Evidence of Absence or Absence of Evidence? A Re-analysis of the Effects of Low-Dose Aspirin in IVF

Marcus D. Ruopp, National Institutes of Health
Tara C. Collins, National Institutes of Health
Brian W Whitcomb, University of Massachusetts - Amherst
Enrique F. Schisterman, National Institutes of Health

Available at: https://works.bepress.com/brian_whitcomb/8/
Evidence of Absence or Absence of Evidence? A Re-analysis of the Effects of Low-Dose Aspirin in IVF

Marcus D. Ruopp, B.A., Tara C. Collins, M.D., Brian W. Whitcomb, Ph.D., and Enrique F. Schisterman, Ph.D.

Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Human Development, National Institutes of Health, DHHS, Bethesda, MD

Abstract

Objective—To assess the conflicting evidence whether low-dose aspirin (LDA) is beneficial in in-vitro fertilization (IVF), and evaluate the meta-analysis performed by Gelbaya et al. in the March 2007 issue of Human Reproduction Update, in which they found no effects of LDA and recommended discontinuing its use in IVF. We present a re-analysis of the effects of LDA in IVF and raise methodological questions regarding the analysis by Gelbaya et al.

Design—A meta-analysis of prospective randomized trials evaluating the effects of low-dose aspirin in IVF.

Patients—Women undergoing in-vitro fertilization/intracytoplasmic sperm injection

Interventions—Low-dose acetylsalicylic acid (aspirin)

Main Outcome Measure(s)—Pregnancy rates, implantation rates, miscarriage rates.

Result(s)—Ten randomized clinical trials were included in the analysis. Clinical pregnancy rate per embryo transfer was significant comparing LDA to non-treatment (RR: 1.15, 95% CI: 1.03 – 1.27). Non-significant estimates comparing LDA to non-treatment were found for implantation and miscarriage rates.

Conclusion(s)—Our results suggest that aspirin may increase clinical pregnancy rates and that more data is needed to resolve the issue. At this point, there is no reason to change clinical management and discontinue the use of aspirin.

Keywords

Aspirin; In-vitro fertilization; Pregnancy; Implantation; Fixed-effects
Introduction

In-vitro fertilization (IVF) has become a technique widely utilized in assisted reproductive technology (ART). Despite its current popularity, success rates in IVF remain low with estimates suggesting around a twenty-eight percent live birth rate for each IVF cycle started at best (1). Fertilization rates have reached high success rates; however, pregnancy rates per embryo transfer are still relatively low at around 34 percent (1). In addition, women undergoing IVF often face substantial costs for each failed cycle. As a result, any intervention with even a small relative effect on implantation, pregnancy, and live birth rates in IVF has potential significant quality of life and cost implications.

One such potential intervention is low-dose acetylsalicylic acid (aspirin). Aspirin has been a primary target of interest because of its anti-inflammatory, vasodilatory and platelet aggregation inhibition properties. Aspirin inhibits the enzyme cyclooxygenase-1 (COX-1), which facilitates the synthesis of multiple prostaglandins (2). By inhibiting COX-1, aspirin increases levels of the vasodilator prostacyclin and prevents thromboxane regulated vasoconstriction (3 and 4). In this regard, low-dose aspirin blocks the vascular-constricting compound while not affecting the compound that promotes placental blood flow.

While these properties suggest the possibility of beneficial effects of aspirin in IVF, randomized clinical trials have produced conflicting results in this regard. The March issue of the Human Reproduction Update included a meta-analysis by Gelbaya et al. that systematically evaluated the current state of evidence regarding the effects of aspirin in IVF (5). The authors evaluated data from 6 studies that met their inclusion criteria and, having found estimates with confidence intervals overlapping the null hypothesis, concluded that aspirin has no effect on pregnancy in IVF. Accordingly, Gelbaya et al. advocate a change in standard-of-care, stating “on the basis of up-to-date evidence, low-dose aspirin has no substantial positive effect on likelihood of pregnancy and, therefore, it should not be routinely recommended for women undergoing IVF/ICSI” (5).

This line of reasoning raises several methodological issues regarding use of, and interpretation of results from, meta-analysis, and statistical analysis in general. Specifically, questions regarding use of fixed-effects versus random effects models, selection of studies for consideration, proper interpretation of hypothesis tests and confidence intervals, and the link between research to policy are raised. This paper discusses each of the above issues in the context of a re-analysis of these data.

