Estimation of the Contribution of Non–Assisted Reproductive Technology Ovulation Stimulation Fertility Treatments to US Singleton and Multiple Births

Laura A. Schieve*, Owen Devine, Coleen A. Boyle, Joann R. Petrini, and Lee Warner

* Correspondence to Dr. Laura A. Schieve, National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, MS E-86, 1600 Clifton Road, Atlanta, GA 30333 (e-mail: lschieve@cdc.gov).

Initially submitted April 29, 2009; accepted for publication August 10, 2009.

Infertility treatments that include ovulation stimulation, both assisted reproductive technologies (ARTs) and non-ART ovulation stimulation, are associated with increased risks of multiple birth and concomitant sequelae and adverse outcomes, even among singletons. While a US surveillance system for ART-induced births is ongoing, no population-based tracking system exists for births resulting from non-ART treatments. The authors developed a multistage model to estimate the uncertain proportion of US infants born in 2005 who were conceived by using non-ART ovulation treatments. Using published surveillance data, they estimated proportions of US multiple births conceived naturally and by ART and assumed that the remainder were conceived with non-ART treatments. They used Bayesian meta-analyses to summarize published clinical studies on the multiple-gestation risk associated with non-ART ovulation treatments, applied a fetal survival factor, and used this multiple-birth risk estimate and their own estimate of the proportion of US multiple births attributable to non-ART ovulation stimulation to estimate the total (and, through subtraction, singleton) proportion of infants conceived with such treatments. On the basis of the model, the authors estimate that 4.6% of US infants born in 2005 (95% uncertainty range: 2.8%–7.1%) resulted from non-ART ovulation treatments. Notably, this figure is 4 times greater than the ART contribution.

birth rate; infertility; meta-analysis; Monte Carlo method; ovulation induction

Abbreviations: ART, assisted reproductive technology; NASS, National ART Surveillance System; triplet/+; triplet or higher order.

The striking increase in multiple births in the United States in recent decades is attributed primarily to infertility treatments that include ovulation stimulation medications (1,2). Such treatments are generally divided into 1) assisted reproductive technology (ART) in which, after stimulation, multiple oocytes are typically retrieved from a woman and fertilized in vitro, and one or more resulting embryos are transferred to the woman’s uterus (or fallopian tubes); and 2) non-ART treatments in which ovulation medications are used in conjunction with either timed intercourse or assisted insemination. The high multiple-gestation and multiple-birth rates associated with both treatment types are concerning because of accompanying adverse sequelae, including markedly higher risks than for singletons of pregnancy complications, preterm delivery, infant death, and neurological impairments in survivors (3–6). Beyond the multiple-birth health risks, numerous studies indicate that ART-conceived singletons have higher risks than naturally conceived singletons of pregnancy complications (7,8), low birth weight and preterm delivery (7–9), birth defects (10,11), and possibly neurodevelopmental delays and disabilities (12). Studies also suggest that singletons conceived through non-ART ovulation treatments are at increased risk of such outcomes (13–15).

The Centers for Disease Control and Prevention maintains the National ART Surveillance System (NASS), a population-based registry of ART treatments (2,16). Annually, US providers submit information on all ART procedures initiated and any resultant pregnancies and births. NASS data indicate that ART treatments currently account for 1.2% of total US livebirths, 16% of US twins, and 38% of US triplet or higher-order (triplet/+; triplet or higher order) liveborn infants (2). No comparable data system exists for non-ART infertility treatments.
Estimates from the 2002 National Survey of Family Growth indicate that, among US reproductive-age women with one or more previous births, lifetime use of ovulation stimulation medications (4.6%) was 23 times higher than lifetime ART use (0.2%) (17). However, lifetime prevalence estimates provide merely a snapshot of women ever using ovulation medications, regardless of whether the treatment resulted in conception and birth. Because infertility treatments are often used in sequence, with women undergoing unsuccessful non-ART treatments before turning to ART, the disparity between lifetime non-ART and ART treatment use cannot be used to infer the magnitude of a birth differential. We used information from the published literature to estimate the contribution of non-ART infertility treatments that included ovulation stimulation to the annual US birth cohort.

MATERIALS AND METHODS

We used a multistep modeling process to estimate the proportions of US multiple and singleton livebirths in 2005 conceived by using non-ART ovulation stimulation. This process is described in the following paragraphs.

Estimation of multiple-birth infants

Data on total US twin and triplet/+ livebirths were obtained from the 2005 US natality data file and published report (18). More than 99% of US births are registered and thus are included in the natality files. Each infant represents a separate observation; data on multiple-birth infants from the same pregnancy are not linked.

Using data from both 2005 and 1971 natality files (National Center for Health Statistics, 1969–1971 Public Use Natality File 9 http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm), we calculated the proportion of multiple-birth infants in 2005 who were likely conceived naturally. Nearly all multiple births in 1971 were conceived naturally, secular trends in multiple births were relatively stable until the mid-1970s, and the relatively small fluctuations observed before 1971 have been largely attributed to variations in the maternal age distribution (19). Maternal age is the factor most strongly associated with twinning in natural conceptions. We thus applied age-specific twin and triplet/+ livebirth rates from 1971 to the maternal age distribution of births in 2005 to obtain the age-adjusted expected number of naturally conceived multiple births in 2005.

