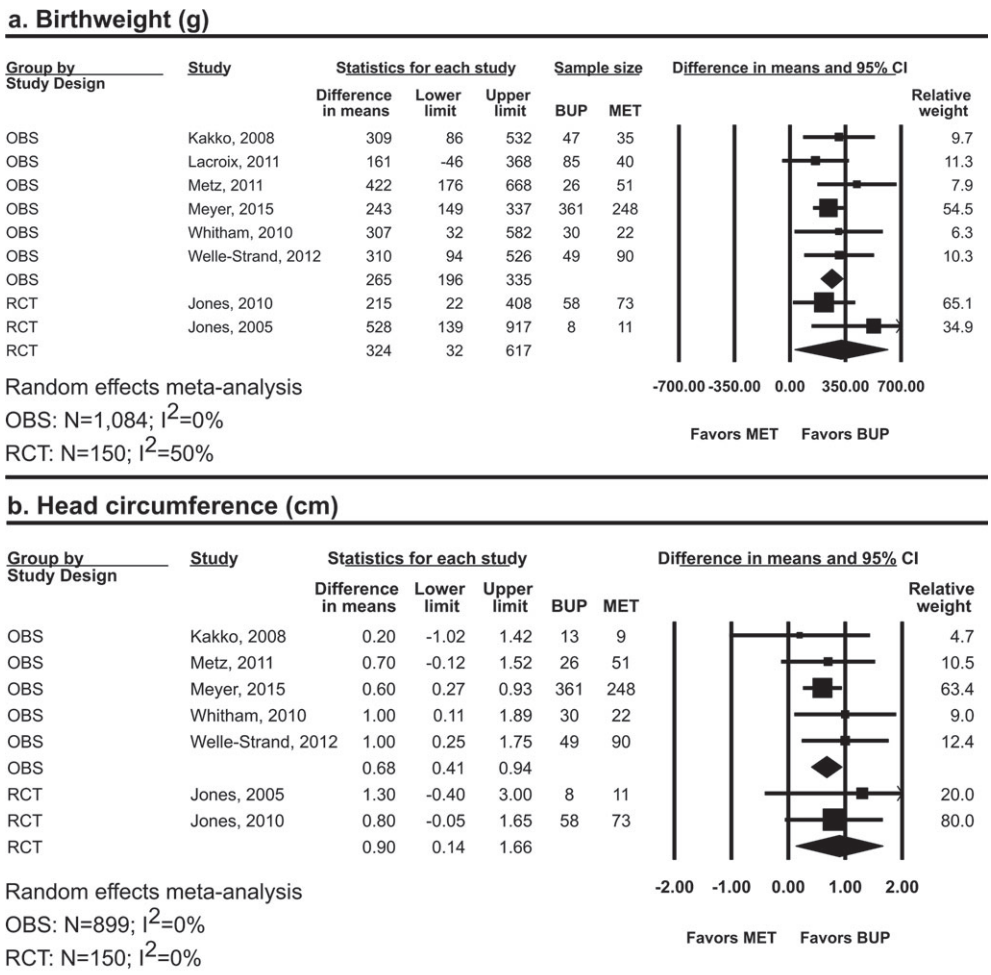
Fetal/child neurodevelopment

As gestation progresses, coupling of fetal movement and heart rate increases, reflecting coordination of autonomical and somatic nervous systems and general fetal wellbeing [55]. Neurobehavior can be monitored non-invasively with non-stress testing (NST) that assesses heart rate variability and associated fetal movement. A high ROB substudy of a RCT assessed maternal blood levels and NST after daily dosing of OMAT during the third trimester [51]. Peak OMAT blood levels were associated with significantly less suppression of fetal heart rate variability and movement and more favorably reactive NSTs in BUP- versus MET-treated women (Supporting information, Table S8). Another RCT substudy found similar differential treatment-related effects on heart rate and movement in the early versus late third trimester [50].

Two medium ROB cohort studies assessed visual development in infants and children exposed prenatally to OMAT (Supporting information, Table S8). One found no significant difference in visual selective attention among 31 children of mothers treated with BUP versus MET [40]. The second study found significantly prolonged visual latency in 22 infants of MET-treated mothers compared with 30 infants of BUP-treated mothers [49].