Materials and Methods

In order to review the selection of studies by Gelbaya et al., we performed a thorough search of multiple databases by two independent reviewers identifying trials utilizing aspirin during IVF. The databases that were electronically searched in February 2007 were: MEDLINE (1966 – 2007), Web of Science (1980 – 2007), EMBASE (1974 – 2007), TOXLINE (1900 – 2007), DART (1900 – 2007), and the Cochrane Database of Systematic Reviews (1991 – 2007). In addition, the references of all selected studies were searched manually. The keywords used in a cross search with aspirin or salicylates to identify relevant papers were: in-vitro fertilization (IVF), perfusion, implantation, fertility, infertility, secondary infertility, sub fertility, fecundity, endometrium, and endometrial thickness.

Study selection was based on prospective human population trials that used low-dose aspirin (less than 150 mg) during IVF. All trials included were randomized or matched controlled trials. Only abstracts written in English were considered. Excluded studies included retrospective and uncontrolled studies, editorials, reviews, and animal studies. In addition, studies that used aspirin therapy in conjunction with other drug therapies were excluded. The
screening process was independently performed by the two reviewers, with disputes settled with assessment by a third reviewer. After selection of studies that met the specified criteria, data extraction was performed using 2 x 2 structured forms. The analysis was conducted using the Cochrane Review Manager Software (version 4.1, Update Software, Oxford) to calculate odds ratios, 95 percent confidence intervals, and heterogeneity using the Mantel-Haenszel method, assuming fixed effects. The fixed-effects model pools data across studies weighting the studies based on their sample size.

Results

Having performed the above search strategy, 6608 citations were identified for review. Titles and abstracts were evaluated and 33 studies found to meet basic inclusion criteria and were selected for detailed review by two independent researchers. Of these 33 studies, 23 studies were excluded for reasons including study design — retrospective cohort studies (n = 5), prospective observational studies (n = 2); use of aspirin treatment including other factors (e.g. heparin/prednisone; n =15); and for presentation of study results as medians rather than means (n = 1). After exclusions, 10 studies (6, 7, 8, 9, 10, 11, 12, 13, 14 and 15) met the inclusion criteria [Table 1].

We performed analyses including all ten studies together as well as subgroup analyses to address important distinctions among the studies. In the first subgroup analysis, we included only studies that used fresh embryo transfer (n = 9). In the second subgroup analysis, we included all studies evaluated by Gelbaya et al., and thus excluded conference abstracts (n = 2), subgroups of infertile patients (n = 1), and frozen embryo transfer (n = 1).

Clinical pregnancy rates per embryo transfer were reported by all ten studies (n = 2801 IVF cycles). Utilizing the Mantel-Haenszel fixed-effect model, a significant risk ratio of 1.15 (95 percent confidence interval:1.03 – 1.27) was found comparing clinical pregnancy rates per embryo transfer in the low-dose aspirin to non-treatment groups (Refer to figure 1). Implantation rates were reported by three studies (n = 612 embryos transferred) with a risk ratio of 1.08 (95 percent confidence interval: 0.69 – 1.71) comparing low-dose aspirin to non-treatment groups. Miscarriage rate per clinical pregnancy was reported by 4 studies (n = 671 IVF cycles) with a risk ratio of 1.19 (95 percent confidence interval: 0.86 – 1.65) comparing low-dose aspirin to non-treatment groups.

In the subgroup analysis of data from studies that utilized fresh embryo transfer, clinical pregnancy rates were reported by 9 studies (n = 2765 IVF cycles). A significant risk ratio of 1.16 (95 percent confidence interval: 1.04 – 1.29) was found comparing clinical pregnancy rates per embryo transfer in the low-dose aspirin to non-treatment groups. Implantation rates were reported by two studies (n = 480 embryos transferred) with a risk ratio of 1.32 (95 percent confidence interval: 0.81 – 2.16) comparing low-dose aspirin to non-treatment groups. Miscarriage per clinical pregnancy was the same as the full analysis.