We did not adjust for 2 other factors implicated as potential risk factors for natural dizygotic twinning: African-American race and maternal obesity. However, review of available data suggests that adjusting for either factor would not substantially alter our estimates. While slight increases in twinning among African-American versus white women in the United States were noted before fertility therapies were introduced (19), race represents a heterogeneous social and biologic construct, and any potential association between race and twinning might not be consistent for all African-American women. Moreover, although we did not adjust for race in the analyses presented here, we conducted sensitivity analyses with available data and found that race adjustment would have only negligibly impacted our final estimates (data not shown).

While recent studies suggest obesity may be moderately associated with dizygotic twinning (20–22), we could not directly assess its impact because body mass index data were not included in the 1971 or 2005 natality files. However, review of the aforementioned studies suggests that the independent association between obesity and the total twinning rate among naturally conceived births is much lower than the well-established maternal age association (19). Moreover, because obesity is a hallmark feature of polycystic ovary syndrome, we cannot discount the possibility that infertility treatments are initiated earlier in obese women of all ages.

We obtained the numbers of twin and triplet/+ liveborn infants in 2005 who were conceived with ART from published NASS reports (2, 16). Because NASS is federally mandated, it captures an estimated 90%–95% of ART births in the United States. The numbers of twins and triplet/+ infants that exceeded our estimates for naturally and ART-conceived infants were assumed to indicate those most likely conceived with non-ART ovulation treatments.

Estimation of singletons

Estimation of singletons conceived with non-ART ovulation treatments was conceptualized in 3 steps. We estimated the probability of twin, triplet/+, and total multiple gestations among pregnancies conceived after non-ART ovulation treatment. We relied on published data from clinical efficacy studies to inform this estimate. Next, we applied a survival factor to the multiple-gestation estimates to obtain multiple-birth risk estimates. Finally, we used this multiple-birth risk estimate and our estimate of the proportion of US multiple-birth infants conceived by using non-ART treatments to estimate the total number of infants and, through subtraction of multiple births, the number of singleton infants conceived after non-ART ovulation treatment. There were varying degrees of uncertainty in the parameter estimates used as inputs in each step of the modeling process. We used Monte Carlo simulation methods to propagate uncertainty in the model inputs through the final estimates. The model equations are described in detail in the Appendix. The specifics of the modeling process follow.

Step 1. Estimate the probability that a fetus was part of a multiple-gestation pregnancy given that a woman became pregnant using non-ART ovulation stimulation. Although our objective was to estimate the composite impact of non-ART ovulation treatments, we considered 2 general ovulation medication types with different modes of action and multiple-birth risks. Clomiphene, approved by the US Food and Drug Administration in 1967, is an oral medication available under various brand names that inhibits the negative feedback of estrogen on the pituitary gland, which results in comparatively mild ovarian stimulation. Because clomiphene is less expensive than other (injectable) ovulation medications and does not require intensive monitoring, it is prescribed by many provider types and is often a first-line treatment. Letrozole, approved in 2008, has a mode of action similar to that of clomiphene and was considered
a clomiphene-related medication in this study. Gonadotropin medications are injectable and contain natural or synthesized follicle-stimulating hormone either alone or in combination with luteinizing hormone. The first gonadotropin medication, Pergonal (Serono, Inc., Geneva, Switzerland), was approved by the Food and Drug Administration in 1975. Currently, a wide range of preparations are used. Gonadotropin medications directly stimulate the ovaries, which results in a much stronger stimulation than with clomiphene. These medications are expensive, require monitoring, and are mainly prescribed by infertility specialists.

Estimates of the proportions of singleton, twin, and triplet/+ gestations among non-ART pregnancies conceived with each medication type were derived based on a meta-analysis of peer-reviewed literature published from May 1997 to May 2007. Our search terms and strategy are provided in Appendix Table 1.

The 342 studies initially identified were systematically reviewed to determine whether they qualified for full review and inclusion in meta-analyses. Studies that qualified were those with 1) original data collection beyond case reports and series, 2) one or more study groups that received a clearly defined non-ART ovulation treatment, 3) data presentation of at least one distinct ovulation treatment group that resulted in 50 or more pregnancies, and 4) data presented on twin- and triplet/+ gestation pregnancy risk. When 2 or more studies used the same patient population with overlapping time periods, the most recent was selected. Three studies with 5 separate groups of women treated with clomiphene or related medication (hereafter referred to as “clomiphene treatments”) (23–25) and 16 studies with 17 separate groups of women treated with gonadotropin medications or a protocol of clomiphene plus gonadotropin medications (referred to as “gonadotropin treatments”) (25–40) were completely reviewed and retained in the analysis. Treatment groups based on a combined clomiphene and gonadotropin protocol were analyzed together with gonadotropin-only treatment groups because multiple-gestation risks were similar. Data on treatment specifications, sample size, and multiple-gestation risks from each identified treatment group were abstracted by 2 study authors, and discrepancies were resolved.

