

## LETTER TO THE EDITOR

## META-ANALYSIS OF ASPIRIN FOR THE PREVENTION OF PREECLAMPSIA: DO THE MAIN RANDOMIZED CONTROLLED TRIALS SUPPORT AN ASSOCIATION BETWEEN LOW-DOSE ASPIRIN AND A REDUCED RISK OF DEVELOPING PREECLAMPSIA?

Tiago V. Pereira, Martina Rudnicki, Júlia M. P. Soler, and José E. Krieger

The meta-analysis by Ruano et al<sup>1</sup> evaluating 17 randomized controlled trials considering the pooled relative risk (RR) of 16,898 women at high-risk of preeclampsia receiving acetylsalicylic acid (8679 treatment and 8219 placebo) suggested that low-dose aspirin (50-150 mg/day) has a small but significant protective effect on the risk of preeclampsia: Pooled RR = 0.87; 95% confidence intervals (CI) = 0.79 to 0.96,  $P = 0.004$ . This conclusion relies uniquely on the estimate from the fixed-effects model using the Mantel-Haenszel method (MH).<sup>2</sup> However, we would like to raise some issues that allow different interpretations from the same meta-analysis.

First, results of a meta-analysis are model-dependent.<sup>3</sup> There are potential differences between models in meta-analysis, since they have distinct properties and assumptions,<sup>2,4</sup> and can model for different effects.<sup>3</sup> Second, because the choice between fixed- and random-effect models remains a controversial issue,<sup>5</sup> results from both models should be considered when drawing conclusions from a meta-analysis,<sup>6</sup> especially in cases where heterogeneity among treatment effects is observed.<sup>7</sup> Third, when evidence for publication bias is not properly considered, results from meta-analytic studies can be seriously overestimated, yielding misleading conclusions.<sup>8</sup>

In this regard, we extended further the analysis by Ruano et al<sup>1</sup> and provide an alternative interpretation from the same data set, illustrating practical consequences of fitting different types of statistical modeling in meta-analysis.

Furthermore, we illustrate that changes in model selection substantially affect the outcome of a meta-analysis, particularly when the null hypothesis of homogeneity among treatment effects is rejected.

It is well established that in the fixed-effects model, the true treatment difference is often considered to be the same for all trials.<sup>2</sup> In other words, fixed-effects models assume homogeneity among study estimates, whereas the variance of each trial estimate is based uniquely on sampling variation within the trial (within-study variance). In the random-effects model, however, the true treatment difference in each trial is itself assumed to be a random variable with a normal distribution.<sup>4</sup> In addition, a second component of the total variance, the variance between-study, is estimated. Thus, in the random-effects model both within-study and between-study variances are taken into account in the assessment of the uncertainty (confidence intervals). While estimates based on fixed-effect models are specific to the particular trials included in the meta-analysis, results of the random-effects models are more amenable to generalization in terms of population.<sup>6</sup> When there is no heterogeneity between trials, both models tend to provide the same overall estimate and standard error.<sup>5</sup> However, this is not true when heterogeneity is present.<sup>5</sup> As the heterogeneity increases, the standard error of the common estimate from the random effects model becomes wider than that from the fixed-effects model.<sup>4</sup> Thus, the difference between the overall estimates from the two approaches depends strongly on the amount of between-study variation among trials.<sup>6</sup> Such properties have led some authors to recommend the incorporation of the between-study variance through an implementation of the random-effects model in meta-analysis when heterogeneity is observed.<sup>9</sup> In fact, published evidence suggests the broad use of random-effects models for the combination of both homogeneous and mild heterogeneous trials,<sup>7</sup> whereas the combination of highly heterogeneous studies should be avoided.<sup>6</sup>

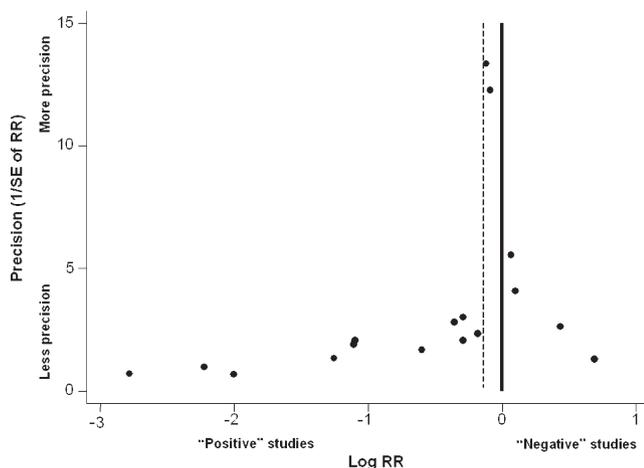
Department of Biochemistry, Paulista School of Medicine, Federal University of São Paulo - São Paulo/SP, Brazil.  
Clinical and Toxicological Analysis Department, São Paulo University of Pharmaceutical Sciences - São Paulo/SP, Brazil.  
Institute of Mathematics and Statistics, São Paulo University - São Paulo/SP, Brazil.  
Heart Institute (InCor), São Paulo University Medical School, São Paulo/SP, Brazil.  
Email: t27026t@yahoo.com.br

