Human Reproduction, Vol.32, No.10 pp. 2117–2122, 2017
Advanced Access publication on August 29, 2017 doi:10.1093/humrep/dex273

human reproduction **ORIGINAL ARTICLE** ***Reproductive epidemiology***

Selective outcome reporting and sponsorship in randomized controlled trials in IVF and ICSI

M. Braakhekke[1,*], I. Scholten[1], F. Mol[1], J. Limpens[2], B.W. Mol[3], and F. van der Veen[1]

[1]Department of Obstetrics & Gynaecology, Centre for Reproductive Medicine, Q3-119, Academic Medical Centre, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands [2]Medical Library, Research Support, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands [3]The Robinson Research Institute, School of Paediatrics and Reproductive Health, University of Adelaide and the South Australian Health and Medical Research Institute, Adelaide, Australia

*Correspondence address. Tel: +31-20 5663557; Fax: +31-20 5669044; E-mail: m.w.braakhekke@amc.nl

Submitted on December 17, 2016; resubmitted on July 25, 2017; accepted on August 2, 2017

STUDY QUESTION: Are randomized controlled trials (RCTs) on IVF and ICSI subject to selective outcome reporting and is this related to sponsorship?

SUMMARY ANSWER: There are inconsistencies, independent from sponsorship, in the reporting of primary outcome measures in the majority of IVF and ICSI trials, indicating selective outcome reporting.

WHAT IS KNOWN ALREADY: RCTs are subject to bias at various levels. Of these biases, selective outcome reporting is particularly relevant to IVF and ICSI trials since there is a wide variety of outcome measures to choose from. An established cause of reporting bias is sponsorship. It is, at present, unknown whether RCTs in IVF/ICSI are subject to selective outcome reporting and whether this is related with sponsorship.

STUDY DESIGN, SIZE, DURATION: We systematically searched RCTs on IVF and ICSI published between January 2009 and March 2016 in MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials and the publisher subset of PubMed. We analysed 415 RCTs.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Per included RCT, we extracted data on impact factor of the journal, sample size, power calculation, and trial registry and thereafter data on primary outcome measure, the direction of trial results and sponsorship.

MAIN RESULTS AND THE ROLE OF CHANCE: Of the 415 identified RCTs, 235 were excluded for our primary analysis, because the sponsorship was not reported. Of the 180 RCTs included in our analysis, 7 trials did not report on any primary outcome measure and 107 of the remaining 173 trials (62%) reported on surrogate primary outcome measures. Of the 114 registered trials, 21 trials (18%) provided primary outcomes in their manuscript that were different from those in the trial registry. This indicates selective outcome reporting. We found no association between selective outcome reporting and sponsorship. We ran additional analyses to include the trials that had not reported sponsorship and found no outcomes that differed from our primary analysis.

LIMITATIONS, REASONS FOR CAUTION: Since the majority of the trials did not report on sponsorship, there is a risk on sampling bias.

WIDER IMPLICATIONS OF THE FINDINGS: IVF and ICSI trials are subject, to a large extent, to selective outcome reporting. Readers should be aware of this to avoid implementation of false or misleading results in clinical practice.

STUDY FUNDING/COMPETING INTERESTS: No funding received and there are no conflicts of interest.

TRIAL REGISTRATION NUMBER: N/A

Key words: selective outcome reporting / IVF / ICSI / outcome measures / RCT / sponsorship

Introduction

Clinical research is one of the fundamental features of evidence based medicine. The highest level of evidence in clinical research is generated by aggregated evidence of randomized controlled trials (RCTs; Guyatt *et al.*, 2000). Pooling the data of trials involves careful and meticulous analysis of the internal validity of the included trials in order to know if the research question has been answered correctly and is free from bias (Higgins and Green, 2011). Bias in clinical trials may be described as systematic errors that favour one outcome over the other. Implementation of biased trial results leads to the use of ineffective drugs and devices, resulting in poor patient care and waste of money (Gluud, 2006). Potential sources of bias may be classified into selection bias (the selection and randomization of patients), performance bias (blinding of participants and personnel), detection bias (blinding of outcome assessment), attrition bias (incomplete outcome data caused by withdrawals) and reporting bias, caused by selective outcome reporting (Higgins and Green, 2011).

