European Journal of Obstetrics & Gynecology and Reproductive Biology 146 (2009) 138–148

Contents lists available at ScienceDirect

European Journal of Obstetrics & Gynecology and Reproductive Biology

journal homepage: www.elsevier.com/locate/ejogrb

Review

Preterm birth and low birth weight among *in vitro* fertilization singletons: A systematic review and meta-analyses

Sarah D. McDonald [a,*], Zhen Han [b], Sohail Mulla [c], Kellie E. Murphy [d,e], Joseph Beyene [e,f], Arne Ohlsson [d,e,g] on behalf of the Knowledge Synthesis Group[1]

[a] Division of Maternal-Fetal Medicine, Departments of Obstetrics & Gynecology, and Diagnostic Imaging, McMaster University, Hamilton, Canada
[b] Department of Obstetrics and Gynecology, First Affiliated Hospital of Xi'an Jiao Tong University, Shaanxi Province, PR China
[c] Faculty of Health Sciences, McMaster University, Hamilton, Canada
[d] Department of Obstetrics and Gynecology, Mount Sinai Hospital, Toronto, Ontario, Canada
[e] Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
[f] Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, Ontario, Canada
[g] Department of Paediatrics, University of Toronto, Toronto, Ontario, Canada

ARTICLE INFO

Article history:
Received 12 January 2009
Received in revised form 21 April 2009
Accepted 25 May 2009

Keywords:
In vitro fertilization
Preterm birth
Low birth weight
Meta-analysis
Singletons

ABSTRACT

Our objective was to determine the risks of preterm birth (PTB) and low birth weight (LBW) in singletons conceived through *in vitro* fertilization (IVF) ± intracytoplasmic sperm injection (ICSI) compared to spontaneously conceived singletons after matching or controlling for at least maternal age. The MOOSE guidelines for meta-analysis of observational studies were followed. Medline and Embase were searched using comprehensive search strategies. Bibliographies of identified articles were reviewed. English language studies examining LBW or PTB in singletons conceived by IVF or IVF/intracytoplasmic sperm injection, compared with spontaneously conceived singletons, that matched or controlled for at least maternal age. Two reviewers independently assessed titles, abstracts, full articles and study quality and extracted data. Dichotomous data were meta-analyzed using relative risks (RR) as measures of effect size with a random effects model and for continuous data weighted mean difference was calculated. Seventeen studies were included with 31,032 singletons conceived through IVF (±ICSI) and 81,119 spontaneously conceived singletons. After matching or controlling for maternal age and often other factors, compared to spontaneously conceived singletons, IVF singletons had increased risks of our two primary outcomes, PTB (RR 1.84, 95% CI 1.54, 2.21) and LBW (<2500 g, RR 1.60, 95% CI 1.29, 1.98). Singletons conceived through IVF or IVF/ICSI were at increased risk for late PTB (32–36 weeks, RR 1.52, 95% CI 1.01, 2.30), moderate PTB <32–33 weeks (RR 2.27, 95% CI 1.73, 2.97), very LBW (<1500 g, RR 2.65, 95% CI 1.83, 3.84), and intrauterine growth restriction (RR 1.45, 95% CI 1.04, 2.00), lower birth weights (−97 g, 95% CI −161 g, −33 g) and shorter mean gestations (−0.6 weeks, 95% CI −0.9 weeks, −0.4 weeks). In conclusion, IVF singletons have significantly increased risks of PTB, LBW and other adverse perinatal outcomes compared to spontaneously conceived singletons after matching or controlling for maternal age at least.

Contents

* Corresponding author at: McMaster University, Division of Maternal-Fetal Medicine, Departments of Obstetrics & Gynecology and Diagnostic Imaging and Clinical Epidemiology & Biostatistics, 1200 Main St. West, HSC 3N52B, Hamilton, Ontario L8N 3Z5, Canada. Tel.: +1 905 525 9140x26622; fax: +1 905 524 2911.
E-mail address: mcdonals@mcmaster.ca (S.D. McDonald).

[1] See Appendix C.

0301-2115/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.ejogrb.2009.05.035

1. Introduction

In vitro fertilization (IVF) provides an opportunity for an increasing proportion of couples who were unable to conceive spontaneously to become parents. Following the first successful IVF birth in 1978, the proportion of infants born as a result of this procedure has steadily increased to approximately 1 in 30 in Finland [1] and 1 in 80 in the USA [2].

The last meta-analysis (with the literature search ending in 2003) was published by our group and identified that singletons born after IVF had increased perinatal risks compared to spontaneously conceived singletons after accounting for maternal age [3]. However, there have been a number of studies published since that time, and a current, unbiased, systematic review of the literature is required. Our previous study reported only pooled crude data, whereas we now wanted to also examine data which accounted for confounders. We have focused on singletons given that they make up the largest proportion of IVF infants. This review will concentrate on the two most important determinants of neonatal morbidity and mortality, namely preterm birth (PTB) and low birth weight (LBW) births [4].

2. Materials and methods

We followed the MOOSE consensus statement on the conduct of meta-analysis of observational studies [5].

2.1. Search strategy

Medline (1978 [the year of the first IVF birth] – June 8, 2008) and EMBASE (since its inception in 1980 – June 8, 2008) were searched with the help of an experienced librarian using individual comprehensive search strategies. The search terms are included in Appendix A. We also reviewed the reference lists of identified articles.

2.2. Study selection

2.2.1. Inclusion criteria

We included case–control and cohort studies which examined PTB and LBW if they compared singletons conceived either by IVF or IVF/intracytoplasmic sperm injection (ICSI) with spontaneously conceived singletons with matching or adjustment for maternal age at least. We included information available from the publications only and did not contact primary authors. We included English language articles only.

2.2.2. Exclusion criteria

We excluded studies involving multifetal or selective reduction. We excluded reviews, editorials, case reports, letters to editors, duplicate publications and studies published only as abstracts.

2.3. Outcome measures

Our primary outcomes were PTB (<37 weeks gestation) and LBW (<2500 grams [g]). Our secondary outcomes were

(i) late PTB (32–36 weeks gestation) and early PTB (<32 weeks gestation)
(ii) VLBW (<1500 g) and ELBW (<1000 g)
(iii) intrauterine growth restriction (IUGR, birth weight <10% for gestational age)
(iv) birth weight (in grams)
(v) gestational age at birth (in weeks).

2.4. Study process

Two independent reviewers (two of SDM, ZH and SM) screened the titles and abstracts of all citations identified in the search. The full-text article was retrieved if either reviewer considered the citation potentially relevant. Each full text article was independently evaluated by two reviewers (two of SDM, ZH and SM). Disagreements were resolved by discussion and consensus. An independent adjudicator was available should the need have arisen.