Summary estimates for the probabilities of singleton, twin, and triplet/+ fetuses were derived by using Bayesian meta-analysis methods, in which we assumed a multinomial likelihood for the gestation outcomes. The summary estimates were developed separately for each medication type. Prior means for the outcome probabilities were modeled by using a common term across studies for gestation outcome plus a random effect specific to each combination of study and outcome type. The random effects were included in the model to account for the observed heterogeneity in the singleton, twin, and triplet/+ probabilities across studies (41). We did not preferentially weight studies with any particular treatment strategy because empirical data are not available on the proportionate use of various protocols in clinical practice. However, we evaluated the potential impact of study design (observational vs. randomized controlled trial) and date of publication (on or after 2000) by including terms for these effects in the prior expectation models for the

<table>
<thead>
<tr>
<th>Authors, Year (Reference No.)</th>
<th>Study Type</th>
<th>Time Period</th>
<th>No. of Women Included</th>
<th>No. of Total Treatments</th>
<th>No. of Pregnancies</th>
<th>Twin Pregnancies</th>
<th>Triplet/+ Pregnancies</th>
<th>Total Multiple Pregnancies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Park et al., 2007 (23)</td>
<td>Obs/R</td>
<td>1998–2000</td>
<td>254</td>
<td>585</td>
<td>65</td>
<td>1</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Legro et al., 2007 (24)</td>
<td>RT</td>
<td>NS</td>
<td>209</td>
<td>NS</td>
<td>50</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Legro et al., 2007 (24)</td>
<td>RT</td>
<td>NS</td>
<td>209</td>
<td>NS</td>
<td>65</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Mitwally et al., 2005 (25)</td>
<td>Obs/P</td>
<td>1999–2001</td>
<td>NS</td>
<td>432</td>
<td>61</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Mitwally et al., 2005 (25)</td>
<td>Obs/P</td>
<td>1999–2001</td>
<td>NS</td>
<td>994</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Posterior mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.26</td>
<td>0.52</td>
<td>95% Credible interval</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ART, assisted reproductive technology; NS, not stated; Obs/P, observational, prospective; Obs/R, observational, retrospective; RT, randomized trial; triplet/+ or triplet or higher order.

*From each study, data for each non-ART ovulation treatment group that included ≥50 pregnancies were abstracted separately. Thus, some studies contributed more than one treatment group to summary calculations.

**Exact numbers on twin pregnancies were not provided, but twinning rates were estimated from bar graphs. Triplet/+ pregnancies were stated as 0.
<table>
<thead>
<tr>
<th>Authors, Year (Reference No.)</th>
<th>Study Type</th>
<th>Time Period</th>
<th>No. of Women Included</th>
<th>No. of Total Treatments</th>
<th>No. of Pregnancies</th>
<th>Twin Pregnancies</th>
<th>Triplet/+ Pregnancies</th>
<th>Total Multiple Pregnancies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedaiwy et al., 2007 (26)</td>
<td>Obs/P</td>
<td>2003–2006</td>
<td>389</td>
<td>630</td>
<td>94</td>
<td>16</td>
<td>17.0</td>
<td>3</td>
</tr>
<tr>
<td>Ragni et al., 2006 (27)</td>
<td>Obs/R</td>
<td>2001–2004</td>
<td>621</td>
<td>1,259</td>
<td>116</td>
<td>11</td>
<td>9.5</td>
<td>0</td>
</tr>
<tr>
<td>Matorras et al., 2006 (28)</td>
<td>Obs/R</td>
<td>2002–2003</td>
<td>NS</td>
<td>328</td>
<td>54</td>
<td>9</td>
<td>16.7</td>
<td>2</td>
</tr>
<tr>
<td>Gorry et al., 2006 (29)</td>
<td>Obs/R</td>
<td>1990–2002</td>
<td>199</td>
<td>916</td>
<td>91</td>
<td>3</td>
<td>3.3</td>
<td>0</td>
</tr>
<tr>
<td>Tur et al., 2005 (30)</td>
<td>Obs/P</td>
<td>2001–2002</td>
<td>NS</td>
<td>1,542</td>
<td>207</td>
<td>33</td>
<td>15.9</td>
<td>5</td>
</tr>
<tr>
<td>Mitwally et al., 2005 (25)b</td>
<td>Obs/P</td>
<td>1999–2001</td>
<td>NS</td>
<td>671</td>
<td>95</td>
<td>NS</td>
<td>–12</td>
<td>0</td>
</tr>
<tr>
<td>Mitwally et al., 2005 (25)b</td>
<td>Obs/P</td>
<td>1999–2001</td>
<td>NS</td>
<td>358</td>
<td>57</td>
<td>NS</td>
<td>–13</td>
<td>0</td>
</tr>
<tr>
<td>Dickey et al., 2005 (31)</td>
<td>Obs/R</td>
<td>1987–2002</td>
<td>2,272</td>
<td>4,067</td>
<td>587</td>
<td>108</td>
<td>18.4</td>
<td>38</td>
</tr>
<tr>
<td>Ibérico et al., 2004 (32)</td>
<td>Obs/R</td>
<td>2000–2002</td>
<td>470</td>
<td>1,010</td>
<td>93</td>
<td>NS</td>
<td>8.6</td>
<td>0</td>
</tr>
<tr>
<td>Calaf Alsina et al., 2003 (33)</td>
<td>Obs/P</td>
<td>1998–1999</td>
<td>343</td>
<td>945</td>
<td>136</td>
<td>8</td>
<td>5.9</td>
<td>0</td>
</tr>
<tr>
<td>Tur et al., 2001 (34)</td>
<td>Obs/R</td>
<td>1988–1998</td>
<td>NS</td>
<td>NS</td>
<td>1,878</td>
<td>294</td>
<td>15.7</td>
<td>107</td>
</tr>
<tr>
<td>Schachter et al., 2001 (35)</td>
<td>Obs/P</td>
<td>1997–1999</td>
<td>220</td>
<td>480</td>
<td>129</td>
<td>11</td>
<td>8.5</td>
<td>3</td>
</tr>
<tr>
<td>Ragni et al., 1999 (38)</td>
<td>RT</td>
<td>NS</td>
<td>273</td>
<td>449</td>
<td>51</td>
<td>12</td>
<td>23.5</td>
<td>1</td>
</tr>
</tbody>
</table>

Posterior mean 12.32 1.99
95% Credible interval 9.05, 16.16 0.90, 3.54

Abbreviations: ART, assisted reproductive technology; NS, not stated; Obs/P, observational, prospective; Obs/R, observational; retrospective; RT, randomized trial; triplet/+, triplet or higher order.