Adverse *effects*

Three included studies reported non-fatal maternal AEs [19,41,44] (Supporting information, Table S9), while none reported any maternal deaths. One RCT ($n = 175$) [19] assessed AEs weekly and graded them as serious or non-serious. The RCT observed a lower risk of non-serious AEs in BUP-treated women but no difference in the risk of serious AEs. Two high ROB OBSs that did not describe how AEs were collected or assessed had disparate findings. One study ($n = 90$) [44] reporting selected AEs that are



Weights are from random effects meta-analysis. Size of the data markers reflects the study weight.
Note that 'Favors BUP' and 'Favors MET' are on opposite sides of the line of no effect compared to the forest plots for harm outcomes.

Figure 4 (a) Birth weight; (b) head circumference associated with buprenorphine compared with methadone

typically considered serious found no significant treatment-related difference for BUP–naloxone versus MET. The other study ($n = 135$) [41] reported all AEs and found an increased risk of AEs overall for BUP.

DISCUSSION

We synthesized the evidence from three RCTs and 15 OBSs that compared buprenorphine and methadone treatment of pregnant women with OUD. We calculated treatment-related risk estimates for eight pregnancy-related outcomes, including four without previous published meta-analysis. Our work confirms and extends previous treatment risk estimates from limited RCT evidence by also synthesizing the available, larger body of observational study evidence.

Consistent with previous RCT-based meta-analyses [5,22], we identified no statistically or clinically significant difference between BUP and MET in the risk estimates for spontaneous fetal death among the OBSs and across all studies. However, the paucity of events and small sample sizes limited the precision of estimates, ability to stratify by early versus late pregnancy losses and the confidence in our estimates of the relationship between fetal death and OMAT. The overall frequency of spontaneous fetal deaths in women with OUD among both study types was substantially lower than the estimated 15–20% in the general population. Further, most occurred after the first trimester, in contrast to three-fourths of spontaneous losses during the first trimester in the general population [56,57]. This apparent underestimate is probably related to a delay by many opioid-dependent pregnant women in seeking prenatal care until after completing the high-risk first trimester [35,58,59] and insufficient reporting of time of enrollment in several studies [36,37,44–46]. Our ability to assess the relationship between fetal death and OMAT was also limited substantially, given sparse and inconsistent patient-level reporting of gestational timing (onset

