human reproduction update

The effects of acupuncture on rates of clinical pregnancy among women undergoing *in vitro* fertilization: a systematic review and meta-analysis

Eric Manheimer[1],*, Daniëlle van der Windt[2], Ke Cheng[3], Kristen Stafford[4], Jianping Liu[5], Jayne Tierney[6], Lixing Lao[1], Brian M. Berman[1], Patricia Langenberg[4], and Lex M. Bouter[7]

[1]Center for Integrative Medicine, University of Maryland School of Medicine, Baltimore, MD 21201, USA [2]Institute for Health and Primary Care Sciences, Keele University, Staffordshire ST5 5BG, UK [3]College of Acupuncture-Moxibustion and Tuina, Shanghai University of Traditional Chinese Medicine, Shanghai Research Center for Acupuncture and Meridians, Shanghai 201203, China [4]Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD 21201, USA [5]Center for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing 100029, China [6]UK MRC Clinical Trials Unit, Meta-Analysis Group, London WC2B 6NH, UK [7]Executive Board of the VU University, 1081 HV Amsterdam, The Netherlands

Correspondence address. Tel: +1-410-706-6169; E-mail: emanheimer@compmed.umm.edu

Submitted on January 3, 2013; resubmitted on April 18, 2013; accepted on April 30, 2013

TABLE OF CONTENTS

BACKGROUND: Recent systematic reviews of adjuvant acupuncture for IVF have pooled heterogeneous trials, without examining variables that might explain the heterogeneity. The aims of our meta-analysis were to quantify the overall pooled effects of adjuvant acupuncture on IVF clinical pregnancy success rates, and evaluate whether study design-, treatment- and population-related factors influence effect estimates.

METHODS: We included randomized controlled trials that compared needle acupuncture administered within 1 day of embryo transfer, versus sham acupuncture or no adjuvant treatment. Our primary outcome was clinical pregnancy rates. We obtained from all investigators additional methodological details and outcome data not included in their original publications. We analysed sham-controlled and no adjuvant treatment-controlled trials separately, but since there were no large or significant differences between these two subsets, we pooled all trials for subgroup analyses. We prespecified 11 subgroup variables (5 clinical and 6 methodological) to investigate sources of heterogeneity, using single covariate meta-regressions.

RESULTS: Sixteen trials (4021 participants) were included in the meta-analyses. There was no statistically significant difference between acupuncture and controls when combining all trials [risk ratio (RR) 1.12, 95% confidence interval (CI), $0.96 - 1.31$; $I^2 = 68\%$; 16 trials; 4021 participants], or when restricting to sham-controlled (RR 1.02, $0.83 - 1.26$; $I^2 = 66\%$; 7 trials; 2044 participants) or no adjuvant treatment-controlled trials (RR 1.22, $0.97 - 1.52$; $I^2 = 67\%$; 9 trials; 1977 participants). The type of control used did not significantly explain the statistical heterogeneity (interaction $P = 0.27$). Baseline pregnancy rate, measured as the observed rate of clinical pregnancy in the control group of each trial, was a statistically significant effect modifier (interaction $P < 0.001$), and this covariate explained most of the heterogeneity of the effects of adjuvant acupuncture across all trials (adjusted $R^2 = 93\%$; I^2 residual $= 9\%$). Trials with lower control group rates of clinical pregnancy showed larger effects of adjuvant acupuncture (RR 1.53, $1.28 - 1.84$; 7 trials; 1732 participants) than trials with higher control group rates of clinical pregnancy (RR 0.90, $0.80 - 1.01$; 9 trials; 2289 participants). The asymmetric funnel plot showed a tendency for the intervention effects to be more beneficial in smaller trials.

CONCLUSIONS: We found no pooled benefit of adjuvant acupuncture for IVF. The subgroup finding of a benefit in trials with lower, but not higher, baseline pregnancy rates (the only statistically significant subgroup finding in our earlier review) has been confirmed in this update, and was not explained by any confounding variables evaluated. However, this baseline pregnancy rate subgroup finding among published trials requires further confirmation and exploration in additional studies because of the multiple subgroup tests conducted, the risk of unidentified confounders, the multiple different factors that determine baseline rates, and the possibility of publication bias.

Key words: acupuncture / assisted conception / complementary medicine / *in vitro* fertilization / systematic review

Introduction

Some 10–15% of couples have difficulty conceiving at some point in their reproductive lives, and seek specialist fertility care (Evers, 2002). One of the most commonly used treatment options is IVF and related expansions of it [e.g. ICSI (ESHRE Capri Workshop Group, 2007)]. For the year 2003, it was estimated that 932 000 IVF and related cycles were performed worldwide, resulting in an estimated 232 000 babies being born (Nygren et al., 2011). The USA has among the highest IVF success rates worldwide, which is likely due to a combination of factors, including few restrictions on the number of embryos that can be transferred and the routine use of embryo selection (Abdalla, 2010; Nygren et al., 2011). European countries generally have lower IVF success rates because they are increasingly moving towards single embryo transfers (Gerris, 2005; Maheshwari et al., 2011) which carry lower health risks (Pinborg, 2005); the rates of multiple births and the associated health risks for mother and baby in Europe are approximately half of those in the USA (Nygren et al., 2011).

Despite many recent technological advances, average success rates with IVF remain low, with only 30% of treatment cycles resulting in a live birth in the USA (Centers for Disease Control and Prevention, 2011). Consequently, there is a need to investigate new laboratory techniques and drug therapies to improve the success rates of IVF, by means other than increasing the number of embryos transferred. However, progress in developing such safe and effective therapies has been limited (Harper et al., 2012), driving patients to consider IVF adjuvant complementary and alternative medical (CAM) therapies (Smith et al., 2010), many of which are unproved and inadequately investigated. Acupuncture is the most commonly used adjuvant CAM fertility treatment among couples seeking fertility care in US fertility clinics (Smith et al., 2010; Domar et al., 2012). 'Fertility problems' is the second most common health condition, following pain-related conditions, for which people seek acupuncture treatment in the UK (Hopton et al., 2012). A review of the effects of adjuvant acupuncture on IVF is warranted because acupuncture has been found to be relatively safe among general patient populations (Witt et al., 2011) and also among women at various stages of pregnancy (Smith et al., 2002; Kvorning et al., 2004; Elden et al., 2005), it is low cost, and, if effective in increasing IVF success rates, could potentially reduce the need for an additional high cost IVF cycle. Furthermore, qualitative research suggests that adjuvant acupuncture may help IVF patients deal with the psychological and emotional issues that accompany both subfertility and IVF treatment (de Lacey et al., 2009; Kovarova et al., 2010; Smith et al., 2011; Isoyama et al., 2012).

We updated our previous systematic review and meta-analysis (Manheimer et al., 2008) with nine new trials including 2672 new randomized participants, and carried out an *a priori* defined set of subgroup analyses to evaluate whether estimates of the effects of adjuvant acupuncture on IVF success rates is influenced by study design-, treatment- and population-related factors.

Methods

The systematic review protocol was registered in the PROSPERO database of prospectively registered systematic reviews (identification number: CRD42011001309) in May 2011 (Manheimer et al., 2011). The systematic review follows PRISMA reporting guidelines (Moher et al., 2009).

Identification of studies

We searched Medline (OVID), Embase (Elsevier Sciences), CENTRAL (2012, Issue 4) (OVID), and the Chinese databases Sino-Med (previously called the Chinese Biomedical Database), Chinese National Knowledge

Infrastructure, and VIP Database for Chinese Technical Periodicals, from inception to April 2012 (the MEDLINE search strategy is available from the author). We also searched the proceedings of the following three major annual conferences on assisted reproductive technology for 2001–2012: American Society for Reproductive Medicine; European Society of Human Reproduction and Embryology; and Pacific Coast Reproductive Society. We also searched for previous systematic reviews on this topic, and reviewed their reference lists (Cheong *et al.*, 2008, 2010; El-Toukhy *et al.*, 2008; Manheimer *et al.*, 2008; Ng *et al.*, 2008; El-Toukhy and Khalaf, 2009; Sunkara *et al.*, 2009; Qu *et al.*, 2012; Zheng *et al.*, 2012a, b). In addition, we sent the provisional list of included and excluded trials to experts in the field (see Acknowledgements), and asked if they were aware of any potentially eligible trials, published or unpublished, that were not on our list. To identify trials that may be relevant for future updates of this review, we also searched the following databases of ongoing trials: the US National Institutes of Health's clinicaltrials.gov, the World Health Organization's International Clinical Trials Registry Platform, and controlled-trials.com.

Selection criteria, data extraction and quality assessment

We sought randomized controlled trials (RCTs) that compared acupuncture with sham acupuncture or no adjuvant treatment. Because we evaluated acupuncture as a complement to embryo transfer, we considered only RCTs in which acupuncture was administered within 1 day of the procedure, with the objective of improving IVF success rates. Trials that included ICSI as part of the IVF procedure were eligible.

We excluded RCTs which evaluated electrical acupuncture as an alternative to conventional anaesthesia for the surgical procedure of removing the oocytes from the woman's ovaries. We considered these two sets of RCTs to be fundamentally different, in terms of the aim of the treatment with acupuncture, the timing of acupuncture administration (i.e. during oocyte retrieval versus embryo transfer) and the acupuncture protocol used. That is, when acupuncture is used during oocyte retrieval, the primary purpose is to relieve pain, and correspondingly, electrical stimulation of the needles is always used (Ulett *et al.*, 1998) and the points are selected to effect pain reduction. In contrast, when acupuncture is used to accompany embryo transfer, the primary purpose is to assist conception, electrical stimulation of the needles may or may not be used, and the points are selected to improve blood flow to the uterus to make it more receptive to the embryo (Stener-Victorin *et al.*, 1996). In addition, two systematic reviews (Kwan *et al.*, 2013; Stener-Victorin, 2005) have already evaluated electroacupuncture as an alternative to anaesthesia during oocyte retrieval, and pooled analyses found no statistically significant differences in later pregnancy rates between the electroacupuncture and conventional anaesthesia groups.

We included only trials in which acupuncture involved the insertion of needles into traditional meridian points. The needles could be inserted into tender points in addition to the traditional meridian points, and the needles could also be electrically stimulated.

We included RCTs in any language, published as either full articles or abstracts. We also included any unpublished trials that we identified, but only if we were able to obtain sufficient and reliable information on their methods and outcomes.

For trials to be eligible, we had to be able to extract data on at least one of the following outcomes, as recommended (Daya, 2003; Griesinger *et al.*, 2004; Arce *et al.*, 2005): clinical pregnancy [i.e. presence of at least one gestational sac or fetal heartbeat, confirmed by transvaginal ultrasound (Zegers-Hochschild *et al.*, 2006)], ongoing pregnancy (i.e. pregnancy beyond 12 weeks of gestation, as confirmed by fetal heart activity on ultrasound) or live birth.

Two authors independently selected articles and extracted data, with disagreements resolved by discussion. We extracted data pertaining to the quality of the methods, participants, interventions and outcomes. We contacted corresponding authors with specific questions related to the design and outcomes of their trials and asked them to review the information we extracted from their trials and clarify any ambiguities.

We evaluated the methodological quality of the trials using the Cochrane risk of bias tool (Higgins and Altman, 2011), supplemented with an additional item that assessed whether there was a co-intervention that was applied unequally across treatment groups. We added 'unequal co-intervention' to the Cochrane tool because this item can indicate an important source of bias and has been included in other quality assessment tools (van Tulder *et al.*, 2003), including an assessment tool used in subfertility meta-analyses (Cochrane Menstrual Disorders and Subfertility Group, 2010).

Data synthesis and analysis

The pooled risk ratio (RR) of achieving a clinical pregnancy for women in the acupuncture group compared with women in the control group was the primary outcome measure. We prespecified clinical pregnancy instead of live birth as our primary outcome because it would allow for more RCTs and more data to contribute to our primary analyses, as only 22% of subfertility RCTs that report the clinical pregnancy outcome data also report the live birth outcome data (Clarke *et al.*, 2010). Therefore, including clinical pregnancy as the primary outcome will avoid the risk of including in the primary analysis only a subset of the RCTs (i.e. only those RCTs that report live births), which may not be representative of all RCTs included in the review (Manheimer *et al.*, 2008). In addition, clinical pregnancy is a practical and clinically relevant surrogate for live birth because ∼84% of clinical pregnancies achieved through IVF would be expected to result in a live birth, based on data from IVF RCTs (Clarke *et al.*, 2010). Also, the effect estimates on the clinical pregnancy outcome have been shown to be nearly identical to the effect estimates on the live birth outcome for IVF RCTs that report both outcomes [i.e. ratio of odds ratios: 0.99 (95% confidence interval (CI), 0.87–1.13)] (Clarke *et al.*, 2010). Finally, the clinical pregnancy outcome may allow for easier control of potentially confounding factors, which might be introduced after the first trimester, once the women are no longer under the care of the IVF clinic. That is, collecting reliable information about live births after patients have left the IVF clinic might be difficult because patients, who may travel long distances to reach the clinic, might be less accessible to the clinic staff after the clinical pregnancy is achieved (Daya, 2003; Arce *et al.*, 2005; Clarke *et al.*, 2010). Ongoing pregnancy, live birth and spontaneous abortion rates were analysed as secondary outcomes.

