Title

Effectiveness and safety of intrauterine insemination vs. assisted reproductive technology: emulating a target trial using an observational database of administrative claims.

Permalink

https://escholarship.org/uc/item/2c3538bk

Journal

Fertility and Sterility, 117(5)

Authors

Chiu, Yu-Han

Yland, Jennifer

Rinaudo, Paolo

et al.

Publication Date

2022-05-01

DOI

10.1016/j.fertnstert.2022.02.003

Peer reviewed

HHS Public Access

Author manuscript

Fertil Steril. Author manuscript; available in PMC 2023 May 01.

Effectiveness and safety of intrauterine insemination versus assisted reproductive technology: emulating a target trial using an observational database of administrative claims

Yu-Han Chiu, M.D., Sc.D.[1,2,3], **Jennifer J Yland, M.Sc.**[2,4], **Paolo Rinaudo, M.D., Ph.D.**[5], **John Hsu, M.D., M.B.A.**[3,6], **Sean McGrath, A.M.**[7], **Sonia Hernández-Díaz, M.D., Dr. P.H.**[1,2], **Miguel A. Hernán, M.D., Dr.P.H.**[1,2,7]

[1]CAUSALab, Harvard T.H. Chan School of Public Health, Boston, MA

[2]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

[3]Mongan Institute, Massachusetts General Hospital, Boston, MA.

[4]Department of Epidemiology, Boston University School of Public Health, Boston, MA

[5]Center for Reproductive Health, Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California, San Francisco, CA.

[6]Department of Health Care Policy, Harvard Medical School, Boston, MA

[7]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA.

Abstract

Objectives: To compare the effectiveness and safety of one cycle of assisted reproductive technology (ART) versus three cycles of intrauterine insemination (IUI).

Design: Target trial emulation using observational data.

Setting: MarketScan Claims Database (2011–2015).

Correspondence to: Yu-Han Chiu. Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. Tel: +1 857-200-6269. yuc187@mail.harvard.edu.

Patient(s): 29,021 women 18–45 years with an infertility diagnosis and no history of IUI or ART for at least 12 months.

Interventions: (1) one ART cycle immediately, no more cycles of ART or IUI for 4 months, or (2) one IUI cycle immediately, 2 additional consecutive cycles of IUI within 4 months unless pregnancy occurs.

Main outcome measure: Live births, multiple births, congenital malformation, preterm births, small-for-gestational-age, large-for-gestational-age, newborn intensive care unit (NICU) admission, gestational diabetes, preeclampsia, and gestational hypertension.

Results: The probability of live birth was 27.3% under ART and 26.3% under IUI. The observational analog of per-protocol risk difference (95% CI) for ART compared with IUI was 1.0% (−0.1%, 2.2%) for live birth, 4.3% (3.7%, 4.9%) for multiple births, 3.4% (2.8%, 4.0%) for preterm births, 1.5% (0.9%, 2.1%) for NICU admissions, and 0.6% (0.2%, 1.0%) for gestational diabetes. The risk differences for the other outcomes were <0.5%. Results were similar under both strategies in women 40 years, but the probability of live birth was greater for ART (14.4%) than IUI strategy (7.4%) in women >40 years.

Conclusions: Compared with three cycles of IUI, one cycle of ART was estimated to have a similar probability of live birth but slightly higher risks of multiple gestations, preterm birth, and NICU admission.

Capsule:

Compared with three cycles of IUI, one cycle of ART has similar probability of live birth but higher risks of multiple gestations, preterm birth, and NICU admission.

Keywords

Introduction

The use of fertility treatments is increasing, particularly in developed countries (1–9). Intrauterine insemination (IUI) with ovarian stimulation and assisted reproductive technology (ART), including in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI), are commonly used treatments for infertility. Because each attempt of IUI is less costly and invasive than ART, IUI is often the first-line treatment in couples with unexplained or mild male infertility (10). However, guidelines are inconsistent: the UK National Institute for Health and Care Excellence (NICE) recommended ART as the first-line therapy for people with unexplained infertility, mild endometriosis, or mild male factor infertility (11), whereas the American Society of Reproductive Medicine (ASRM) recommended IUI before proceeding to ART for couples with unexplained infertility (10). These recommendations were based on randomized trials and observational studies that assessed pregnancy and live birth rates, but not maternal and neonatal complications such as preeclampsia and congenital anomalies (12–17).

In the only trial (INeS) that included neonatal outcomes, 3 cycles of ART with single embryo transfer and 6 cycles of stimulated IUI resulted in similar probabilities of live birth, multiple births, and a composite outcome of healthy infants (12). However, the sample size of the INeS trial (602 couples and 333 live births) was insufficient to study specific maternal and neonatal outcomes (12). Observational studies with larger sample sizes were not designed to assess the comparative effectiveness of multiple IUI cycles versus ART (18–24).

Therefore, we need improved observational approaches that extend the results from randomized trials to the study of maternal and neonatal complications, and that can be used for ongoing evaluation of a variety of IUI and ART dynamic strategies, given that there is no guarantee that additional randomized trials will be conducted for many contrasts of interest. Here, we used a large observational healthcare database to emulate a randomized trial—a target trial (25–27)—of 1 cycle of ART versus 3 cycles of IUI. The outcomes of interest were pregnancy, live birth, pregnancy complications, and neonatal outcomes.

Materials and Methods

The target trial is a (hypothetical) pragmatic randomized trial that would answer our question of interest (25–27). We can conceptualize observational analyses as an attempt to emulate a target trial. The approach has two steps: 1) specifying the protocol of the target trial, and 2) emulating the components of that protocol using the observational data. We now describe both steps (25–27).

Target Trial Specification

The protocol of the target trial (Table 1) would include the following components. Eligibility criteria: Women aged 18–45 years between January 2011 and August 2014 with an infertility diagnosis, no IUI or ART treatments in the previous year, no history of diabetes or hypertension, insurance coverage for IUI and ART, and who have been enrolled in any of the insurance plans included in the IBM MarketScan Commercial Claims and Encounters Database (MarketScan), for at least 12 months. A variation of this target trial might include women with at least 18 months of enrollment and no IUI/ART during that period.