Selective outcome reporting occurs if some outcomes are reported while others are not, which threatens the validity of trials, by increasing the likelihood that the results were caused by chance, and potentially renders published results misleading (Higgins and Green, 2011; Dwan *et al.*, 2013). Selective outcome reporting is mainly determined by the choice of the primary outcome measure (Chan *et al.*, 2004). There are two types of selective outcome reporting. The first is that the primary outcome measure is a surrogate outcome measure that is not the most relevant for answering the research question; the second is that there are inconsistencies between planned and reported primary outcome measures (Grimes and Schulz, 2005; Lassere, 2008).

Selective outcome reporting is particularly relevant to reproductive medicine, especially IVF and ICSI, since there is a wide range of possible primary outcome measures to choose from. Outcome measures can vary from number of oocytes retrieved and hormone levels to singleton live birth and birth emphasizing a successful singleton at term (Braakhekke *et al.*, 2014a,b). In trials aiming to assess differential effectiveness, the number of oocytes retrieved for instance, represents an obvious surrogate outcome that may reach significance, whereas live birth rate does not.

Selective outcome reporting overlaps with publication bias, i.e. bias that arises when studies are published or not, depending on their results. One of the drivers behind publication bias is that studies that report positive or significant results are more likely to be published and this, in turn, may lead to selective outcome reporting of the significant outcomes in favour of the non-significant ones (Dwan *et al.*, 2013).

Selective outcome reporting may be related to sponsorship, since sponsorship by industry is known to negatively influence valid interpretation of trial results by more often reporting on outcomes favourable to the sponsor's products (Vandenbroucke *et al.*, 2000; Lundh *et al.*, 2012; Flacco *et al.*, 2015). It is, at present, unknown whether RCTs on IVF and ICSI are subject to selective outcome reporting and whether this is related to sponsorship. The aim of this study was therefore to evaluate this by performing a systematic review of RCTs in couples undergoing IVF and ICSI.

Materials and Methods

Literature search

A medical information specialist (JL) performed a systematic search in OVID MEDLINE, OVID EMBASE, the Cochrane Central Register of Controlled Trials (CENTRAL) and the publisher subset of PubMed to identify RCTs on IVF and/or ICSI in humans. The search strategy consisted of MeSH terms and text words for the concepts IVF and ICSI combined with a methodological search filter for RCTs that was adapted from the Cochrane MEDLINE RCT-filter (Glanville *et al.*, 2006). We included human RCTs, published in English, that compared (cost) effectiveness of interventions in IVF and ICSI and were published between January 2009 until March 2016. At first, we chose a five-year interval from 2009 to 2014, that we updated before preparing this manuscript. The reason to choose a recent time period is that trial registries and guidelines on trial reporting have become essential developments in the last decade. We excluded duplicates and trials with study designs other than RCTs. We excluded studies if the subject was not IVF or ICSI, if the record was only an abstract submitted for a conference without a published RCT, if the study turned out not to be a RCT, if pregnancy was not an outcome, if it was not written in English and if we could not retrieve the full text.

Data extraction

Two researchers (MB and IS) independently extracted the data from trials published between January 2009 and March 2016. First, we extracted some general characteristics of the included trials: journal category (high impact factor versus medium to low impact factor with a cut-off level of 3.5), sample size (categorical) and power calculation (yes or no). For the analysis of sample sizes, we included trials that randomized women, and excluded trials that randomized oocytes or embryos.

Second, we classified the primary outcome measures as clinical or surrogate outcome measures. We defined a clinical outcome as 'an outcome of major clinical importance', i.e. clinical/ongoing pregnancy or live birth and a surrogate outcome as an outcome measure that 'substitutes for a clinical event of true importance', e.g. biochemical pregnancy or number of oocytes retrieved (Grimes *et al.*, 2010). We consulted trial registries to rule out a discrepancy between planned and reported primary outcome measures. The trial registries we consulted were clinicaltrials.gov, ISRCTN, NTR, IRCT, KHUCTR, Eudra, DRKS and ChiCTR.

Third, because of the overlap between selective outcome reporting and publication bias, we categorized the results of the trials in 'positive' and 'negative'. We defined positive results as those that were statistically significant in favour of the tested intervention, negative results as those that were statistically significant in favour of the standard intervention and non-significant results as those where no evidence of a difference was found between the interventions.

Fourth, we recorded the presence of sponsorship statements. If information on sponsorship was provided, the trials were categorized into commercial sponsorship, public sponsorship and no sponsorship. We defined commercial sponsorship as sponsorship by pharmaceutical companies and companies that develop medical devices, and public sponsorship as sponsorship by the government or medical research councils and charities. Trials were categorized as being not sponsored if this was explicitly stated.