For pooled data, summary test statistics were calculated using the DerSimonian and Laird model (DerSimonian and Laird, 1986; Deeks and Higgins, 2010), which is the random effects model used in the RevMan software, version 5.1 (The Nordic Cochrane Centre. The Cochrane Collaboration, 2011). This model estimates the average treatment effect by incorporating heterogeneity among clinically diverse trials with different, but related, treatment effects. When heterogeneity exists, the model assigns smaller studies more weight than they would receive in a fixed effects model (Deeks *et al.*, 2011). We used the random effects model because of the expected heterogeneity of the studies' acupuncture protocols and settings. We evaluated heterogeneity using both the I^2 statistic, which indicates the proportion of variability across trials not explained by sampling variation alone, and the P-value of the χ^2 test of heterogeneity (Deeks *et al.*, 2011). Although interpreting the importance of inconsistency depends on other factors in addition to the I^2 values (e.g. P-value from χ^2 test, magnitude and direction of effects), the Cochrane Handbook suggests the following rough guide to interpreting I^2 values: '0–40% might not be important; 30–60% may represent moderate heterogeneity; 50–90% may represent substantial heterogeneity; and 75–100% may represent considerable heterogeneity' (Deeks *et al.*, 2011).

We included in the meta-analyses all randomized women who began the IVF process including those who did not complete the treatment (i.e. had no embryo transfer), even if these participants were excluded in the author's analysis of the trial. Although inclusion of women without an embryo transfer will tend to underestimate the effect of acupuncture (Daya, 2003), it is the more conservative and appropriate analytical approach (Daya, 2003; Griesinger et al., 2004; Deeks et al., 2006; Dias et al., 2006) because it preserves the groups created by the randomization and reduces the chance of a type I error (Daya, 2003). The only randomized participants we did not include were those for whom the clinical pregnancy or live birth outcome data were missing.

We analysed the sham-controlled and no adjuvant treatment-controlled trials separately, but if there was no large or significant difference in pooled effect estimates between these two subgroups of trials, we pooled all available trials.

If at least 10 trials were available for a meta-analysis, we assessed for the likelihood of publication bias by constructing funnel plots (Sterne et al., 2011).

Subgroup analyses

We conducted subgroup analyses on five clinical characteristics that might influence the effect of adjuvant acupuncture on clinical pregnancy success rates: (i) two acupuncture sessions or more than two; (ii) selection of meridian acupuncture points the same as the points selected in the first published trial (Paulus et al., 2002) that evaluated acupuncture as an adjuvant to embryo transfer, and which showed a large effect, or a modified version of this trial's acupuncture point selection protocol; (iii) control group clinical pregnancy rate (as an estimate of the baseline clinical pregnancy rate) dichotomized as higher [32% or greater, which is the European average of pregnancy rate per embryo transfer (de Mouzon et al., 2012)] or lower; the control group clinical pregnancy rate was also analysed as a continuous variable to test whether the relation was linear and consistent with the findings of the categorical analysis; (iv) explanatory trials conducted to test the effects of adjuvant acupuncture under controlled conditions in which the acupuncture was administered onsite at the IVF clinic or pragmatic trials conducted to test the effects of adjuvant acupuncture delivered off-site, which might better approximate every day, 'real life' conditions since most IVF clinics do not have onsite acupuncturists (Arce et al., 2005); and (v) trials that involved a treating acupuncturist who was judged as adequately experienced or not adequately experienced, with such judgments made by acupuncturist assessors who were blinded to the identities and results of the trials.

In addition, to assess whether the effects varied with the risk of bias domains of the trials, we also conducted subgroup analyses (Deeks et al., 2011) on the following six 'risk of bias' domains: random allocation sequence generation; concealment of allocation of randomization sequence; blinding of patients (i.e. use of sham control); blinding of embryo transfer physicians; incomplete outcome data; and unequal co-intervention.

For each subgroup analysis, we performed a single covariate weighted random effects meta-regression (Thompson and Higgins, 2002; Knapp and Hartung, 2003; Higgins and Thompson, 2004) in Stata version 11 (StataCorp) (Harbord and Higgins, 2008) to investigate whether differences in effects of adjuvant acupuncture between the covariate's two subgroups were statistically significant. In interpreting the importance of subgroup effects, we also considered the difference in the magnitude or direction of effect between the two subgroups as well as whether or not the CIs of the subgroups overlapped (Deeks et al., 2011). For each single covariate meta-regression subgroup analysis, we calculated the P-value of the test for interaction; the percentage of the between-study variance explained by the covariate (adjusted R^2); and the percentage of the residual variation that is attributable to between-study heterogeneity (I^2_{res}). A random effects meta-regression model was used rather than a fixed effects model or univariate subgroup

testing because the random effects meta-regression model allows for potential residual heterogeneity and it is therefore a more conservative analytical approach. We did not attempt to develop a multivariate meta-regression model, primarily because of the relatively small number of studies relative to the number of subgroup variables prespecified, but also because our objective was to identify modifiers of the effects of adjuvant acupuncture on IVF, rather than to build an optimal 'prediction' model. A major limitation of subgroup analyses is the potential bias by confounding (Thompson and Higgins, 2002). Therefore, we also examined whether the apparent effect modification due to a specific subgroup variable might be explained by an association between that subgroup variable of interest and a potentially confounding subgroup variable, by including both variables in a meta-regression model.

Sensitivity analyses

For our primary analyses, we excluded participants with unrecorded clinical pregnancy data; however, for sensitivity analyses, we imputed the rate of clinical pregnancies in participants with unrecorded outcomes using available case rates: (i) from the corresponding trial group that these participants were assigned to (i.e. either acupuncture or control); and (ii) from the corresponding trial's acupuncture plus control group combined. We also conducted a third sensitivity analysis which assumed that no randomized participants with an unrecorded outcome achieved a clinical pregnancy.

Results

Supplementary data, Fig. S1 shows the study selection process. The list of studies that were excluded with the reasons for exclusion and the list of characteristics of the ongoing trials are available from the author.

Sixteen RCTs with a total of 4038 participants met the inclusion criteria (Table I; full details of the trials are available from the author). All trials were published in English since 2002, and were conducted in seven different countries. Ten were published as full reports (Paulus et al., 2002; Dieterle et al., 2006; Smith et al., 2006; Westergaard et al., 2006; Domar et al., 2009; So et al., 2009, 2010; Andersen et al., 2010; Madaschi et al., 2010; Moy et al., 2011) and six were published as abstracts (Paulus et al., 2003; Craig et al., 2007; Fratterelli et al., 2008; Arnoldi et al., 2010; Omodei et al., 2010; Feliciani et al., 2011). We obtained unpublished methodological information for all trials, and unpublished live birth outcome data for seven trials (Paulus et al., 2002, 2003; Dieterle et al., 2006; Westergaard et al., 2006; Craig et al., 2007; Arnoldi et al., 2010; Omodei et al., 2010).

Trial characteristics

The only differences in trial eligibility criteria were that: two German trials (Paulus et al., 2002, 2003) included only women with good-quality embryos whereas the other trials included women with embryos of varying quality; one trial (Arnoldi et al., 2010) restricted eligibility to women with an unfavourable reproductive prognosis; one trial (So et al., 2010) restricted eligibility to frozen−thawed embryo transfer cycles while the others used fresh embryos; and one trial (Madaschi et al., 2010) used ICSI for all participants, whereas all other trials reported use of ICSI for only some participants.

The timing of the acupuncture sessions relative to embryo transfer differed somewhat among trials (Table I). In all trials, however, women received acupuncture on the day of embryo transfer, before and/or after the embryo transfer procedure. In all but one trial (Craig et al., 2007), the acupuncture was administered directly on-site of the IVF clinic. The number of acupuncture treatments ranged from one to

Table I Characteristics of included trials.

Study, Country (reference)	n^a	Acupuncture	Control type[b]
Andersen, Denmark (Andersen *et al.*, 2010)	635	Two sessions: (1) 30 m before ET; (2) immediately after ET. [To maintain blinding, needles not manipulated to obtain de qi needling sensation]; (*n* = 314.)	Non-penetrating blunt needles, placed on the real acupoints [and not manipulated to achieve de qi sensation]; (*n* = 321)
Arnoldi, Italy (Arnoldi *et al.*, 2010)	204	Three sessions: (1) d 5 of ovarian stimulation; (2) 30 m before ET; (3) immediately after ET. [Needles stimulated manually by rotating, lifting and thrusting the handle of the needle in order to maintain de qi sensation, both during initial insertion and after 10 m]; (*n* = 102.)	No adjuvant treatment (*n* = 102)
Craig, USA (Craig *et al.*, 2007)	[113]	Two sessions, both at an off-site location: (1) within 1–2 h before ET; (2) within 1–2 h after ET. [Needles simulated manually at insertion, and then manually rotated after 10 m to maintain de qi]; (*n* = 57.)	No adjuvant treatment (*n* = 56)
Dieterle, Germany (Dieterle *et al.*, 2006)	225	Two sessions: (1) immediately after ET; (2) 3 days after ET. Needles rotated at the start and after 15 m, to evoke de qi (*n* = 116)	Needle acupuncture, with needles inserted in real acupoints not expected to influence fertility (*n* = 109)[c]
Domar, USA (Domar *et al.*, 2009)	150	Two sessions: (1) 25 m before ET; (2) immediately after ET (*n* = 78)[d]	No adjuvant treatment (*n* = 68)
Feliciani, Italy (Feliciani *et al.*, 2011)	46	Three sessions: (1) 5–7 d before egg retrieval; (2) 2–3 d before oocyte retrieval; (3) within 1 h after ET. [Needles manipulated to obtain de qi]; (*n* = 23.)	No adjuvant treatment (*n* = 23)
Fratterelli, USA (Fratterelli *et al.*, 2008)	397	Two sessions: (1) 25 m before ET; (2) after ET (*n* = 200)[d]	No adjuvant treatment (*n* = 197)[e]
Madaschi, Brazil (Madaschi *et al.*, 2010)	[455][f]	Two sessions: (1) 25 m before ET; (2) immediately after ET. Needles manipulated until de qi sensation obtained (*n* = 230)	No adjuvant treatment (*n* = 225)
Moy, USA (Moy *et al.*, 2011)	161	Two sessions: (1) 25 m before ET; (2) immediately after ET. [Needles rotated to obtain de qi sensation]; (*n* = 87.)	Needle acupuncture, with body acupuncture insertion sites close to, but not on, the real acupoints, with same manipulation of needles, as performed in true acupuncture group (*n* = 74)
Omodei, Italy (Omodei *et al.*, 2010)	168	Two sessions: (1) 25 m before ET; (2) immediately after ET (*n* = 44)[d]	No adjuvant treatment (*n* = 124)
Paulus, Germany (Paulus *et al.*, 2002)	160	Two sessions: (1) 25 m before ET; (2) immediately after ET. Needles rotated at the start and after 10 m, to evoke de qi (*n* = 80)	No adjuvant treatment (*n* = 80)
Paulus, Germany (Paulus *et al.*, 2003)	200	Two sessions: (1) 25 m before ET; (2) immediately after ET. Needles rotated after 10 m to evoke de qi (*n* = 100)	Non-penetrating blunt needles, placed on the real acupoints (*n* = 100)[c]
Smith, Australia (Smith *et al.*, 2006)	228	Three sessions: (1) d 9 of stimulating injections; (2) 25 m before ET; (3) immediately after ET. [Needles manually stimulated to obtain de qi sensation]; (*n* = 110.)	Non-penetrating blunt needles, placed close to, but not on, the real acupoints, with the sham needles manipulated by lifting and thrusting the handle of the needles and running a fingernail along the handle, but de qi not sought (*n* = 118)
So, Hong Kong/China (So *et al.*, 2009)	370	Two sessions: (1) 25 m before ET; (2) immediately after ET. Needles stimulated manually by rotating, lifting and thrusting the handle of the needle in order to maintain de qi sensation, both during initial insertion and after 10 m (*n* = 185)	Non-penetrating blunt needles, placed on the real acupoints, and manipulated in the same way as the true acupuncture needles to give patients a pricking, penetrating sensation (*n* = 185)
So, Hong Kong/China (So *et al.*, 2010)	226	One session immediately after ET. Needles stimulated manually by rotating, lifting and thrusting the handle of the needle in order to maintain de qi sensation, both during initial insertion and after 10 m (*n* = 113)	Non-penetrating blunt needles, placed on the real acupoints, and manipulated in the same way as the true acupuncture needles to give patients a pricking, penetrating sensation (*n* = 113)

Continued

Table I *Continued*

Study, Country (reference)	*n*[a]	Acupuncture	Control type[b]
Westergaard, Denmark (Westergaard *et al.*, 2006)	300	Acupuncture group 1: (1) 25 m before ET; (2) immediately after ET (*n* = 100); Acupuncture group 2: (1) 25 m before ET; (2) immediately after ET; (3) 2 d after ET. Needles rotated at the start and after 10 m, to evoke de qi (*n* = 100)[g]	No adjuvant treatment (*n* = 100)

Additional data obtained from RCT authors are enclosed in brackets to allow such data to be differentiated from the data included only in the publications. Acupoints, Acupuncture points; d, day; ET, embryo transfer; h, hour; m, minute.