Treatment strategies:

(1) one cycle of ART (conventional IVF or ICSI without preimplantation genetic testing) and no subsequent cycles of ART or IUI for 4 months, and 2) one cycle of IUI plus 2 additional consecutive cycles of IUI within 4 months unless pregnancy occurs. Cycles are considered consecutive if they occur within 60 days of each other, which would allow participants some flexibility to rest for one cycle in between. A variation of this target trial could consider cycles to be consecutive if they occur within 35 days of each other. Women would be allowed to discontinue IUI cycles if/when they develop serious medical conditions (e.g., cancer, psychosis) within 4 months of the first cycle. Under the ART strategy, insemination and embryo transfer procedures would be used as clinically indicated. Under both strategies, there would be specific instructions to avoid spontaneous pregnancy during cycles without fertility treatment, ovulation stimulation protocol would be used as

clinically indicated, a pregnancy test would be performed 2 weeks after IUI or embryo transfer, and ongoing pregnancy would be confirmed by ultrasonography. A variation of this target trial might allow spontaneous pregnancies during the first 4 months.

Treatment assignment:

Participants would be randomly assigned to either one of the strategies and would be aware of the treatment they are receiving.

Outcomes:

Pregnancy, live birth, and maternal and neonatal outcomes. Maternal outcomes include gestational diabetes, preeclampsia, and gestational hypertension. Neonatal outcomes include multiple births, major congenital malformations, preterm birth (delivery before 37 weeks of gestation), small-for-gestational-age (SGA), large-for-gestational-age (LGA), and admission to neonatal intensive care unit (NICU).

Follow-up:

For each woman, a follow-up would start at treatment assignment and end at live birth, 14 months after randomization, loss of insurance eligibility, or administrative end of follow-up on October 1, 2015 (which allows for sufficient follow up of a full-term gestation after the last eligible IUI in 2014), whichever occurs first.

Causal contrasts:

The intention-to-treat effect and the per-protocol effect, that is, the effect if all individuals had adhered to their assigned treatment strategy (28).

Statistical Analysis

The intention-to-treat analysis estimates the risks of each outcome in each group according to the assigned strategy and their differences. For maternal outcomes, we would estimate the probability of pregnancy lasting 20 weeks with and without pregnancy complications. For neonatal outcomes, we would estimate the probability of live births with and without an infant event (29). In the presence of loss to follow-up, each of these four probabilities can be estimated using the Kaplan-Meier method (a product of conditional probabilities) to take into account when individuals are censored. As selection bias may arise (e.g., older women have a higher risk of loss of follow up and adverse pregnancy outcomes), adjustment for covariates may be needed. Specifically, the probability of, say, pregnancy with a particular complication would be estimated as the product of the probability of pregnancy times the probability of the complication among those with pregnancy, where both probabilities are standardized to the distribution of the baseline covariates (see Supplemental Materials). The probability of pregnancy would be estimated via a pooled logistic model with indicators of treatment assignment, cycle, and their product terms, and baseline covariates age, calendar year, infertility diagnosis, overweight or obesity, polycystic ovarian syndrome (PCOS), and region of residence. A logistic regression model with the same covariates would be fit separately for each outcome to estimate the conditional risk of maternal outcomes among pregnancies (and of neonatal outcomes among live births). For comparison with previous

studies, we would also present the latter conditional risks. Nonparametric bootstrapping with 500 samples can be used to construct a percentile-based 95% confidence interval (CI).

The per-protocol analysis would be the same as the intention-to-treat analysis with baseline covariates, except that individuals would be censored if/when they deviate from the protocol. In the ART group, women are censored at a second ART cycle or an IUI cycle before 4 months (i.e., ART outcomes from everyone's first ART cycle are included in the analysis). In the IUI group, women are censored at ART, 60 days since the last IUI cycle, a fourth IUI cycle, or 4 months if they do not complete the 3 IUI cycles in the absence of medical conditions. In both groups, women are censored if they become spontaneously pregnant after discontinuation of fertility treatment. These women would not be censored, however, in the variation of protocol that allows spontaneous pregnancies.

Subgroup analyses would be conducted by age (18–34, 35–37, 38–40, 41–45 years), in participants with unexplained infertility (a condition for which IUI is recommended in the US (10)), by insurance plan, and in participants living in states with an insurance mandate to cover fertility treatments. To examine whether the estimates were sensitive to loss to follow-up, we would conduct a sensitivity analysis restricted to women with complete follow-up.

Target Trial Emulation

We emulated the above target trial using observational data from the MarketScan Commercial Claims and Encounters Database (MarketScan). MarketScan contains individual-level, de-identified healthcare claims data including clinical utilization, insurance enrollment/plan benefit for inpatient, outpatient, and prescription drug for individuals and their dependents (about 100 million individuals between 2011 and 2015) with employer-provided commercial insurance in the United States (30, 31). The types of insurance plans included are preferred and exclusive provider organizations (PPOs and EPOs), comprehensive (COMP) plans, point-of-service (POS) plans, health maintenance organizations (HMOs), consumer-directed health plans (CDHP) and high deductible health plans (HDHP).

Women were linked to their male partners and liveborn infants through a family ID variable. We used International Classification of Disease (ICD)-9 codes to identify female infertility diagnosis, PCOS, tobacco use disorder, alcohol abuse or dependence, obesity or overweight, and male infertility diagnosis (Supplemental Table 1). We identified IUI and ART by current procedural terminology (CPT) codes in the inpatient and outpatient files (32) (Supplemental Table 2). Cancellations of ART cycles were identified by the procedure codes "in vitro fertilization procedure canceled before aspiration", where the outcome was coded as no live birth. Because cancellation of IUI procedures was not recorded in the claims data, the study did not include canceled IUI cycles (See Supplementary Materials for details).

We identified abortion (spontaneous, elective, unspecified), stillbirth, and live births, using validated algorithms (33) based on ICD-9 code, CPT, and Healthcare Common Procedure Coding System, *and* Diagnosis Related Group codes. For women who disenrolled from

MarketScan within 280 days after initiation of treatment, pregnancy outcomes were coded as missing. For each treatment cycle, we searched for pregnancies that ended before the next treatment or a pre-specified window (<45 weeks if the outcome was live births/stillbirths and <20 weeks if the outcome was abortions), whichever occurred earliest. Pregnancy outcomes not linked to any ART or IUI cycles were considered as pregnancies from natural conception.