Statistical analysis

We described the general characteristics of the trials, the presence or absence of sponsorship and the direction of the trial results as counts and percentages. We calculated odds ratios and confidence intervals using IBM® SPSS® Statistics version 23, to test differences between the studied trial characteristics related to the sponsorship of the trial. To test if direction of the trial results was related to the sponsorship of the trial, we used the Pearson-Chi-square test, IBM® SPSS® Statistics version 23.

Since the majority of the papers were secondarily excluded because of the non-reporting of sponsorship, we added an secondary analysis. We ran our analysis again, comparing the trials that did not report sponsorship with the trials that did report sponsorship. We also ran a scenario analysis

in which we presumed that the trials not reporting sponsorship were publicly sponsored.

Results

Our search identified 1879 records. The flowchart is shown in Fig. 1. After screening on subject, study design, language and duplicates in title and abstract, we selected 541 articles for full text reading, of which we excluded 126 articles because of various reasons specified in the flowchart. We included 415 articles in our database of which 72 trials (17%) reported commercial sponsorship, 79 trials (19%) reported public sponsorship, 29 trials (7%) reported no sponsorship and 235 trials (57%) did not report on sponsorship. The 235 trials that did not report on sponsorship trials were excluded from further analysis.

The characteristics of the included trials are summarized in Table I. Commercially sponsored trials were significantly more often published in high impact factor journals compared to trials with public or no sponsorship. Commercially sponsored trials significantly more often had a sample size of over 500 participants compared to trials with public sponsorship. Power calculations were significantly more often performed in commercially sponsored trials compared to trials with public or no sponsorship.

Data on primary outcome measures are summarized in Table II. The majority of the trials had a surrogate outcome as the primary outcome measure. There was no evidence of a difference between commercially, public and not sponsored trials.

Data on discrepancies between registered and reported primary outcome measures are summarized in Table III. Commercially sponsored trials were significantly more often registered in a trial registry compared to trials with public or no sponsorship. Of the 114 registered trials, 21 trials (18%) provided different primary outcomes in their manuscript compared to the registry. Of these 21 trials, 12 registered a clinical primary outcome while a surrogate primary outcome was reported in the manuscript. The other nine trials registered a surrogate primary outcome, while a clinical primary outcome was reported in the manuscript. There was no difference between commercially, public and not sponsored trials.

Data on direction of trial results are summarized in Table IV. A minority of the trials reported negative results which were statistically significant in favour of the standard intervention. There was no association between the direction of trial results and sponsorship.

The additional analyses are summarized in Supplementary Tables S1–IV. Trials that did not report on sponsorship were significantly less often published in journals with a high impact factor and significantly less often performed power calculations compared to trials that did report on sponsorship. Also, the trials that did not report on sponsorship were significantly less often registered in a trial

