Introduction

Intrauterine growth is usually estimated using gestational age and weight at birth, which are routinely recorded on all births in most countries where assisted reproductive technologies (ART) have been introduced. It is therefore not surprising that the possible effects of ART on these two measures were reported soon after the first ART birth in 1978. For example, in 1985, the Australian In Vitro Fertilisation Collaborative Group [1] reported that 19% of 108 ART singletons in Australia were born before 37 weeks of gestation and 19% had a birth weight <2500 g, about three times the population rates at that time of preterm birth (PTB) and low birth weight (LBW). Although there were also early reports on birth defects following ART, they were based on small numbers of affected children, as birth defects are relatively rare. Hence, risk estimates were imprecise and often interpreted as showing no increase because the difference was not statistically significant. It took until the mid-2000s for the increased risk of birth defects in ART-conceived infants to be generally acknowledged [2,3].

ART has changed greatly since those early days. Techniques have altered, advanced and multiplied, pregnancy success rates have increased considerably, and more couples are using ART and for an increasing list of indications – there are now estimated to be more than five million ART children worldwide. This article surveys the systematic reviews and meta-analyses on intrauterine growth and birth defects in ART compared with non-ART singletons and discusses the possible reasons for the differences found, again using evidence from systematic reviews and meta-analyses where available. We report on recent reductions in risk of these outcomes and discuss two emerging issues: (i) the shift towards frozen–thawed embryos, and (ii) the risk of excessive intrauterine growth.

Systematic reviews and meta-analyses

2.1 Intrauterine growth

Poor intrauterine growth is a predictor of adverse perinatal (and later) outcomes and is usually estimated using infant weight and gestation at the time of birth. Whereas the advent of prenatal ultrasound scanning has allowed serial measurement of growth during pregnancy, all the population studies included in the systematic reviews have used weight and gestation at birth to assess intrauterine growth.
There have now been six systematic reviews (with meta-analyses) including comparisons of weight and gestation at birth in ART and non-ART singletons (Table 1) [4–9].

2.2 Gestation

The pooled odds ratios (ORs) from each of the reviews show an increase in preterm birth (<37 weeks of gestation), with around a two-fold increase in all but the most recent review, in which the OR was 1.5 (Table 1) [9]. All these estimates of effect have 95% confidence intervals (CIs) that exclude unity and are therefore all statistically significant.

The pooled ORs for very preterm birth (<32 weeks) were even more elevated – around two-and-a-half to three times that of non-ART singletons in the earliest reviews and again lower (1.7) but still statistically significant in the most recent review [9].

2.3 Birth weight

The findings for LBW and small for gestational age (SGA) show increased ORs (1.4–1.8 for LBW; 1.4–1.6 for SGA) and the ORs in the Pandey et al. [9] review were similar. For very LBW the pooled ORs were around 3, falling to around 2 in the recent review (Table 1) [9].

2.4 Birth defects

There have been six systematic reviews with meta-analyses of birth defects and ART in singletons (or singletons and multiples combined), compared with their non-ART counterparts (Table 2) [2,3,7–9,11]. The pooled ORs show a 30–70% increase in birth defects (ORs: 1.3–1.7), with slightly lower pooled ORs for studies examining singletons and multiples combined (ORs: 1.3–1.4). Although some individual studies (the earlier ones in particular) had methodological limitations and most of them were small, there are now many large, well-conducted studies examining the risk of birth defects in ART infants, and the pooled ORs from each meta-analysis are similar. This may be partly due to the inclusion of very large population-based studies from Scandinavia, Germany and Australia that carry considerable weight in the analyses. Data from the large Swedish ART cohort [12], for example, contributed 34% of all ART infants in our recent meta-analysis [10].

| Table 2 | Pooled estimates (95% confidence intervals) derived from the meta-analyses of six systematic reviews examining birth defects in ART compared with non-ART singletons (or singletons and multiples together). |  |
| --- | --- | --- | --- | --- | --- | --- |
| Singleton | | | | | | |
| No. of studies in meta-analysis | IVF: 8 | 15 | 7 | 7 | – | 23 |
Increased risks of anatomically grouped birth defects (e.g. cardiovascular, musculoskeletal, gastrointestinal, urogenital) in ART compared with non-ART singletons were found in pooled data from a meta-analysis [11]. There have been no meta-analyses of individual birth defects, although there was a review of the increased association between ART and rare imprinting disorders such as Beckwith-Wiedemann and Angelman syndromes, consistent with animal studies demonstrating alteration in gene imprinting of embryos cultured in vitro [10].