We estimated the gestational age at birth as {delivery date – (date of IUI or date of first ART procedure) + 14 days} or (delivery date – date of embryo transfer + 17 days) when the date of embryo transfer was available (Supplemental Figure 1) (34).

We identified maternal and neonatal outcomes using validated algorithms (35–37). Preeclampsia and gestational hypertension were defined as the presence of at least 2 inpatient ICD-9 codes after 140 gestational days and within 30 days after the delivery date (35, 36, 38) Gestational diabetes was defined as 1) at least one inpatient or outpatient ICD-9 code for gestational diabetes or CPT code for glucose tolerance test, 2) at least one diabetes code after 140 gestational days and the delivery date, and 3) an absence of diabetes codes or non-metformin antidiabetics before 140 gestational days.

SGA and LGA were defined as the presence of at least one of the ICD-9 codes for poor fetal growth or excess fetal growth, respectively, in maternal or infant claims from delivery until 30 days after delivery (35–37). NICU admission was identified using CPT codes in maternal and infant claims within 30 days of birth. A non-chromosomal structural major malformation was defined by at least two inpatient or outpatient ICD-9 codes indicating a birth defect, or a diagnostic code and a corrective surgical procedure in infant claims during the 90 days after birth or maternal claims during 30 days after delivery (35, 37).

Women in the database who met the eligibility criteria of the target trial were classified into the treatment strategy with which their data were compatible at baseline. Because pregnancy test results are not systematically recorded, we defined pregnancy as the presence of codes for either a completed or terminated pregnancy, e.g., spontaneous or therapeutic abortion within 20 weeks or delivery within 45 weeks following ART or IUI. We assumed that the treatment strategies were randomly assigned within levels of the baseline covariates age, calendar year of cycle initiation, infertility diagnosis, PCOS, overweight or obesity, and region of residence.

To estimate the observational analogs of the intention-to-treat effects and the per-protocol effects, we conducted analyses identical to the ones described for the target trial with one modification: In the target trial, we would study all outcomes in both the intention-to-treat and per-protocol analyses. In this emulation, however, we restricted the intention-to-treat analysis to pregnancy outcomes because the high nonadherence to the strategies assigned at baseline would make it hard to interpret intention-to-treat effect estimates for maternal and neonatal outcomes.

The study received ethics board approval from the Harvard T.H. Chan School of Public Health and the MassGeneral Brigham Healthcare System Institutional Review Boards (Boston, Massachusetts).

Results

Of 29,021 eligible women (Supplemental Figure 1), 18,495 initiated IUI and 10,526 initiated ART. Compared with IUI initiators, ART initiators were older, less likely to have a diagnosis of PCOS or anovulation, and more likely to have infertility of tubal origin (Supplemental Table 3). Figure 1 shows treatment trajectories for the first 3 cycles in IUI and ART initiators. The proportions of women who adhered to the protocol was 35% for the IUI group and 50% for the ART group. The probabilities of pregnancy per cycle were about 13–15% for IUI and about 33–40% for ART (Figure 1).

Intention-to-treat analysis

The probability of live birth was 28.4% among IUI initiators (21.8% from the assigned IUI treatment, 0.9% from additional IUI cycles, 3.2% from switching to ART, and 2.6% from natural conception) and 41.5% among ART initiators (27.4% from the assigned ART, 11.8% from additional ART cycles, 0.03% from switching to IUI, and 2.3% from natural conception).

Per-protocol analysis

The estimated probability of live birth was 26.3% under the IUI strategy and 27.4% under the ART strategy; risk difference 1.0% (95% CI: −0.1%, 2.2%). The risk difference was 4.3% (95% CI: 3.7%, 4.9%) for multiple births, 3.4% for preterm births (95% CI: 2.8%, 4.0%), and 1.5% (95% CI: 0.9%, 2.1%) for NICU admission. The absolute risk differences for other neonatal outcomes were less than 0.5% (Table 2).

The estimated probability of pregnancy lasting 20 weeks was 26.7% under the IUI strategy and 28.0% under the ART strategy; risk difference 1.3% (0.2% to 2.4%). The risk differences were 0.6% (95% CI: 0.2%, 1.0%) for gestational diabetes and less than 0.5% for preeclampsia and gestational hypertension (Table 2). The differences of conditional risks among live births (Table 2) and women with pregnancy lasting at least 20 weeks were in the same direction.

The probability of live birth was similar under both strategies in women aged 40 years, but greater for ART (14.4%) than IUI (7.4%) in women aged 41–45 years (Figure 2); the risk difference was 7.0% (95% CI: 4.6%, 9.1%). The risk differences for neonatal and maternal outcomes by age group were consistent with those in the main analysis, but with wider 95% confidence intervals (Supplementary Tables 4-8). Risk differences were similar when restricting the analysis to women living in States with an insurance mandate (Supplementary Table 9), women with unspecified infertility origin (Supplementary Table 10), with at least 18 months of enrollment before baseline (Supplementary Table 11), without loss to follow-up (Supplementary Table 12), and when redefining consecutive cycles as cycles occurring within 35 days from each other (Supplementary Table 13). Results were also similar in subgroup analyses by type of insurance plan (Supplementary Tables 14-17). When natural pregnancies were permitted in a variation of the protocol, the probability of live birth was 29.8% under ART versus 31.0% under IUI; risk difference −1.3% (95% CI: −2.3%, −0.3%).

The risk differences for neonatal and maternal outcomes were similar to those in the main analysis (Supplementary Table 18).

Discussion

We emulated a target trial of two fertility treatment strategies using observational data from 29,021 women in a large healthcare database. Had all participants adhered to the protocol, we estimated that three cycles of IUI would result in a similar probability of live birth than one cycle of ART (26.3% vs. 27.3%), but in lower risks of multiple births, preterm, NICU admission, and gestational diabetes. Results were similar in women with unspecified infertility origin and in women aged between 18 and 40 years. However, in women aged 41 to 45 years, a single ART cycle resulted in a higher probability of live birth than did 3 IUI cycles.

Our estimates of probability of live birth are compatible with those from randomized trials that compared 3 cycles of stimulated IUI with one cycle of ART in women ≤ 38 years. The pilot of the INeS trial in the Netherlands (n=116) found 21% for IUI compared with 22% for ART (16) and a UK randomized trial (n=207) found 24.7% for IUI vs 31.1% for ART (13). Note that in our study, live birth rates were higher for ART when not adjusting for adherence to the 1-cycle strategy because the analysis then included all pregnancies within 4 months from additional ART or other fertility treatments.