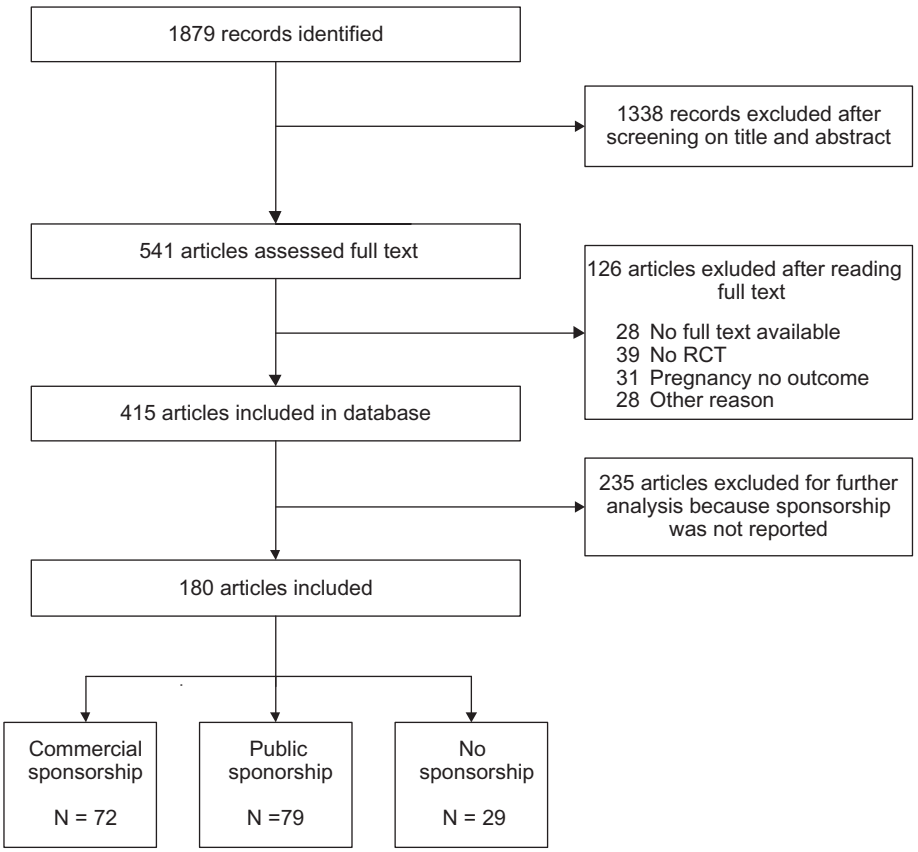

Figure 1 Flowchart of included trials.

Table I Baseline characteristics of included trials.

Sponsorship	Commercial (*n* = 72)	Public (*n* = 79)	No sponsorship (*n* = 29)	OR (95% CI)	
				Comm vs. public	Comm vs. no sponsor
Journal of publication (% high impact factor)	56 (78%)	31 (39%)	11 (38%)	5.42 (2.50–11.86)	5.73 (2.05–16.32)
Sample size, median (minimum–maximum)	160 (13–2042)	134 (30–1446)	140 (20–604)		
Categorical					
1–100 pts*	22 (31%)**	19 (25%)**	9 (32%)**	1.30 (0.60–2.85)	0.92 (0.33–2.63)
100–300 pts	28 (40%)	46 (61%)	15 (54%)	0.40 (0.20–0.82)	0.55 (0.21–1.45)
300–500 pts	6 (8%)	7 (10%)	2 (7%)	0.88 (0.25–3.13)	1.18 (0.20–9.11)
>500 pts	15 (21%)	3 (4%)	2 (7%)	6.55 (1.66–30.07)	3.55 (0.69–24.28)
Power calculation performed (%)	55 (76%)	45 (57%)	13 (45%)	2.44 (1.15–5.26)	3.98 (1.46–11.01)

*pts = participants.
**Studies in which oocytes or embryos are randomized are not included in this analysis.

Table II Primary outcome measures and the association with sponsorship.

Sponsorship	Commercial (*n* = 72)	Public (*n* = 74)*	No sponsorship (*n* = 27)*	OR (95% CI)	
				Comm vs. public	Comm vs. no sponsor
Primary outcome measure					
Clinical	26 (36%)	29 (39%)	11 (41%)		
Live birth	3	6	3		
Ongoing pregnancy	11	3	3		
Clinical pregnancy	12	20	5		
Surrogate	46 (64%)	45 (61%)	16 (59%)	1.14 (0.55–2.36)	1.22 (0.45–3.30)
Fertilization	4	4	2		
Number of oocytes	13	8	2		
Biochemical pregnancy	0	3	0		
Other	29	30	12		

*Not all studies noted a primary outcome.

Table III Discrepancies between registered and reported primary outcome measure.

Sponsorship	Commercial (*n* = 72)	Public (*n* = 74)	No sponsorship (*n* = 27)	OR (95% CI)	
				Comm vs. public	Comm vs. no sponsor
Registered in trial registry	59 (82%)	43 (58%)	12 (44%)	3.27 (1.44–7.51)	5.68 (1.95–16.86)
No discrepancy	50	33	10		
Discrepancy	9	10	2	0.59 (0.20–1.80)	0.90 (0.14–7.08)

registry. We found no other differences between trials that did not report on sponsorship and trials that did report on sponsorship. When presuming that all trials that did not report sponsorship were publicly sponsored, we found no outcomes that differed from our primary analysis

Discussion

The majority of the analysed trials in this systematic review reported on surrogate primary outcome measures, while a considerable proportion reported primary outcomes different from that listed in the

Table IV Trial results and the association with sponsorship.