2.5. Data abstraction

Two reviewers (two of SDM, ZH and SM) independently abstracted data from full text articles on country of origin, period of study, study design, characteristics of participants, outcomes and information on bias using a piloted data extraction form. Disagreements were resolved through the above consensus process.

2.6. Data synthesis

Statistical analyses were performed on available data using the Review Manager software (Revman 5.0; the Cochrane Collaboration, Oxford, England). Dichotomous data were meta-analyzed with relative risk (RR) and continuous data with a weighted mean difference. Weighting of studies was based on the inverse variance of the study. The random effects model was chosen because it accounts for both random variability and the variability in effects among the studies as we expected a degree of clinical and statistical heterogeneity among the studies. Crude, matched, and adjusted data were initially each pooled separately and then data that were matched and/or adjusted were pooled together. Where required and when the incidence of the outcome was rare, in order to be able to pool data, adjusted RR were calculated from adjusted OR [6]. As is typical in meta-analyses there was no adjustment for multiple analyses. Clinical heterogeneity was reported in the table of included studies. Statistical heterogeneity was assessed with the

I-squared (I^2) value, which represents the percentage of total variation across studies due to heterogeneity rather than chance [7]. I^2 values of 25%, 50% and 75% have been related to low, moderate and high heterogeneity [7].

Sensitivity analyses were performed using a *priori* chosen groups to examine the effects of: (1) ICSI, (2) cryopreservation and (3) low quality studies with high bias (see Section 2.7).

2.7. Quality assessment

Two reviewers (two of SDM, ZH and SM) independently performed the quality assessment by assessing six types of bias: selection, exposure, outcome, confounding, analytic and attrition. Bias was classified as minimal, low, moderate, high or not reported (Appendix B). We attempted to minimize selection bias by ensuring that studies matched for the key known confounder of maternal age. If "all" or "consecutive" patients or a "random" selection of controls were included selection bias was assessed as "minimal". Exposure and outcome assessment were "minimal" bias if from the hospital record or direct questioning. Confounding bias was assessed as "low" if only age and parity were matched or controlled for and "minimal" if age and two other variables were addressed. Analytic bias was "moderate" if no sample size calculation was done and only a subsample studied, and "high" if inappropriate analyses done. Attrition was "minimal" if <10% were lost to follow-up. Studies with: (1) high risk of bias or "not reported" in three or more domains or (2) an overall assessment of bias as "high" were excluded using a sensitivity analysis. Selection bias and confounding were given predominance in the overall assessment of bias due to their importance in this meta-analysis. Funnel plots were assessed for the possibility of publication bias.

3. Results

Three hundred and sixty-one non-duplicate titles and abstracts were identified, 117 full articles were reviewed and 17 articles met inclusion criteria [8–24] with 31,032 singletons conceived through IVF and IVF/ICSI and 81,119 spontaneously conceived singletons (Fig. 1). Although many studies were labeled as case–control studies and others as retrospective cohort studies, all 17 studies involved retrospectively studying infants conceived following IVF (exposure) and spontaneously conceived infants, to determine rates of PTB and LBW, the two primary outcomes. Hence, these studies were felt to be retrospective cohort studies, and data from all these studies were pooled. Characteristics of included studies are shown in Table 1.

Generally, the magnitude and direction of the increased perinatal risk in singletons born after IVF was similar whether crude data, matched data, adjusted data or adjusted and/or matched data were examined (Table 2). The number of confounders which were matched or adjusted for varied between studies (Table 3). After matching or controlling for maternal age and often other factors, compared to spontaneously conceived singletons, IVF singletons had increased risks of both primary outcomes: PTB (<37 weeks, RR 1.84, 95% CI 1.54, 2.21; Table 2, Fig. 2) and LBW (<2500 g, RR 1.60, 95% CI 1.29, 1.98, Fig. 3). They were at increased risk for late PTB (32–36 weeks, RR 1.52, 95% CI 1.01, 2.30), moderate PTB <32 or 33 weeks (RR 2.27, 95% CI 1.73, 2.97; some studies presented results <32 weeks, and others <33 weeks), and VLBW (<1500 g, RR 2.65, 95% CI 1.83, 3.84). One study reported on the risk of ELBW (<1000 g, RR 3.02, 95% CI 0.12, 74.66). The risk of spontaneous PTB was also higher in IVF infants (RR 2.05, 95% CI 1.44, 2.91, data not shown). The risk of IUGR was higher among IVF infants (RR 1.45, 95% CI 1.04, 2.00). IVF infants had lower birth weights (−97 g, 95% CI −161 g, −24 g) and shorter mean gestations (−0.6 weeks, 95% CI −0.9 weeks, −0.4 weeks).

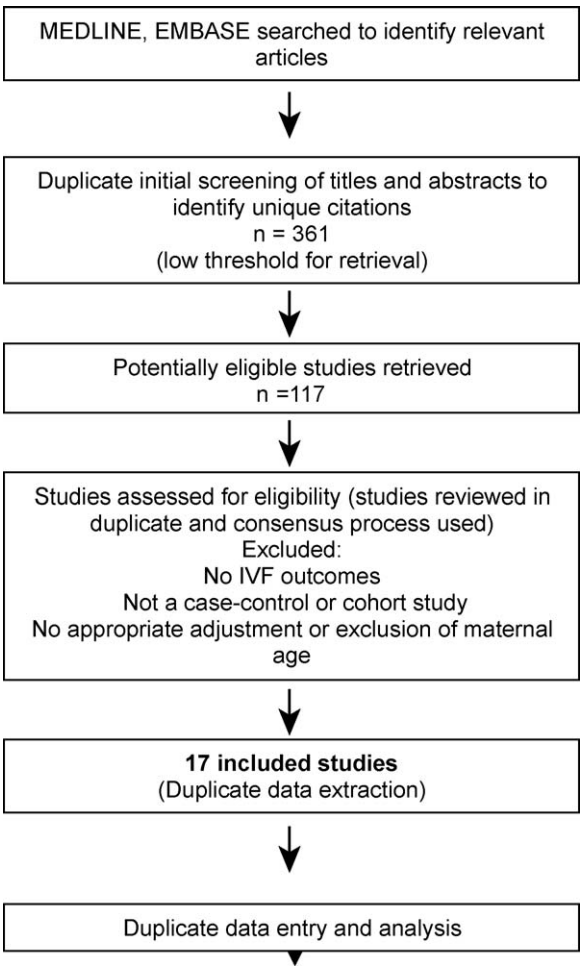

Fig. 1. Study process.