Two studies have reported on the association of birth defects and ART over time [12,14] and both have found a reduction in risk among more recent births. In the Swedish study, the risk ratio decreased from 1.5 (1986–2001) to 1.3 (2001–2006), whereas in our Australian study the OR decreased from 1.9 (1994–1998) to 1.3 (1999–2002). A study from Finland using a more recent birth cohort (2006–2010) reported an even lower risk of birth defects in ART singletons (OR: 1.1; 95% CI: 1.0–1.3) [15].

### 3 Possible reasons for differences in intrauterine growth and birth defects between ART and non-ART outcomes

The possible reasons for ART singletons having poorer outcomes than their non-ART counterparts relate to: differences in characteristics of the parents; the ART procedures used; pregnancy factors; methodological issues; or a combination of these.

#### 3.1 Parental characteristics

Parental demographic characteristics that may play a role include older parental age, higher socio-economic status, and lower parity in ART compared with non-ART parents. These factors are now usually accounted for in the design and/or analysis of studies.

Underlying infertility is frequently raised as a possible reason for poor outcomes. A substantial body of research, summarized in a recent systematic review and meta-analysis, indicates that subfertile couples who conceive naturally after a prolonged time to pregnancy (TTP >1 year) have greater risk of PTB (pooled OR: 1.3; 95% CI: 1.2–1.4), LBW (1.3; 1.2–1.5) and SGA (1.2; 1.0–1.3) compared with fertile couples (with TTP <1 year) [16]. Subfertility has a variety of underlying pathologies, however, so that simply grouping all ‘underlying infertility’ together may be combining specific causes of subfertility that have substantially different risks of adverse outcomes. A recent study found that singletons born to women with tubal factor infertility were 30% more likely to be born preterm and LBW compared with couples without male factor infertility using ART [17]. Risks were 70% greater for more severe outcomes (birth weight <1500 g and PTB <32 weeks gestation). Wang et al. [18] also found that PTB and LBW were more frequent among couples with female-factor infertility. Further research on more homogeneous subgroups of infertility is needed [16].

Another meta-analysis [19] showed increased ORs for PTB, comparing ART with non-ART singletons born to subfertile couples (pooled OR: 1.6; 95% CI: 1.3–1.8) and comparing ART and non-ART siblings born to the same mother (1.3; 1.1–1.5). These last two analyses suggest that although subfertility may be an important contributor to adverse perinatal outcome in ART infants, some aspect of ART treatment itself further increases that risk [19].

There is also growing evidence that subfertile couples who conceive without ART have an increased risk of birth defects compared with fertile couples [20–22], with ORs ranging from 1.2 to 1.4. However, comparing ART with non-ART siblings born to the same subfertile mother, Davies et al. [20] found an OR of 1.5 (1.1–2.1), suggesting that an excess risk of ART also remains for birth defects.

Subfertile comparison groups have been suggested as being more appropriate in ART health outcome studies to identify whether ART treatments have any effect over and above adverse effects associated with subfertility. However, infertile couples who do conceive naturally may differ in important ways from infertile couples who proceed to ART. These may include differences in underlying causes of infertility and in health behaviours that may affect both fertility and child health outcomes. A recent study found that couples who conceived naturally after a TTP >2 years had substantially worse health behaviours as measured by the prevalence of obesity, smoking before and during pregnancy, and by alcohol consumption before pregnancy than those conceiving with ART [23].
Parental conditions that predispose to infertility as well as poor intratuerine growth and/or birth defects, such as maternal diabetes and obesity, need to be considered. Some studies [15,20,24] have accounted variously for diabetes and obesity, and risks remain elevated (at the population level), but these factors may still be important when counselling individual affected couples. The importance of preconception health is gaining greater attention and could provide an important mechanism for improving health outcomes with relatively small investment [25]. An example of a recent national public education campaign to improve community knowledge about factors that impact on fertility and the health of future babies is the Your Fertility campaign, funded by the Australian Government (www.yourfertility.org.au).