[a]Number randomized.

[b]For all sham-controlled trials, the sham acupuncture procedure was given on the same schedule as that used for the true acupuncture group.

[c]These two trials did not report whether the sham needles were manipulated in the same way as the true acupuncture needles.

[d]These three trials did not report whether or not the needles used in the true acupuncture group were manipulated to achieve the de qi sensation.

[e]We excluded from our meta-analysis the following three additional control groups included in this trial: laser acupuncture, sham laser acupuncture and relaxation.

[f]The 39 participants (22 acupuncture; 17 control) who did not proceed to embryo transfer were excluded from the trial authors' analysis. The approach for our meta-analysis was to re-include these participants [i.e. acupuncture group = 230 (i.e. 208 + 22) and the no acupuncture group = 225 (i.e. 208 + 17)], because it could be assumed that these participants without an embryo transfer did not achieve the clinical pregnancy outcome.

[g]For this trial, acupuncture groups 1 and 2 were combined together for the meta-analysis.

three. In all trials, the acupuncture sessions lasted 25–30 min. In all trials, the acupuncture protocol and selection of acupuncture points was designed for the sole purpose of improving rates of pregnancy.

Twelve trials reported that the 'de qi' needling sensation (i.e. a pain, achiness, stinging or dullness at the needle insertion site, which is an indicator that the acupuncture needle has been correctly placed) was sought (Paulus *et al.*, 2002, 2003; Dieterle *et al.*, 2006; Smith *et al.*, 2006; Westergaard *et al.*, 2006; Craig *et al.*, 2007; So *et al.*, 2009, 2010; Arnoldi *et al.*, 2010; Madaschi *et al.*, 2010; Feliciani *et al.*, 2011; Moy *et al.*, 2011), whereas one sham-controlled trial reported that there was no attempt to manipulate the needles to achieve the 'de qi' needling sensation, in order to avoid unblinding trial participants (Andersen *et al.*, 2010). The three other trials did not report on de qi (Fratterelli *et al.*, 2008; Domar *et al.*, 2009; Omodei *et al.*, 2010). No trial used electroacupuncture. For all trials, the mean numbers of embryos transferred was similar between the randomized groups.

Methodological quality of included trials

A summary of the risks of bias is presented in Table II; full details of the quality assessments are available from the authors. For one trial (Madaschi *et al.*, 2010), the method of generating the random allocation sequence was not clear; however, the randomization appeared to be successful in creating similar groups, as there were no baseline differences in prognostic factors between treatment groups. For the other 15 trials, adequate methods of sequence generation were used. For three trials (Arnoldi *et al.*, 2010; Madaschi *et al.*, 2010; Feliciani *et al.*, 2011), there was inadequate allocation concealment (Table II), but in these trials there was also baseline similarity between the two groups. In three trials (Westergaard *et al.*, 2006; Craig *et al.*, 2007; Moy *et al.*, 2011), there were small numbers of randomized women with missing clinical pregnancy outcomes (Fig. 1, footnotes). In four trials (Smith *et al.*, 2006; Westergaard *et al.*, 2006; Craig *et al.*, 2007; Madaschi *et al.*, 2010), some randomized women began the IVF process but did not complete the treatment (that is, no embryo transfer); however, as noted above, these women were still included in the meta-analyses.

Seven of the trials used a sham acupuncture control (Paulus *et al.*, 2003; Dieterle *et al.*, 2006; Smith *et al.*, 2006; So *et al.*, 2009, 2010;

Andersen *et al.*, 2010; Moy *et al.*, 2011), of which four (Paulus *et al.*, 2003; So *et al.*, 2009, 2010; Andersen *et al.*, 2010) used the Streitberger non-penetrating sham needles (Streitberger and Kleinhenz, 1998; White *et al.*, 2003) placed on the true acupuncture points (Table I). The remaining nine trials used a no adjuvant treatment control. For 10 trials, the embryo transfer physicians were blinded to the treatment assignments (Paulus *et al.*, 2002, 2003; Dieterle *et al.*, 2006; Craig *et al.*, 2007; Domar *et al.*, 2009; So *et al.*, 2009, 2010; Andersen *et al.*, 2010; Madaschi *et al.*, 2010; Moy *et al.*, 2011); for the other six trials the physicians were not blinded (Smith *et al.*, 2006; Westergaard *et al.*, 2006; Fratterelli *et al.*, 2008; Arnoldi *et al.*, 2010; Omodei *et al.*, 2010; Feliciani *et al.*, 2011; Table II).

Efficacy analysis

There were no statistically significant pooled benefits of adjuvant acupuncture relative to either control, for the clinical pregnancy, ongoing pregnancy or live birth outcomes; however, the overall statistical heterogeneity was substantial for each outcome ($I^2 = 68$, 69 and 68%; χ^2 *P*-value = 0.0001, 0.0004 and 0.0003, respectively; Fig. 1). The type of control used did not significantly explain this statistical heterogeneity (interaction $P = 0.27$, for clinical pregnancy outcome; Fig. 1).

Of the 11 variables planned for subgroup analyses for the clinical pregnancy outcome, only 10 could be tested because no trial had a high risk of bias from incomplete outcome data. In addition, only the Craig *et al.* (2007) trial used an 'unequal co-intervention', which was driving to and from the off-site acupuncturist's office, before and after the embryo transfer procedure, in only the acupuncture group. Therefore, the 'unequal co-intervention' and 'site of acupuncture administration' variables classified trials in the same subgroups, and therefore had the same subgroup effect results.

We combined all 16 trials for the meta-regression subgroup analyses. Of the 10 subgroup variables tested, only baseline clinical pregnancy rate was a statistically significant effect modifier of adjuvant acupuncture (Fig. 2). Baseline pregnancy rate was also the only variable tested that explained a large proportion of the between-study variance. That is, in the meta-regression model fit with the single covariate of baseline clinical pregnancy rate, the adjusted $R^2 = 90\%$ when baseline clinical pregnancy rate was operationalized as a continuous variable (interaction $P < 0.001$)

Table II Risk of bias summary of included trials.

Study, Country (reference)	Allocation sequence generation	Allocation concealment	Blinding of patients	Blinding of physicians	Incomplete outcome data addressed: Clinical pregnancy	Selective reporting	Unequal co-intervention
Andersen, Denmark (Andersen et al., 2010)	Low risk	[Low risk]	Low risk	Low risk	Low risk	Low risk	Low risk
Arnoldi, Italy (Arnoldi et al., 2010)	Low risk	[High risk][a]	High risk	[High risk]	[Low risk][b]	Low risk	Low risk
Craig, USA (Craig et al., 2007)	[Low risk]	[Low risk]	High risk	Low risk	[Low risk]	Low risk	High risk
Dieterle, Germany (Dieterle et al., 2006)	[Low risk]	[Low risk]	Low risk	Low risk	Low risk	Low risk	Low risk
Domar, USA (Domar et al., 2009)	Low risk	[Low risk]	High risk	[Low risk]	[Low risk][c]	Low risk	Low risk
Feliciani, Italy (Feliciani et al., 2011)	[Low risk]	[High risk]	High risk	[High risk]	[Low risk]	Low risk	Low risk
Fratterelli, USA (Fratterelli et al., 2008)	Low risk	Low risk	High risk	High risk	Low risk	Low risk	Low risk
Madaschi, Brazil (Madaschi et al., 2010)	[Unclear][d]	[High risk][e]	High risk	[Low risk]	[Low risk][f]	Low risk	Low risk
Moy, USA (Moy et al., 2011)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Omodei, Italy (Omodei et al., 2010)	[Low risk]	[Low risk]	High risk	[High risk]	[Low risk]	Low risk	Low risk
Paulus, Germany (Paulus et al., 2002)	Low risk	[Low risk]	High risk	Low risk	[Low risk]	Low risk	Low risk
Paulus, Germany (Paulus et al., 2003)	[Low risk]	[Low risk]	Low risk	[Low risk]	[Low risk]	Low risk	Low risk
Smith, Australia (Smith et al., 2006)	[Low risk]	[Low risk]	Low risk	[High risk]	[Low risk]	Low risk	Low risk
So, Hong Kong/China (So et al., 2009)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
So, Hong Kong/China (So et al., 2010)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Westergaard, Denmark (Westergaard et al., 2006)	[Low risk]	[Low risk][g]	High risk	High risk	[Low risk]	Low risk	Low risk

Additional data obtained from RCT authors are enclosed in brackets to allow such data to be differentiated from data included only in the publications.

[a]Although sealed envelopes were used, the envelopes were not sequentially numbered and the trial's investigators could not recall whether or not the envelopes were opaque.

[b]Although there was a large imbalance in the number of women who did not proceed to an embryo transfer (i.e. 6/102 in acupuncture group and 20/102 in control group), which the trial authors postulated to be due to the acupuncture increasing the likelihood of a viable embryo being available for transfer, the outcomes for the randomized participants without embryo transfer were known (i.e. not pregnant), so we did not consider this as a bias due to incomplete outcome data.

[c]The treatment assignment and the outcomes for 4/150 randomized participants were not recorded by the trial authors. Therefore, the 146 participants analysed were instead used in the authors' analysis (Domar et al., 2009), and for this meta-analysis. However, we scored this criterion as 'low risk' for incomplete outcome data because the reasons for missing outcome data were unlikely to be related to the outcomes and the proportion of missing outcomes was not likely to have a clinically relevant impact on the effect estimates (Higgins and Altman, 2011).

[d]A computer-generated randomization list was used to assign patients to treatment groups. However, this trial was judged as unclear for 'allocation sequence generation' because it was not clear how the trial authors assigned to treatment groups 39 new participants who replaced the 39 participants excluded because of no embryo transfer.

[e]This trial used an open randomization list, so the investigators enrolling participants could possibly foresee assignments and thus introduce selection bias.

[f]In this trial, the randomization occurred prior to the start of the ovarian stimulation, although the participants were not informed of their treatment assignment until the start of the ovarian stimulation. Approximately 15% of participants were randomized, but then decided not to get IVF, primarily because of its costs and withdrew prior to the start of the ovarian stimulation. Because these participants withdrew from the study before they were told whether they had been randomized to acupuncture or control, their decision to withdraw from the trial could not have been affected by knowledge of the randomized intervention (i.e. acupuncture or control), and these withdrawals would not be expected to cause an important bias due to missing outcome data (Higgins and Altman, 2011).

[g]For this trial, the randomization treatment assignments were placed in sealed, opaque envelopes, which were shuffled and deposited in a cardboard box, from which each participant selected only one. This procedure has handled by an independent nurse not responsible for obtaining information about patients and enrolling them. Although the envelopes were not sequentially numbered, we considered the safeguards used in the randomization process to have provided adequate assurance of allocation concealment.