3.1. Quality assessment

Quality assessment was based on the evaluation of six types of bias: (1) selection, (2) exposure, (3) outcome, (4) confounding, (5) analytic and (6) attrition (Table 3). (1) Selection bias was likely in some studies given that IVF and spontaneously conceived infants were often from different populations. (2) Exposure bias was unlikely in most studies, given that data were obtained from charts or registries. (3) There was little outcome bias given that the outcomes of interest have standard definitions (for instance, IUGR was always defined as birth weight <10%) and are objectively measured (such as birth weight). (4) Confounding was minimized by ensuring that studies matched for or controlled for at least maternal age, the key confounder. Other potentially important confounders were variably assessed from study to study. (5) Many of the studies did not calculate a sample size or power calculation. Few studies used matched analyses when indicated. (6) Attrition bias was rare given the short duration of follow-up (the duration of gestation). Because much of the outcome information was from hospital records during the peripartum admission, follow-up was within hospital, and hence information was readily available.

There does not appear to be publication bias as assessed by a funnel plot in either of our primary outcomes, PTB or LBW.

3.2. A priori defined sensitivity analyses

For our *a priori* defined sensitivity analyses (Table 4), we had hoped to be able to compare outcomes in infants born following

Table 1
Characteristics of studies included in meta-analysis of *in vitro* fertilization singletons.

Author, year	Years study spanned	Place of study	Population	Setting	Exposure information ascertained from	Number of IVF infants	Number of spontaneously conceived infants
Bergh et al., 1999 [8]	1982–1995	Sweden	IVF: all children; control: >28 wks	IVF: Swedish Cancer Registry; spontaneous: SMBR	IVF clinics	3305	NR
Dhont et al., 1999 [9]	1992–1997	Dutch-speaking part of Belgium	All ART babies, ≥500 g	Regional register	Direct questioning/survey	3048	3048
Howe et al., 1990 [10]	NR	Delaware and Pennsylvania, USA	First 100 pregnancies >20 wks	IVF: University of Pennsylvania; spontaneous: Medical Centre of Delaware	Telephone calls, medical records	54	54
Isaksson et al., 2002 [11]	1993–1999	Helsinki, Finland	Pregnancies ending in birth, unexplained infertility (IVF/ICSI)	Helsinki University Centre Hospital	NR	69	345
Koivurova et al., 2002 [12]	1990–1995	Provinces of Oulu and Lapland, Northern Finland	Fresh embryos	IVF: outpatient clinic; spontaneous: FMBR	IVF registry	153	287
Koudstaal et al., 2000 [13]	Before 1992	Amsterdam, Leiden, Nymegen, Utrecht, The Netherlands	Ongoing pregnancies >16 wks before 1992; spontaneous: antenatal care provided by same hospital as IVF group	University hospitals of Amsterdam, Leiden, Nikmegen, and Utrecht	NR	307	307
Perri et al., 2001 [14]	1996	Tel Aviv, Israel	All ART at authors' centre; spontaneous: selection from all singleton pregnancies	Rabin Medical Centre	ART database	95	190
Poikkeus et al., 2007 [15]	1997–2003	Helsinki, Finland	Out of all frozen embryo transfers, only viable singletons, ≥500 g	Helsinki University Centre Hospital	Hospital charts	499	15037
Reubinoff et al., 1997 [16]	1983–1993	Jerusalem, Israel	Singleton IVF pregnancies, >25 wks or ≥500 g	Hadossah Medical Centre	Charts	260	260
Schieve et al., 2002 [17]	1996–1997	United States	Liveborn, IVF	Centre for Disease Control database	Medical records	18408	20369
Schieve et al., 2007 [18]	1997–1998	Massachusetts, United States	Conceived by ART based on a positive link with ART records	US ART surveillance system, Massachusetts birth and death files	ART clinic records	1400	1400
Shevell et al., 2005 [19]	1999–2002	United States	Patients enrolled in the First And Second Trimester Evaluation of Risk (FASTER) trial (singletons)	Centres participating in the multi-centre FASTER trial	NR	554	34286
Tan et al., 1992 [20]	IVF: 1978–1987; control: 1988–1989	Great Britain, UK	Liveborns, stillborns (>28 wks)	IVF: Bourne Hall Clinic, Hallam Medical Centre; spontaneous: St. Thomas Hospital, Queen's Mary Hospital	Charts	494	978
Verlaenen et al., 1995 [21]	1988–1994	Brussels, Belgium	First time pregnancies, >20 wks, attended antenatal clinic	Authors' hospital (not specified further)	NR	743	2915
Wennerholm et al., 1997 [22]	1990–1995	Göteborg, Sweden	IVF with cryopreserved embryos; all liveborn, stillborn (>28 wks)	Sahlgrenska University Hospital and Carlandeska Hospital	Charts	160	160
Westergaard et al., 1999 [23]	1994–1996	Denmark	All ART women (public/private) were reviewed but only a sample were analyzed; spontaneous: randomly selected within strata	IVF: Danish IVF Registry; spontaneous: DMBR	IVF registry	1298	1298
Zadori et al., 2003 [24]	1995–2002	Szeged, Hungary	Following IVF	University of Szeged	NR	185	185

ART: assisted reproductive technology; IVF: in vitro fertilization; NR: not reported; g: grams; DMBR: Danish Medical Birth Registry; FMBR: Finnish Medical Birth Registry; wks: weeks.

Table 2
Summary table of outcomes in singletons conceived through *in vitro* fertilization compared to spontaneously conceived singletons.

Outcomes	Total *N* of studies	Pooled crude data		Pooled matched data		Pooled adjusted data		Pooled adjusted or matched data		
		N of studies	RR (95% CI)	*N* of studies	RR (95% CI)	*N* of studies	RR (95% CI)	*N* of studies	RR (95% CI)	I^2 value (%)
PTB <37 wks	15	2	1.89 (0.94–3.81)	13	1.87 (1.51–2.33)	2	2.04 (1.27–3.26)	15	1.84 (1.54–2.21)	75
PTB 32–36 wks	3	1	1.67 (0.74–3.79)	3	1.69 (1.27–2.25)	NA	NA	3	1.52 (1.01–2.30)	34
PTB <32–33 wks	5	2	2.13 (1.22–3.71)	4	2.91(2.02–4.19)	1	1.91 (1.55–2.36)	5	2.27 (1.73–2.97)	18
LBW <2500 g	12	2	1.72 (0.80–3.71)	8	1.64 (1.32–2.03)	4	1.58 (1.07–2.35)	12	1.60 (1.29–1.98)	87
VLBW <1500 g	8	1	1.75 (0.71–4.30)	6	3.22 (2.23–4.67)	2	1.95 (1.59–2.38)	8	2.65 (1.83–3.84)	45
ELBW <1000 g	1	NA	NA	1	3.00 (0.12–73.09)	NA	NA	NA	NA	NA
IUGR (<10% for gestation)	8	NA	NA	8	1.45 (1.04–2.00)	NA	NA	NA	NA	23
Mean birth weight (g)	12	2	−174.1 (−226.6, −121.5)	10	−97.0 (−161.2, −33.0)	NA	NA	NA	NA	78
Gestational age (weeks)	9	1	−0.70 (−0.91, −0.49)	8	−0.61 (−0.85, −0.36)	NA	NA	NA	NA	48

N: number; LBW: low birth weight; VLBW: very low birth weight; ELBW: extremely low birth weight; g: grams; wks: weeks; RR: relative risk; 95% CI: 95% confidence interval; NA: no available studies.