Our findings also support ASRM's recommendation of ART for women over 38 years (10). This recommendation was based on the FORT-T trial (n=154), which showed that in women aged 38–42 years, two cycles of ART had a greater probability of live birth compared with two cycles of stimulated IUI (ART vs clomiphene-IUI vs gonadotropin-IUI group: 31.4% vs 13.5% vs 15.7%) (17). We found that compared with two cycles of IUI, one cycle of ART was as effective in women between 38 and 40 years (18.2% vs. 18.1% of live births; Figure 2) and more effective in women older than 40 (14.4% vs. 6.4% of live births; Figure 2). Further, because the probability of live birth decreases with age, the shorter time to pregnancy after a single cycle can make ART more attractive than 3 IUI cycles even for younger women when the time to a live birth is an important consideration.

Our target trial differs from these trials in two aspects. First, our target trial evaluated the comparative safety of the two strategies while the sample size of most trials only allowed them to evaluate the probability of live birth (13, 15, 16). One exception is the INeS trial, which found similar probabilities of having a healthy infant for 3 cycles of ART with single embryo transfer versus 6 cycles of stimulated IUI, although the 95% confidence intervals were wide (12). Second, the proportion of ART procedures with live births that resulted in multiple births was 27.7% in our study compared with 6% to 8.3% in the trials (12, 13). Our data thus reflects the higher use of multiple embryo transfers in the US, where the proportion of multiple births (from fresh non-donor ART cycles) was about 28.8% in 2011 and 22.7% in 2015 (3, 7). Despite this higher use of multiple embryo transfers, we did not estimate ART to be more effective than 3 IUI consecutive cycles in women ≤ 40 years. The lower probability of multiple births in the ART group of the INes trial may explain that it did not find a higher risk of preterm delivery for ART (7%) compared with IUI (11%). In

contrast, the probability of multiple births in the IUI group of our study (12.8%) and of trials (6% to 13%) was comparable (13).

Our target trial emulation is not directly comparable with previous observational studies for three reasons. First, previous observational studies used comparators that are less relevant for decision making (e.g., natural conception from fertile couples, natural conception from infertile couples, non-ART groups including live births from natural or non-ART treatments) (18–24, 39). Second, most studies did not consider sustained treatment strategies involving multiple attempts, which are more relevant for decision making. That is, children were classified according to the conception method while disregarding that women who delivered a child after the first treatment attempt likely differ from women who delivered after several failures. Third, most studies only reported associations between treatment and neonatal outcomes conditional on live birth, which can introduce selection bias (40). In contrast, we presented risks of live birth with and without an offspring event (29).

Our study has several limitations. First, the outcome ascertainment was based on diagnostic and procedure codes and thus is subject to misclassification. However, most outcomes were identified through validated algorithms with good positive predictive value (33, 35–38). Second, we cannot rule out some pregnancies that linked to treatments may have resulted from a natural conception, especially those deliveries after 41 weeks from a cycle. Nonetheless, such proportion was small (Supplemental Figure 2). In addition, our estimate of pregnancy, miscarriage, and live birth percentages following ART treatments in Marketscan closely matched the national reports (Supplemental Tables 19-20). Third, the information available on the actual procedures in claims allowed us to exclude individuals with codes for preimplantation genetic testing, but not those who might have paid out of pocket (thus no claims) for preimplantation genetic testing. Also, because the proportion of single embryo transfers in the U.S. increased from 25%–38% between 2012 and 2014 (our study period) to 64% in 2017 (4, 6, 9), the risk of multiple births and associated complications (e.g., preterm, NICU admission) in the ART group is expected to be lower today than during the study period. Fourth, we could not identify women who might have decided to cancel IUI after ovarian stimulation (in contrast, we could identify women who decided to cancel ART). However, this likely small proportion of women would have been excluded from our per-protocol analyses anyway (See Supplemental Materials). Fifth, our population only included women with insurance coverage for fertility treatments, which may limit its generalizability to women who paid out of pocket if factors that affect the source of payment also affect the biological relations studied. However, the distributions of age, infertility diagnosis, and tobacco use of our study population are comparable to those of ART patients nationwide (Supplemental Tables 19 and 21) and previous studies suggested that, these and other factors (e.g., age, ethnicity, types of subfertility, duration of subfertility, and body mass index) do not predict who would benefit from immediate ART over IUI (14, 41). Therefore, our findings are probably transportable to other women seeking fertility treatments in the US. Last, as in any observational analyses, confounding is always a possibility However, we adjusted for known predictors of the outcomes and restricted the cohort to women without records of IUI or ART for at least one year. Potential residual confounding by aspects not recorded in the data (e.g., infertility duration; lifetime treatment

history; reasons for infertility), if any, may channel more severe infertility to ART group, which would bias the results towards more favorable findings for IUI.

In summary, our findings support the effectiveness and safety of three IUI cycles compared with one ART cycle as a first-line treatment for women ≤ 40 years. In women > 40 years, one ART cycle had a higher probability of live birth than three IUI cycles. Our target trial emulation replicates the findings from randomized trials, expands the inference to maternal and neonatal complications, and provides a blueprint for ongoing evaluation of fertility treatment strategies. As practice patterns evolve, and in the absence of new randomized trials, target trial emulation using real world data will be needed to inform future guidelines for fertility treatments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding:

References

1. Ferraretti AP, Nygren K, Andersen AN, de Mouzon J, Kupka M, Calhaz-Jorge C et al. Trends over 15 years in ART in Europe: an analysis of 6 million cycles. Hum Reprod Open 2017;2017:hox012. [PubMed: 31486803]

2. De Geyter C, Calhaz-Jorge C, Kupka MS, Wyns C, Mocanu E, Motrenko T et al. ART in Europe, 2015: results generated from European registries by ESHRE. Hum Reprod Open 2020;2020:hoz038. [PubMed: 32123753]

3. Sunderam S, Kissin DM, Crawford SB, Folger SG, Jamieson DJ, Barfield WD. Assisted reproductive technology surveillance--United States, 2011. MMWR Surveill Summ 2014;63:1–28.