and the adjusted $R^2 = 93\%$ when the baseline pregnancy rate was operationalized as a dichotomous variable (interaction $P < 0.001$). To evaluate whether any of the other prespecified subgroup variables was a confounder of the effects of baseline pregnancy on IVF clinical pregnancy success rates, we also fit nine other separate meta-regression models which included both baseline pregnancy rate as a continuous variable and one of each of the nine other prespecified subgroup variables. In none of these nine other models did the correlation coefficient of the baseline clinical pregnancy rate variable substantially change (i.e. the maximum percentage change of the coefficient was 7%). To investigate whether the effect modification of baseline pregnancy rate was maintained across trial subsets, we also fit six separate single covariate (i.e. baseline pregnancy rate) meta-regression models, restricting to subsets of trials, according to type of control (sham or no adjuvant treatment), blinding of embryo transfer physician (yes or no) and allocation concealment (yes or no). In these meta-regression models, the magnitude of the baseline pregnancy rate subgroup effect was maintained when restricting to: sham-controlled trials [76% difference in RR estimates between sham-controlled trials with higher ($RR = 0.89$), compared with sham-controlled trials with lower ($RR = 1.65$), baseline pregnancy rates; interaction $P = 0.022$; adjusted $R^2 = 100\%$]; no adjuvant treatment-controlled trials (63% difference in RR; $P = 0.012$;
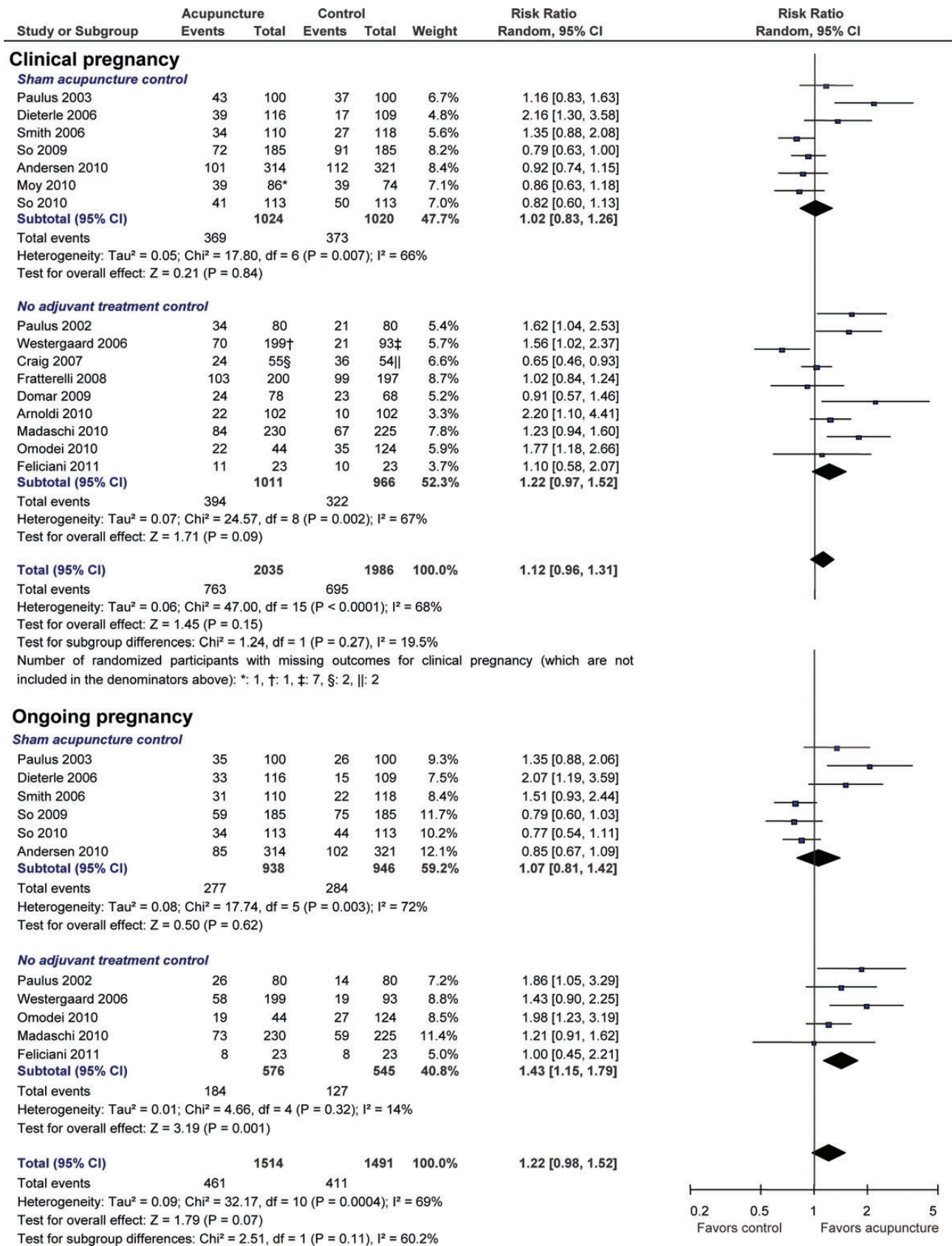
Figure 1 Effects of acupuncture on clinical pregnancy, ongoing pregnancy and live birth outcomes. The centres of the squares represent estimates from individual trials, the centres of the quadrilaterals represent pooled estimates and the horizontal lines represent 95% CIs.
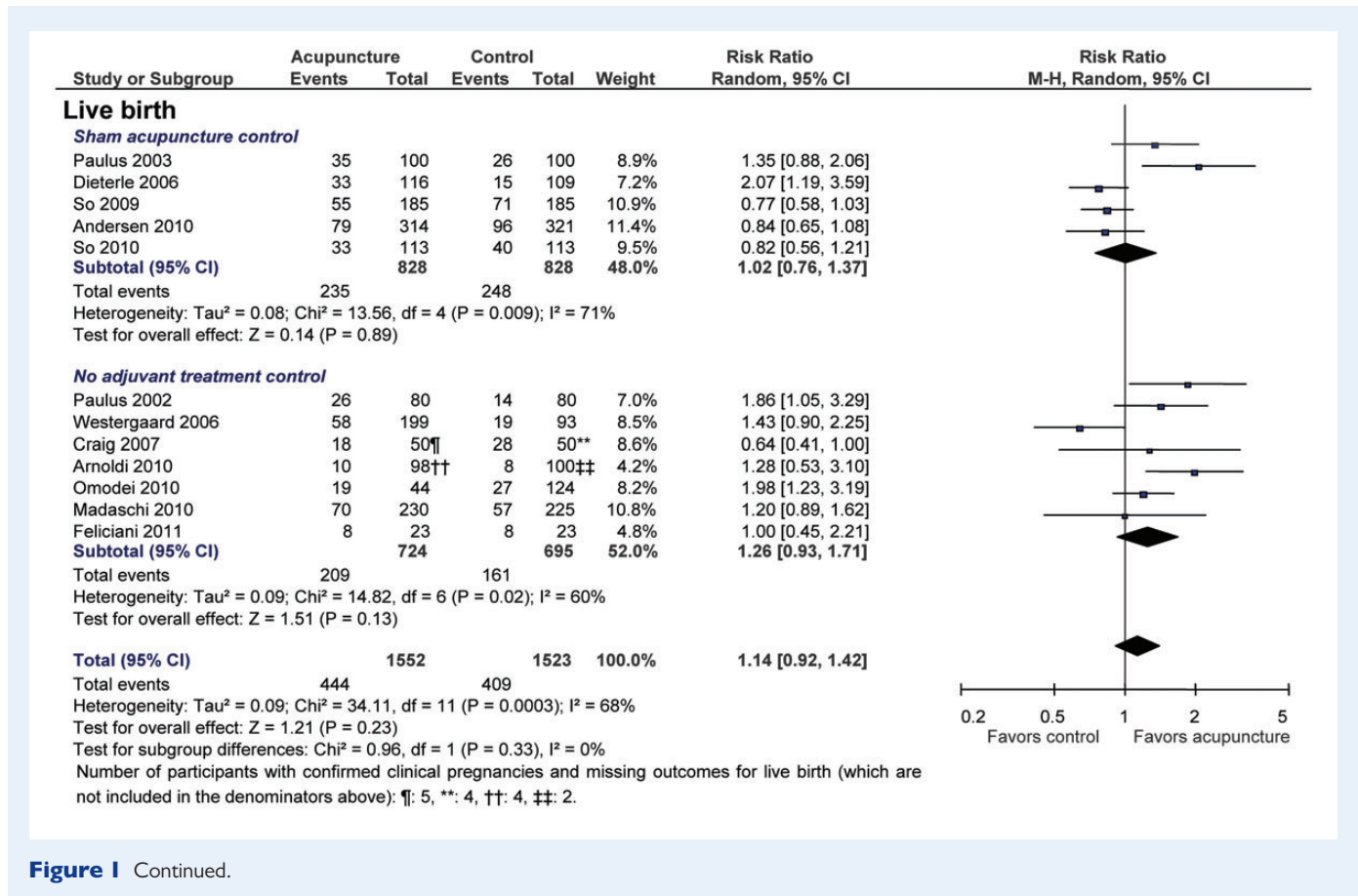
Figure 1 Continued.

adjusted $R^2 = 79\%$); trials with embryo transfer blinding (60% difference in RR; $P = 0.005$; adjusted $R^2 = 92\%$); trials without embryo transfer blinding (59% difference in RR; $P = 0.039$; adjusted $R^2 = 100\%$); trials with adequate allocation concealment (76% difference in RR; $P < 0.001$; adjusted $R^2 = 98\%$); and trials without adequate allocation concealment (40% difference in RR; $P = 0.661$; adjusted R^2 could not be calculated because only three trials were included in this meta-regression model).

Figure 3 shows a graph of the meta-regression line, fit with the single continuous covariate of baseline pregnancy rate. This meta-regression model showed that a 20 and 40% increase in baseline pregnancy rate was associated, respectively, with a 33 and 55% decrease in the RR of clinical pregnancy with adjuvant acupuncture. A 20 and 40% decrease in baseline clinical pregnancy rate was associated, respectively, with a 49 and 122% increase in the RR of clinical pregnancy.

The subgroup variables 'blinding of embryo transfer physician' and 'site of acupuncture administration' were borderline statistically significant across all trials (interaction $P = 0.084$ and 0.096, respectively); however, neither variable explained a large proportion of the between trial variability (adjusted $R^2 = 22$ and 19%, respectively). Because the trials that blinded the embryo transfer physician also tended to blind the patients (i.e. use a sham control; see Table II), we were concerned about confounding between these two variables. Therefore, we also fit a meta-regression model which included the patient blinding variable in addition to the embryo transfer physician blinding variable. In this model which included both variables, the

correlation coefficient of the embryo transfer physician blinding variable was not substantially changed (i.e. from −0.158 to −0.150), and the interaction *P*-value of the embryo transfer blinding covariate only increased slightly (i.e. from $P = 0.084$ to $P = 0.15$). Finally, because the only trial that evaluated acupuncture administered off-site of the IVF clinic (Craig *et al.*, 2007) used a no adjuvant treatment control, we also fit a meta-regression model using the single covariate of site of acupuncture administration, but restricting to the no adjuvant treatment-controlled trial subset. In this meta-regression model, the site of acupuncture administration covariate was statistically significant (interaction $P = 0.04$), and this covariate explained some of the heterogeneity within this subset of trials (adjusted $R^2 = 61\%$). The embryo transfer blinding variable was not a statistically significant effect modifier when restricting to the sham-controlled (interaction $P = 0.476$) or no adjuvant treatment-controlled ($P = 0.246$) trial subsets. Although we initially planned subgroup analyses stratified by control type, we combined all trials for our final subgroup analyses shown in Fig. 2 because of the largely similar subgroup effects across the two control groups.

The funnel plot (Fig. 4) showed small study effects, with the intervention effects estimated in smaller studies showing more benefit than the effects estimated in larger studies. The Egger statistical test for funnel plot asymmetry found that the association between estimated intervention effects and the standard error of the intervention effect (as a measure of the trial's size) was greater than would be expected to occur by chance ($P = 0.032$).
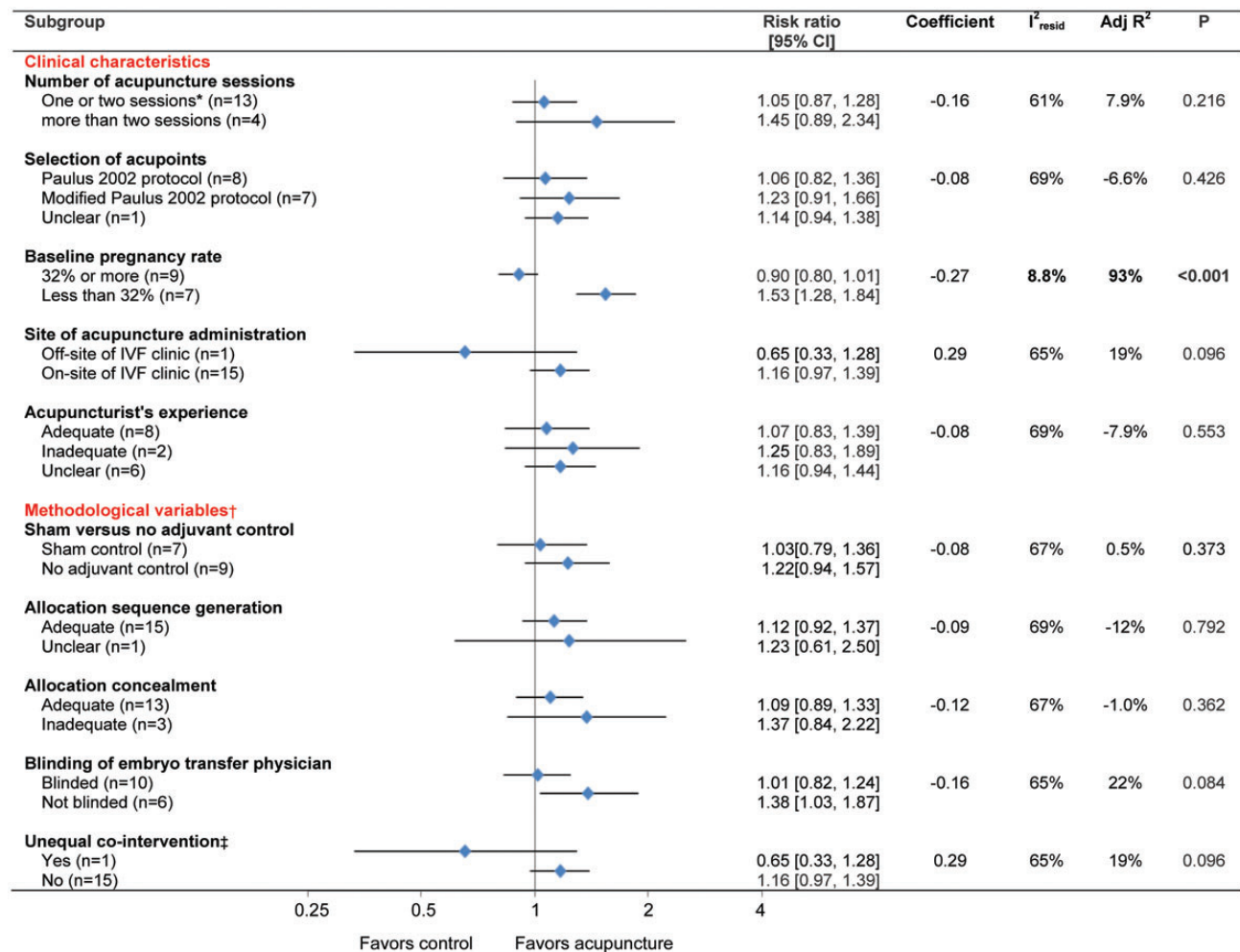
Figure 2 Meta-regression subgroup analyses for the primary outcome (i.e. clinical pregnancy). The quadrilaterals represent pooled estimates from the trials included in the given subgroup. The estimated regression coefficient from each model was obtained by a weighted least squares meta-regression with RR of pregnancy as the dependent variable, using the modification to the variance of the estimated coefficient suggested by Knapp and Hartung (2003), and supported by Higgins and Thompson (2004). I^2_{resid} values indicate the proportion of the residual variation that is attributable to between-study heterogeneity. The adjusted R^2 values indicate the proportion of between-study variance explained by the covariate. The P-values for test of interaction indicate whether the observed differences in results of trials within a given subgroup are compatible with chance alone. *The So 2010 trial was the only trial that used only one acupuncture treatment session. The Westergaard 2006 trial had two acupuncture treatment arms, one arm received two sessions and the other arm received three, and these arms were grouped separately for this subgroup analysis. †Incomplete outcome data were also prespecified as a methodological subgroup variable, but no trial had a high risk of bias due to incomplete outcome data for the clinical pregnancy outcome. ‡Only the Craig et al. trial was judged to have a co-intervention that was applied unequally across treatment group. Namely, in this trial, which evaluated off-site acupuncture, patients in the acupuncture group only were required to drive to and from the off-site acupuncturist's office both before and after the embryo transfer procedure.