In subgroup analysis looking only at studies included by Gelbaya et al. that excluded conference abstracts, subgroups of infertile patients, and frozen embryo transfer, data regarding clinical pregnancy rates reported by 6 studies (n = 2515 IVF cycles) were considered (5). A risk ratio of 1.12 (95 percent confidence interval: 1.00 – 1.25) was found comparing the clinical pregnancy rates per embryo transfer in the low-dose aspirin to non-treatment groups. Implantation rates were reported by one study (n = 348 embryos transferred) with a risk ratio of 0.89 (95 percent confidence interval: 0.48 – 1.66) comparing low-dose aspirin to non-treatment groups. Looking at miscarriage per clinical pregnancy (n = 658 IVF cycles), a risk ratio of 1.17 (95 percent confidence interval: 0.84 –1.63) was observed comparing low-dose aspirin to non-treatment groups. The test for heterogeneity was non-significant in all analyses.
(p > .05) except for implantation rates for the full analysis and subgroup analysis utilizing only fresh embryo transfer. Table 2 displays summary statistics and confidence intervals for the full analysis and subgroup analysis using the fixed-effects (Mantel-Haenszel) and random-effects (DerSimonian and Laird) model to compare the differences in results produced by the models.

**Discussion**

To evaluate the effect of modeling assumptions and inclusion criteria on meta-analysis of the effect of aspirin on IVF outcomes, we compared results of the recent paper by Gelbaya et al. with those reached in our re-analysis. Our literature search identified the same initial set of studies for consideration as that of Gelbaya et al.; however, our conclusions are somewhat different, even for analyses using identical data. We observed statistically significant results regarding relative pregnancy rates per embryo transfer between the aspirin treatment group and the non-treatment group, and non-significant findings for other outcomes considered. Based on our results, the overall affect of low-dose aspirin in pregnancy appears uncertain. Reasons for these divergent findings include approaches to statistical modeling, study selection, and statistical inference.

Conventionally, meta-analysis uses either fixed- or random-effects models to generate a test statistic and confidence interval. The decision whether to use a fixed- or random-effects model is frequently misunderstood. The differences between these two approaches are philosophical, and reflect fundamentally different conceptualizations of the research question. As a result, while both models produce an estimate for the effect of treatment, the interpretation of the regression coefficient is subtly different. In this paper, the Mantel-Haenszel fixed-effects model was used, which weights the studies by the inverse of the variance of estimates. The fixed-effects model produces a confidence interval that takes into account the random variation within each trial, and attempts to answer the question: *Was an effect seen in the trials included in the analysis?* Fixed-effects model weighting thus constitutes a direct synthesis of the literature (16). Radhakrishna showed that the test based on the Mantel-Haenszel chi-square is the uniformly most powerful test and has optimal statistical properties (17).

Gelbaya et al. used a random-effects model, which introduces a between study component of variance. This approach is sometimes thought to be a more conservative approach, since confidence intervals produced from the random-effects model are necessarily wider than the fixed-effects model due to the contribution of the inter-study variability to total variability. Random-effects models use study power to estimate inter-study covariance, and the weighting of each study in the model is determined by both intra- and inter-study variance. Accordingly, the random-effects model attempts to answer the question: *Will an effect be seen in a given trial in the future?* (16) Beyond the loss of precision inherent to random-effects models, there are other drawbacks as well. Peto (18) called use of random-effects meta-analysis “wrong” because it answers a question that is “abstruse and uninteresting”. Moreover, utilization of random-effects models requires strong assumptions that are unlikely to be valid in practice. Most notably, random-effect analysis is based on “the peculiar premise that the trials done are representative of some hypothetical population of trials, and... that the heterogeneity can be represented by a single variance.” (19)

The results of this analysis and the analysis by Gelbaya et al. both indicate the heterogeneity between studies to be non-significant, yet, there is still certainly much variation between study results (5). It is tempting to assume that through the use of a random-effects model, one is able to adequately address the effects of heterogeneity between studies by introducing a between-study component of variance. However, in reality, the inter-study component of variance of the random-effects model may not be able to account for actual inter-study heterogeneity. Petitti (16) suggests that the random-effects model should only be used when the absence of inter-
study heterogeneity can be assumed, because any significant between-study heterogeneity dominates the weights assigned to the studies. As a consequence, in the presence of substantial between-study heterogeneity, small and large trials become weighted the same, and the summary statistic is greatly affected by the inclusion of small trials into the analysis. The fixed-effect model, however, takes into account only between-study variances and thus weights studies according to sample size.