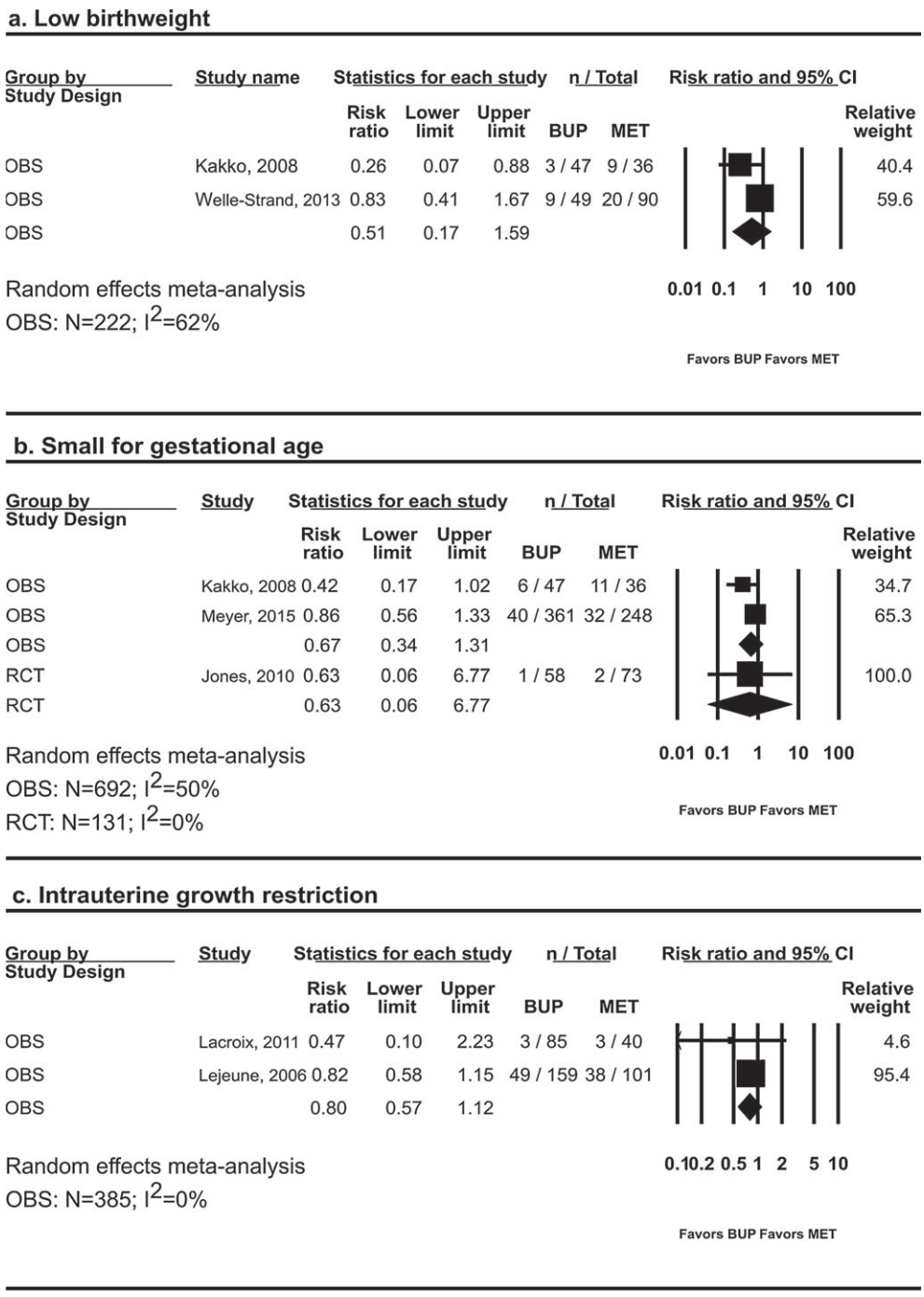


Figure 5 (a) Low birth weight; (b) small for gestational age; (c) intrauterine growth restriction associated with buprenorphine compared with methadone

and duration) of BUP or MET treatment and important potential confounders [60,61].

We found no difference in the risk of fetal/congenital anomalies by maternal treatment. The frequency and type of reported anomalies were broadly similar to what would be expected in the general population, with no particular patterns noted by treatment group. However, most studies characterized the reported defects poorly and failed to collect or describe relevant confounders adequately or even details of exposure to the opioid agonist, particularly during the critical first trimester. Moreover, no included study was powered to detect an increase in specific congenital anomalies, which occur rarely ($\leq 1/1000$ births), and sparse events and exposed pregnancies limited the precision of estimates. In addition, defects not readily apparent at birth may be under-ascertained, as no studies evaluating

congenital anomalies followed infants through the entire first year. Therefore, we have limited confidence in our effect estimates.

The risk of preterm birth was lower in BUP-exposed infants compared with MET-exposed infants. The risk reduction found was consistent between study types and with a previous meta-analysis of the same RCTs examined in the present analysis [5]. However, potentially confounding influences were not reported or adjusted for in any treatment effect estimates. Notwithstanding these limitations, we have moderate confidence in our findings.

Unadjusted birth weight and head circumference were significantly greater in infants of BUP-treated mothers compared with MET-treated mothers. The findings were consistent between study designs and with previous meta-analyses [5,22]. We have moderate confidence in these treatment estimates, but adequate adjustment of confounding factors, particularly gestational age, and larger sample sizes would probably provide more stable and valid estimates of treatment effect. A very small body of observational studies showed no association between prenatal BUP or MET and LBW, IUGR or SGA. The sparse body of evidence limits confidence in the LBW, SGA and IUGR findings.

Fetal growth and birth parameters are influenced by sex, gestational age, multi-fetal pregnancy, maternal cigarette smoking and use of other substances, and placental and anatomical factors [62]. Studies in this review were inconsistent in describing whether they included multi-fetal pregnancies and preterm births (both tend to be smaller) in analyses of growth parameters. Multi-fetal pregnancies were infrequent and unlikely to significantly impact effect estimates differentially. However, failure to adjust for gestational age or exclude preterm births from growth parameter analyses may overestimate the effects of maternal BUP treatment due to BUP's associated significantly lower risk of preterm birth. We were unable to explore fully this confounding effect without patient-level data. Aggregated source data for birth weight and head circumference also limited the clinical interpretation of treatment effect estimates because established norms, and thus minimally important differences, are sex- and gestational age-dependent [63].

Data from three small studies provided preliminary and insufficient evidence that maternal BUP treatment may be associated with more favorable fetal neurobehavior than MET treatment. The developing fetal nervous system appeared more vulnerable to opioid-related suppression earlier versus later in pregnancy with significantly less suppression of fetal heart rate and movement by BUP compared with MET, at least transiently at peak maternal exposure associated with once-daily dosing [64]. Split-dose administration of MET has been associated with less fetal suppression [65].