* From each study, data for each non-ART ovulation treatment group that included ≥50 pregnancies were abstracted separately. Thus, some studies contributed more than one treatment group to summary calculations.

* Exact numbers on twin pregnancies were not provided, but twinning rates were estimated from bar graphs. Triplet/+ pregnancies were stated as 0.
multinomial probabilities. Evaluation of the posterior distributions of both of the study-design and time-of-publication effects and the estimated outcome probabilities indicated that inclusion of these terms had a negligible impact on the estimated posterior distributions for the singleton, twin, and triplet/+ probabilities. Additional detail on the methods used to derive the posterior estimates is provided in the Appendix.

The probability that a fetus was part of a multiple-gestation pregnancy given use of either drug was estimated as

Table 3. Estimation of Multiple and Singleton Births in 2005 Conceived With a Non-ART Ovulation Treatment

<table>
<thead>
<tr>
<th>Data Used to Estimate Parameters</th>
<th>Estimate</th>
<th>Estimated Uncertainty Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multiple birth estimation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of US multiple births (twins and triplets/+ in 2005)</td>
<td>2005 US natality data</td>
<td>139,816</td>
</tr>
<tr>
<td>No. (proportion) of US multiple births in 2005 conceived naturally</td>
<td>Age-specific multiple birth rates from 1971 US natality data multiplied by age-stratified no. of births from 2005 US natality data</td>
<td>83,748.54 (59.90)</td>
</tr>
<tr>
<td>No. (proportion) of US multiple births in 2005 conceived with ART</td>
<td>National ART Surveillance System data on no. of ART multiple births in US divided by the total no. of US multiple births in 2005</td>
<td>24,165 (17.28)</td>
</tr>
<tr>
<td><strong>Model output results</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of US multiple births in 2005 conceived with a non-ART ovulation treatment</td>
<td></td>
<td>31,902.46</td>
</tr>
<tr>
<td>Proportion of US multiple births in 2005 conceived with a non-ART ovulation treatment</td>
<td></td>
<td>22.82</td>
</tr>
<tr>
<td><strong>Total birth and singleton estimation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of twin gestations among pregnancies conceived with non-ART clomiphene treatment</td>
<td>Systematic review of clinical efficacy studies assessing clomiphene treatments (refer to Table 1)</td>
<td>8.26</td>
</tr>
<tr>
<td>Proportion of triplet/+ gestations among pregnancies conceived with non-ART clomiphene treatment</td>
<td>Systematic review of clinical efficacy studies assessing clomiphene treatments (refer to Table 1)</td>
<td>0.52</td>
</tr>
<tr>
<td>Proportion of total multiple gestations among fetuses in pregnancies conceived with non-ART clomiphene treatment</td>
<td>Convert twin and triplet gestations from per-pregnancy proportions to per-fetus proportions: 2(twin) + 3(triplet/+)/singleton + 2(twin) + 3(triplet/+))</td>
<td>16.22</td>
</tr>
<tr>
<td>Proportion of twin gestations among pregnancies conceived with non-ART gonadotropin treatment</td>
<td>Systematic review of clinical efficacy studies assessing gonadotropin treatments (refer to Table 2)</td>
<td>12.32</td>
</tr>
<tr>
<td>Proportion of triplet/+ gestations among pregnancies conceived with non-ART gonadotropin treatment</td>
<td>Systematic review of clinical efficacy studies assessing gonadotropin treatments (refer to Table 2)</td>
<td>1.99</td>
</tr>
<tr>
<td>Proportion of total multiple gestations among fetuses in pregnancies conceived with non-ART gonadotropin treatment</td>
<td>Convert twin and triplet gestations from per-pregnancy proportions to per-fetus proportions: 2(twin) + 3(triplet/+)/singleton + 2(twin) + 3(triplet/+))</td>
<td>26.25</td>
</tr>
</tbody>
</table>

Table continues
a weighted combination of the medication-specific singleton, twin, and triplet/þ probability estimates, with the proportion of women taking clomiphene as the clomiphene weight and 1 minus this proportion as the gonadotropin weight. Although these proportions are virtually unknown on a population ba-

Table 3. Continued

<table>
<thead>
<tr>
<th>Data Used to Estimate Parameters</th>
<th>Estimate</th>
<th>Estimated Uncertainty Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportions of clomiphene vs. gonadotropin treatment among pregnancies conceived with non-ART ovulation treatments</td>
<td>National Birth Defects Prevention Study control group births—self-reported maternal medication use (extrapolation of this estimate was subject to various assumptions and data uncertainty; thus, a large uncertainty range was incorporated in the model)</td>
<td>71 vs. 29</td>
</tr>
<tr>
<td>Proportion of total multiple gestations among fetuses in pregnancies conceived with non-ART ovulation treatment (clomiphene and gonadotropin treatments combined)</td>
<td>Proportions of clomiphene vs. gonadotropin above used to weight multiple-gestation estimates</td>
<td>19.13</td>
</tr>
<tr>
<td>Fetal loss correction factor—proportion of multiple infant births expected from a population of multiple gestations</td>
<td>National ART Surveillance System data on fetal loss among ART singleton and multiple-gestation pregnancies (the estimate was adjusted to account for the likelihood that maternal age was slightly lower in non-ART than ART births)</td>
<td>0.93</td>
</tr>
<tr>
<td>Proportion of total multiple births among infants conceived with non-ART ovulation treatment (clomiphene and gonadotropin treatments combined)</td>
<td>Apply fetal loss correction factor</td>
<td>18.04</td>
</tr>
<tr>
<td>No. (proportion) of US multiple births in 2005 conceived with a non-ART ovulation treatment</td>
<td>Refer to the multiple-birth estimation model output results above</td>
<td>31,902.46 (22.82)</td>
</tr>
<tr>
<td>Total no. of US singleton births in 2005</td>
<td>2005 US natality data</td>
<td>3,998,553</td>
</tr>
</tbody>
</table>