Sponsorship	Commercial ($n = 72$)	Public ($n = 79$)	No sponsorship ($n = 29$)	*P*-value
Trial result				0.19
Positive	34 (47%)	41 (52%)	17 (58%)	
Negative	3 (4%)	4 (5%)	4 (14%)	
Neutral	35 (49%)	34 (43%)	8 (28%)	

trial registry and some trials did not mention a primary outcome measure at all. A minority of the trials reported on negative results. We did not find an association between selective outcome reporting and sponsorship. We did find differences between some general characteristics of the trials, related to sponsorship. The trials that reported commercial sponsorship were more often published in high impact factor journals, more often had large sample sizes, more often performed power calculations and were more often registered in trial registries compared to trials with public or no sponsorship.

Our study has certain limitations. We restricted our sample to English-language RCTs and did not include cohort and case-control studies, while selective outcome reporting might also play a role in trials with these study designs. In our study, we focused on reporting bias only, although selection bias, performance bias, detection bias and attrition bias might also be influenced by sponsorship. Because of our small dataset, we made no geographic stratification, although there might be differences in the general regulations for the registration of trials. The majority of the trials did not report on sponsorship and we excluded these trials, which created a risk on sampling bias and a smaller sample size of trials with useful information. Our analysis might therefore be underpowered. However in our additional scenario analysis, we presumed that all trials that did not report on sponsorship were publicly sponsored. We think this is a more realistic scenario than the other option of all such trials being commercially sponsored, since in our primary analysis the majority of the trials were publicly sponsored. Furthermore, we did not found any striking differences after running the additional analysis.

From 2009 to thepresent, less than half of the published RCTs on IVF and ICSI reported sponsorship. This is not in line with the CONSORT statement (Schulz *et al.*, 2010). We did find a positive trend in the reporting of sponsorship, from 26% in 2009 to 49% in 2015. To be able to perform high quality analyses on this topic in the future, we feel that all submitted manuscripts should report on sponsorship and that this should be taken into account in the peer review process. We considered contacting the authors of the manuscripts that did not report on sponsorship, but we rejected this idea for two reasons. First, we wanted to report on published data and not on unpublished data and second, we felt that, asking the authors for a statement on sponsorship would not necessarily result in a trustworthy insight in sponsorship, or in other words, in asymmetrical data, which would be unfortunate, especially on this topic.

The fact that only 79 of 180 trials (44%) report public sponsorship is an issue that needs attention. The trials we selected were designed to improve clinical outcome in couples with subfertility and therefore have the potential to guide towards better clinical decision making. If reproductive health is seen as an important collective heath issue, this low public contribution is worrisome.

Our analysis of the trials that did report on sponsorship showed that commercially sponsored trials are more often published in journals with a high impact factor. This is not surprising, since they report on larger sample sizes and more often perform sample size calculations that result in better underpinned conclusions of greater impact. Next to the origin of the sponsorship, these differences may also be due to the study budget, but information on budget of the studies was not available.

Still, conclusions from all evaluated trials, regardless of sponsorship, were mainly based on surrogate primary outcome measures, which is worrisome since clinical decisions cannot be made based on trials that report on surrogate primary outcome measures. Surrogate outcome measures can be of interest to explain the main results of a clinical trial and thereby provide insight into the underlying biological mechanisms. However the most relevant primary outcome measure for clinical trials in reproductive medicine, in our opinion, is ongoing pregnancy (Braakhekke *et al.*, 2014a,b). Discrepancies between the registered and reported primary outcome measure such as that found in our analysis, count for selective outcome reporting for two reasons. As primary outcomes are used for calculation of the sample size, this mechanism is no longer functional if predefined outcomes are subsequently changed (Chan *et al.*, 2004). Also, pre-defining primary outcome measures is necessary to prevent the researchers from cherry picking significant results and presenting these as the main findings of the study (Andrade, 2015). Secondary findings may be of relevance and it is not wrong to report these findings and to discuss them, if only it is clearly stated that these findings are secondary, thus they do not correspond to the original power calculation and should be considered as hypothesis generating rather than for rejecting the null hypothesis. Discrepancies between the registered and reported primary outcome measure such as we found leads to false, or at least misleading, trial results.

A recent review has revealed similar discrepancies. The review included studies which compared published trial outcomes to trial outcomes documented in a publicly accessible clinical trials registry. The proportion of trials with an identified discrepancy between the registered and published primary outcome was 31% (Jones *et al.*, 2015).