Table 3
Quality assessment based on evaluation of bias.

Author, year	Selection bias	Exposure bias	Outcome assessment bias	Confounding factor bias	Analytical bias	Attrition bias	Overall likelihood of bias based mainly on selection and confounding
Bergh et al., 1999 [8]	Minimal (consecutive, unselected)	Minimal (National Board of Health and Welfare and all 14 IVF clinics)	Low (Swedish Medical Birth Registry)	Minimal; *adjusted for*: maternal age[a], parity, education, smoking, #fetuses (sub-analysis adjusted for infertility)	Minimal (all available eligible patients)	Low (<10% loss to follow-up)	Minimal
Dhont et al., 1999 [9]	Low (consecutive, unselected patients in Dutch-speaking Belgium)	Minimal (questionnaire to MD)	Minimal (questionnaire to MD)	Minimal; *matched for*: maternal age (±2 years), parity, gender, duration of gestation	High (unmatched analyses)	Minimal (<10% loss)	Moderate
Howe et al., 1990 [10]	Low (consecutive, unselected from university hospital)	Minimal (from medical records)	Minimal (from medical records, call to obstetrician)	Minimal; *matched for*: maternal age (±5 years), parity, year delivered, race, DES exposure, medical problems, insurance status	High (unmatched analyses)	Minimal (<10% loss)	Low
Isaksson et al., 2002 [11]	Medium (IVF from 1 university centre, spontaneous from Finnish Medical Birth Registry)	NR	NR	Minimal; *matched for*: maternal age, parity, year delivered, mother's residence, # of fetuses; assessed, but not different: DM, hypertension, race; confounders assessed, different, and not controlled: smoking, marital status	High (unmatched analyses)	Minimal (<10% loss)	Moderate
Koivurova et al., 2002 [12]	Minimal (consecutive, unselected, included private and public)	Minimal (IVF registry)	Minimal (hospital chart by resident)	Minimal; *matched for*: maternal age, parity, social class, year delivered, area of residence, gender, plurality	Low (sample size calculation done, but no adjustment for multiple analyses)	Minimal (<10% loss)	Low

Koudstaal et al., 2000 [13]	Low (IVF, spontaneous from same 4 Dutch university centres)	NR	Low (administrative database)	Minimal; *matched for*: maternal age (±2 years), parity, smoking, year delivered (±2 years), race, height (±10 cm), weight (±10 kg), BMI, obstetric histories, medical disorders; assessed, but not different: gender, GDM	Low (no sample size calculation, but all available patients, no adjustment for multiple analyses)	Minimal (<10% loss)	Low
Perri et al., 2001 [14]	Low (from 1 centre only)	Minimal (assisted reproductive technologies database)	Minimal (labor and delivery log book, assisted reproductive technologies database)	Minimal; *matched for*: maternal age (±2 years), parity, ethnic origin, gravidity	High (unmatched analyses)	Minimal (<10% loss)	Low
Poikkeus et al., 2007 [15]	Minimal (1 hospital)	Minimal (from the chart)	Minimal (from the chart)	Minimal; *adjusted for*: maternal age, parity, socioeconomic class confounders assessed, different, and not controlled: preeclampsia, GDM	Low (sample size calculation not performed, but all available)	Minimal (<10% loss)	Low
Reubinoff et al., 1997 [16]	Low (IVF and spontaneous from same single centre)	Minimal (from the chart)	Minimal (from the chart)	Minimal; *matched for*: maternal age (±2 years), parity, year delivered, race, location of birth assessed, but not different: DM, GDM, gravidity, obstetric complications	Low (matched analyses, sample size calculation not performed, all available)	Minimal (<10%)	Low
Schieve et al., 2002 [17]	Low (consecutive, unselected, exclusion for controls not defined)	Minimal (from patient records)	Low (National Health Centre Statistics)	Low; *adjusted for*: age, parity	Low (sample size calculation not performed, all available)	Minimal (<10% loss)	Low
Schieve et al., 2007 [18]	Low (assisted reproductive technology surveillance system for Massachusetts)	Minimal (from patient record)	Low (from files)	Minimal; *adjusted for*: education; *matched for*: maternal age, parity, year and month delivered, race, birth hospital assessed, but not different: GDM confounders assessed, different, and not controlled smoking, PIH, DM	High (unmatched analyses)	Minimal (<10% loss)	Low
Shevell et al., 2005 [19]	Minimal	NR	Low	Minimal; *adjusted for*: maternal age, education, smoking, race, marital status, BMI, bleeding in current pregnancy; assessed, but not different: GDM; confounders assessed, different, and not controlled: preeclampsia	Low	Minimal (<10%)	Low
Tan et al., 1992 [20]	High	Minimal	Low	Low; *matched for*: maternal age, multiplicity; assessed, but not different: gender	Moderate	Minimal (<10%)	Moderate
Verlaenen et al., 1995 [21]	Low	NR	NR	Minimal; *matched for*: maternal age, parity, year delivered, height, weight; assessed, but not different: GDM	Low	Minimal (<10%)	Low
Wennerholm et al., 1997 [22]	Low	Minimal	Minimal	Minimal; *matched for*: maternal age (±5 years), parity, year delivered, plurality, locality; assessed, but not different: smoking, gender	Low	Minimal (<10%)	Low
Westergaard et al., 1999 [23]	Minimal	Minimal	Low	Minimal; *adjusted*[a] *for*: maternal age, parity, year delivered, multiplicity	Low	Minimal (<10%)	Low
Zadori et al., 2003 [24]	Medium	NR	NR	Minimal: *matched for*: maternal age, parity, gravidity, previous obstetric outcome; assessed, but not different: preeclampsia, GDM, BMI, congenital anomalies of uterus; confounders assessed, different, and not controlled: education	High	Minimal (<10%)	Moderate

NR: not reported; DES: diethylstilbestrol exposure; DM: diabetes mellitus; GDM: gestational diabetes mellitus; BMI: body mass index; PIH: pregnancy induced hypertension.

[a] Authors use the term stratified for regression.