One medium ROB RCT that collected and analyzed maternal AEs systematically found significantly fewer non-serious AEs but no difference in serious AEs among BUP-treated women versus MET-treated women. Differential cardiovascular effects are plausible due to the established risk of QT-interval prolongation and serious arrhythmia associated with MET [66,67]. Two high ROB OBSs with poorly characterized methods of collecting and analyzing AEs had discordant findings. In one study, with significantly more AEs in BUP-treated women, the cohorts were comparable for several confounders but the MET-treated women had more frequent study visits (a confounding co-intervention) [41]. The evidence regarding AEs is insufficient to draw a clinically meaningful conclusion in either direction. Future studies should collect and analyze treatment-associated AEs during pregnancy in a systematic and standardized fashion and use an established system to code and analyze AEs descriptively [68].

A strength of this review is the inclusion of all available evidence regarding opioid agonist treatment during pregnancy, including data from well-conducted observational cohort studies [69–71]. Previous published systematic reviews and meta-analyses included only three RCTs, with a total combined sample size of 223 drawn from seven university treatment centers in the United States and Austria. The addition of 1923 participants in 15 cohort studies conducted in six additional countries and among a wider range of clinical settings increased the precision, statistical power and generalizability of our findings. Furthermore, most outcomes examined were largely objective, documented routinely in clinical obstetric practice and thus less prone to detection bias from measurement error and lack of blinding. Concordance between the treatment-related risk estimates from both the RCTs and OBSs bolstered confidence in the strength of the evidence for spontaneous fetal death, fetal/congenital anomalies, preterm birth, birth weight, head circumference and SGA.

The main limitations were the uneven quality of the studies and limited number of events and sample sizes, potentially providing low statistical power to detect between-group differences. RCTs provide the most consistent and unbiased estimates of treatment effects, but high-quality RCTs often are not available, particularly in vulnerable populations such as the one under study [68,69]. The complexities of both OUD and pregnancy present daunting challenges in the design, recruitment and conduct of rigorous clinical studies. Moreover, RCTs are generally not designed with sufficient sample size (especially for rare outcomes such as fetal death and congenital anomalies), follow-up duration or population variability for results generalizable to the population at large. The RCTs included in this review were conducted rigorously, but suffered from relatively high levels of

overall and differential attrition that increased the risk of selection bias and were not accounted for optimally in the published analyses [28,70] Twelve of the 15 observational studies assessed baseline comparability of the cohorts for a few confounders or co-interventions but did not adjust effect estimates for imbalances. For example, many studies did not assess, report or adjust for concomitant substance use during pregnancy as evaluated by urine toxicology or self-report. Information on substance use would inform the interpretation of the OMAT-related effects in terms of possible differences between the study participants.

Finally, maternal AEs and fetal/congenital anomalies (and, to a lesser extent, SGA) were defined inconsistently or ascertained among the studies that reported them, increasing clinical heterogeneity and limiting the opportunity to pool results among studies. In clinical studies, AEs are often not collected, analyzed or reported in a standardized and systematic fashion [68].

CONCLUSION

BUP treatment of maternal opioid use disorder during pregnancy was not associated with greater harms than MET treatment, and moderately strong evidence indicated lower risk of preterm birth, greater birth weight and larger head circumference with BUP. Our results confirm and extend previous RCT evidence and further inform benefit/risk assessment in clinical decision-making regarding treatment of pregnant women with OUD, although evidence is currently insufficient to establish superior safety of either opioid agonist during pregnancy for all maternal, fetal and child outcomes examined.

Declaration of interests

Partial funding for this review was provided by Indivior PLC (formerly Reckitt Benckiser Pharmaceuticals) to Venebio Group, LLC. The review was conceived, designed, executed and reported by the authors, who had sole control over the literature selected, data analysis, interpretation and manuscript preparation. Indivior PLC was asked to review the final manuscript for proprietary information. The opinions and conclusions of the authors are their own and do not necessarily reflect the position of Indivior. At the time the work was conducted, B.K.Z., A.L.M., M.M.K., H.R. A., A.R.J. and E.L.M. were paid consultants of Venebio Group, LLC, which has had research and consulting agreements with Indivior PLC. H.E.J. has no financial ties to either Indivior PLC or Venebio Group, LLC, and did not receive any form of remuneration in the preparation or writing of this paper. All authors report no other potential conflicts of interest with the gaming, pharmaceutical, alcohol or tobacco industries.