4. Sunderam S, Kissin DM, Crawford SB, Folger SG, Jamieson DJ, Warner L et al. Assisted reproductive technology surveillance—United States, 2012. Morbidity and Mortality Weekly Report: Surveillance Summaries 2015;64:1–29.

5. Sunderam S, Kissin DM, Crawford SB, Folger SG, Jamieson DJ, Warner L et al. Assisted Reproductive Technology Surveillance - United States, 2013. MMWR Surveill Summ 2015;64:1–25.

6. Sunderam S, Kissin DM, Crawford SB, Folger SG, Jamieson DJ, Warner L et al. Assisted Reproductive Technology Surveillance - United States, 2014. MMWR Surveill Summ 2017;66:1–24.

7. Sunderam S, Kissin DM, Crawford SB, Folger SG, Boulet SL, Warner L et al. Assisted Reproductive Technology Surveillance - United States, 2015. MMWR Surveill Summ 2018;67:1–28.

8. Sunderam S, Kissin DM, Zhang Y, Folger SG, Boulet SL, Warner L et al. Assisted Reproductive Technology Surveillance - United States, 2016. MMWR Surveill Summ 2019;68:1–23.

9. Sunderam S, Kissin DM, Zhang Y, Jewett A, Boulet SL, Warner L et al. Assisted Reproductive Technology Surveillance - United States, 2017. MMWR Surveill Summ 2020;69:1–20.

10. Evidence-based treatments for couples with unexplained infertility: a guideline. Fertil Steril 2020;113:305–22. [PubMed: 32106976]

11. National Collaborating Centre for Women's and Children's Health. Fertility: Assessment and Treatment for People with Fertility Problems. National Institute for Health and Clinical Excellence (NICE) Guideline In. London: Royal College of Obstetricians & Gynaecologists, 2013.

12. Bensdorp AJ, Tjon-Kon-Fat RI, Bossuyt PM, Koks CA, Oosterhuis GJ, Hoek A et al. Prevention of multiple pregnancies in couples with unexplained or mild male subfertility: randomised controlled trial of in vitro fertilisation with single embryo transfer or in vitro fertilisation in modified natural cycle compared with intrauterine insemination with controlled ovarian hyperstimulation. Bmj 2015;350:g7771. [PubMed: 25576320]

13. Nandi A, Bhide P, Hooper R, Gudi A, Shah A, Khan K et al. Intrauterine insemination with gonadotropin stimulation or in vitro fertilization for the treatment of unexplained subfertility: a randomized controlled trial. Fertil Steril 2017;107:1329–35.e2. [PubMed: 28501361]

14. Tjon-Kon-Fat RI, Tajik P, Zafarmand MH, Bensdorp AJ, Bossuyt PMM, Oosterhuis GJE et al. IVF or IUI as first-line treatment in unexplained subfertility: the conundrum of treatment selection markers. Hum Reprod 2017;32:1028–32. [PubMed: 28333222]

15. Goverde AJ, McDonnell J, Vermeiden JP, Schats R, Rutten FF, Schoemaker J. Intrauterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis. Lancet 2000;355:13–8. [PubMed: 10615885]

16. Custers IM, König TE, Broekmans FJ, Hompes PG, Kaaijk E, Oosterhuis J et al. Couples with unexplained subfertility and unfavorable prognosis: a randomized pilot trial comparing the effectiveness of in vitro fertilization with elective single embryo transfer versus intrauterine insemination with controlled ovarian stimulation. Fertil Steril 2011;96:1107–11.e1. [PubMed: 21890134]

17. Goldman MB, Thornton KL, Ryley D, Alper MM, Fung JL, Hornstein MD et al. A randomized clinical trial to determine optimal infertility treatment in older couples: the Forty and Over Treatment Trial (FORT-T). Fertil Steril 2014;101:1574–81.e1–2. [PubMed: 24796764]

18. Stanford JB, Simonsen SE, Baksh L. Fertility treatments and adverse perinatal outcomes in a population-based sampling of births in Florida, Maryland, and Utah: a cross-sectional study. Bjog 2016;123:718–29. [PubMed: 26148540]

19. Malchau SS, Loft A, Henningsen AK, Nyboe Andersen A, Pinborg A. Perinatal outcomes in 6,338 singletons born after intrauterine insemination in Denmark, 2007 to 2012: the influence of ovarian stimulation. Fertil Steril 2014;102:1110–6.e2. [PubMed: 25064412]

20. Poon WB, Lian WB. Perinatal outcomes of intrauterine insemination/clomiphene pregnancies represent an intermediate risk group compared with in vitro fertilisation/intracytoplasmic sperm injection and naturally conceived pregnancies. J Paediatr Child Health 2013;49:733–40. [PubMed: 23758121]

21. De Sutter P, Veldeman L, Kok P, Szymczak N, Van der Elst J, Dhont M. Comparison of outcome of pregnancy after intra-uterine insemination (IUI) and IVF. Hum Reprod 2005;20:1642–6. [PubMed: 15790611]

22. Ombelet W, Martens G, Bruckers L. Pregnant after assisted reproduction: a risk pregnancy is born! 18-years perinatal outcome results from a population-based registry in Flanders, Belgium. Facts Views Vis Obgyn 2016;8:193–204. [PubMed: 28210479]

23. Bahadur G, Homburg R, Bosmans JE, Huirne JAF, Hinstridge P, Jayaprakasan K et al. Observational retrospective study of UK national success, risks and costs for 319,105 IVF/ICSI and 30,669 IUI treatment cycles. BMJ Open 2020;10:e034566.

24. Luke B, Brown MB, Wantman E, Forestieri NE, Browne ML, Fisher SC et al. The risk of birth defects with conception by ART. Hum Reprod 2021;36:116–29. [PubMed: 33251542]

25. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol 2016;183:758–64. [PubMed: 26994063]

26. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol 2016;79:70–5. [PubMed: 27237061]

27. Hernán MA. Methods of Public Health Research - Strengthening Causal Inference from Observational Data. N Engl J Med 2021;385:1345–8. [PubMed: 34596980]

28. Hernán MA, Robins JM. Per-Protocol Analyses of Pragmatic Trials. N Engl J Med 2017;377:1391–8. [PubMed: 28976864]

29. Chiu YH, Stensrud MJ, Dahabreh IJ, Rinaudo P, Diamond MP, Hsu J et al. The Effect of Prenatal Treatments on Offspring Events in the Presence of Competing Events: An Application to a Randomized Trial of Fertility Therapies. Epidemiology 2020;31:636–43. [PubMed: 32501812]

30. Butler AM, Nickel KB, Overman RA, Brookhart MA. IBM MarketScan Research Databases. In: Databases for Pharmacoepidemiological Research: Springer, 2021:243–51.