Another reason for discordant findings is in regard to the choice of studies for inclusion in the analyses. Gelbaya et al. excluded studies using frozen embryo transfer, subgroups of infertile patients, and conference abstracts. The justification for these exclusions is unclear. Likewise, the inclusion of studies that appear to meet inclusion criteria is likewise curious, such as Check et al. which is included because it is a matched-controlled study (7). Moreover, there is no evidence in the literature to suggest that aspirin may act differently in these subgroups than in those included in the analysis by Gelbaya et al.. Few studies have examined the effects of aspirin use with frozen embryo transfer in IVF (20 and 7) or among infertile patients (21 [only reported medians] and 15). Similarly, there is no clear justification for the exclusion of conference abstracts, which clearly show positive benefits of the use of aspirin in IVF. In this re-analysis, we have illustrated the effect of exclusion of these trials on study findings.

The impact of these exclusions would be less of an issue if the excluded studies comprised a small proportion of available evidence, but this is not the case for this research question. In weighing the evidence to find the balance, one ought first to consider whether total weight justifies the exercise. Gelbaya et al. observed non-significant effects of aspirin for all considered outcomes in IVF and concluded that aspirin use in IVF should be abandoned. Confidence intervals including the null are observed not only in the face of a true null hypothesis (i.e. \( P_{\text{aspirin}} = P_{\text{no treatment}} \)) but also when study power is inadequate to detect a true alternative hypothesis, (i.e., \( \text{SE}(P_{\text{aspirin}}) \) is relatively large). A power calculation using the trials included by Gelbaya et al. on clinical pregnancy rates per embryo transfer and a fixed-effect model determines a power of 0.56 assuming the estimated effect is equal to the true effect (Hedges and Pigott, 2001). The random-effects model, which is more conservative than the fixed-effects model, has even less power to detect an effect. The hypothesis of a positive effect of aspirin use on IVF outcomes is clearly of interest, but even great interest cannot make up for an absence of evidence. One must question whether a meta-analysis for this question is premature. Recommendation of policy changes to abandon treatment based on coverage of the null may be unwarranted, especially in this case with such few studies in the analysis.

Moreover, even acceptance of the null hypothesis of no effect over the alternative hypothesis of an effect is not an argument against aspirin use. Rather, it suggests either the absence of utility, or failure to detect utility that exists. In the face of an insufficiency of evidence, it is unclear that policy revision is prudent. Demonstration of adverse effects of aspirin that outweigh beneficial effects would provide an argument against starting or continuing aspirin use for this purpose.

In multiple large trials, aspirin has been used without significant complications, though a small increased risk of gastrointestinal effects as well as bleeding events has been observed (22, 23 and 24). Meta-analyses regarding the use of aspirin during pregnancy found neither toxicity nor serious negative maternal/fetal/infant side effects from the use of low-dose aspirin preconceptionally or during pregnancy (25). Similarly, no side effects of low-dose aspirin treatment were observed in meta-analysis of high-risk pregnancies (26). Increased risks for gastroschisis and cleft lip and palate have been observed with self-reported high doses of aspirin (325 mg) (27 and 26). It is suggested that misclassification of exposure, inadequate control of confounding, and reporting bias are likely to have affected investigations of aspirin use in pregnancy, and that available data do not support a teratogenic effect of aspirin (28).
Systematic reviews and meta-analysis are intended to consolidate available literature and provide an overview of the state of evidence and, as such, are a valuable tool for evidenced-based medicine. By combining results from multiple studies, meta-analysis may overcome statistical power issues limiting the individual studies; however, this combination does not inure the method from issues with lack of statistical power itself. When meta-analysis is performed using a small set of trials and has low statistical power, the resultant absence of evidence regarding the possible effect of a treatment of interest should not be confused as evidence that there is no effect of that treatment. This re-analysis attempts to illustrate this point and shows how conflicting results and conclusions can be reached under differing modeling approaches. Our results suggest that any recommended alterations of current practice may be premature, and that more trials are required for meta-analysis on the effects of low-dose aspirin on outcomes in IVF to have adequate power. Until that point, conclusions regarding an effect or lack thereof are wanting of solid support and clinicians should continue their current practice with regard to aspirin use until more information comes to light.

Acknowledgements

This research was supported by the Intramural Research Program of the NIH, Epidemiology Branch, DESPR, NICHD.