Acknowledgements

The authors wish to thank Drs Georgiy Bobashev, Dale Glasser, and Pradeep Rajan for helpful comments on drafts of the text, Dr Laura Morgan for contributions to the assessment of study quality and preparation of forest plots and Dr Pradeep Rajan and Molly Ronayne Sherwood for assistance in the literature search.

References

1. Substance Abuse and Mental Health Services Administration (SAMHSA). Results from the 2010 National Survey on Drug Use and Health: Summary of National Findings. Rockville, MD: SAMHSA, 2011 Contract No.: HHS Publication no. (SMA) 11–4658.
2. Maeda A., Bateman B. T., Clancy C. R., Creanga A. A., Leffert L. R. Opioid abuse and dependence during pregnancy: temporal trends and obstetrical outcomes. *Anesthesiology* 2014; **121**: 1158–65.
3. Gyarmathy V. A., Giraudon I., Hedrich D., Montanari L., Guarita B., Wiessing L. Drug use and pregnancy—challenges for public health. *Euro Surveill* 2009; **14**: 33–36.
4. De Leon G., Jainchill N. Residential therapeutic communities for female substance abusers. *Bull NY Acad Med* 1991; **67**: 277–90.
5. World Health Organization (WHO). Guidelines for identification and management of substance use and substance use disorders in pregnancy 2014. Available at: http://www.who.int/substance_abuse/publications/pregnancy_guidelines/en/ (accessed 1 September 2015) (Archived by WebCite® at http://www.webcitation.org/6gOmIM2FB).
6. Zuspan F., Gumpel J., Mejia-Zelaya A., Madden J., Davis R. Fetal stress from methadone withdrawal. *Am J Obstet Gynecol* 1975; **122**: 42–6.
7. Dashe J. S., Jackson G. L., Olscher D. A., Zane E. H., Wendel G. D. Jr. Opioid detoxification in pregnancy. *Obstet Gynecol* 1998; **92**: 854–8.
8. Luty J., Nikolaou V., Bearn J. Is opiate detoxification unsafe in pregnancy? *J Subst Abuse Treat* 2003; **24**: 363–7.
9. Young J. L., Martin P. R. Treatment of opioid dependence in the setting of pregnancy. *Psychiatr Clin North Am* 2012; **35**: 441–60.
10. American College of Obstetrics and Gynecology The American College of Obstetrics and Gynecology (ACOG) Committee Opinion No. 524: Opioid abuse, dependence, and addiction in pregnancy. *Obstet Gynecol* 2012; **119**: 1070–76.
11. Jones H. E., Finnegan L. P., Kaltenbach K. Methadone and buprenorphine for the management of opioid dependence in pregnancy. *Drugs* 2012; **72**: 747–57.
12. Metz V., Köchl B., Fischer G. Should pregnant women with substance use disorders be managed differently? *Neuropsychiatry* 2012; **2**: 29–41.
13. Winklbaur B., Kopf N., Ebner N., Jung E., Thau K., Fischer G. Treating pregnant women dependent on opioids is not the same as treating pregnancy and opioid dependence: a knowledge synthesis for better treatment for women and neonates. *Addiction* 2008; **103**: 1429–40.
14. Winklbaur-Hausknost B., Jagsch R., Graf-Rohrmeister K., Unger A., Baewert A., Langer M., *et al.* Lessons learned from a comparison of evidence-based research in pregnant opioid-dependent women. *Hum Psychopharmacol* 2013; **28**: 15–24.