31. The Henry J. Kaiser Family Foundation (2018) Health insurance coverage of the total population http://kff.org/other/state-indicator/total-population/. Accessed 4 Nov 2021.

32. Correct coding for laboratory procedures during assisted reproductive technology cycles. Fertil Steril 2016;105:e5–8. [PubMed: 26746136]

33. MacDonald SC, Cohen JM, Panchaud A, McElrath TF, Huybrechts KF, Hernández-Díaz S. Identifying pregnancies in insurance claims data: Methods and application to retinoid teratogenic surveillance. Pharmacoepidemiol Drug Saf 2019;28:1211–21. [PubMed: 31328328]

34. American College of Obstetricians and Gynecologists. Committee opinion no 611: method for estimating due date. Obstet Gynecol 2014;124:863–6. [PubMed: 25244460]

35. Palmsten K, Huybrechts KF, Kowal MK, Mogun H, Hernández-Díaz S. Validity of maternal and infant outcomes within nationwide Medicaid data. Pharmacoepidemiol Drug Saf 2014;23:646–55. [PubMed: 24740606]

36. He M, Huybrechts KF, Dejene SZ, Straub L, Bartels D, Burns S et al. Validation of algorithms to identify adverse perinatal outcomes in the Medicaid Analytic Extract database. Pharmacoepidemiol Drug Saf 2020;29:419–26. [PubMed: 32124511]

37. Cooper WO, Hernandez-Diaz S, Gideon P, Dyer SM, Hall K, Dudley J et al. Positive predictive value of computerized records for major congenital malformations. Pharmacoepidemiol Drug Saf 2008;17:455–60. [PubMed: 18081215]

38. Practice ACoO. Practice bulletin# 33: diagnosis and management of preeclampsia and eclampsia. Obstetrics & Gynecology 2002;99:159–67. [PubMed: 16175681]

39. Wessel JA, Mol F, Danhof NA, Bensdorp AJ, Tjon-Kon Fat RI, Broekmans FJM et al. Birthweight and other perinatal outcomes of singletons conceived after assisted reproduction compared to natural conceived singletons in couples with unexplained subfertility: follow-up of two randomized clinical trials. Hum Reprod 2021;36:817–25. [PubMed: 33347597]

40. Hernán MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004;15:615–25. [PubMed: 15308962]

41. Tjon-Kon-Fat RI, Bensdorp AJ, Bossuyt PM, Koks C, Oosterhuis GJ, Hoek A et al. Is IVF-served two different ways-more cost-effective than IUI with controlled ovarian hyperstimulation? Hum Reprod 2015;30:2331–9. [PubMed: 26269539]
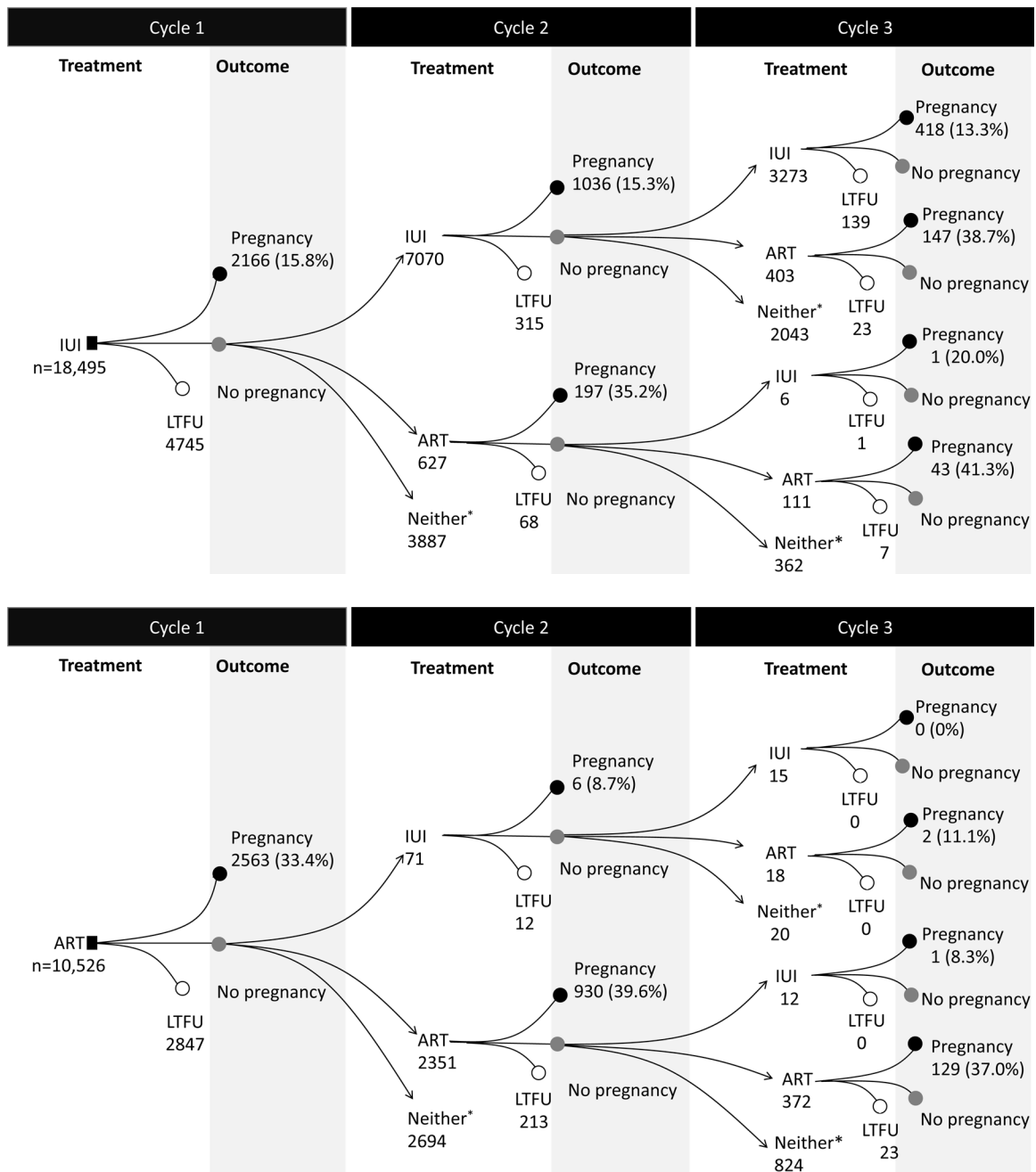
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Treatment trajectories and outcomes for the first 3 consecutive cycles in the IBM MarketScan Commercial Claims and Encounters Database, USA, January 2011-October 2015

Abbreviations: LTFU, loss to follow up. ART, assisted reproductive technology; IUI, intrauterine insemination.