Hence, although Ruano et al<sup>1</sup> correctly tested the heterogeneity among effect sizes of individual studies, they have not considered adequately this finding. According to these authors, the homogeneity among treatment effects was formally assessed through the Cochran's Q statistic,<sup>4</sup> which is a weighted sum of squares of the deviations of individual study RR estimates from the overall estimate. This statistic estimates the probability that the observed pattern of results may have occurred simply by chance. When the RRs are homogeneous, the Q statistic follows a chi-squared distribution with  $r - 1$  degrees of freedom (df), where  $r$  is the number of studies. Unfortunately, heterogeneity tests in meta-analysis lack power, and statistical significance set at 10% is advocated.<sup>7,10</sup> Running Ruano et al<sup>1</sup> analysis again, we observed a significant evidence for heterogeneity both in the asymptotic ( $\chi^2 = 25.86$ ,  $df = 16$ ,  $P = 0.056$ ) and the parametric bootstrap version (1000 replications,  $P = 0.049$ ) of the Q statistic. To further explore heterogeneity, we derived the proportion of the total variance of the pooled effect measure<sup>7</sup> ( $R_i$  statistic) due to between-study variance (random-effects model), and quantified the  $I^2$  metric,<sup>11</sup> which takes values between 0% and 100%, with higher values denoting greater degree of heterogeneity. Derivation of the  $R_i$  statistic suggests that almost half (46%) of the total variance from the meta-analysis of Ruano et al.<sup>1</sup> came from between-study variance. In addition, the  $I^2$  metric denotes the presence of moderate heterogeneity (38.1%). Thus, we provide compelling evidence for a mild to moderate heterogeneity among study results, suggesting that the conclusion of Ruano et al<sup>1</sup> should not be based uniquely on a fixed-effects model. In this respect, we applied to the same data set a random-effects model using the DerSimonian-Laird method (DL).<sup>4</sup> Surprisingly, we found no convincing evidence for a protective effect of low-dose aspirin under this model at 5% (DL Common RR = 0.835; 95% CI = 0.697 to 1.001,  $P = 0.051$ ). In order to scrutinize the heterogeneity, we fitted a random-effects regression model (meta-regression) of the log RR.<sup>12</sup> This regression model investigates whether any study feature influences the magnitude of the relative risk in individual studies. Year of publication, impact factor (according to the Journal of Citation Reports 2004), aspirin dose, sample size, and gestation age of inclusion were assigned as study-level covariates. An exploratory stepwise backward random-effects meta-regression reveals that year of publication and total sample size were associated with reduced magnitudes of RR (year of publication slope coefficient = 0.07; 95% CI = 0.03 to 0.11,  $P < 0.001$  and total sample size slope coefficient =  $4 \times 10^{-5}$ ; 95% CI =  $6.9 \times 10^{-6}$  to  $7 \times 10^{-5}$ ,  $P = 0.02$ ). These meta-regression coefficients are estimates of the expected increment in the

log RR per unit increase in the covariate. In other words, in the studies of low-dose aspirin for prevention of preeclampsia (high-risk women) the log RR is estimated to rise 0.07 per year increase and to increase =  $4 \times 10^{-5}$  per individual added in each the study. These findings led us to a perform sensitivity analysis considering total sample size.<sup>13</sup> In this analysis we separately explored studies with sample size  $>1000$  and  $\leq 1000$ . By combining RR from these 14 small trials (1549 treatment and 1704 placebo) with sample sizes equal or smaller than 1000 subjects, we obtained a marginally significant but stronger protective effect of low-dose aspirin in the prevention of preeclampsia (DL common RR = 0.67; 95% CI = 0.48 to 0.93,  $P = 0.02$ , Q statistic,  $\chi^2 = 23.15$ ,  $df = 13$ ,  $P = 0.04$ ). In contrast, by combining the 3 larger studies totaling 13645 individuals (7130 treatment and 6515 placebo) we found no evidence for a role of low-dose aspirin in preeclampsia prevention (MH common RR = 0.91; 95% CI = 0.82 to 1.02,  $P = 0.094$ , Q statistic parametric bootstrap version with 1000 replications,  $P = 0.34$ ). Likewise, stratification of studies by samples sizes  $>500$  and  $\leq 500$  yielded virtually the same results.

