

- 21 Pinborg A, Loft A, Ziebe S, Nyboe Andersen A. What is the most relevant standard of success in assisted reproduction? Is there a single "parameter of excellence"? *Hum Reprod* 2004;19:1052-4.
- 22 Daya S. Pitfalls in the design and analysis of efficacy trials in subfertility. *Hum Reprod* 2003;18:1005-9.
- 23 White A, Hayhoe S, Hart A, Ernst E. Adverse events following acupuncture: prospective survey of 32 000 consultations with doctors and physiotherapists. *BMJ* 2001;323:485-6.
- 24 MacPherson H, Thomas K, Walters S, Fitter M. The York acupuncture safety study: prospective survey of 34 000 treatments by traditional acupuncturists. *BMJ* 2001;323:486-7.
- 25 American Society for Reproductive Medicine. *Frequently asked questions about infertility*. www.asrm.org/Patients/faqs.html.

Accepted: 17 December 2007

Commentary: Good, but not perfect

Mike Clarke

EDITORIAL by Pinborg and colleagues

UK Cochrane Centre, Oxford
OX2 7LG

mclarke@cochrane.co.uk

BMJ 2008;336:549

doi:10.1136/bmj.39490.520046.25

In commenting on the systematic review by Eric Manheimer and colleagues of the effects of acupuncture for women undergoing in vitro fertilisation (IVF),¹ I have focused on the methods used in the review. My commentary stems from issues raised when the manuscript was refereed and, perhaps, a wish to have an independent opinion on the reliability of the findings. Having accepted this challenge, I set out to assess whether the review provides knowledge of a sufficient standard to influence decisions.

The eligible interventions were specific types of acupuncture, used close to the time of embryo transfer, compared with sham acupuncture or no adjuvant treatment. Other aspects of care were the same for women within each trial. The trials were randomised and the population studied was women trying to get pregnant through IVF. Whether it is appropriate to combine trials using sham acupuncture and no adjuvant treatment is dealt with by presenting results for the two types of trial separately and together. This showed little difference in the point estimates for the effects of acupuncture or the finding of significance, whichever way the analyses were done.

No reviewers can search absolutely everywhere for potentially eligible studies. This would be a never ending task, accompanied by diminishing returns of eligible studies. The compromise is a balance between the pragmatic and the perfect by searching various sources likely to provide a reasonable yield of eligible studies while minimising the impact of publication bias. Manheimer et al did this, as they have in other systematic reviews.² And, although they might still be missing some studies, this is a problem for all reviews and will remain so until trial registration and the availability of trial findings become the norm.³

Each trial was assessed in a standard way. Most were judged to be satisfactory for methodological features related to the risk of bias. These features included concealment of allocation, which was "adequate" in six of the seven included trials, although the reviewers do express some concerns about the preference for sealed envelopes, rather than more secure off-site processes.

The authors sought to supplement published information with data from the original researchers. They were successful to some extent—for example, they obtained unpublished data on live births from three trials. They conducted their analyses in a standard way (odds ratios and a random effects model), and their findings would have been similar had they used other approaches, such as risk ratios or the fixed effect model. One potential problem is with their subgroup analysis

based on the proportion of women in the control group who became pregnant. Although this analysis was prespecified and used a predefined threshold of 28%, splitting a meta-analysis on the basis of outcome data from one intervention group to investigate the comparative effect against the other group can lead to bias. Focusing on trials that found good prognosis for the control group tends to produce a lower effect estimate than using the prognosis for both groups combined. Hence, it might be preferable to apply the pregnancy threshold to each trial as a whole. If this is done, I calculate that five trials would be in the subgroup analysis for "higher pregnancy" and the odds ratio for clinical pregnancy would be 1.52 (95% confidence interval 1.13 to 2.05, $P=0.006$).

The review supplements the calculated odds ratios, which are difficult to interpret, with a number needed to treat to estimate how many women would need to receive acupuncture during one cycle of IVF to become pregnant. The authors veer on the side of caution by basing some of the discussion on the upper end of the confidence interval and note that 17 women would need to be treated for one more to become pregnant. Whether or not 17 is "too many" for acupuncture to be judged to be a clinically useful intervention is debatable.

So is this review by Manheimer and colleagues a well conducted review, worthy of consideration when making decisions about IVF? Yes. Is it perfect? No. However, several thousand systematic reviews are published each year in health care,⁴ and none of them is likely to be perfect. This one seems as good as many. Unless, of course, you know differently?

Competing interests: None declared.

Provenance and peer review: Commissioned; not peer reviewed.

- 1 Manheimer E, Zhang G, Udoff L, Haramati A, Langenberg P, Berman BM, et al. Effects of acupuncture on pregnancy and live birth rates among women undergoing in vitro fertilisation: systematic review and meta-analysis. *BMJ* 2008 doi: 10.1136/bmj.39471.430451.BE.
- 2 Lim B, Manheimer E, Lao L, Ziea E, Wisniewski J, Liu J, et al. Acupuncture for treatment of irritable bowel syndrome. *Cochrane Database Syst Rev* 2006;(4):CD005111.
- 3 Lemmens T, Bouchard RA. Mandatory clinical trial registration: rebuilding public trust in medical research. In: *Global forum update on research for health*. Vol 4. *Equitable access: research challenges for health in developing countries*. London: Pro-Brook Publishing, 2007:40-6.
- 4 Moher D, Tetzlaff, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Med* 2007;4:e78.

Accepted: 8 February 2008