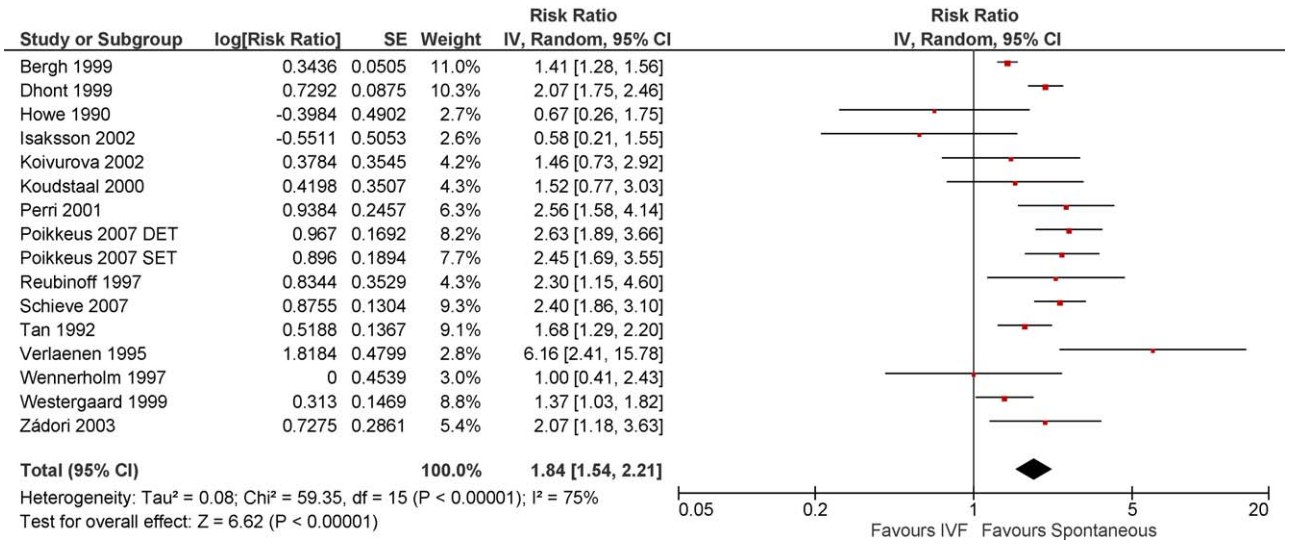

Fig. 2. Forest plot of preterm birth in singletons conceived with *in vitro* fertilization compared to spontaneously conceived singletons in studies which matched or controlled for at least maternal age.
Preterm birth is defined as birth <37 weeks gestation. Poikkeus 2007 DET refers to singletons conceived following double embryo transfer and Poikkeus 2007 SET refers to singletons conceived following single embryo transfer. Sizes of data markers indicate the weights of each study in the analysis. CI indicates confidence interval. Random indicates that the random effects model was used for statistical pooling.

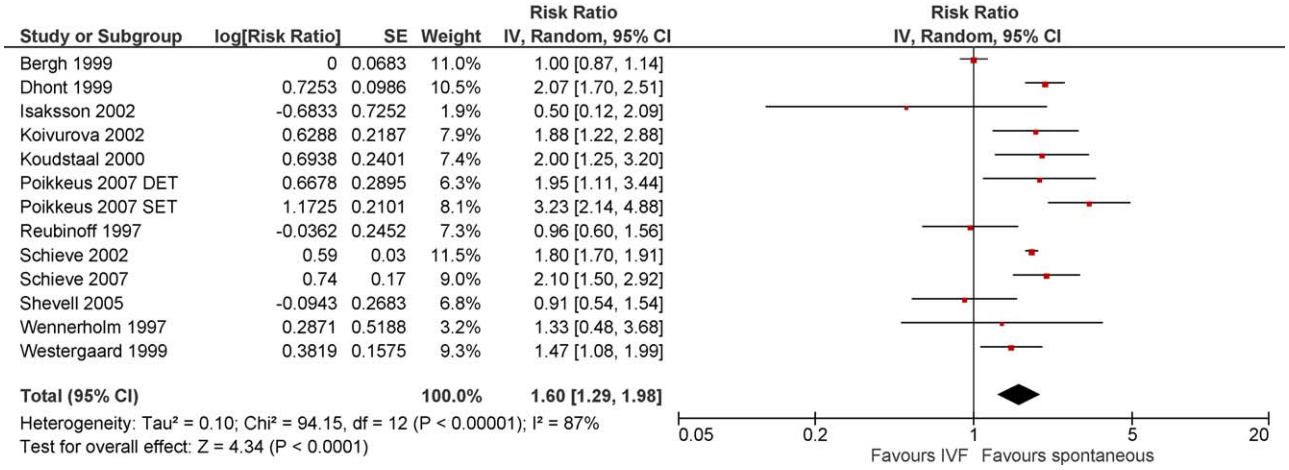

Fig. 3. Forest plot of low birth weight in singletons conceived with *in vitro* fertilization compared to spontaneously conceived singletons in studies which matched or controlled for at least maternal age.
Low birth weight is defined as birth weight <2500 g. Poikkeus 2007 DET refers to singletons conceived following double embryo transfer and Poikkeus 2007 SET refers to singletons conceived following single embryo transfer. Sizes of data markers indicate the weights of each study in the analysis. CI indicates confidence interval. Random indicates that the random effects model was used for statistical pooling.

Table 4
A priori defined sensitivity analyses of studies including only IVF (without ICSI) compared to studies including both IVF/ICSI and IVF, fresh versus frozen embryos, and low quality versus remainder of the studies.

Comparison group	Outcome	Number of studies	Number of participants	OR [95% CI]
Studies including only IVF versus studies including IVF and IVF/ICSI	PTB in IVF	8	6484	2.14 [1.58, 2.89]
	PTB IVF/ICSI and IVF	7	60207	1.87 [1.34, 2.62]
	LBW in IVF	6	6006	1.85 [1.40, 2.43]
	LBW in IVF/ICSI and IVF	6	59922	1.71 [1.26, 2.32]
Studies examining outcomes in fresh compared to studies examining outcomes in frozen embryos	PTB in fresh embryos	2	1054	2.23 [1.27, 3.93]
	PTB in frozen embryos	1	320	1.00 [0.39, 2.59]
	LBW in fresh embryos	1	440	1.93 [0.75, 4.97]
	LBW birth in frozen embryos	1	320	1.35 [0.46, 3.99]
Studies of low quality[a] versus remainder	PTB in low quality studies	1	370	2.27 [1.16, 4.45]
	PTB in remainder of studies	14	66321	1.97[1.57, 2.48]
	LBW in low quality studies	0	0	Not estimable
	LBW in remainder of studies	12	65928	1.78 [1.46, 2.12]

N/A: not applicable as a single study.
[a] Low quality studies included those with a high overall assessment of bias (based mainly on selection bias and confounding) or "high bias" or "not reported" in 3 or more bias categories.