Sensitivity analyses

The use of the different methods of imputing the unrecorded outcome data values did not affect the results of the individual trials, any of the pooled results or the results of any of the subgroup analyses. In trials that had participants with unrecorded clinical pregnancy outcomes (Westergaard et al., 2006; Craig et al., 2007; Moy et al., 2011), imputing these values using available case rates from the corresponding trial group that these participants were assigned to resulted in very similar pooled estimates for the primary outcome measure (RR = 1.12; 95% CI, 0.96–1.30; $I^2 = 68\%$; 16

trials). Imputing the values using the corresponding trial's acupuncture plus control group combined resulted in identical pooled estimates. Assuming that no participants with an unrecorded outcome achieved a clinical pregnancy resulted in very similar pooled estimates (RR = 1.13; 95% CI, 0.96–1.31; $I^2 = 69\%$; 16 trials).

Safety of acupuncture

Figure 5 shows the spontaneous abortion rates calculated as the number of losses between clinical pregnancy and live birth divided by the number of clinical pregnancies. Similar results were obtained when calculating

spontaneous abortion rates as losses between clinical pregnancy and ongoing pregnancy (across all trials, RR = 1.01; 95% CI, 0.76–1.35; $I^2 = 0\%$; 11 trials). A total of seven trials (Dieterle et al., 2006; Smith et al., 2006; Craig et al., 2007; So et al., 2009, 2010; Omodei et al., 2010; Moy et al., 2011) included descriptions of other adverse events, but none of these trials reported any serious adverse events associated with acupuncture.
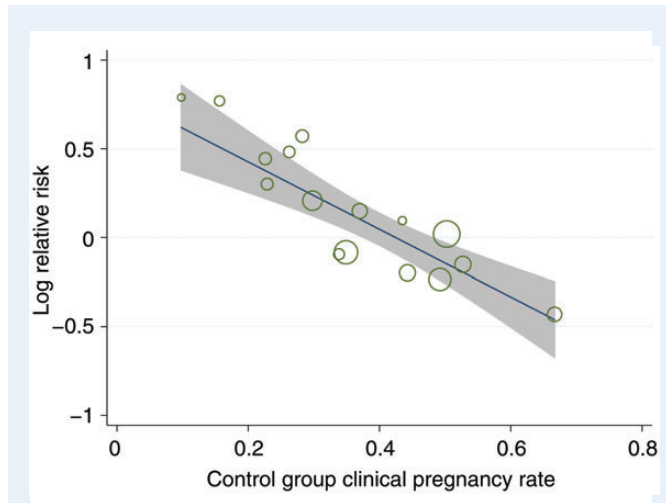
Figure 3 Meta-regression with baseline clinical pregnancy rate as single continuous covariate. Outcome was RR of clinical pregnancy. Circle size represents weight each study was given in the meta-regression analysis. For this meta-regression, the adjusted $R^2 = 90\%$ and the $I^2_{resid} = 3.3\%$.

Discussion

Summary of main results

We found no statistically significant pooled benefits of one to three sessions of acupuncture adjuvant to IVF, across all trials, or when restricting to sham-controlled or no adjuvant treatment-controlled trials. However, there was substantial (Deeks et al., 2011) heterogeneity. Of the 11 variables prespecified for subgroup analyses, only the baseline pregnancy rate variable showed a statistically significant subgroup effect and appeared to explain most of the heterogeneity in the results for the primary outcome measure (clinical pregnancy).

Subgroup analyses based on clinical characteristics

Baseline pregnancy rate was the only statistically significant subgroup finding in our previous review on this topic (Manheimer et al., 2008) and this has been confirmed in this review update. The magnitude of this subgroup effect is large and there is a low likelihood that chance explains this apparent subgroup effect. In addition, this subgroup effect does not appear to be due to an association between baseline pregnancy rate and either blinding- or randomization-related covariates, or any of the other prespecified covariates.

The reasons for differences in baseline (i.e. control group) pregnancy rates across trials are complex and driven by several factors, although these differences may be largely explained by differences in IVF regulations across countries, in particular, regarding the number of embryos transferred per cycle (Ludwig et al., 2000; Dal Canto et al., 2006; Abdalla, 2010; Maheshwari et al., 2011). That is, European countries have lower IVF pregnancy rates than does the USA, and a major
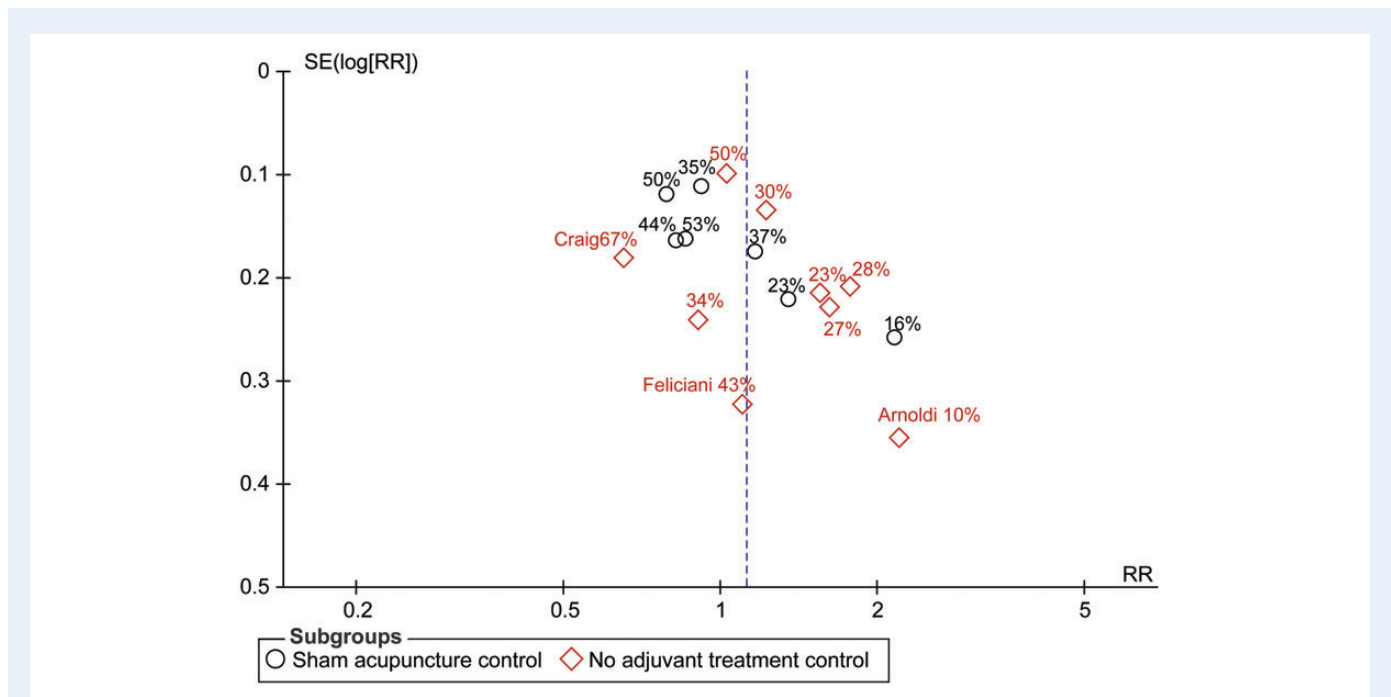
Figure 4 Funnel plot of trials meeting inclusion criteria. The intervention effect estimated from individual trials is plotted on the horizontal scale and a measure of the standard error of the intervention effect is plotted on the vertical axis. The control group success rate is included for each trial on the plot. The Egger test *P*-value for funnel plot asymmetry is 0.032.
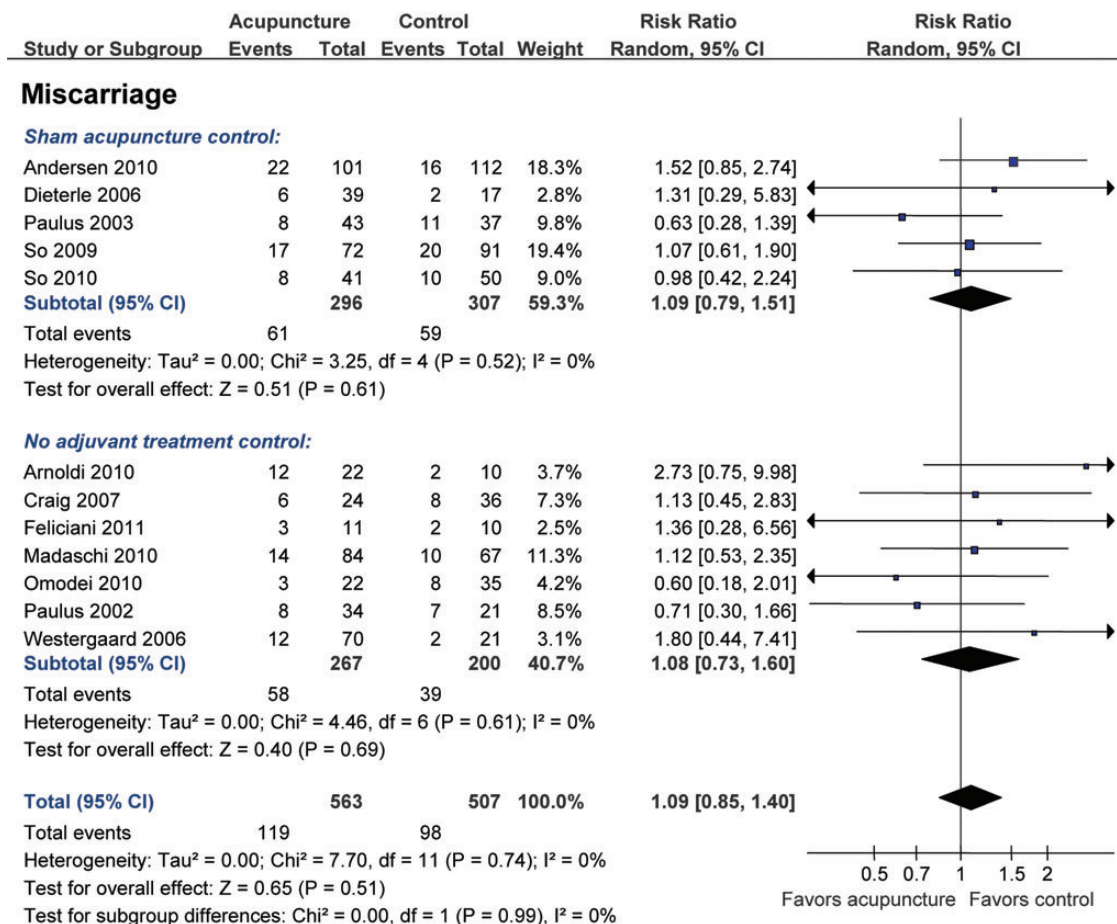
Figure 5 Effects of acupuncture on spontaneous abortion outcome. The centres of the squares represent estimates from individual trials, the centres of the quadrilaterals represent pooled estimates, and the horizontal lines represent 95% CIs.