Paternal birth defects, such as undescended testes and hypospadias, may be associated with infertility and, as they are also in part genetically determined, may lead to an increased risk of these birth defects in male offspring. Both these conditions have been reported as increased in ART-conceived boys [14,26]. New genetic methods are revolutionizing gene discovery for complex disorders such as infertility [27] and ultimately may provide new insights into the association between different types of infertility, ART, and pregnancy outcomes.

3.2 ART treatments

The complexity of ART treatment, important regional differences in clinical practice, and the often substantial and rapid changes in treatment over time make it a challenge to disentangle specific aspects of treatment that may adversely affect health outcomes. In the following section we summarize current evidence, from meta-analyses where available, for the effects of widely used ART techniques on intratuerine growth and birth defects.

In a meta-analysis of five studies [19], ovulation induction was associated with an increased risk of PTB when compared with natural conceptions (OR: 1.4; 95% CI: 1.2–1.7). Individual studies have generally shown slightly better outcomes for ovulation induction when compared with ART conceptions, although direct comparisons between the two groups are rarely made [28–30].

Comparison of frozen–thawed (FET) versus fresh embryo transfer can also be used to assess the potential effects of ovulation induction (if the former embryos are transferred in a natural cycle), although any direct effects of freezing and thawing are also included in such analyses. Risks of PTB, SGA, and LBW were all lower for FET compared with fresh transfer in two meta-analyses [19,31]. However, PTB (ORs: 1.2–1.4) and LBW (OR: 1.5) were still more likely in FET compared with non-ART singletons [9,19]. No significant differences were seen between FET and fresh embryo transfers for birth defects in one meta-analysis of three studies [31] (OR: 1.1; 95% CI: 0.8–1.4) and in other individual studies [14,20,32]. The data in these studies relate almost exclusively to slow-freezing techniques. We consider data examining rapid freezing (vitrification) below.

Measures of intratuerine growth and birth defects for intracytoplasmic sperm injection (ICSI) compared with IVF singleton infants have generally been considered similar, although a recent meta-analysis of five studies indicated that ICSI singletons may have a lower risk of PTB (OR: 0.8; 95% CI: 0.7–0.9) than IVF singletons [19]. One suggested reason for a better outcome for ICSI pregnancies is a healthy female partner [33]. Although no differences for birth defects in ICSI compared with IVF have been reported from meta-analyses [11,34], an increased risk of PTB in ICSI infants was recently found in an individual study [20]. Since ICSI is now being used in many countries for indications other than severe male factor infertility, it is becoming more difficult to compare these two techniques without having additional information about indications for use.

In a systematic review of five studies, the risk of birth defects in children born following ICSI with ejaculated versus non-ejaculated sperm (epididymal or testicular extracted samples) was similar although the studies were small and heterogeneous and no pooled ORs were calculated [35]. No differences in PTB or LBW were found between ejaculated and non-ejaculated sperm groups in a study published since the systematic review, but the risk of major birth defects in liveborn infants conceived with the use of non-ejaculated sperm (singletons and multiples combined) was increased (OR: 1.4; 95% CI: 0.9–2.2) and major genital malformations were more frequent in this group (2.2; 0.7–6.8) [36]. Finally, a Danish study comparing 290 singleton births following ICSI with non-ejaculated sperm to 5866 births following ICSI with ejaculated sperm found no differences in PTB, LBW, or birth defects overall. However, cardiac and genital defects (hypospadias and undescended testes) increased in prevalence across comparison groups from non-ART to IVF to ICSI with ejaculated sperm, and were highest following ICSI with non-ejaculated sperm [37].