IVF to ones born following IVF combined with ICSI. However, all of the studies reporting on IVF combined with ICSI also included some infants born after IVF alone. Hence, for the sensitivity analyses we were forced to examine studies of IVF infants compared to studies including both IVF/ICSI and IVF infants, and found no significant differences in PTB or LBW.

Unfortunately, there were too few studies which examined outcomes in fresh versus frozen embryos to draw meaningful conclusions. There were only two studies which assessed PTB in fresh embryos, and a single study of frozen embryos. A single study of each looked at LBW.

A sensitivity analysis was performed to evaluate the effect of excluding studies that were defined *a priori* as either an overall assessment of high bias or three or more "high" bias or "not reported" categories. Excluding the single study with three or more "high bias" or "not reported" (Zadori et al. [24]) did not appear to have any effect on the PTB estimate, and this study did not report on LBW.

3.3. Heterogeneity assessment

Clinical heterogeneity is shown in Table 1 in the characteristics of the included studies. Statistical heterogeneity varied from low (I^2 = 18% early PTB, I^2 = 23% IUGR) to high for both our primary outcomes (I^2 = 75% for PTB and I^2 = 87% for LBW) and is shown in Table 2.

4. Comment

In this systematic review and meta-analyses, we determined that IVF singletons have higher rates of our primary outcomes, PTB and LBW, the two most important determinants of neonatal morbidity and mortality [4]. In addition, IVF singletons have increased risks of moderate and late PTB, VLBW, IUGR, lower mean birth weights and shorter mean gestations. The risks of the most extreme preterm birth cannot be assessed due to the inclusion criteria of the original studies. Two studies included births more than 28 weeks, another study included births more than 25 weeks or >500 g, and two studies included infants weighing >500 g.

We have presented (1) crude, (2) matched, (3) adjusted and (4) matched or adjusted pooled analyses. Although pooling of matched or adjusted data is controversial given that studies address varying confounders, we and others [25] believe that unadjusted estimates are likely to be subject to confounding and where possible, examining both crude and adjusted data is ideal. Because of the importance of maternal age on all our outcomes, our inclusion criteria required that maternal age be addressed either through adjustment or matching. The degree of adjustment and matching varied widely between studies. Given that women who underwent IVF *in a particular population* sometimes did not differ significantly from those who conceived spontaneously for a given factor (for instance, preeclampsia, gestational diabetes, body mass index, congenital anomalies of the uterus in Zadori et al. [24]), adjustment or matching were not always necessary for some of the factors. Dealing with matched and adjusted data in meta-analyses is a topic in evolution. In our previous meta-analysis [3], we combined matched and crude data, whereas in this iteration, we felt it was more informative to view them separately. The three other previous meta-analyses to our knowledge on IVF outcomes (besides our own) either did not state if they used unadjusted data or adjusted data [26,27] or used adjusted OR from studies that provided them [28].

The literature searches of the other three meta-analyses besides our prior one ended in 2000 [27] and 2002 [26,28]. We were able to include several additional, large studies [15,17–19] which allowed for a larger number of participants in this meta-analysis (31,032 IVF singletons and 81,119 spontaneously conceived singletons) and hence a more accurate risk estimation than in our previous meta-analysis, for which the literature ended in 2003 [3] (8237 IVF singletons and at least 8454 spontaneously conceived singletons [number not clear in some of cohort studies]). The other meta-analyses used different methodologies, including studies with multifetal reduction, which has been shown to be associated with worse perinatal outcomes and was hence excluded in our meta-analyses. In addition, they included studies in which pregnancies were conceived with intrauterine insemination and ovarian stimulation using the pharmaceutical Clomid or gonadotropins, not IVF. They provided relatively little discussion of the methodologic issues, particularly bias assessment, which we felt was critical in meta-analysis of observational studies.

Possible explanations for the increased risk of poor perinatal outcomes in IVF gestations include: (1) residual confounding from factors such as infertility itself or (2) medical illnesses causing it, (3) the IVF procedure, (4) vanishing twins, (5) increased surveillance and/or intervention prompting PTB, or (6) combinations of these factors. We required that all studies address the key confounder, maternal age, either by controlling for it or matching. However, infertility itself, or as a result of medical complications might also explain the poor perinatal outcomes that we found. Only one study controlled for infertility, Bergh et al. [8], and still found increased risks for most outcomes. Some studies, such as Howe et al. [10] and Koudstaal et al. [13], controlled for medical problems but other studies did not. Studies in mice and sheep suggest that IVF itself may increase perinatal risks by affecting the imprinting of genes important in fetal growth and development, such as Igf2 which is normally maternally imprinted [29]. One possible explanation for the increased risk of PTB is that relaxin concentrations following gonadotropin-stimulation remain higher than expected throughout gestation [30]. It has been postulated that relaxin, which stimulates collagen breakdown, may affect cervical integrity.

Future research needs to be undertaken to try to address the preventable causes of infertility (such as pelvic inflammatory disease) and outcomes after single embryo transfer. All IVF centres should be reporting outcomes of all their IVF cycles using standard variable definitions.

In conclusion, IVF singletons are at significantly increased risk of PTB, LBW, VLBW, IUGR, shorter mean gestations and lower birth weights compared to spontaneously conceived singletons, after matching or controlling for maternal age at least. All couples considering IVF should be counseled about the increased perinatal risks so that they are truly informed when they consent for the procedures.

Acknowledgement

We thank Elizabeth Uleryk, Chief Librarian, Hospital for Sick Children, Toronto, for developing the search strategy.

Appendix A

Embase:

1 Low Birth Weight/
2 Birth Weight, Low/or Infant, Low Birth Weight/or Infant, Low Birth Weight/or LBW Infant/or LBW Neonate/or LBW Newborn/ or Low Birth weight/or Low Birth Weight Infant/or Neonatal Underweight/
3 (low birth: weight: or birth weight:, low).mp.
4 1 or 2 or 3
5 Premature Labor/
6 Labor, Premature/or Obstetric Labor, Premature/or Premature Delivery/or Premature Labor/or Preterm Delivery/or Preterm Labor/