reason for this difference is that fewer embryos are transferred per cycle in European countries than in the USA (Maheshwari et al., 2011), in an effort to reduce twins and triplets (Abdalla, 2010). In Italy and Germany, the country setting of six of the eight included European RCTs, the average IVF pregnancy rates are lower than in other European countries because of Italian/German laws which set a limit of using only three ooctyes per IVF cycle, all of which must be transferred, regardless of whether these oocytes develop into high-quality embryos (Ludwig et al., 2000; Gerris, 2005; Dal Canto et al., 2006; Maheshwari et al., 2011). In the USA, on the other hand, embryo selection (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011) is routine and the number of embryos that can be transferred is less strict and less well-regulated (Abdalla, 2010). Another reason for higher baseline pregnancy rates in the USA is that US patients with poor ovarian reserve (i.e. those who produce fewer oocytes and end up with fewer embryos to transfer or chose from) are more likely to be diverted from undergoing IVF (Abdalla, 2010). Differences in baseline pregnancy rates can also be due to differences in inclusion criteria across trials. For example, one Italian trial in this review (Arnoldi et al., 2010) restricted trial inclusion to women with an unfavourable reproductive prognosis, which might explain why the baseline pregnancy rates in this Italian trial were lower than Italian national

averages (Nygren et al., 2011). Two German trials (Paulus et al., 2002, 2003) restricted trial inclusion to only women with high-quality embryos available to transfer, which might explain why the baseline pregnancy rates in these German trials were higher than German national averages (Nygren et al., 2011). Finally, differences in baseline pregnancy rates across included trials can also be due to participant selection in the individual trials, in terms of participants' predictors for success in IVF (van Loendersloot et al., 2010). Namely, for two of the trials with lower baseline pregnancy rates (Dieterle et al., 2006; Smith et al., 2006), the trial authors reported that the trials' baseline pregnancy rates were lower than the average pregnancy rates at the trials' clinics because of the higher mean age of the participants included in the trials relative to the mean age of patients seen at the clinic site. And for one of these two trials (Smith et al., 2006), the trial authors also reported an overrepresentation by women with poor pregnancy outcomes from previous cycles. Thus, there are multiple factors that can explain differences across trials in baseline pregnancy rates. One potential explanation for the differential effect of acupuncture in trials with higher versus lower baseline rates may be that in IVF settings where the baseline pregnancy rates are already high, the relative added value of additional co-interventions, such as acupuncture, may be lower (Sherman et al., 2009).

Although it was not a statistically significant subgroup finding, there was a trend suggesting that acupuncture administered on-site of the IVF clinic had more positive effects than acupuncture administered off-site of the IVF clinic. However, inferences regarding this observed differential effect are weak because only a single trial (Craig et al., 2007) contributes to the off-site subgroup. The qualitatively and quantitatively different results in the off-site Craig et al. trial, which had very high baseline pregnancy rates (i.e. 67%), could have been driven by another variable, other than the off-site acupuncture administration. One potential biological rationale for the negative results of the Craig et al. trial, postulated by the trial's principal investigator (Sullivan, 2007), is that the co-intervention, in only the acupuncture group, of driving to and from the off-site acupuncturist's office, both before and after the embryo transfer, rather than resting in bed, may have increased patients' stress, which may have affected pregnancy success rates negatively through a psychobiological effect mechanism (Ferin and Vande Wiele, 1984; Stener-Victorin et al., 1996; Matthiesen et al., 2011). However, studies of the effects of stress on IVF success rates have had inconsistent results, and the influence of psychological stress on IVF success appears limited, at best (Matthiesen et al., 2011). In addition, while resting in bed for 2 or 3 days after the embryo transfer procedure has been recommended by some IVF physicians on the theory that bed rest can prevent the embryo from being expelled from the uterus, there is no evidence to support this recommendation (Li et al., 2011). Thus, the support from a biological rationale for this putative subgroup effect regarding on- versus off-site acupuncture is unconvincing. However, because only a single trial has evaluated off-site acupuncture, and because Craig et al. conducted this trial specifically to address whether the positive findings from the earlier Paulus et al. (2002, 2003) trials, which involved on-site acupuncture could be confirmed in 'real world' off-site settings (Sullivan, 2007), and because this Craig et al. trial found a reduction in pregnancy rates with off-site adjuvant acupuncture compared with no IVF adjuvant treatment, IVF physicians and patients may wish to consider this trial's findings in deciding whether IVF patients should have off-site acupuncture on the day of embryo transfer.

Subgroup analyses based on risk of bias

Seven of the trials (Paulus et al., 2003; Dieterle et al., 2006; Smith et al., 2006; So et al., 2009, 2010; Andersen et al., 2010; Moy et al., 2011) used a sham control to blind participants. However, the necessity to blind participants is arguable when the outcomes are entirely objective (that is, pregnancy and birth) (Hrobjartsson and Gotzsche, 2004; Kaptchuk et al., 2006; Manheimer et al., 2008; Wood et al., 2008). Indeed, the Cochrane Menstrual Disorders and Subfertility Group website states that 'a study may not be blinded but if follow-up is complete and outcomes are unequivocal the lack of blinding may be assessed not to increase risk of bias' (Cochrane Menstrual Disorders and Subfertility Group, 2010). As previously argued (Manheimer et al., 2008; Manheimer, 2011), it seems unlikely that a woman's knowledge of whether or not she was receiving acupuncture would affect her ability to become pregnant. Even if adjuvant acupuncture were to increase IVF success rates strictly through a non-needling-related psychobiological placebo effect mechanism (i.e. women who believe they are receiving acupuncture have reduced stress, which thereby subtly influences their pregnancy-related hormone production), this effect would be integral to the working mechanism by which adjuvant acupuncture increases

IVF pregnancy success rates. Therefore, it seems inappropriate to control for and equalize any such stress-reduction placebo effect by using a sham control to blind trial participants, unless the purpose of the trial is to study the mechanism responsible for adjuvant acupuncture's effects on IVF success rates.

The risk of using a sham control is that some sham interventions may influence the pregnancy outcome through the same putative mechanism of true acupuncture (e.g. by needle insertion, stimulation of true acupuncture points; Manheimer, 2011). For example, the sham intervention used in four of the sham-controlled trials (Paulus et al., 2003; So et al., 2009, 2010; Andersen et al., 2010) was the non-penetrating Streitberger sham needles placed at the true acupuncture points. The two validation trials (Streitberger and Kleinhenz, 1998; White et al., 2003) of this Streitberger sham needle found that these Streitberger sham needles elicit the de qi needling sensation in a large proportion of patients, suggesting that these sham needles may have acupuncture-specific needling effects, particularly if the Streitberger sham needles are manipulated in the same way as the true acupuncture needles, as was the case in two of the included sham-controlled trials (So et al., 2009, 2010). We initially attempted to assess whether or not the sham control intervention used in each trial was likely to have a risk of affecting the pregnancy outcome through the same putative mechanism of true acupuncture; however, because no consensus could be reached in making these assessments, the assessments could not be used in the analyses. Instead, our subgroup analysis test involved grouping all sham-controlled trials together in a single subgroup; however, some sham control interventions may be physiologically active, and others physiologically inert, in terms of their effects on the pregnancy outcome, and therefore this subgroup test (i.e. no adjuvant treatment versus any type of sham control) may not detect an important effect of control type. Finally, although this review found no statistically significant subgroup effect for the type of control, the no adjuvant treatment-controlled subgroup showed a slightly larger pooled effect than the sham-controlled subgroup, particularly if the outlying Craig et al. trial was removed [RR 1.31 (1.08–1.58); $I^2 = 49\%$; 8 trials]. Therefore, the possibility that some of the sham interventions may have had acupuncture-specific effects cannot be ruled out.

Although it was not a statistically significant subgroup finding, there was a trend suggesting that IVF physicians who are not blinded may perform more successful embryo transfers when adjuvant acupuncture is used, at least in a trial setting. This subgroup finding was unexpected because, considering the cost of embryo transfer and the importance of successful transfers to maintaining high pregnancy success rates at clinics, we *a priori* expected that physicians would be motivated primarily to perform a successful embryo transfer for all trial participants, rather than to show that acupuncture, a non-proprietary treatment, is an effective adjuvant procedure. Indeed, in IVF trials, where outcomes are objective, blinding of either patients or physicians is 'infrequently attempted' (Arce et al., 2005), and such blinding components are often not considered as critical elements related to the evaluation of risk of bias (Vail and Gardener, 2003; Dias et al., 2006). Yet, this borderline significant subgroup finding suggests that physician blinding may decrease bias, and therefore this variable may possibly be appropriate for inclusion in scales that assess risk of bias of IVF trials. However, blinding the embryo transfer physician still does not require the use of a sham acupuncture control (Paulus et al., 2002; Craig et al., 2007; Domar et al., 2009;

Madaschi *et al.*, 2010) because the adjuvant acupuncture is not administered at the same time as the embryo transfer procedure.

Interpretation of small study effects

The funnel plot indicated that estimates of the intervention effect were more beneficial in smaller studies, which may be due to smaller studies with statistically significant, positive results being more likely to be published (i.e. publication bias). Funnel plot asymmetry can also be due to other types of reporting bias besides publication bias (e.g. language bias, multiple publication bias, publication of negative results only as conference abstracts). However, in this review, these other types of reporting bias seem less likely than publication bias, given our extensive searches (including Chinese databases and conference proceedings), and our contacts with the authors of included trials, which confirmed that none of these trials are duplicate publications. Although we attempted to locate any unpublished trials by contacting experts in the field, none were located, and the possibility of unidentified, unpublished studies with negative results is one possible explanation for the asymmetric funnel plot. However, the Cochrane Collaboration's guidelines for addressing reporting biases and interpreting funnel plot asymmetry stress that 'an asymmetric funnel plot should not be equated with publication bias', and that 'publication bias should be considered as only one of a number of possible explanations' (Sterne *et al.*, 2011). Also according to Cochrane guidelines, 'Funnel plot asymmetry . . . may also result from clinical heterogeneity between studies (for example different control event rates) . . . ' (Alderson and Green, 2002). We therefore added the control event rates to the funnel plot, and the resulting plot shows that the trials with the lowest control group success rates show the most positive results (Fig. 4). Therefore, it is impossible to know whether the funnel plot asymmetry is due to publication bias, or whether this pattern is due to greater benefits in the trials with lower control group success rates (Alderson and Green, 2002).

We also tried to identify other potential explanations for the funnel plot asymmetry. Namely, we prepared separate funnel plots (and conducted corresponding separate Egger tests) for the sham-controlled trial subgroup and the no adjuvant treatment-controlled trial subgroup. These separate tests found that the funnel plot asymmetry seemed to be driven largely by the sham-controlled trials (Egger $P = 0.028$ for sham-controlled trials and $P = 0.303$ for no adjuvant treatment-controlled trials). Cochrane guidelines suggest that funnel plot asymmetry may be due to larger 'negative' trials (conducted later to confirm the earlier 'positive' trials' results) using a more effective control treatment or a less thoroughly implemented test intervention (Sterne *et al.*, 2011). Therefore, this issue was examined among the sham-controlled trials that contributed to the funnel plot asymmetry. Namely, three of the largest sham-controlled trials (So *et al.*, 2009, 2010; Andersen *et al.*, 2010), which all had 'negative' results, used non-penetrating Streitberger sham needles placed at the true acupuncture points as the control intervention. For two of these trials (So *et al.*, 2009, 2010), these sham needles were manipulated in the same way as the true acupuncture needles. The author of these two trials discussed that this sham control intervention used was likely an effective treatment due to its acupuncture-specific effects, and a conclusion of their 2009 trial was that 'Placebo acupuncture may not be inert' (So *et al.*, 2009). For the third and largest trial (Andersen *et al.*, 2010), for which the acupuncture was performed by nurse-acupuncturists and other nurses trained by these nurse-acupuncturists, the true acupuncture needles were not manipulated, in order to avoid unblinding the patients to treatment assignment. Therefore, in these three recent, large sham-controlled trials with 'negative' results, manipulating the needles in the sham acupuncture group (So *et al.*, 2009, 2010) and/or not manipulating the needles in the true acupuncture group (Andersen *et al.*, 2010) may have caused the control treatment to be an effective treatment and/or the test intervention to be less effective. This possibility is particularly relevant in light of a recent trial which found that manipulating the acupuncture needles to obtain the 'de qi' sensation resulted in a much greater therapeutic effect compared with not manipulating the needles (odds ratio = 4.16, for primary outcome; Xu *et al.*, 2013). The use of a potentially more effective control and/or less thoroughly implemented test intervention in these three large sham-controlled trials with 'negative' results is only one of many possible explanations for the funnel plot asymmetry among the sham-controlled trials. This example is provided primarily to illustrate the challenges in interpreting funnel plot asymmetry in meta-analyses of complex interventions that used heterogeneous treatments and controls, when the funnel plot test was devised on the premise that the studies come from a single underlying population (Terrin *et al.*, 2003; Sterne *et al.*, 2011).