There are still insufficient data on the health of babies born following blastocyst versus cleavage-stage transfer. Pinborg et al. [19] identified two studies for inclusion in their meta-analysis – one showed a higher rate of PTB [38] and the other a lower rate [39]. Two other studies (not included in the review) have found an increased risk for PTB (ORs: 1.3–1.4) [40,41] and a third found that blastocyst culture may increase the risk of babies born large for gestational age (LGA) [42].

Assisted hatching is a procedure in which a small hole is made in the outer shell of an embryo (the zona pellucida) through which the embryo can ‘hatch’ [43]. Some embryos seem to have a thicker or harder shell that may decrease their ability to hatch and therefore reduce the likelihood that they will implant. Various techniques have been used to breach the zona including mechanical, chemical, and laser-assisted hatching. A recent systematic review [19] failed to identify any studies on intratuerine growth for singleton babies born following assisted hatching despite the introduction of this technique to ART clinical practice 20 years ago and its more recent routine use in cycles involving preimplantation genetic diagnosis or screening. Data on birth defect prevalence are also lacking.

Studies of the effects of culture media are limited, despite experimental evidence that blastocyst quality and gene expression are critically dependent on embryo culture environment [44] and the increase in imprinting disorders in ART infants [13]. One study found a significant (250 g) difference in mean birth weight between two media [45]. Others have not found a difference [46,47], but these are all small studies and there are many possible variations in culture media.

Single embryo transfer (SET) reduces the risk of multiple pregnancies and vanishing twins, hence it is no surprise that meta-analyses of trials comparing infants born following elective SET (eSET) vs double embryo transfer (DET) found significantly better perinatal outcomes for eSET babies [48,49]. Singleton births following SET also had better perinatal outcomes than singletons born following DET in a large population-based study [50]. However, increased perinatal risks remain for ART singletons...
3.3 Pregnancy factors

A recent study has shown significant increases in anaemia in pregnancy, placenta praevia, abruption, and pre-eclampsia for all IVF compared with all other births in the population in 2006–2010 [15]. These complications of pregnancy increase the risk of poor outcomes and hence may account for some of the increase in PTB and LBW.

The increased risk of caesarean section [4,5,7,9] (in particular elective early term caesarean) in ART pregnancies has been suggested as a potential contributor to increased rates of PTB and LBW, although the increased rates of other adverse outcomes such as SGA argue against it playing a major role.

Vanishing twins, which account for up to 10% of IVF singletons born following DET, have been associated with increased risk of PTB [51,52], LBW [51,53] and SGA [53–55], with ORs ranging from 1.7 to 2.9. A 2.4-fold increase in birth defects has also been shown in comparisons of non-ART singletons from a vanishing twin pregnancy with non-ART singletons from a singleton pregnancy [56]. The contribution of vanishing twins to adverse perinatal outcome in ART infants will decline in countries such as Australia and Sweden that have made a strong shift towards eSET and may in part explain the improved perinatal and birth defect outcomes seen in these countries over time [50,57].

4.4 Methodological issues

Concerns about differential ascertainment of outcomes, particularly birth defects, in ART and non-ART cohorts have been discussed fully elsewhere [58] and include issues related to the use of standard definitions of birth defects, inclusion and exclusion criteria, age at diagnosis, inclusion of pregnancies terminated for fetal anomaly and increased scrutiny of ART pregnancies and infants. Pooled ORs based on the quality of studies included in meta-analyses have found higher OR for birth defects in ART singletons from more methodologically sound studies [10]. We assessed the possibility of increased clinical scrutiny of ART infants and found that it did not explain the excess of birth defects in ART infants [59].

In summary, although early research findings were influenced by methodological issues, more recent research on large datasets has overcome many of these problems. Recent meta-analyses suggest that underlying infertility contributes to the increased risk of adverse perinatal outcome in ART infants, and that ART techniques themselves also contribute to this risk. There are limited data assessing the influence of parental health and maternal health during pregnancy and few studies have attempted to disentangle combinations of these factors on intrauterine growth or birth defects.