*Neither were women who did not undergo IUI or ART within 60 days from the previous cycle. This group included women who underwent next IUI or ART after 60 days from the

previous cycle, had natural pregnancy, or discontinued the treatment within 4 months after initiation of treatment.
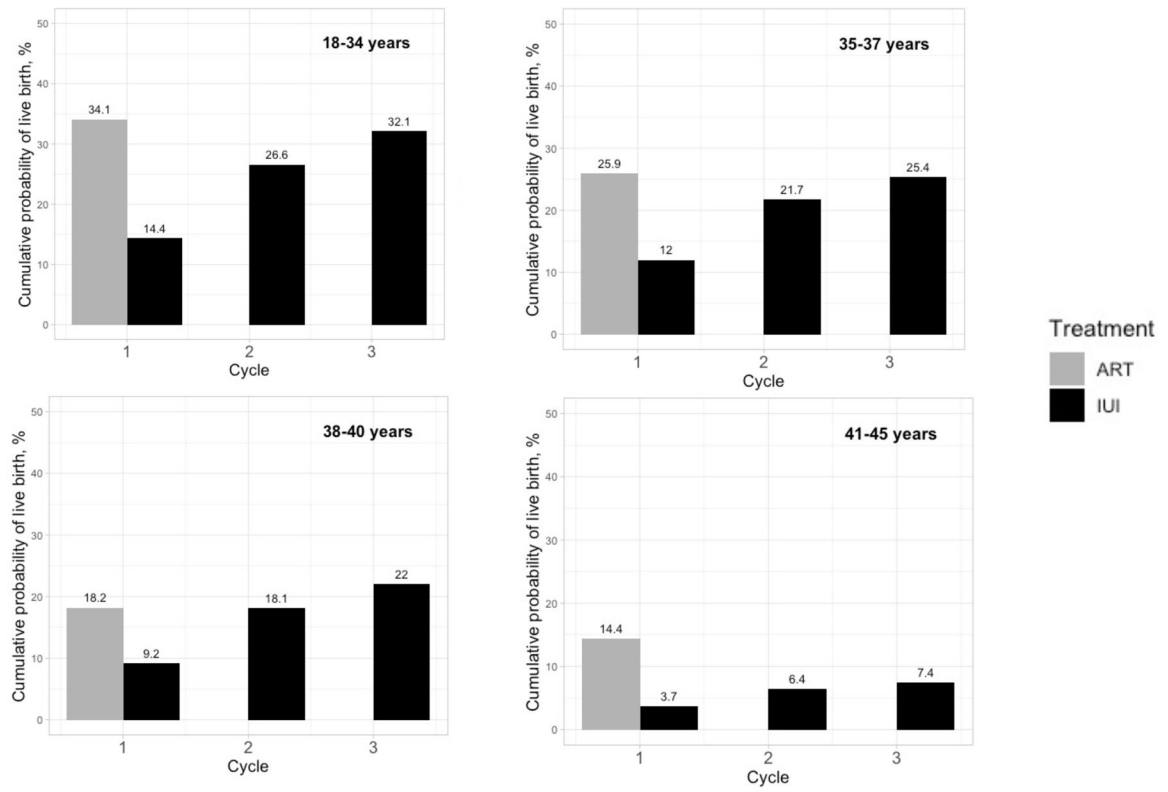
Figure 2.
Estimated cumulative probability of live birth under 1 cycle of ART versus 3 cycles of IUI, stratified by age groups, IBM MarketScan Commercial Claims and Encounters Database, USA, January 2011-October 2015

Table 1.

Specification and emulation of a target trial of 3 cycles of IUI versus one cycle of ART and pregnancy outcomes using IBM MarketScan Commercial Claims and Encounters Database, United States, January 2011-Oct 2015.