References


Figure 1.
Forest plot of the effect of aspirin versus placebo or no treatment on clinical pregnancy rate per embryo transfer in full analysis using fixed-effects model.
### Table 1
Authors, study type, year and treatment regimen for studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>Author</th>
<th>Study Type</th>
<th>Year</th>
<th>Aspirin Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bordes et al.</td>
<td>Randomized, Prospective Study</td>
<td>2003</td>
<td>Treatment (100mg aspirin/placebo) starting at day 21 of cycle preceding the first dose of gonadotropin</td>
</tr>
<tr>
<td>Check et al.</td>
<td>Matched Controlled Study</td>
<td>1998</td>
<td>Treatment (81mg aspirin/no placebo) starting at day 2 of frozen embryo transfer cycle until delivery</td>
</tr>
<tr>
<td>Duvan et al.</td>
<td>Randomized, Prospective Study</td>
<td>2006</td>
<td>Treatment (100mg aspirin/placebo) starting at embryo transfer until clinical pregnancy</td>
</tr>
<tr>
<td>Lentini et al.</td>
<td>Randomized, Prospective Study</td>
<td>2003</td>
<td>Treatment (100mg aspirin/no placebo) starting a month before the first dose of gonadotropin</td>
</tr>
<tr>
<td>Pakkila et al.</td>
<td>Randomized, Prospective Study</td>
<td>2005</td>
<td>Treatment (100mg aspirin/placebo) starting at first dose of gonadotropin until delivery</td>
</tr>
<tr>
<td>Rubinstein et al.</td>
<td>Randomized, Prospective Study</td>
<td>1999</td>
<td>Treatment (100mg aspirin/placebo) starting on day 21 of cycle before IVF until 12 weeks gestation</td>
</tr>
<tr>
<td>Urman et al.</td>
<td>Randomized, Prospective Study</td>
<td>2000</td>
<td>Treatment (80mg aspirin/ no placebo) starting at first day of stimulation and continued until fetal cardiac activity</td>
</tr>
<tr>
<td>Van Dooren et al.</td>
<td>Randomized, Prospective Study</td>
<td>2004</td>
<td>Treatment (100mg aspirin/placebo) starting at day 16 of cycle until 10 weeks gestation</td>
</tr>
<tr>
<td>Waldenstrom et al.</td>
<td>Randomized, Prospective Study</td>
<td>2006</td>
<td>Treatment (75mg aspirin/no placebo) starting at day of embryo transfer until pregnancy test</td>
</tr>
<tr>
<td>Weckstein et al.</td>
<td>Randomized, Prospective Study</td>
<td>1997</td>
<td>Treatment (81mg aspirin/no placebo) starting at day 1 of estrogen treatment until 9 weeks post embryo transfer</td>
</tr>
</tbody>
</table>

*Fertil Steril. Author manuscript; available in PMC 2009 July 1.*
Table 2
Summary statistics and confidence intervals for fixed and random effects models for full and subgroup analysis.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Fixed Effects</th>
<th>Random Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Analysis: All Studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy Rate</td>
<td>1.15 (1.03, 1.27)</td>
<td>1.14 (0.95, 1.35)</td>
</tr>
<tr>
<td>Implantation Rate</td>
<td>1.08 (0.69, 1.71)</td>
<td>1.00 (0.34, 2.93)</td>
</tr>
<tr>
<td>Miscarriage Rate</td>
<td>1.19 (0.86, 1.65)</td>
<td>1.18 (0.85, 1.64)</td>
</tr>
<tr>
<td>Subgroup Analysis: Studies Including Only Fresh Embryo Transfer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy Rate</td>
<td>1.16 (1.04, 1.29)</td>
<td>1.15 (0.98, 1.36)</td>
</tr>
<tr>
<td>Implantation Rate</td>
<td>1.32 (0.81, 2.16)</td>
<td>1.49 (0.50, 4.46)</td>
</tr>
<tr>
<td>Miscarriage Rate</td>
<td>1.19 (0.86, 1.65)</td>
<td>1.18 (0.85, 1.64)</td>
</tr>
<tr>
<td>Subgroup Analysis: Studies Included by Gelbaya et al.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy Rate</td>
<td>1.12 (1.00, 1.25)</td>
<td>1.09 (0.92, 1.29)</td>
</tr>
<tr>
<td>Implantation Rate</td>
<td>0.89 (0.48, 1.66)</td>
<td>0.89 (0.48, 1.66)</td>
</tr>
<tr>
<td>Miscarriage Rate</td>
<td>1.17 (0.84, 1.63)</td>
<td>1.17 (0.84, 1.63)</td>
</tr>
</tbody>
</table>