4 Recent reductions in risks of preterm birth, low birth weight and birth defects

Recent data from the Nordic countries and Australia indicate improvements in perinatal outcomes and fewer birth defects in ART singletons [14,15,57,60]. The reasons for these encouraging reductions in risk are not clear, but could be due to a number of factors. The use of ART is increasing worldwide, so patient characteristics and the causes of their infertility may be changing, such that couples with an inherently lower risk of adverse outcomes may be undergoing ART. Laboratory culture conditions and ART techniques change frequently and may be leading to improved outcomes. Examples include the increased use of SET and FET. Changes in culture media and milder ovarian stimulation regimens may also be having an effect, although data are limited. Studies that stratify by years of birth rather than pooling data over a wide time-period are now required to assist in these investigations.

5 Should we be shifting to a freeze-all approach?

In meta-analyses, perinatal outcomes are generally similar or better following FET compared with fresh transfer although outcomes remain poorer when compared with non-ART singletons [9,19]. However, these meta-analyses include studies mainly using slow-freezing techniques and the transfer of largely cleavage-stage embryos, so they will not accurately reflect the more recently used techniques of vitrification (a rapid freezing technique) and transfer of embryos at the later blastocyst stage of development.

There is limited information available on the health of children born following vitrification. One small study found no significant differences in PTB or LBW between children born following the transfer of vitrified blastocysts (n = 106), fresh blastocysts (n = 207), and slow-freezing cleavage-stage embryos (n = 206) [61] but a higher rate of SGA in singletons born after transfer of fresh compared with vitrified blastocysts. Kato et al. [82] compared results for singletons born following SET of vitrified (n = 4092) versus fresh embryos (n = 2531), but the vitrified embryos were mostly blastocysts (89%) and the fresh embryos mostly cleavage stage (96%), making it difficult to separate any potential effects of freezing method from culture length. In this study, singletons born following vitrification had the same rates of PTB (6.9%), were less likely to be born SGA (OR: 0.4; 95% CI: 0.3–0.6) and LBW (0.6; 0.5–0.8), and more likely to be born LGA (1.2; 0.9–1.7) or have a birth defect (1.4; 1.0–2.1) compared with singletons born following fresh embryo transfer [62].

Possible explanations for better perinatal outcome following FET include the replacement of embryos in a more natural uterine environment; physical effects of freezing and thawing that may filter out ‘weaker’ and thus poorer quality embryos; and a potential bias towards better prognosis patients in the frozen-thawed group (i.e. those with sufficient good quality embryos available for freezing). Maheshwari et al. [31] caution that further research is needed before elective cryopreservation can be routinely recommended in preference to fresh embryo transfer because outcomes have generally been reported per ongoing pregnancy rather than per treatment cycle begun and, despite improvements in

3.3 Pregnancy factors

A recent study has shown significant increases in anaemia in pregnancy, placenta praevia, abruption, and pre-eclampsia for all IVF compared with all other births in the population in 2006–2010 [15]. These complications of pregnancy increase the risk of poor outcomes and hence may account for some of the increase in PTB and LBW.

The increased risk of caesarean section [4,5,7,9] (in particular elective early term caesarean) in ART pregnancies has been suggested as a potential contributor to increased rates of PTB and LBW, although the increased rates of other adverse outcomes such as SGA argue against it playing a major role.

Vanishing twins, which account for up to 10% of IVF singletons born following DET, have been associated with increased risk of PTB [51,52], LBW [51,53] and SGA [53–55], with ORs ranging from 1.7 to 2.9. A 2.4-fold increase in birth defects has also been shown in comparisons of non-ART singletons from a vanishing twin pregnancy with non-ART singletons from a singleton pregnancy [56]. The contribution of vanishing twins to adverse perinatal outcome in ART infants will decline in countries such as Australia and Sweden that have made a strong shift towards eSET and may in part explain the improved perinatal and birth defect outcomes seen in these countries over time [50,57].
cryopreservation, most clinics still have better pregnancy rates following fresh embryo transfer.