Model output results

| No. of US singleton births in 2005 conceived with non-ART ovulation treatments | 159,380 | 82,876–260,807c |
| Proportion of US singleton births in 2005 conceived with non-ART ovulation treatments | 3.98 | 2.07–6.52c |
| No. of total US births in 2005 conceived with non-ART ovulation treatments | 191,286 | 114,782–292,713c |
| Proportion of total US births in 2005 conceived with non-ART ovulation treatments | 4.62 | 2.77–7.08c |

Abbreviations: ART, assisted reproductive technology; triplet/þ, triplet or higher order.

a Details of the methodology to estimate the uncertainty intervals for each parameter are described in the Materials and Methods section of the text.

b 95% credible interval.
c 5th and 95th percentiles of Monte Carlo–based uncertainty distribution.
d Assumed range of uncertainty based on available information.

Am J Epidemiol  2009;170:1396–1407
clomiphene medications were reported by 71%. (42). Because of methodological limitations posed by a small sample size, ambiguities in infertility questions, noncoverage of many US areas, and an extended study time period (1997–2004), we assumed a large degree of uncertainty in this estimate. We also assumed that 1) 71% was the midpoint of uniform distribution, and 2) because clomiphene is often a first-line treatment prescribed by many different providers, the proportion of clomiphene pregnancies was unlikely lower than the proportion of gonadotropin pregnancies. We thus estimated that the true clomiphene proportion likely was 50%–90%.

We used Monte Carlo simulation to incorporate the uncertainty associated with the medication-specific gestation outcome estimates and the estimated proportion using clomiphene in the step 1 results. Uncertainty propagation was conducted by running 20,000 simulations of the step 1 models with repeated sampling of both the posterior distributions of the singleton, twin, and triplet/+ probabilities from the meta-analysis and the uncertain value for the proportion of women prescribed clomiphene from the assumed uncertainty distribution for this parameter. When we used this approach, the step 1 result was 20,000 estimates for the uncertain probability that a fetus was part of a multiple-gestation pregnancy given use of either clomiphene or gonadotropin. We refer to these distributions as uncertainty distributions, as opposed to Bayesian credible intervals, to reflect the fact that they are a combination of both posterior estimates of the singleton, twin, and triplet/+ probabilities and what is more akin to prior uncertainty regarding the prevalence of clomiphene use.

Step 2. Estimate the probability that an infant was one of a multiple livebirth given that a woman became pregnant using non-ART ovulation stimulation. Because data were not available on the proportion of multiple fetuses surviving to birth for women taking non-ART ovulation stimulation medications, we based our pregnancy-to-birth survival estimates on ART pregnancies included in NASS (2, 16). In NASS, the proportion of infants delivered as multiple births from a given cohort of pregnancies is consistently lower than the proportion of multiple gestations because of complete pregnancy loss and induced and spontaneous reduction of multiple-gestation pregnancies to singleton pregnancies. Among liveborn infants conceived with ART in 2001–2005, the proportion delivered as multiple births was 92%–93% of the proportion of initial multiple-gestation fetuses.

To translate this ART-based estimate to non-ART treatment survival, we assumed that the survival rates of singleton and multiple-gestation fetuses were not substantially different by mode of conception, with one possible exception. Because women who conceive with non-ART treatments might be younger than those who conceive with ART, the fetal loss rate might be slightly lower for pregnancies conceived with non-ART treatments. Hence, the differences between the multiple births and multiple gestations might not be as great as that observed for ART. To reflect this possibility, we modeled uncertainty associated with the non-ART treatment survival factor using a triangular distribution ranging from 0.90 (allowing for only a small possibility of lower survival among pregnancies conceived after non-ART treatment) to 1.00, with the most likely value set to 0.93. The step 2 result was derived by multiplying 20,000 randomly sampled values from the estimated step 1 value set to 0.93. The step 2 result was derived by multiplying 20,000 randomly sampled values from the estimated step 1 uncertainty distribution by a random sample from the assumed uncertainty distribution for the survival factor.

Step 3. Estimate the percentage of all US infants and US singletons resulting from non-ART ovulation stimulation. In the final step, we used the results generated in step 2 and our previous estimates of US multiple births in 2005 likely conceived by using a non-ART treatment. Using the equations in the Appendix, we combined these data to produce 10,000 possible values for the uncertain proportion of all infants and all singleton infants born in the United States in 2005 from pregnancies that resulted from non-ART ovulation stimulation.