7 Prematurity/
8 Infant, Premature/or Infant, Premature, Diseases/or Neonate, Premature/or Premature/or Premature Baby/or Premature Birth/or Premature Child/or Premature Childbirth/or Premature Infant/or Premature Neonate/or Premature Newborn/or Premature Syndrome/or Prematuritas/or Preterm Infant/
9 (preterm: or prematur:).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
10 5 or 6 or 7 or 8 or 9
11 Fertilization in Vitro/
12 Extracorporeal Fertilization/or In Vitro Fertilisation/or In Vitro Fertilization/or Testtube Baby/
13 Micromanipulation/
14 (Test tube or test-tube or testtube).mp.
15 (Extracorp: fertili: or in vitro fertili:).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name]
16 fertili: in vitro.mp.
17 or/11–16
18 4 or 10
19 17 and 18
20 Cohort analysis/
21 Case Control Study/
22 (case–control or case control or control or cohort).mp.
23 20 or 21 or 22
24 19 and 23

Medline:

1 Fertilization in Vitro/
2 fertilization in vitro/or fertilizations in vitro/or in vitro fertilization/or in vitro fertilizations/or test-tube fertilization/or fertilization, test-tube/or fertilizations, test-tube/or test tube fertilization/or test-tube fertilizations/or test-tube babies/or babies, test-tube/or baby, test-tube/or test tube babies/or test-tube baby/
3 (fertili: in vitro or in vitro fertili: or test-tube:).mp.
4 1 or 2 or 3
5 Infant, Low Birth Weight/
6 infant, low birth weight/or low-birth-weight infant/or infant, low-birth-weight/or infants, low-birth-weight/or low birth weight infant/or low-birth-weight infants/or
7 (low: birth: weight: or birth weight:, low).mp.
8 5 or 6 or 7
9 Infant, Premature/or Obstetric Labor, Premature/or Premature Birth/
10 infant, premature/or infants, premature/or premature infant/or premature infants/or prematurity/or birth, premature/or births, premature/or premature births/or preterm birth/or birth, preterm/or births, preterm/or preterm births/or obstetric labor, premature/or labor, premature obstetric/or labor, premature/or premature labor/or premature obstetric labor/or preterm labor/or labor, preterm/
11 (preterm: or prematur:).mp.
12 10 or 11 or 12
13 8 or 13
14 4 and 14
15 cohort studies/
16 case–control studies/
17 (case–control or case control or control or cohort).mp.
18 16 or 17 or 18
19 15 and 19

Appendix B

Bias	NR	Minimal	Low	Moderate	High
Selection	☐	☐ Consecutive unselected population	☐ Sample selected from large population but selection criteria not defined	☐ Sample selection ambiguous but sample may be representative	☐ Sample selection ambiguous and sample likely not representative
		☐ Sample selected from general population rather than a select group		☐ Eligibility criteria not explained	☐ Comparative groups differ in baseline characteristics
		☐ Eligibility criteria explained	☐ A select group of population (based on race, ethnicity, residence etc.) studied	☐ Rationale for case and controls not explained	
		☐ Rationale for case and control selection explained		☐ Follow up or assessment time not explained	☐ A very select population studied making it difficult to generalize findings
		☐ Follow up or assessment time explained			
Exposure	☐	☐ Direct questioning (interview) or completion of survey by mother at the time of exposure or close to the time of exposure	☐ Assessment of exposure from global dataset	☐ Recall 1–5 years after birth	☐ Recall >5 years after birth
		☐ Direct measurement of exposure (laboratory)	☐ Indirect assessment (postal survey, mailed question)	☐ Extrapolating data from population exposure sample (with some assumptions) and not direct assessment at any time	☐ Indirect method of assessment (obtaining data from others and not from mother or father)
		☐ Exposure from the chart	☐ Recall <1 year after birth		
Outcome	☐	☐ Assessment from hospital record, birth certificate or from direct question to mother about birth weight	☐ Assessment from administrative database	☐ Assessment from ‘‘close-ended’ questions (was your baby early? or premature? or small? or before due date)	☐ Assessment from non-validated sources or generic estimate from overall population
			☐ Direct question to mother regarding gestational age		
Confounding	☐	☐ Assessed for common confounders	☐ Only certain confounders assessed	☐ Not assessed for confounders	

Appendix B (*Continued*)

Bias	NR	Minimal	Low	Moderate	High
Analytical	☐	☐ Analyses appropriate for type of sample (if matched: paired *t* test, McNemar) ☐ Analytical method accounted for sampling strategy in cross-sectional study ☐ Sample size calculation performed and adequate sample studied	☐ Analyses not accounting for common statistical adjustment (e.g. multiple analyses e.g. Bonferroni) when appropriate ☐ Sample size calculation not performed, but all available eligible patients studied ☐ Sample size calculated and reasons for not meeting sample size given	☐ Sample size estimation unclear or only sub-sample of eligible patients studied	☐ Analyses inappropriate for type of sample/study
Attrition	☐	☐ None or <10% attrition & reasons for loss of follow up explained ☐ All subjects from initiation of study to final outcome assessment were accounted for	☐ <10% attrition AND reasons for loss of follow up not explained ☐ 11–20% attrition, reasons for loss of follow up explained	☐ 11–20% attrition but reasons for loss of follow up not explained ☐ >20% attrition but reasons for loss of follow up explained ☐ All subjects from initiation of study to final outcome assessment not accounted for	☐ >20% attrition, reasons for loss of follow up not explained

Appendix C. Members of Knowledge Synthesis Group on determinants of LBW/preterm births

Prakesh Shah, Assistant Professor, Department of Paediatrics, Mount Sinai Hospital and Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada.

Arne Ohlsson, Professor Emeritus, Department of Paediatrics, Mount Sinai Hospital and Departments of Paediatrics, Obstetrics and Gynaecology, and Health Policy, Management and Evaluation, University of Toronto, Canada.

Vibhuti Shah, Associate Professor, Department of Paediatrics, Mount Sinai Hospital and Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada.

Kellie E. Murphy, Associate Professor, Department of Obstetrics and Gynecology, Mount Sinai Hospital and University of Toronto, Canada.

Sarah D. McDonald, Assistant Professor, Division of Maternal-Fetal Medicine, Departments of Obstetrics & Gynecology and Diagnostic Imaging, McMaster University, Hamilton, Canada.

Eileen Hutton, Associate Professor, Department of Obstetrics and Gynecology, McMaster University, Hamilton, Canada.

Christine Newburn-Cook, Associate Professor & Associate Dean Research, Faculty of Nursing, University of Alberta, Edmonton, Canada.

Corine Frick, Director, Alberta Perinatal Health Program and Adjunct Professor, Faculty of Nursing, University of Calgary, Calgary, Canada.

Fran Scott, Associate Professor, Dalla Lana School of Public Health, University of Toronto and Toronto Public Health, Toronto, Canada.

Victoria Allen, Associate Professor, Department of Obstetrics and Gynaecology, Dalhousie University, Halifax, Canada.

Joseph Beyene, Assistant Professor, Research Institute of The Hospital for Sick Children and Dalla Lana School of Public Health, University of Toronto, Toronto, Canada.