15. Kaltenbach K., Berghella V., Finnegan L. Opioid dependence during pregnancy. Effects and management. *Obstet Gynecol Clin North Am* 1998; **25**: 139–51.
16. World Health Organization (WHO). Guidelines for the Psychosocially Assisted Pharmacological Treatment of Opioid Dependence. Geneva: WHO; 2009.
17. Joseph H., Stancliff S., Langrod J. Methadone maintenance treatment (MMT): a review of historical and clinical issues. *Mt Sinai J Med* 2000; **67**: 347–64.
18. Lejeune C., Simmat-Durand L., Gourarier L., Aubisson S. Prospective multicenter observational study of 260 infants born to 259 opiate-dependent mothers on methadone or high-dose buprenophine substitution. *Drug Alcohol Depend* 2006; **82**: 250–57.
19. Jones H. E., Kaltenbach K., Heil S. H., Stine S. M., Coyle M. G., Arria A. M., *et al.* Neonatal abstinence syndrome after methadone or buprenorphine exposure. *N Engl J Med* 2010; **363**: 2320–31.
20. Fischer G., Ortner R., Rohrmeister K., Jagsch R., Baewert A., Langer M., *et al.* Methadone versus buprenorphine in pregnant addicts: a double-blind, double-dummy comparison study. *Addiction* 2006; **101**: 275–81.
21. Jones H. E., Johnson R. E., Jasinski D. R., O'Grady K. E., Chisholm C. A., Choo R. E., *et al.* Buprenorphine versus methadone in the treatment of pregnant opioid-dependent patients: effects on the neonatal abstinence syndrome. *Drug Alcohol Depend* 2005; **79**: 1–10.
22. Minozzi S., Amato L., Bellisario C., Ferri M., Davoli M. Maintenance agonist treatments for opiate-dependent pregnant women. *Cochrane Database Syst Rev* 2013; **12**: CD006318.
23. Liberati A., Altman D. G., Tetzlaff J., Mulrow C., Gotzsche P. C., Ioannidis J. P., *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLOS Med* 2009; **6**e1000100.
24. Viswanathan M., Ansari M. T., Berkman N. D., Chang S., Hartling L., McPheeters M., Santaguida P. L., *et al.* Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions; Agency for Healthcare Research and Quality Methods Guide for Comparative Effectiveness Reviews. March 2012. AHRQ Publication No. 12-EHC047-EF. Available at: www.effectivehealthcare.ahrq.gov/
25. Guyatt G., Busse J. Methods Commentary: Risk of Bias in Cohort Studies: Evidence Partners; 2015. Available at: https://distillercer.com/resources/methodological-resources/risk-of-bias-in-cohort-studies/ (accessed 10 September 2015) (Archived at http://www.webcitation.org/6gOpc4uqj on 1 September 2015).
26. Wells G., Shea B., O'Connell D., Peterson J., Welch V., Losos M., *et al.* The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed 1 September 2015) (Archived at http://www.webcitation.org/6gOpODPiP).
27. DerSimonian R., Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177–88.
28. Higgins J. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. Oxford, UK: The Cochrane Collaboration; 2011. Available at: http://handbook.cochrane.org/ (accessed 1 September 2015).
29. Welsh A. W. Randomised controlled trials and clinical maternity care: moving on from intention-to-treat and other simplistic analyses of efficacy. *BMC Pregnancy Childbirth* 2013; **13**: 15.
30. Guyatt G. H., Oxman A. D., Kunz R., Woodcock J., Brozek J., Helfand M., *et al.* GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol* 2011; **64**: 1294–302.
31. West S., Gartlehner G., Mansfield A., Poole C., Tant E., Lenfestey N. *et al.* Comparative Effectiveness Review Methods: Clinical Heterogeneity. Methods Research Paper (AHRQ Publication No 10-EHC070-EF) (internet). 2010. Available at: http://effectivehealthcare.ahrq.gov/ (accessed 1 September 2015) (Archived at http://www.webcitation.org/6gOoAtvlc).
32. Sterne J. A., Egger M., Smith G. D. Systematic reviews in health care: investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001; **323**: 101–5.
33. Berkman N. D., Lohr K. N., Ansari M., McDonagh M., Balk E., Whitlock E. *et al.* Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update Methods Guide for Comparative Effectiveness Reviews (internet). 2013; AHRQ Publication No. 13 (14)-EHC130-EF. Available at: http://effectivehealthcare.ahrq.gov/ehc/products/457/1752/methods-guidance-grading-evidence-131118.pdf (accessed 1 September 2015) (Archived at http://www.webcitation.org/6gOn0I4Tx).
34. Basselin C., Boucher A., Carlier P., Marie-Noelle B., Vial T., Descotes J. Follow-up of pregnancies exposed to buprenorphine maintenance treatment. *Fundam Clin Pharmacol* 2011; **25**: 87.
35. Binder T., Vavrinkova B. Prospective randomised comparative study of the effect of buprenorphine, methadone and heroin on the course of pregnancy, birthweight of newborns, early postpartum adaptation and course of the neonatal abstinence syndrome (NAS) in women followed up in the outpatient department. *Neuro Endocrinol Lett* 2008; **29**: 80–6.
36. Brulet C., Chanal C., Ravel P., Mazurier E., Boulot P., Faucherre V. Multidisciplinary monitoring and psychosocial support reduce complications of opiate dependence in pregnant women: 114 pregnancies. *Presse Med* 2007; **36**: 1571–80.
37. Colombini N., Elias R., Busuttil M., Dubuc M., Einaudi M. A., Bues-Charbit M. Hospital morphine preparation for abstinence syndrome in newborns exposed to buprenorphine or methadone. *Pharm World Sci* 2008; **30**: 227–34.
38. Gordon A. L., Lopatko O. V., Stacey H., Pearson V., Woods A., Fisk A., *et al.* Patterns of symptom reporting during pregnancy comparing opioid maintained and control women. *J Addict Med* 2012; **6**: 258–64.
39. Kakko J., Heilig M., Sarman I. Buprenorphine and methadone treatment of opiate dependence during pregnancy: comparison of fetal growth and neonatal outcomes in two consecutive case series. *Drug Alcohol Depend* 2008; **96**: 69–78.
40. Konijnenberg C., Melinder A. Visual selective attention is impaired in children prenatally exposed to opioid agonist medication. *Eur Addict Res* 2014; **21**: 63–70.
41. Lacroix I., Berrebi A., Garipuy D., Schmitt L., Hammou Y., Chaumerliac C., *et al.* Buprenorphine versus methadone in pregnant opioid-dependent women: a prospective multicenter study. *Eur J Clin Pharmacol* 2011; **67**: 1053–59.
42. Metz V., Jagsch R., Ebner N., Würzl J., Pribasnig A., Aschauer C., *et al.* Impact of treatment approach on maternal and neonatal outcome in pregnant opioid-maintained women. *Hum Psychopharmacol Clin Exp* 2011; **26**: 412–21.
43. Meyer M. C., Johnston A. M., Crocker A. M., Heil S. H. Methadone and buprenorphine for opioid dependence during