RESULTS

Results of our systematic review of non-ART ovulation treatments are presented in Tables 1 and 2. Expanded versions of these tables that include patient selection criteria, ovulation stimulation dose and protocol, and cancellation criteria are posted on the Journal’s website (http://aje.oupjournals.org/) as Web Tables 1 and 2. Among the 5 study groups assessing clomiphene treatments (Table 1), the empirical twin and triplet/+ gestation rates ranged from 3.1% to 22% and from 0.0% to 2.0%, respectively. Given the small number of studies and the high variability in reported twin rates, the estimated Bayesian credible interval for twin probability was fairly wide (2.01%, 20.00%). The posterior credible interval for the clomiphene triplet/+ gestation probability ranged from 0.01% to 2.04%.

Among the 17 study groups assessing gonadotropin treatments (Table 2), the empirical twin gestation rate ranged from 3.3% to 23.5%; however, for most (n = 12) study groups, the twinning rate was 10% or higher. Thus, the Bayesian credible interval reflected less uncertainty in twin rates associated with gonadotropin treatments (9.05%, 16.16%) than for clomiphene treatments. The empirical triplet/+ gestation rate for gonadotropin treatment groups ranged from 0.0% to 8.8%, and the corresponding Bayesian credible interval was 0.90%, 3.54%.

Data estimates and uncertainty ranges that informed our multiple-birth and singleton estimation models are summarized in Table 3. We estimate that in 2005, 31,902 (22.8%) of the infants born as part of multiple-birth deliveries in the United States were conceived by using non-ART ovulation treatments. Because this estimate is derived by using data from population-based data sets with very complete reporting, we estimate negligible uncertainty. We estimate that 159,380 (3.98%) of the 2005 US singleton births were conceived by using non-ART ovulation treatments. Numerous data sources informed this estimate, some with a great deal of uncertainty. We estimate that the true number of singletons is likely 82,876–260,807 (2.07%–6.52%). Finally, we estimate that 114,782–292,713 (2.77%–7.08%) of the total US infants in 2005 were conceived with non-ART ovulation treatments.
DISCUSSION

We estimate that currently 3%–7% of infants born in the United States are conceived with a non-ART ovulation treatment, equating to more than 190,000 livebirths annually. Notably, this figure is 2–6 times higher than the contribution of ART to US births. Moreover, the level of exposure to either ART or non-ART ovulation treatments can be estimated as approximately 6%. In the United States, ongoing, population-based surveillance of birth outcomes is currently in place for ART but not non-ART treatments. Consequently, the total impact of ovulation stimulation treatments is substantially underestimated.

Although both ART and non-ART treatments have benefited many couples facing infertility, potential risks should also be monitored. Most well defined is the risk of multiple-gestation pregnancy, which is strongly associated with both types of ovulation treatment and is additionally associated with numerous adverse outcomes (3–6). Studies suggest that even twin pregnancies with early loss of one fetus and subsequent singleton birth (“vanishing twin” pregnancies) are at higher risk than singleton births that originated as single-gestation pregnancies of preterm delivery, intrauterine growth restriction (43–45), and possibly neurodevelopmental outcomes (46, 47). Here, we demonstrate that in the United States, non-ART ovulation induction is associated with a higher multiple-birth-attributable risk than ART; thus, non-ART treatments are very likely associated with an even higher proportion of multiple-gestation pregnancies overall.

Our estimate of US singleton births conceived with non-ART ovulation treatments is more than 6 times higher than published estimates for ART singleton births (2). Beyond health issues related to survival of a vanishing twin pregnancy, hormonal stimulation has been associated with impaired oocyte quality and increased DNA methylation errors in animals (48). Such effects could potentially affect the long-term health and development of singletons as well as multiple births.

Although, in the current study, we could not identify the specific infants conceived with non-ART treatments and thus could not study health outcomes, population estimation of exposure level is nonetheless an important first step. Greater awareness of the increase in higher-order multiple births during the 1980s and 1990s led in part to the updated clinical recommendations of the American Society for Reproductive Medicine for fewer embryos transferred during ART (49, 50). To date, no accurate population-based tracking system exists for births resulting from non-ART treatments.

The data we used to inform our model should be scrutinized. The greatest degree of uncertainty was in the proportion of twin gestations associated with clomiphene treatment. Despite an extensive systematic review, we found few published studies with sufficient data on this parameter; thus, the actual and estimated ranges were large. In particular, one study reported an unusually high (22%) twinning rate associated with clomiphene pregnancies (25). We chose not to exclude this study because it met our a priori selection criteria and might adequately reflect the diversity in the unknown “true” rate of twinning in various clomiphene-treated population subgroups. However, to the extent that these data may be erroneous, the impact is toward a more conservative estimate. Sensitivity analyses in which the model for the percentage of all US births resulting from non-ART ovulation stimulation was rerun by excluding the findings for this treatment group resulted in a mean estimate of 5.30% (95% uncertainty interval: 3.01, 8.21) births as opposed to 4.62% (95% uncertainty interval: 2.77%, 7.08%) when the treatment group was included.

Non-ART ovulation treatments substantively contribute to the US birth cohort. Continual population-based tracking would be more complex for non-ART than ART treatments given the much wider range of providers who dispense these treatments. Nonetheless, it is important to consider feasible methods of conducting sentinel site surveillance to ascertain more detail on use and efficacy of non-ART ovulation medications and any associated health risks for the many women treated and the many children annually conceived with these treatments.

ACKNOWLEDGMENTS

Author affiliations: National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, Georgia (Laura A. Schieve, Owen Devine, Coleen A. Boyle); Danbury Hospital, Danbury, Connecticut (Joann R. Petrini); March of Dimes Foundation, White Plains, New York (Joann R. Petrini); Department of Obstetrics and Gynecology, and Women’s Health, Albert Einstein College of Medicine, Yeshiva University, New York, New York (Joann R. Petrini); and National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, Georgia (Lee Warner).