6 Should we be concerned about excessive fetal growth?

Macrosomic babies have a range of increased perinatal risks including stillbirth, birth asphyxia, respiratory distress, and perinatal mortality. Hence, despite the evidence of better outcomes for infants following FET, concerns have been raised about an increased prevalence of LGA in these infants. Combined data from the Nordic countries have shown increases in LGA [defined as birth weight >2 standard deviations above the reference value for gestational age and sex; OR: 1.4 (95% CI: 1.3–1.6)] and macrosomia [birth weight ≥4500 g; OR 1.6 (1.4–1.8)] when FET cycles were compared with fresh [63]. For infants born following FET compared with non-ART infants, the ORs for LGA and macrosomia were both 1.3 (1.2–1.4). This suggests that cryopreservation induces changes in the early embryo and in growth potential, perhaps related to disturbance in genomic imprinting, although we are yet to determine whether this excessive growth has negative consequences for the children such as is seen in animals, where IVF is associated with the ‘large offspring syndrome’ [64].

7 Conclusions

There are still many gaps in our knowledge of the consequences of current ART practice. Further research is required to examine mechanisms of epigenetic modification in human embryos, how cryopreservation, including the new cryopreservation techniques, may play a role, and the effect of extended culture on developing embryos. Using large datasets, it should be possible to start disentangling the inter-related effects of different types of infertility, the multiple aspects of ART treatment and parental characteristics, as well as investigating risks of individual birth defects. It may also be instructive to examine growth trajectories during pregnancy, rather than relying on gestational age and weight at birth, to improve our understanding of the effects of ART on both poor and excessive intrauterine growth. These research endeavours, including an exploration of trends in outcomes over time, should lead to a better understanding of the causes of adverse ART outcomes and help us to identify modifiable risk factors that may further reduce the disparities in outcome between ART and non-ART infants.

Practice points

- ART singletons have increased risks of LBW, PTB, small for gestational age, and birth defects compared with non-ART singletons.
- ART procedures and underlying infertility contribute to these increased risks.
- There is insufficient evidence on risks associated with different subgroups of ART exposure and different types of underlying infertility to allow for the identification of the safest treatment options for optimizing intrauterine growth and minimizing birth defect risk.

Research directions

- Using large datasets:
  - investigate effects of different types of infertility, different ART treatments, and their inter-relationships on intrauterine growth and birth defects;
  - investigate risks of individual birth defects;
  - examine growth trajectories during pregnancy to investigate effects of ART on both poor and excessive fetal growth;
  - explore trends in ART outcomes over time.
- Investigate epigenetic mechanisms in ART.

Conflict of interest statement

None declared.

Funding sources

This work was supported by an Australian National Health and Medical Research Council (NHMRC) program grant #572742 (M.H.) and NHMRC Research Fellowship #634341 (C.B.).
References


A. Maheshwari, S. Pandey, A. Shetty, M. Hamilton and S. Bhattacharya, Obstetric and perinatal outcomes in singleton pregnancies resulting from the transfer of frozen thawed versus fresh embryos generated through in vitro fertilization...


[60] A. Macaldowie, Y.A. Wang, G.M. Chambers and E.A. Sullivan, Assisted reproductive technology in Australia and New Zealand 2011, 2013, National Perinatal Epidemiology and Statistics Unit, the University of New South Wales; Sydney.
Queries and Answers

Query: Please note that as per the journal style, if there are more than 6 authors/editors, the first 6 author names are listed followed by ‘et al.’ if the author group consists of 6 authors or fewer, all author names should be listed. Therefore, in Ref(s). [11]; [15]; [19]; [20]; [24]; [26]; [32]; [36]; [41]; [45]; [49]; [61]; [62]; [63]; please list all names for up to 6 authors/editors. For more than 6 authors/editors, use ‘et al.’ after the first 6 authors/editors.

Answer: all corrected

Query: Please confirm that given names and surnames have been identified correctly.

Answer: these are correct