pregnancy: a retrospective cohort study. *J Addict Med* 2015; **9**: 81–6.
44. Prasad M., Calaway B., Abdel-Rasoul M. Obstetric and neonatal sequelae of buprenorphine/naloxone (B/N) exposure versus methadone maintenance (MM). *Am J Obstet Gynecol* 2013; **208**: S201.
45. Rohrmeister K., Bernert G., Langer M., Fischer G., Weninger M., Pollak A. Opiate addiction in gravidity—consequences for the newborn. Results of an interdisciplinary treatment concept. *Z Geburtshilfe Neonatol* 2001; **205**: 224–30.
46. Sandtorv L., Reigstad H., Bruaroy S., Elgen I., Laegreid L. M. Substitution treatment of drug addicts during pregnancy: consequences for the children? *Tidsskr Nor Laegeforen* 2009; **129**: 287–90.
47. Siedentopf J. P., Nagel M., Esser M., Casteleyn S., Dudenhausen J. W. Experience with buprenorphine induction and subsequent dose reduction as compared to treatment with L-methadone in pregnant opiate addicts. *Geburtshilfe Frauenheilkd* 2004; **64**: 711–18.
48. Welle-Strand G. K., Skurtveit S., Jones H. E., Waal H., Bakstad B., Bjarko L., *et al.* Neonatal outcomes following *in utero* exposure to methadone or buprenorphine: a National Cohort Study of opioid-agonist treatment of pregnant women in Norway from 1996 to 2009. *Drug Alcohol Depend* 2013; **127**: 200–6.
49. Whitham J. N., Spurrier N. J., Sawyer M. G., Baghurst P. A., Taplin J. E., White J. M., *et al.* The effects of prenatal exposure to buprenorphine or methadone on infant visual evoked potentials. *Neurotoxicol Teratol* 2010; **32**: 280–8.
50. Jansson L. M., Dipietro J. A., Velez M., Elko A., Williams E., Milio L., *et al.* Fetal neurobehavioral effects of exposure to methadone or buprenorphine. *Neurotoxicol Teratol* 2011; **33**: 240–3.
51. Salisbury A. L., Coyle M. G., O'Grady K. E., Heil S. H., Martin P. R., Stine S. M., *et al.* Fetal assessment before and after dosing with buprenorphine or methadone. *Addiction* 2012; **107**: 36–44.
52. Holbrook A. M., Baxter J. K., Jones H. E., Heil S. H., Coyle M. G., Martin P. R., *et al.* Infections and obstetric outcomes in opioid-dependent pregnant women maintained on methadone or buprenorphine. *Addiction* 2012; **107**: 83–90.
53. Larance B., Lintzeris N., Ali R., Dietze P., Mattick R., Jenkinson R., *et al.* The diversion and injection of a buprenorphine–naloxone soluble film formulation. *Drug Alcohol Depend* 2014; **136**: 21–7.
54. Martin J. Pregnancy and Buprenorphine Treatment. Providers' Clinical Support System for Medication Assisted Treatment 2014. Available at: http://pcssmat.org/opioid-resources/clinical-tools/ (accessed 27 March 2016) (Archived at http://www.webcitation.org/6gQTmOQHT).
55. DiPietro J. A., Hodgson D. M., Costigan K. A., Hilton S. C., Johnson T. R. Fetal neurobehavioral development. *Child Dev* 1996; **67**: 2553–67.
56. Farquharson R., Exalto N. Early pregnancy—models of healthcare. In: Farquharson R., Stephenson M., editors. Early Pregnancy. New York: Cambridge University Press; 2010. pp. 1-8.
57. Cousens S., Blencowe H., Stanton C., Chou D., Ahmed S., Steinhardt L., *et al.* National, regional, and worldwide estimates of stillbirth rates in 2009 with trends since 1995: a systematic analysis. *Lancet* 2011; **377**: 1319–30.
58. Kandall S. R., Albin S., Gartner L. M., Lee K. S., Eidelman A., Lowinson J. The narcotic-dependent mother: fetal and neonatal consequences. *Early Hum Dev* 1977; **1**: 159–69.