The authors acknowledge the contributions of Dr. Maurizio Macaluso of the Division of Reproductive Health, Centers for Disease Control and Prevention, for his advice during the development phase of this project; and Tomoko Yamada-Kushiri of the March of Dimes Perinatal Data Center for her assistance in reviewing references.

The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the Centers for Disease and Control and Prevention. Conflict of interest: none declared.

REFERENCES


APPENDIX

Estimation of the Percentage of Total and Singleton Livebirths in the United States in 2005 Resulting From Non-ART Ovulation Treatments

**Step 1.** Estimate the probability that a fetus was part of a multiple-gestation pregnancy given that a woman became pregnant using non-ART ovulation stimulation. Because US natality data on livebirths are organized and presented with each infant considered a separate observation, and data on infants from multiple births are not linked, in the estimation of multiple- versus single-gestation pregnancies associated with non-ART ovulation treatments, we similarly considered each fetus within a given pregnancy as a separate observation.

Let $S$ represent the event during which a woman conceives a single fetus, $Tw$ the event during which she conceives twins, and $Tr$ the event during which she conceives triplets or more. Let $C$ represent a pregnancy conceived with a clomiphene medication and $G$ represent a pregnancy conceived with a gonadotropin medication ($G$ includes pregnancies conceived while taking both $C$ and $G$). In the meta-analysis, we estimated the posterior distributions for $P[S/C]$, $P[Tw/C]$, $P[Tr/C]$, $P[S/G]$, $P[Tr/G]$, and $P[Tw/G]$ on the basis of a summary of available literature. These estimates were obtained by using a 2-stage hierarchical model that was fit separately within each medication type. At the likelihood stage, the $S$, $Tw$, and $Tr$ outcomes were assumed to be realizations from a collection of study-specific multinomial distributions with $p_{ik}$ being the probability of outcome type $k$, $k = S$, $Tw$, or $Tr$ in study $i$. The multinomial probabilities were modeled as

$$p_{ik} = \frac{e^{\phi_{ik}}}{\sum_{k=1}^{3} e^{\phi_{ik}}}$$

where

$$\phi_{ik} = \mu_k + e_{ik}$$

and $e_{ik}$ is the common effect of outcome $k$, the parameter of interest, and $e_{ik}$ is a random effect for outcome $k$ specific to study $i$. Under this model, the mean outcome probabilities can be estimated as

$$p_k = \frac{e^{\mu_k}}{\sum_{k=1}^{3} e^{\mu_k}}$$

The model was fit by using noninformative multivariate Normal priors for both the $\mu_k$'s and the random effects. The random effects were included in the models to allow for more heterogeneity across studies in the observed probabilities than could be accounted for by assuming, for example, fixed Dirichlet priors for the probabilities. Two chains, starting at disparate initial values, were run in a Markov chain Monte Carlo algorithm, using WinBUGS 1.4 software (http://www.mrc-bsu.cam.ac.uk/bugs/), to obtain posterior estimates of the probabilities. Based on Gelman-Rubin statistics (51), convergence was assumed after a burn-in of 20,000 iterations, and 10,000 samples from each chain were subsequently selected, resulting in 20,000 posterior samples for $P[S/C]$, $P[Tr/C]$, $P[Tw/C]$, and $P[S/G]$, $P[Tr/G]$, and $P[Tw/G]$. Additional sampling along the chains indicated no impact of within-chain autocorrelation on the posterior estimates when sampling was repeated with a lag of 10 iterations. Model adequacy was assessed by using an evaluation of the chi-square posterior deviations (41). These predictive checks should be close to 0.5 for acceptable models and were 0.39 and 0.56 for the clomiphene and gonadotropin models, respectively, indicating adequate modeling of the variation in the published results across selected studies. The potential impact of study type (observational or randomized controlled trial) and year of study (published in or after 2000) were evaluated by altering the models for $\phi_{ik}$ to include terms for these effects with assumed noninformative Normal priors. Comparison of the deviance information criteria (41) and posterior 95% credible intervals for both the study-type and year-of-publication effects and the probability estimates indicated that inclusion of these terms resulted in negligible change to model adequacy.
When the posterior samples of \( S, Tw, \) and \( Tr \) probabilities were used, the probability that a fetus was included in a multiple-fetus (MF) pregnancy given conception with \( C \) (which we designate as \( P[MF|C] \)) was estimated as

\[
P[MF|C] = (2 \times P[Tw|C] + 3 \times P[Tr|C]) / (P[S|C] + 2 \times P[Tw|C] + 3 \times P[Tr|C]).
\]

Similarly, the probability that a fetus was included in a multiple-fetus pregnancy given conception with \( G \) was estimated as

\[
P[MF|G] = (2 \times P[Tw|G] + 3 \times P[Tr|G]) / (P[S|G] + 2 \times P[Tw|G] + 3 \times P[Tr|G]).
\]

The probability that a fetus was included in a multiple-fetus pregnancy given conception with either \( C \) or \( G \) was then estimated as

\[
P[MF|C \text{ or } G] = P[MF|C] \times P[C] + P[MF|G] \times (1 - P[C]),
\]

where \( P[C] \) is the probability that a woman conceived by using \( C \) given that she was taking either \( C \) or \( G \), and \( P[G] = 1 - P[C] \) is the probability that she conceived by using \( G \). \( P[C] \) was estimated as having a mean value of 0.71 and a wide 95\% uncertainty interval ranging from approximately 50\% to 90\%.