Comparison with other studies or reviews

Although there have been multiple systematic reviews of acupuncture as an adjuvant to IVF (Cheong *et al.*, 2008, 2010; El-Toukhy *et al.*, 2008; Manheimer *et al.*, 2008; Ng *et al.*, 2008; El-Toukhy and Khalaf, 2009; Sunkara *et al.*, 2009; Qu *et al.*, 2012; Zheng *et al.*, 2012a), only the 2012 reviews by Zheng *et al.* (2012a) and Qu *et al.* (2012) included some of the more recent trials. And therefore, comparisons with the Zheng *et al.* review and Qu *et al.* review are most relevant. The Zheng *et al.* review concluded that adjuvant acupuncture improves clinical pregnancy rates and live birth rates, but only after excluding the Streitberger sham needle-controlled trials. Zheng *et al.* excluded these trials in drawing their conclusions because they contended that the Streitberger sham intervention is not an inert control; however, Zheng *et al.* did not provide evidence to support this contention. Also, because the Zheng *et al.* review was not based on a publicly available protocol, it cannot be determined whether the decision to exclude the trials using the Streitberger shams was prespecified, or whether instead this exclusion was driven by the Streitberger sham-controlled trials' results. In addition, the Zheng *et al.* review, as well as all other previous systematic reviews, were based on only the published data, with the exception of one previous review (El-Toukhy *et al.*, 2008) for which two RCT authors (Craig *et al.*, 2007; Domar *et al.*, 2009) were contacted and supplied additional information. We have successfully obtained unpublished data on methods and/or outcomes from authors of all eligible RCTs, which allowed for the inclusion of more RCTs and more complete information from these included RCTs. For example, Zheng *et al.*, excluded two methodologically sound RCTs (Omodei *et al.*, 2010; Moy *et al.*, 2011) because the 'data for the exact pregnancy events and totals were not available' from the publication; however, we were able to obtain precise outcome data for these two RCTs by contacting the authors. Finally, Zheng *et al.*, included both the early results from one trial (Benson *et al.*, 2006), as well as the final results (Fratterelli *et al.*, 2008), thus double counting some of the same participants from this

trial. The other 2012 review, by Qu et al., concluded that no benefits of adjuvant acupuncture were apparent from the effect estimate pooled from all trials. However, the Qu et al. review did not include six of the RCTs included in our review (Fratterelli et al., 2008; Arnoldi et al., 2010; Madaschi et al., 2010; Omodei et al., 2010; So et al., 2010; Feliciani et al., 2011), four of which (Fratterelli et al., 2008; Arnoldi et al., 2010; Omodei et al., 2010; So et al., 2010) were published prior to October 2010, the end date of Qu et al.'s search window. Also, Qu et al. did not examine potential subgroup variables to understand their impact on the effects of adjuvant acupuncture on IVF, on the clinically and methodologically heterogeneous trials reviewed. In contrast, the detailed information we obtained allowed us not only to evaluate potential effect modification due to the prespecified subgroup variables, but also to examine whether other subgroup variables were confounders of any apparent subgroup effects.

Implications for research

Future studies might further investigate the relationship between baseline rate of pregnancy and the efficacy of adjuvant acupuncture, and also further investigate which variables are responsible for this relationship. Because baseline pregnancy rate is possibly a proxy for the number of embryos transferred, although other factors are also involved (Abdalla, 2010), an individual patient data meta-analysis (Broeze et al., 2010) may be helpful to further investigate the relationship between the number of embryos transferred, as well as other individual participant level variables (e.g. age (Westergaard et al., 2006)), and the efficacy of adjuvant acupuncture.

For future trials, if the objective is to investigate the mechanism of IVF adjuvant acupuncture, and specifically whether any effect is mediated through the point-specific needle placement, or alternatively through a non-needling-related psychobiological placebo mechanism, both a sham control arm and a no adjuvant treatment arm would be necessary. A sham control arm might also be necessary for recruitment, if potential trial participants would be unwilling to be randomized to a no adjuvant treatment control (Smith et al., 2012). However, because it is difficult to assess whether or not a sham control intervention has acupuncture-specific effects that may increase IVF success rates, sham-controlled trials in this area can also potentially complicate the interpretation of the overall evidence (Manheimer, 2011). For addressing the more clinically relevant question of the total effects of acupuncture (i.e. specific needling effects plus any non-needle-related placebo effects) in contributing to any increases in IVF pregnancy success rates, a sham control seems unnecessary (Manheimer, 2011). This question about the need for sham controls may also apply to other invasive, difficult to blind adjuvant procedures evaluated in IVF RCTs (Holt et al., 2009).

Data availability

All our study data, including full details of the characteristics of each included RCT and all outcome data extracted from each included RCT, is included in a RevMan file, which will be made available to researchers upon request, for reproducing these results.

Supplementary data

Supplementary data are available at http://humupd.oxfordjournals.org/.

Acknowledgements

Authors' roles

E.M. conceived and designed the study, searched the literature to identify studies, selected the articles and extracted the data, conducted and oversaw the analysis and interpretation of the data and wrote the manuscript. D.W., J.T. and L.M.B. contributed to the study conception and design, analysis and interpretation of the data and writing of the manuscript. K.C. contributed to searching the literature to identify studies, article selection and data extraction, and assisted with data entry and data analysis. K.S. contributed to study selection and data extraction. J.L. contributed to searching the literature to identify studies, article selection and data extraction. L.L. rated the adequacy of the acupuncture treatments and the risk of the sham controls affecting the outcome. P.L. contributed to the analysis and interpretation of the data, and oversaw the meta-regression analysis and interpretation. E.M., D.W., C.K., K.S., J.L., J.T., L.L., B.M.B., P.L., L.M.B. provided comments on the drafts of the manuscript, and endorsed its final version.

Funding

Conflict of interest statement

None declared.

References

Abdalla HI. Are US results for assisted reproduction better than the rest? Is it a question of competence or policies? *Reprod Biomed Online* 2010;**21**:624–630.

Alderson P, Green S (eds). *Cochrane Collaboration's open learning material for Cochrane reviewers.* Module 15. Publication Bias. 2002 Available from http://130.226.106.152/openlearning/HTML/mod15.htm (30 December 2012, date last accessed).

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Hum Reprod* 2011;**26**:1270–83.

Andersen D, Lossl K, Nyboe Andersen A, Furbringer J, Bach H, Simonsen J, Larsen EC. Acupuncture on the day of embryo transfer: a randomized controlled trial of 635 patients. *Reprod Biomed Online* 2010;**21**:366–372.

Arce JC, Nyboe Andersen A, Collins J. Resolving methodological and clinical issues in the design of efficacy trials in assisted reproductive technologies: a mini-review. *Hum Reprod* 2005;**20**:1757–1771.

Arnoldi M, Diaferia D, Corbucci MG, De Lauretis L. The role of acupuncture in patients at unfavourable reproductive prognosis in IVF: a prospective randomised study. *Hum Reprod* 2010;**25**(Suppl 1):i257 (Abstract P-363).

Benson MR, Elkind-Hirsch KE, Theall A, Fong K, Hogan RB, Scott RT. Impact of acupuncture before and after embryo transfer on the outcome of in vitro fertilization cycles: a prospective single-blind randomized study. *Fertil Steril* 2006; **86**:S135 (Abstract P-118).

Broeze KA, Opmeer BC, van der Veen F, Bossuyt PM, Bhattacharya S, Mol BW. Individual patient data meta-analysis: a promising approach for evidence synthesis in reproductive medicine. *Hum Reprod Update* 2010;**16**:561–567.

Centers for Disease Control and Prevention. *Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports 2009.* Atlanta, GA: US Department of Health and Human Services. Centers for Disease Control and Prevention, 2011.

Cheong YC, Hung Yu Ng E, Ledger WL. Acupuncture and assisted conception. *Cochrane Database Syst Rev* 2008;CD006920.

Cheong Y, Nardo LG, Rutherford T, Ledger W. Acupuncture and herbal medicine in in vitro fertilisation: a review of the evidence for clinical practice. *Hum Fertil (Camb)* 2010;**13**:3–12.

Clarke JF, van Rumste MM, Farquhar CM, Johnson NP, Mol BW, Herbison P. Measuring outcomes in fertility trials: can we rely on clinical pregnancy rates? *Fertil Steril* 2010; **94**:1647–1651.

Cochrane Menstrual Disorders and Subfertility Group. *Resources for Review Authors.* Auckland: The Cochrane Collaboration, 2010. Available from http://mdsg.cochrane.org/resources-review-authors/ (30 December 2012, date last accessed).

Craig LB, Criniti AR, Hansen KR, Marshall LA, Soules MR. Acupuncture lowers pregnancy rates when performed before and after embryo transfer. *Fertil Steril* 2007;**88**:S–40 (Abstract O-106).

Dal Canto MB, Mignini Renzini M, Brambillasca F, Cepparo H, Comi R, Villa A, Rangoni G, Mastrolilli M, Crippa M, de Ponti E et al. IVM—the first choice for IVF in Italy. *Reprod Biomed Online* 2006;**13**:159–165.

Daya S. Pitfalls in the design and analysis of efficacy trials in subfertility. *Hum Reprod* 2003;**18**:1005–1009.

Deeks JJ, Higgins JPT. Statistical algorithms in Review Manager 5. August 2010. Available from http://www.cochrane.org/sites/default/files/uploads/handbook/Statistical_Methods_in_RevMan5-1.pdf (30 December 2012, date last accessed).

Deeks JJ, Higgins JPT, Altman DG (eds). Chapter 8: analysing and presenting results (section 8.11.3.1). In: Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions Version 4.2.6 [updated September 2006].* The Cochrane Collaboration, 2006. Available from http://www.cochrane.org/sites/default/files/uploads/Handbook4.2.6Sep2006.pdf (30 December 2012, date last accessed).

Deeks JJ, Higgins JPT, Altman DG (eds). Chapter 9: analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].* The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org (30 December 2012, date last accessed).

de Lacey S, Smith CA, Paterson C. Building resilience: a preliminary exploration of women's perceptions of the use of acupuncture as an adjunct to In Vitro Fertilisation. *BMC Complement Altern Med* 2009;**9**:50.

de Mouzon J, Goossens V, Bhattacharya S, Castilla JA, Ferraretti AP, Korsak V, Kupka M, Nygren KG, Andersen AN. Assisted reproductive technology in Europe, 2007: results generated from European registers by ESHRE. *Hum Reprod* 2012; **27**:954–966.

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**:177–188.

Dias S, McNamee R, Vail A. Evidence of improving quality of reporting of randomized controlled trials in subfertility. *Hum Reprod* 2006;**21**:2617–2627.

Dieterle S, Ying G, Hatzmann W, Neuer A. Effect of acupuncture on the outcome of in vitro fertilization and intracytoplasmic sperm injection: a randomized, prospective, controlled clinical study. *Fertil Steril* 2006;**85**:1347–1351.

Domar AD, Meshay I, Kelliher J, Alper M, Powers RD. The impact of acupuncture on in vitro fertilization outcome. *Fertil Steril* 2009;**91**:723–726.

Domar AD, Conboy L, Denardo-Roney J, Rooney KL. Lifestyle behaviors in women undergoing in vitro fertilization: a prospective study. *Fertil Steril* 2012;**97**:697–701.

Elden H, Ladfors L, Olsen MF, Ostgaard HC, Hagberg H. Effects of acupuncture and stabilising exercises as adjunct to standard treatment in pregnant women with pelvic girdle pain: randomised single blind controlled trial. *BMJ* 2005;**330**:761.

El-Toukhy T, Khalaf Y. The impact of acupuncture on assisted reproductive technology outcome. *Curr Opin Obstet Gynecol* 2009;**21**:240–246.

El-Toukhy T, Sunkara SK, Khairy M, Dyer R, Khalaf Y, Coomarasamy A. A systematic review and meta-analysis of acupuncture in in vitro fertilisation. *BJOG* 2008; **115**:1203–1213.

ESHRE Capri Workshop Group. Intracytoplasmic sperm injection (ICSI) in 2006: evidence and evolution. *Hum Reprod Update* 2007;**13**:515–526.

Evers JL. Female subfertility. *Lancet* 2002;**360**:151–159.