59. Boer K. Substance use in pregnancy: do we care? *Acta Paediatr* 1994; **83**: 65–71.
60. Atik R., Hepworth-Jones B., Doyle P. Risk factors for miscarriage. In: Farquharson R., Stephenson M., editors. Early Pregnancy. New York: Cambridge University Press; 2010. pp. 9-18
61. Maconochie N., Doyle P., Prior S., Simmons R. Risk factors for first trimester miscarriage—results from a UK-population-based case–control study. *Br J Obstet Gynaecol* 2007; **114**: 170–86.
62. Reeves S., Bernstein I. Effects of maternal tobacco-smoke exposure on fetal growth and neonatal size. *Exp Rev Obstet Gynecol* 2008; **3**: 719–30.
63. Olsen I. E., Groveman S. A., Lawson M. L., Clark R. H., Zemel B. S. New intrauterine growth curves based on United States data. *Pediatrics* 2010; **125**: e214–24.
64. Konijnenberg C., Melinder A. Prenatal exposure to methadone and buprenorphine: a review of the potential effects on cognitive development. *Child Neuropsychol* 2011; **17**: 495–519.
65. Wittman B., Segal S. A comparison of the effects of the single- and split-dose methadone administration on the fetus: ultrasound evaluation. *Int J Addict* 1991; **26**: 213–8.
66. Anchersen K., Clausen T., Gossop M., Hansteen V., Waal H. Prevalence and clinical relevance of corrected QT interval prolongation during methadone and buprenorphine treatment: a mortality assessment study. *Addiction* 2009; **104**: 993–9.
67. Hanon S., Seewald R. M., Yang E., Schweitzer P., Rosman J. Ventricular arrhythmias in patients treated with methadone for opioid dependence. *J Interv Card Electrophysiol* 2010; **28**: 19–22.
68. Ioannidis J. P., Evans S. J., Gotzsche P. C., O'Neill R. T., Altman D. G., Schulz K., *et al.* Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; **141**: 781–8.
69. Chou R., Aronson N., Atkins D., Ismaila A. S., Santaguida P., Smith D. H., *et al.* AHRQ Series Paper 4: Assessing harms when comparing medical interventions: AHRQ and the Effective Health-Care Program. *J Clin Epidemiol* 2010; **63**: 502–12.
70. Egger M., Schneider M., Davey S. G. Spurious precision? Meta-analysis of observational studies. *BMJ* 1998; **316**: 140–4.
71. Norris S. L., Atkins D., Bruening W., Fox S., Johnson E., Kane R., *et al.* Observational studies in systematic [corrected] reviews of comparative effectiveness: AHRQ and the Effective Health Care Program. *J Clin Epidemiol* 2011; **64**: 1178–86.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.
Figure S2 Funnel plot for studies reporting preterm birth.
Table S1 Literature search strategy.
Table S2 Characteristics of included studies.
Table S3 Risk of bias in included observational studies.
Table S4 Risk of bias in included randomized controlled trials (RCTs).

Table S5 Spontaneous fetal death associated with buprenorphine compared with methadone: individual studies.

Table S6 Summary of findings and grading the strength of evidence for buprenorphine compared with methadone treatment of opioid use disorder during pregnancy.

Table S7 Fetal/congenital anomalies associated with buprenorphine compared with methadone: individual studies.

Table S8 Fetal and child neurodevelopment associated with buprenorphine compared with methadone: individual studies.

Table S9 Maternal adverse events associated with buprenorphine compared with methadone: individual studies.