**Step 2. Estimate the probability that an infant was one of multiple livebirths given that a woman became pregnant by using non-ART ovulation stimulation.** Given the estimate of \( P[MF|C \text{ or } G] \) derived in step 1, the probability of multiple livebirths (MLB) given that a woman became pregnant by taking either \( C \) or \( G \) was estimated as

\[
P[MLB|C \text{ or } G] = P[MF|C \text{ or } G] \times K,
\]

where \( K \) is a survivorship ratio defined as

\[
K = P[MLB|C \text{ or } G] / P[MF|C \text{ or } G].
\]

We assumed that the uncertainty associated with \( K \) could be modeled by using a right triangular distribution with both a lower bound of 0.90 and most likely value of 0.93 but with an upper bound of 1.

**Step 3. Estimate the percentage of all US infants and US singletons resulting from non-ART ovulation stimulation.** We estimated that 59.90\% of all liveborn, multiple-birth infants in 2005 were conceived naturally and that 17.28\% were conceived with ART. Thus, the probability that a liveborn, multiple-birth infant in 2005 was conceived as a result of a non-ART ovulation treatment with either \( C \) or \( G \) can be estimated as

\[
P[C \text{ or } G|MLB] = 1 - 0.599 - 0.1728 = 0.2282.
\]

Given the known number of multiple births that occurred in 2005, \( N_{MUL} \), the number of multiple births resulting from use of either \( C \) or \( G \) can then be estimated as

\[
N_{MUL,C \text{ or } G} = N_{MUL} \times P[C \text{ or } G|MLB] = N_{MUL} \times 0.2282.
\]

Because \( N_{MUL,C \text{ or } G} \) is also defined by the relation

\[
N_{MUL,C \text{ or } G} = N_{C \text{ or } G} \times P[MLB|C \text{ or } G],
\]

where \( N_{C \text{ or } G} \) is the number of all liveborn infants resulting from use of \( C \) or \( G \), \( N_{C \text{ or } G} \) can be estimated as

\[
N_{C \text{ or } G} = N_{MUL,C \text{ or } G} / P[MLB|C \text{ or } G],
\]

where \( N_{MUL,C \text{ or } G} \) is defined by an earlier equation in step 3 and \( P[MLB|C \text{ or } G] \) in step 2.

In addition, given the total number of livebirths in 2005, \( N \), the percentage of livebirths resulting from use of \( C \) or \( G \) can be estimated as

\[
P_{C \text{ or } G} = N_{C \text{ or } G} / N.
\]

The number of singleton livebirths resulting from use of \( C \) or \( G \) in 2005, \( N_{S,C \text{ or } G} \), is estimated by using

\[
N_{S,C \text{ or } G} = N_{C \text{ or } G} - N_{MUL,C \text{ or } G}.
\]

Doing so leads to an estimator for the percentage of all singleton livebirths resulting from use of \( C \) or \( G \) in 2005 of

\[
P_{S,C \text{ or } G} = N_{S,C \text{ or } G} / N_S,
\]

where \( N_S \) is the known number of singleton births in 2005.
### Appendix Table 1. Search Terms and Criteria for Inclusion of Articles in a Systematic Review of Studies Assessing Multiple Gestation Risk Associated With Non-ART Ovulation Stimulation Treatments

<table>
<thead>
<tr>
<th>PubMed search terms: (ovarian stimulation OR ovarian induction OR ovulation stimulation OR ovulation induction OR urofollitropin OR clomiphene) AND (multiple birth OR multiple pregnancy OR twins OR triplets). Limits: published in the last 10 years, human subjects, English language. Search based on ALL FIELDS.</th>
<th>342 Studies identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each abstract reviewed by one of 3 study authors (L. S., C. B., J. P.) and classified as does not qualify for further review OR qualifies/possibly qualifies for further review (cannot discern from abstract). “Qualified for further review” were studies that 64 Individual studies qualified for further review; 14 meta-analyses identified</td>
<td></td>
</tr>
<tr>
<td>Collected original data beyond case reports and series OR used meta-analysis with data presentation. Included ≥1 groups receiving a non-ART ovulation treatment. The 64 studies underwent further screening review to select articles for full review. Criteria for full review: 18 Studies with 22 separate treatment groups selected for final full review</td>
<td></td>
</tr>
<tr>
<td>Separate data presentation for at least one ovulation treatment group clearly defined as treatment with clomiphene or clomiphene-related medication, treatment with gonadotropin, or a specified treatment protocol that included both types of medications. At least one ovulation treatment group with ≥50 pregnancies. Data on both twin and triplet/+ gestation provided. Study not a duplicate or near duplicate of another, more recent study. Each of the 18 studies reviewed and abstracted by 2 of the 4 study authors. For each study treatment group with at least 50 pregnancies, data abstracted on type of study; time period; patient characteristics; treatment protocol; cancellation criteria; and numbers of women, treatments, pregnancies, and twin and triplet/+ gestations. Any discrepancies between the 2 abstractions discussed with the entire author group and resolved through consensus.</td>
<td></td>
</tr>
<tr>
<td>5 Treatment groups examining clomiphene treatment (3 studies); 17 treatment groups examining gonadotropin treatment or gonadotropin + clomiphene treatment protocol (16 studies)</td>
<td></td>
</tr>
<tr>
<td>14 Meta-analyses reviewed further to identify additional studies that met the inclusion criteria. 0 Additional studies identified</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ART, assisted reproductive technologies; triplet/+, triplet or higher order.