In our analysis, only 11 trials (6%) reported negative results. This is possibly caused by publication bias, i.e. non-publication due to lack of submission or rejection of study reports, but can also be a part of selective outcome reporting (Dwan *et al.*, 2013). Commercially sponsored trials are especially at risk for publication bias since the company that owns the research data can decide to publish the data or not. The only way to differentiate between publication bias and selective outcome reporting is to get access to the raw research data. However, participant-level data from health related studies are rarely made available to external researchers (Alsheikh-Ali *et al.*, 2011).

In conclusion, we found inconsistencies in the choice and reporting of primary outcome measures, indicating selective outcome reporting in IVF and ICSI trials. We did not find an association between the existence of selective outcome reporting and sponsorship. Recently the COMPare team has monitored all trials published in the top five medical journals. This team analyses each trial for outcome switching, by comparing the protocol with the trial report. If outcome switching is

found, journals are contacted to correct the record (Goldacre et al., 2016). This might be a helpful step in reducing selective outcome reporting.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Authors' roles

B.W.M., MB and FvdV conceived the idea for this study. JL performed the literature search. MB and IS extracted the data. MB wrote the manuscript. All authors helped to prepare the final manuscript.

Funding

No funding was received.

Conflict of interest

None declared.

References

Alsheikh-Ali AA, Qureshi W, Al-Mallah MH, Ioannidis JPA. Public availibility of published research data in high-impact journals. *PLoS One* 2011;**6**:e24357.

Andrade C. The primary outcome measure and its importance in clinical trials. *J Clin Psychiatry* 2015;**76**:1320–1323.

Braakhekke M, Kamphuis EI, van Rumste MM, Mol F, van der Veen F, Mol BW. How are neonatal and maternal outcomes reported in randomised controlled trials (RCTs) in reproductive medicine? *Hum Rep* 2014a;**29**:211–217.

Braakhekke M, Kamphuis EI, Dancet EA, Mol F, van der Veen F, Mol BW. Ongoing pregnancy qualifies best as the primary outcome measure of choice in trials in reproductive medicine: an opinion paper. *Fertil Steril* 2014b;**101**:1203–1204.

Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials comparison of protocols to published articles. *JAMA* 2004;**291**:2457–2465.

Dwan K, Gamble C, Williamson PR, Kirkham JJ, the Reporting Bias Group. Systematic review of the empirical evidence of study publication bias and outcome reporting bias—an updated review. *PLoS One* 2013;**8**:e66844.

Flacco ME, Manzoli L, Boccia S, Capasso L, Aleksovska K, Rosso A, Scaioli G, Vito de C, Siliquini R, Villari P *et al*. Head-to-head randomized controlled trials are mostly industry sponsored and almost always favor the industry sponsor. *J Clin Epidemiol* 2015;**68**:811–820.

Glanville JM, Lefebvre C, Miles JN, Camosso-Stefinovic J. How to identify randomized controlled trials in MEDLINE: ten years on. *J Med Libr Assoc* 2006;**94**:130–136.

Gluud LL. Bias in clinical intervention research. *Am J Epidemiol* 2006;**163**:493–501.

Goldacre B, Drysdale H, Powell-Smith A, Dale A, Milosevic I, Slade E, Hartley P, Marston C, Mahtani K, Heneghan C. The COMPare Trials Project. www.COMParetrials.org, 2016.

Grimes DA, Schulz KF, Raymond EG. Surrogate end points in women's health research: science, protoscience, and pseudoscience. *Fertil Steril* 2010;**93**:1731–1734. American Society for Reproductive Medicine.

Grimes D a, Schulz KF. Surrogate end points in clinical research: hazardous to your health. *Obstet Gynecol* 2005;**105**:1114–1118.

Guyatt GH, Jaeschke RZ, Naylor CD, Wilson MC, Richardson WS, Page P. Users' guides to the medical literature XXV. Evidence-based medicine: principles for applying the users' guides to patient care. *JAMA* 2000;**284**:1290–1296.

Higgins JPT, Green S (eds). Cochrane Handbook for Systematic Reviews of Interventions *Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration, 2011.

Jones CW, Keil LH, Holland WC, Caughey MC, Platt-Mills TF. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Med* 2015;**13**:282.

Lassere MN. The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multidimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. *Stat Methods Med Res* 2008;**17**:303–340.

Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev* 2012;**12**:MR000033. doi: 10.1002/14651858.MR000033.pub2.

Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Obstet Gynecol* 2010;**115**:1063–1070.

Vandenbroucke JP, Helmerhorst FM, Rosendaal FR. Competing interests and controversy about third generation oral contraceptives. *BMJ* 2000;**320**:381–382.