Disagreements between large and smaller trials are compatible with the presence of publication bias,<sup>8</sup> a tendency on the part of investigators, editors, and reviewers to publish preferentially studies with positive results ("statistically significant results").<sup>6,8,14</sup> In fact, it is well recognized that earlier trials tend to have smaller samples, less methodological quality, and to give more impressive estimates of treatment effect than subsequent research.<sup>13</sup> However, publication bias may lead to overestimation of treatment effects in meta-analysis, since negative, nonstatistically-significant small trials are less likely to be published.<sup>8</sup> Therefore, we tested graphically and formally the presence of publication bias in this set of 14 studies. We carried out analysis of publication bias using the Egger's regression asymmetry statistics<sup>8</sup> and the Begg-Mazumdar adjusted rank correlation test.<sup>15</sup> In addition, funnel plot asymmetry was also used.<sup>6,8</sup> The funnel plot is based on the fact that precision in estimating the underlying treatment effect will increase as the sample size of the included trials increases.<sup>6</sup> Results from small studies will scatter widely at the bottom of the graph, with the spread narrowing among larger studies.<sup>8</sup> In the absence of bias, the plot will resemble a symmetrical inverted funnel. Conversely, if negative small trials are less likely to be published, the graph will tend to be skewed (asymmetric).<sup>6</sup>

Although heterogeneity may affect these statistics,<sup>8,16</sup> both tests showed reasonable evidence for publication bias in the set of 14 small studies (Begg-Mazumdar's test,  $P = 0.02$ ; Egger's statistics,  $P = 0.001$ ). Additionally, as de-



**Figure 1** - Funnel plot of 17 studies investigating the effect of low-dose aspirin for preeclampsia prevention (high-risk women) previously analyzed by Ruano et al.<sup>1</sup> Vertical axis depicts the precision represented here by the inverse of the standard error (SE) of relative risk (RR). Horizontal axis represents the logarithm of RR. Solid line denotes null effect and dotted line is the pooled RR under a fixed-effects model (Mantel-Haenszel method). "Positive" studies denote trials showing a protective effect of aspirin in the prevention of preeclampsia (high-risk women), but not necessarily reflecting "statistically significant" studies. "Negative" studies represent trials showing no beneficial effects of aspirin in preeclampsia (high-risk women). Visual inspection of the funnel plot suggests asymmetry. At the bottom right: are small trials showing no beneficial effects missing?

pictured in Figure 1, the visual inspection of the funnel plot comparing precision (1/Standard Error of RR) against treatment effect (Log RR) reveals noticeable asymmetry (all 17 trials included), reinforcing the probable presence of publication bias.

These results do not completely rule out a role for aspirin in the preeclampsia prevention. However, taken together, the presence of heterogeneity, the conflicting results between large and smaller trials, the lack of robustness between fixed- and random-effects models, and the putative presence of publication bias call into question the reliability of the evidence for a protective effect of low-dose aspirin on the risk of preeclampsia.

In conclusion, despite the 20 years of research since the first report, the relationship between aspirin treatment and a reduced risk of preeclampsia in women considered at high-risk still remains an unresolved issue. Further approaches, such as meta-analysis considering language bias or individual patient data meta-analysis,<sup>17</sup> are required.

## REFERENCES

1. Ruano R, Fontes RS, Zugaib M. Prevention of preeclampsia with low-dose aspirin- a systematic review and meta-analysis of the main randomized controlled trials. *Clinics*. 2005;60(5):407-14.
2. Mantel N, Haenszel WH. Statistical aspects of the analysis of data from retrospective studies of diseases. *J Natl Cancer Inst*. 1959;22:719-48.
3. Brown H, Prescott R. *Applied Mixed Models in Medicine*. New York: Wiley; 2001.
4. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986; 7(3):177-88.
5. Villar J, Mackey ME, Carroli G, Donner A. Meta-analyses in systematic reviews of randomized controlled trials in perinatal medicine: comparison of fixed and random effects models. *Stat Med*. 2001;20:3635-47.
6. Whitehead A. *Meta-analysis of controlled clinical trials*. New York: John Wiley & Sons; 2002.
7. Takkouche B, Cadarso-Suarez C, Spiegelman D. Evaluation of old and new tests of heterogeneity in epidemiologic meta-analysis. *Am J Epidemiol*. 1999;150:206-15.
8. Egger M, Davey SG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629-34.
9. Hunter JE, Schmidt FL. Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*. 2000;8:275-92.
10. Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. *Stat Med*. 1998;17:841-56.

11. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21:1539-58.
12. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med.* 1999;18:2693-708.
13. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med.* 2001;135:982-9.
14. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995;273:408-12.
15. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics.* 1994;50:1088-101.
16. Terrin N, Schmid CH, Lau J, Olkin I. Adjusting for publication bias in the presence of heterogeneity. *Stat Med.* 2003;22:2113-26.
17. The PARIS Collaboration. Antiplatelet agents for prevention of pre-eclampsia and its consequences: a systematic review and individual patient data meta-analysis. *BMC Pregnancy Childbirth.* 2005;5:7.