Feliciani E, Ferraretti AP, Paesano C, Pellizzaro E, Magli MC, Gianaroli L. The role of acupuncture in ART: preliminary results of an ongoing prospective randomized study. *Hum Reprod* 2011;**26**:i230 (Abstract P-281).

Ferin M, Vande Wiele R. Endogenous opioid peptides and the control of the menstrual cycle. *Eur J Obstet Gynecol Reprod Biol* 1984;**18**:365–373.

Fratterelli JL, Leondires MR, Fong K, Theall A, Locatelli S, Scott RT. Laser acupuncture before and after embryo transfer improves ART delivery rates: Results of a prospective randomized double-blinded placebo controlled five-armed trial involving 1000 patients. *Fertil Steril* 2008;**90**:S–105 (Abstract O-285).

Gerris JM. Single embryo transfer and IVF/ICSI outcome: a balanced appraisal. *Hum Reprod Update* 2005;**11**:105–121.

Griesinger G, Dafopoulos K, Schultze-Mosgau A, Felberbaum R, Diedrich K. What is the most relevant standard of success in assisted reproduction? Is BESST (birth emphasizing a successful singleton at term) truly the best? *Hum Reprod* 2004; **19**:1239–1241.

Harbord RM, Higgins JP. Meta-regression in Stata. *Stata J* 2008;**8**:493–519.

Harper J, Magli MC, Lundin K, Barratt CL, Brison D. When and how should new technology be introduced into the IVF laboratory? *Hum Reprod* 2012;**27**:303–313.

Higgins JPT, Altman DG (eds). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org (30 December 2012, date last accessed).

Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. *Stat Med* 2004;**23**:1663–1682.

Holt J, Lord J, Acharya U, White A, O'Neill N, Shaw S, Barton A. The effectiveness of foot reflexology in inducing ovulation: a sham-controlled randomized trial. *Fertil Steril* 2009;**91**:2514–2519.

Hopton AK, Curnoe S, Kanaan M, Macpherson H. Acupuncture in practice: mapping the providers, the patients and the settings in a national cross-sectional survey. *BMJ Open* 2012;**2**:e000456.

Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2004;CD003974.

Isoyama D, Cordts EB, de Souza van Niewegen AM, de Almeida Pereira de Carvalho W, Matsumura ST, Barbosa CP. Effect of acupuncture on symptoms of anxiety in women undergoing in vitro fertilisation: a prospective randomised controlled study. *Acupunct Med* 2012;**30**:85–88.

Kaptchuk TJ, Stason WB, Davis RB, Legedza AR, Schnyer RN, Kerr CE, Stone DA, Nam BH, Kirsch I, Goldman RH. Sham device v inert pill: randomised controlled trial of two placebo treatments. *BMJ* 2006;**332**:391–397.

Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. *Stat Med* 2003;**22**:2693–2710.

Kovarova P, Smith CA, Turnbull DA. An exploratory study of the effect of acupuncture on self-efficacy for women seeking fertility support. *Explore (NY)* 2010;**6**:330–334.

Kvorning N, Holmberg C, Grennert L, Aberg A, Akeson J. Acupuncture relieves pelvic and low-back pain in late pregnancy. *Acta Obstet Gynecol Scand* 2004;**83**:246–250.

Kwan I, Bhattacharya S, Knox F, McNeil A. Pain relief for women undergoing oocyte retrieval for assisted reproduction. *Cochrane Database Syst Rev* 2013;CD004829.

Li B, Zhou H, Li W. Bed rest after embryo transfer. *Eur J Obstet Gynecol Reprod Biol* 2011;**155**:125–128.

Ludwig M, Schopper B, Katalinic A, Sturm R, Al-Hasani S, Diedrich K. Experience with the elective transfer of two embryos under the conditions of the german embryo protection law: results of a retrospective data analysis of 2573 transfer cycles. *Hum Reprod* 2000;**15**:319–324.

Madaschi C, Braga DP, Figueira Rde C, Iaconelli AJ, Borges EJ. Effect of acupuncture on assisted reproduction treatment outcomes. *Acupunct Med* 2010;**28**:180–184.

Maheshwari A, Griffiths S, Bhattacharya S. Global variations in the uptake of single embryo transfer. *Hum Reprod Update* 2011;**17**:107–120.

Manheimer E. Selecting a control for in vitro fertilization and acupuncture randomized controlled trials (RCTs): how sham controls may unnecessarily complicate the RCT evidence base. *Fertil Steril* 2011;**95**:2456–2461.

Manheimer E, Zhang G, Udoff L, Haramati A, Langenberg P, Berman BM, Bouter LM. Effects of acupuncture on rates of pregnancy and live birth among women undergoing in vitro fertilisation: systematic review and meta-analysis. *BMJ* 2008;**336**:545–549.

Manheimer E, van der Windt DA, Stafford K, Liu J, Tierney J, Lao L, Berman B, Bouter L. Effects of adjuvant acupuncture on pregnancy rates among women undergoing in vitro fertilization: systematic review and meta-analysis. *PROSPERO* 2011: CRD42011001309. Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42011001309 (30 December 2012, date last accessed).

Matthiesen SM, Frederiksen Y, Ingerslev HJ, Zachariae R. Stress, distress and outcome of assisted reproductive technology (ART): a meta-analysis. *Hum Reprod* 2011;**26**:2763–2776.

Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;**6**:e1000097.

Moy I, Milad MP, Barnes R, Confino E, Kazer RR, Zhang X. Randomized controlled trial: effects of acupuncture on pregnancy rates in women undergoing in vitro fertilization. *Fertil Steril* 2011;**95**:583–587.

Ng EH, So WS, Gao J, Wong YY, Ho PC. The role of acupuncture in the management of subfertility. *Fertil Steril* 2008;**90**:1–13.

Nygren KG, Sullivan E, Zegers-Hochschild F, Mansour R, Ishihara O, Adamson GD, de Mouzon J. International Committee for Monitoring Assisted Reproductive Technology (ICMART) world report: assisted reproductive technology 2003. *Fertil Steril* 2011;**95**:2209–2222, 2222 e2201–2217.

Omodei U, Piccioni G, Tombesi S, Dordoni D, Fallo L, Ghilardi F. Effect of acupuncture on rates of pregnancy among women undergoing in vitro fertilization. *Fertil Steril* 2010;**94**:S170 (Abstract P-266).

Paulus WE, Zhang M, Strehler E, El-Danasouri I, Sterzik K. Influence of acupuncture on the pregnancy rate in patients who undergo assisted reproduction therapy. *Fertil Steril* 2002;**77**:721–724.

Paulus WE, Zhang M, Strehler E, Seybold B, Sterzik K. Placebo-controlled trial of acupuncture effects in assisted reproduction therapy. *Hum Reprod* 2003;**18**:xviii18–19.

Pinborg A. IVF/ICSI twin pregnancies: risks and prevention. *Hum Reprod Update* 2005;**11**:575–593.

Qu F, Zhou J, Ren RX. Effects of acupuncture on the outcomes of in vitro fertilization: a systematic review and meta-analysis. *J Altern Complement Med* 2012;**18**:429–439.

Sherman KJ, Cherkin DC, Ichikawa L, Avins AL, Barlow WE, Khalsa PS, Deyo RA. Characteristics of patients with chronic back pain who benefit from acupuncture. *BMC Musculoskelet Disord* 2009;**10**:114.

Smith C, Crowther C, Beilby J. Pregnancy outcome following women's participation in a randomised controlled trial of acupuncture to treat nausea and vomiting in early pregnancy. *Complement Ther Med* 2002;**10**:78–83.

Smith C, Coyle M, Norman RJ. Influence of acupuncture stimulation on pregnancy rates for women undergoing embryo transfer. *Fertil Steril* 2006;**85**:1352–1358.

Smith JF, Eisenberg ML, Millstein SG, Nachtigall RD, Shindel AW, Wing H, Cedars M, Pasch L, Katz PP. The use of complementary and alternative fertility treatment in couples seeking fertility care: data from a prospective cohort in the United States. *Fertil Steril* 2010;**93**:2169–2174.

Smith CA, Ussher JM, Perz J, Carmady B, de Lacey S. The effect of acupuncture on psychosocial outcomes for women experiencing infertility: a pilot randomized controlled trial. *J Altern Complement Med* 2011;**17**:923–930.

Smith CA, de Lacey S, Chapman M, Ratcliffe J, Norman RJ, Johnson N, Sacks G, Lyttleton J, Boothroyd C. Acupuncture to improve live birth rates for women undergoing in vitro fertilization: a protocol for a randomized controlled trial. *Trials* 2012;**13**:60.

So EW, Ng EH, Wong YY, Lau EY, Yeung WS, Ho PC. A randomized double blind comparison of real and placebo acupuncture in IVF treatment. *Hum Reprod* 2009;**24**:341–348.

So EW, Ng EH, Wong YY, Yeung WS, Ho PC. Acupuncture for frozen–thawed embryo transfer cycles: a double-blind randomized controlled trial. *Reprod Biomed Online* 2010;**20**:814–821.

Stener-Victorin E. The pain-relieving effect of electro-acupuncture and conventional medical analgesic methods during oocyte retrieval: a systematic review of randomized controlled trials. *Hum Reprod* 2005;**20**:339–349.

Stener-Victorin E, Waldenstrom U, Andersson SA, Wikland M. Reduction of blood flow impedance in the uterine arteries of infertile women with electro-acupuncture. *Hum Reprod* 1996;**11**:1314–1317.

Sterne JAC, Egger M, Moher D (eds). Chapter 10: Addressing reporting biases. In: Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org (30 December 2012, date last accessed).

Streitberger K, Kleinhenz J. Introducing a placebo needle into acupuncture research. *Lancet* 1998;**352**:364–365.

Sullivan MG. Study questions benefit of acupuncture in IVF. Ob.Gyn.News. 2007 November 1, pg. 21. (Available at: http://obgyn.imng.com/fileadmin/content_pdf/obn/archive_pdf/vol42iss21/70921_main.pdf) (30 December 2012, date last accessed).

Sunkara SK, Coomarasamy A, Khalaf Y, El-Toukhy T. Acupuncture and in vitro fertilization: updated meta-analysis. *Hum Reprod* 2009;**24**:2047–2048.

Terrin N, Schmid CH, Lau J, Olkin I. Adjusting for publication bias in the presence of heterogeneity. *Stat Med* 2003;**22**:2113–2126.

The Nordic Cochrane Centre. *The Cochrane Collaboration. Review Manager (RevMan) [Computer program]. Version 5.1*. Copenhagen, Denmark: The Cochrane Collaboration, 2011.

Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 2002;**21**:1559–1573.

Ulett GA, Han S, Han JS. Electroacupuncture: mechanisms and clinical application. *Biol Psychiatry* 1998;**44**:129–138.

Vail A, Gardener E. Common statistical errors in the design and analysis of subfertility trials. *Hum Reprod* 2003;**18**:1000–1004.

van Loendersloot LL, van Wely M, Limpens J, Bossuyt PM, Repping S, van der Veen F. Predictive factors in in vitro fertilization (IVF): a systematic review and meta-analysis. *Hum Reprod Update* 2010;**16**:577–589.

van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. *Spine* 2003;**28**:1290–1299.

Westergaard LG, Mao Q, Krogslund M, Sandrini S, Lenz S, Grinsted J. Acupuncture on the day of embryo transfer significantly improves the reproductive outcome in infertile women: a prospective, randomized trial. *Fertil Steril* 2006;**85**:1341–1346.

White P, Lewith G, Hopwood V, Prescott P. The placebo needle, is it a valid and convincing placebo for use in acupuncture trials? A randomised, single-blind, cross-over pilot trial. *Pain* 2003;**106**:401–409.

Witt CM, Lao L, MacPherson H. Evidence on acupuncture safety needs to be based on large-scale prospective surveys, not single case reports. *Pain* 2011;**152**:2180; auther reply 2184–2186.

Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, Gluud C, Martin RM, Wood AJ, Sterne JA. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;**336**:601–605.

Xu SB, Huang B, Zhang CY, Du P, Yuan Q, Bi GJ, Zhang GB, Xie MJ, Luo X, Huang GY *et al.* Effectiveness of strengthened stimulation during acupuncture for the treatment of Bell palsy: a randomized controlled trial. *CMAJ* 2013;**185**: 473–479.

Zegers-Hochschild F, Nygren KG, Adamson GD, de Mouzon J, Lancaster P, Mansour R, Sullivan E. The ICMART glossary on ART terminology. *Hum Reprod* 2006; **21**:1968–1970.

Zheng CH, Huang GY, Zhang MM, Wang W. Effects of acupuncture on pregnancy rates in women undergoing in vitro fertilization: a systematic review and meta-analysis. *Fertil Steril* 2012a;**97**:599–611.

Zheng CH, Zhang MM, Huang GY, Wang W. The role of acupuncture in assisted reproductive technology. *Evid Based Complement Alternat Med* 2012b; **2012**:543924.