	Target trial specification	Target trial emulation
Eligibility criteria	• Women aged 18–45 years • With an infertility diagnosis • No history of IUI or ART for at least 12 months • Enrolled in MarketScan for at least 12 months • Insurance coverage for IUI and ART treatments • No pre-existing diabetes or hypertension A variation of target trial would require no history of IUI or ART for at least 18 months.	Same.
Treatment strategies	1. one cycle of ART immediately, no more cycles of ART or IUI for 4 months 2. one cycle of IUI immediately, 2 additional consecutive cycles of IUI within 4 months unless pregnancy occurs, and no ART within 4 months. Cycles are considered consecutive if they occur within 60 days of each other. Women would be excused from an additional IUI if/when they develop serious medical conditions (e.g., cancer, psychosis) within 4 months from first cycle. • Under both strategies, there would be specific instructions to avoid spontaneous pregnancy during cycles without fertility treatment. • Under both strategies, ovulation stimulation protocol performed as clinically indicated; a pregnancy test would be performed 2 weeks after an IUI or embryo transfer and ongoing pregnancy would be confirmed by ultrasonography. • ART includes in vitro fertilization and intracytoplasmic sperm injection without preimplantation genetic testing. Insemination, fresh or frozen embryo transfer, and number of embryos transferred as clinically indicated. A variation of the target trial would have no instruction to avoid spontaneous pregnancy.	Same. We defined baseline as the first procedure date of IUI or ART in the Marketscan. Because pregnancy test results are not systematically recorded in the database, we identified pregnancy by the presence of end of pregnancy outcome code (abortion, termination, stillbirth, or livebirth) in the database.
Treatment assignment	Participants are randomly assigned to either strategy and are aware of the strategy to which they have been assigned.	We assumed that women were randomly assigned within levels of baseline covariates: age, calendar year of the cycle, infertility diagnosis, polycystic ovarian syndrome, overweight or obesity, and region of residence.
Outcomes	• **Pregnancy outcomes**: pregnancy, live births. • **Maternal outcomes**: gestational diabetes (GDM), preeclampsia (PE), gestational hypertension (GHTN). • **Neonatal outcomes**: multiple births, non-chromosomal congenital malformation, preterm birth, small-for-gestational-age (SGA), large-for-gestational age (LGA), NICU admission.	Same. Pregnancy outcomes were ascertained by ICD-9 and CPT codes.
Follow-Up	Follow-up starts at assignment to a strategy and ends at live birth, 14 months after randomization, loss of insurance eligibility, or October 1, 2015, whichever comes first	Same
Causal contrasts	Intention-to-treat effect. Per protocol effect.	Observational analog of intention-to-treat and per protocol effect.
Analysis Plan	*Intention-to-treat analysis:* For neonatal outcomes, compare the probabilities of live birth overall, and live birth with and without an infant event, under each strategy. For maternal outcomes, compare the probabilities of pregnancy 20 weeks with and without a maternal event. *Per-protocol analysis:* Same as the intention-to-treat analysis, except that women are censored (and thus excluded) if/when they deviate from protocol Subgroup analyses by age (18–40 years versus 41–45 years), by insurance plan, in women living in states with insurance mandate, and in women with unexplained infertility.	Same intention-to-treat analysis with additional adjustment for baseline covariates. Due to high levels of non-adherence, intention to-treat analysis was conducted for pregnancy outcomes only. Same per-protocol analysis. Same subgroup analyses. "Unspecified origin of infertility diagnosis" in the ICD-9 code was taken as unexplained infertility.

Table 2.

Estimated probabilities of neonatal outcomes and maternal outcomes under adherence to two infertility treatment strategies (3 IUI cycles vs. 1 ART cycle) among women 18–45 years, IBM MarketScan Commercial Claims and Encounters Database, United States, January 2011-Oct 2015.

	Probability/Risk, % (95%CI)		Risk difference, % (95%CI)
All women[*]	**IUI (N=18,495)**	**ART (N=10,526)**	
Pregnancy	34.3 (33.8, 34.8)	34.6 (33.4, 35.6)	0.3 (−1.0, 1.4)
Pregnancy lasting 20 weeks	26.7 (26.2, 27.2)	28.0 (27.0, 29.1)	1.3 (0.2, 2.4)
Live birth	26.3 (25.8, 26.7)	27.4 (26.4, 28.4)	1.1 (−0.1, 2.2)
Singleton live birth	23 (22.5, 23.4)	19.7 (18.9, 20.6)	
Multiple live births	3.4 (3.2, 3.6)	7.7 (7.0, 8.2)	4.3 (3.7, 4.9)
No malformation	25.2 (24.7, 25.6)	26.0 (25.0, 27.0)	
Malformation	1.1 (1.0, 1.2)	1.4 (1.1, 1.6)	0.3 (−0.1, 0.5)
No preterm	22.8 (22.4, 23.2)	20.5 (19.6, 21.4)	
Preterm	3.5 (3.3, 3.7)	6.9 (6.3, 7.4)	3.4 (2.8, 4.0)
No small for gestational age	22.9 (22.4, 23.3)	23.4 (22.5, 24.3)	
Small for gestational age	3.4 (3.2, 3.6)	3.9 (3.5, 4.4)	0.5 (0.1, 1.1)
No large for gestational age	23.2 (22.8, 23.7)	24.7 (23.8, 25.7)	
Large for gestational age	3.1 (2.9, 3.2)	2.7 (2.4, 3.1)	−0.4 (−0.8, 0.1)
No NICU admission	22.5 (22.1, 23.0)	22.1 (21.1, 23.0)	
NICU admission	3.8 (3.5, 4.0)	5.3 (4.8, 5.8)	1.5 (0.9, 2.1)
No GDM	24.4 (24, 24.9)	25.1 (24.2, 26.1)	
Gestational diabetes (GDM)	2.3 (2.1, 2.5)	2.9 (2.5, 3.3)	0.6 (0.2, 1.0)
No preeclampsia	24.6 (24.1, 25.1)	25.6 (24.7, 26.5)	
Preeclampsia	2.1 (1.9, 2.2)	2.4 (2.1, 2.8)	0.3 (0.0, 0.8)
No GHTN	25 (24.5, 25.4)	26.1 (25.1, 27.1)	
Gestational hypertension (GHTN)	1.7 (1.6, 1.9)	1.9 (1.7, 2.3)	0.2 (−0.1, 0.6)
Livebirths only [†]	**IUI (N=2801)**	**ART (N=1992)**	
Multiples	12.6 (11.8, 13.4)	27.7 (25.8, 29.8)	15.1 (13, 17.4)
Malformation	4.1 (3.7, 4.6)	4.9 (4.0, 5.9)	0.8 (−0.3, 1.8)
Preterm	13.1 (12.2, 14.0)	24.9 (23.2, 26.9)	11.8 (9.9, 13.9)
Small for gestational age	13.2 (12.3, 14.0)	14.7 (13.2, 16.1)	1.5 (−0.2, 3.2)
Large for gestational age	12.0 (11.1, 12.8)	10.1 (8.8, 11.4)	−1.9 (−3.2, −0.3)
NICU admission	14.4 (13.4, 15.4)	19.5 (17.5, 21.3)	5.1 (3.1, 7.2)
Women with pregnancy 20 weeks [†]	**IUI (n=2840)**	**ART (n=2038)**	
GDM	8.8 (8.1, 9.6)	10.5 (9.1, 11.9)	1.7 (0.2, 3.3)
Preeclampsia	8.0 (7.3, 8.8)	8.9 (7.5, 10.1)	0.9 (−0.4, 2.3)
GHTN	6.5 (5.9, 7.1)	6.9 (6.0, 8.1)	0.5 (−0.7, 1.7)

[*] Risks are marginal (unconditional) and standardized by baseline age, calendar year of the cycle, infertility diagnosis, history of polycystic ovarian syndrome, overweight or obesity, and region of residence. Risk difference is the observational analog of a per-protocol effect estimate.