References

[1] Stakes: Official statistics of Finland. Assisted reproduction. National Institutes for Health and Welfare 2009. Available at: URL: http://www.stakes.fi/EN/tilastot/statisticsbytopic/reproduction/IVFtreatments.htm.

[2] Centers for Disease Control and Prevention ASfRMSfART. 2006 Assisted reproductive technology success rates: National summary and fertility clinic reports. Centers for Disease Control and Prevention 2009. Available at: URL: http://www.cdc.gov/ART/ART2006/508PDF/2006ARTCover-National.pdf.

[3] McDonald SD, Murphy K, Beyene J, Ohlsson A. Perinatal outcomes of singleton pregnancies achieved by in vitro fertilization: a systematic review and meta-analysis. J Obstet Gynaecol Can 2005;27(May (5)):449–59.

[4] Branum AM, Schoendorf KC. Changing patterns of low birthweight and preterm birth in the United States, 1981–98. Paediatr Perinat Epidemiol 2002;16(January (1)):8–15.

[5] Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283(April (15)):2008–12.

[6] Devereaux PJ, Schunemann HJ, Ravindran N, et al. Comparison of mortality between private for-profit and private not-for-profit hemodialysis centers: a systematic review and meta-analysis. JAMA 2002;288(November (19)):2449–57.

[7] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(September (7414)):557–60.

[8] Bergh T, Ericson A, Hillensjo T, Nygren KG, Wennerholm UB. Deliveries and children born after in-vitro fertilisation in Sweden 1982-95: a retrospective cohort study. Lancet 1999;354(November (9190)):1579–85.

[9] Dhont M, De Sutter P, Ruyssinck G, Martens G, Bekaert A. Perinatal outcome of pregnancies after assisted reproduction: a case-control study. Am J Obstet Gynecol 1999;181(September (3)):688–95.

[10] Howe RS, Sayegh RA, Durinzi KL, Tureck RW. Perinatal outcome of singleton pregnancies conceived by in vitro fertilization: a controlled study. J Perinatol 1990;10(September (3)):261–6.

[11] Isaksson R, Gissler M, Tiitinen A. Obstetric outcome among women with unexplained infertility after IVF: a matched case-control study. Hum Reprod 2002;17(July (7)):1755–61.

[12] Koivurova S, Hartikainen AL, Gissler M, Hemminki E, Sovio U, Jarvelin MR. Neonatal outcome and congenital malformations in children born after in-vitro fertilization [comment]. Hum Reprod 2002;17(5):1391–8.

[13] Koudstaal J, Braat DD, Bruinse HW, Naaktgeboren N, Vermeiden JP, Visser GH. Obstetric outcome of singleton pregnancies after IVF: a matched control study in four Dutch university hospitals. Hum Reprod 2000;15(August (8)):1819–25.

[14] Perri T, Chen R, Yoeli R, et al. Are singleton assisted reproductive technology pregnancies at risk of prematurity? J Assist Reprod Genet 2001;18(May (5)):245–9.

[15] Poikkeus P, Gissler M, Unkila-Kallio L, Hyden-Granskog C, Tiitinen A. Obstetric and neonatal outcome after single embryo transfer. Hum Reprod 2007;22 (April (4)):1073–9.

[16] Reubinoff BE, Samueloff A, Ben-Haim M, Friedler S, Schenker JG, Lewin A. Is the obstetric outcome of in vitro fertilized singleton gestations different from natural ones? A controlled study. Fertil Steril 1997;67(June (6)):1077–83.

[17] Schieve LA, Meikle SF, Ferre C, Peterson HB, Jeng G, Wilcox LS. Low and very low birth weight in infants conceived with use of assisted reproductive technology. N Engl J Med 2002;346(March (10)):731–7.

[18] Schieve LA, Cohen B, Nannini A, et al. A population-based study of maternal and perinatal outcomes associated with assisted reproductive technology in Massachusetts. Matern Child Health J 2007;11(November (6)):517–25.

[19] Shevell T, Malone FD, Vidaver J, et al. Assisted reproductive technology and pregnancy outcome. Obstet Gynecol 2005;106(November (5 Pt 1)):1039–45.

[20] Tan SL, Doyle P, Campbell S, et al. Obstetric outcome of in vitro fertilization pregnancies compared with normally conceived pregnancies. Am J Obstet Gynecol 1992;167(September (3)):778–84.

[21] Verlaenen H, Cammu H, Derde MP, Amy JJ. Singleton pregnancy after in vitro fertilization: expectations and outcome. Obstet Gynecol 1995;86(December (6)):906–10.
[22] Wennerholm UB, Hamberger L, Nilsson L, Wennergren M, Wikland M, Bergh C. Obstetric and perinatal outcome of children conceived from cryopreserved embryos. Hum Reprod 1997;12(August (8)):1819–25.
[23] Westergaard HB, Tranberg Johansen AM, Erb K, Nyboe AA. Danish National In-Vitro Fertilization Registry 1994 and 1995: a controlled study of births, malformations and cytogenetic findings. Hum Reprod 1999;14(7):1896–902.
[24] Zadori J, Kozinszky Z, Orvos H, Katona M, Pal A, Kovacs L. Dilemma of increased obstetric risk in pregnancies following IVF-ET. J Assist Reprod Genet 2003;20(January (6)):216–21.
[25] Peters J, Mengersen K. Selective reporting of adjusted estimates in observational epidemiology studies: reasons for implications for meta-analyses. Eval Health Prof 2008;31(4):370–89.
[26] Helmerhorst FM, Perquin DA, Donker D, Keirse MJ. Perinatal outcome of singletons and twins after assisted conception: a systematic review of controlled studies. BMJ 2004;328(January (7434)):261.
[27] McGovern PG, Llorens AJ, Skurnick JH, Weiss G, Goldsmith LT. Increased risk of preterm birth in singleton pregnancies resulting from in vitro fertilization-embryo transfer or gamete intrafallopian transfer: a meta-analysis. Fertil Steril 2004;82(December (6)):1514–20.
[28] Jackson RA, Gibson KA, Wu YW, Croughan MS. Perinatal outcomes in singletons following in vitro fertilization: a meta-analysis. Obstet Gynecol 2004;103(March (3)):551–63.
[29] Li T, Vu TH, Ulaner GA, et al. IVF results in de novo DNA methylation and histone methylation at an Igf2-H19 imprinting epigenetic switch. Mol Hum Reprod 2005;11(September (9)):631–40.
[30] Mushayandebvu TI, Goldsmith LT, Von HS, Santoro N, Thurston D, Weiss G. Elevated maternal serum relaxin concentrations throughout pregnancy in singleton gestations after superovulation. Obstet Gynecol 1998;92(July (1)):17–20.