Randomized studies in intrauterine insemination

Ian D. Cooke, F.R.C.O.G.

University of Sheffield, Sheffield, United Kingdom

Many factors in IUI might be critical to outcome. What realistically should be tested is another matter. (Fertil Steril® 2004;82:27–9. ©2004 by American Society for Reproductive Medicine.)

What are the critical elements in IUI? If they could be identified, perhaps we could move to a more rational and therefore commonly used protocol. Two reports in this issue of Fertility and Sterility concentrate on the timing of the insemination and the number of inseminations, the former by means of a randomized, controlled trial (RCT) (1) and the latter by a meta-analysis of carefully selected such trials (2). But are these aspects the most important? The meta-analysis revealed heterogeneity, and the contributing trials show great variety in descriptions of their studies. Have all other influencing features in these studies been carefully controlled? A trial is randomized so that bias will be eliminated, but how much bias exists in the clinical environment? We must not assume that practice around the world is as good as that described in papers submitted for publication, so there might be much greater bias in local practice than in the studies reported.

The indications for IUI are not dissimilar to those for IVF and indeed for intracytoplasmic sperm injection (ICSI) for moderate male factor. Is the patient referred specifically for one or the other? Do all practitioners and clinics have equal access to each technique? What influences the decision to have a particular cut-off for the one or the other, say for male factor? How often are there multiple causes contributing to the infertility, such as mild male factor and endometriosis? How variable are the decisions of those recruiting patients to a trial, including the decision to reject patients from a study?

There is wide variation in the detail of indications for IUI. In the United Kingdom, the National Institute for Clinical Excellence (NICE) has recently reviewed the evidence for “Assessment and treatment for people with infertility problems.” It has made draft recommendations, which include the following: “Intrauterine insemination should be offered to couples with unexplained infertility because it is as effective as in vitro fertilisation, it is less invasive and requires fewer resources...” and again, “[c]ouples with male factor infertility should be offered intrauterine insemination with ovulation induction (sic) because it increases the chance of pregnancy” (3). It is also a possible treatment for infrequent ovulation if other treatment has failed. Elsewhere it is stated that a maximum of three cycles of “fresh” IVF should be accessible to all those with unexplained infertility and when the woman is aged ≤39 years, although this will have to be subject to local health funding negotiation. There is no guidance as to the balance between IUI and IVF.

There will be funding issues. Not many patients have had access to public funds. In the United Kingdom, NICE will announce a national “standard price tariff” (4), a reimbursement for all “standardized” treatments. The relative costs of IUI and IVF will be important in the United Kingdom, although the price differential everywhere is significant. For the economic health of the clinic, a balance will need to be struck between the number of IUI cycles and of IVF cycles. How will it be possible to make a dispassionate decision exclusively in the patients’ interest? At what point will one transfer a patient from IUI to IVF?
There is a disparity in thinking between IUI and IVF with respect to the number of follicles developed in IUI and the number of follicles developed/embryos frozen and embryos transferred at IVF. The number of multiple births remains high from non-IVF procedures (5). In some countries efforts have been made to move toward reducing the number of embryos transferred at IVF (6–9), and in Sweden single ET is now the norm. There is little evidence of IUI management moving toward the same objective. A call has been made for lesser degrees of stimulation at IVF (10), but this is not evident for IUI. If cost is such an important element, where is the justification for using a GnRH analogue in IUI without moving to IVF? Should not the use of urinary or recombinant gonadotropin preparations be detailed?

The criteria for follicle development are similarly highly variable. How many mature follicles are desired? At what point is the decision made not to proceed and cancel, aspirate follicles, or convert to IVF? In studies of clinical treatment there should be an intention-to-treat analysis (11); in terms of overall cost to the program, it is important to count the cost of the stimulation. This would be crucial in any cost/benefit analysis (12). This will surely come for IUI vs. IVF if there is a national tariff; clinics will calculate these elements for themselves if their fees are capped centrally by total allocations per patient. There are various reasons that patients withdraw: poor or over-response, stress, and cumulative cost. These are quite different causes and impact differently on whether a particular regimen can be shown to be effective.

The quality of sperm preparations might also vary from patient to patient and from a first to a second insemination. Categories for treatment are quite broad in the case of mild to moderate male factor, and the stress of providing timed ejaculations under clinic conditions also contributes. How is it possible to account for the variation in the fertilizing capacity of compromised sperm? Is the cost of a double insemination factored in?

The number of cycles exceeds considerably the number of patients treated. There are multiple treatment cycles per patient, but the number varies and is not standardized. The pregnancy rate declines in those remaining in treatment, thus a greater number of cycles from individual patients reduces the overall measurable outcome. No weighting is usually carried out (13).

Outcome is usually defined as a positive fetal heart presence on ultrasound. The earlier that this is undertaken, the less representative it is of the final live birth rate, which should be the arbiter of outcome. Furthermore, anything other than a singleton term infant should have a negative weighting. The data on multiple births and on fetal reductions should be published in the same study. The results should be produced as singleton, term, live birth rates, and those cases having fetal reduction should be excluded from that figure. Such a statistic would encompass the effects of the stimulatory regimen and the consequences of follicle development. The total cost of the program would allow a cost per maternity to be estimated, although these costs are not currently available (14). These would be true costs to set against similarly computed costs of IVF. This leaves aside the arguments in favor of calculating the cost of neonatal care of multiple births, let alone longer-term care of handicap engendered by multiple pregnancy and delivery, borne not by the couple alone but by the whole community.

Irrespective of the health care system in which IUI takes place, most of these aspects are relevant. The patient and her partner are highly motivated, but they often have limited resources. Are we truly giving them the optimum management? Although RCTs are helpful in attempting to elucidate differences in therapeutic regimens, the remarkable heterogeneity of studies suggests that there are too many variables. There is insufficient control of these and little careful description of all the elements of a study. Is an RCT the correct model to use? There are not sufficient data from economically sensible models with appropriate outcome measures to merit the carrying out of RCTs. Perhaps more data from cohort studies should be obtained, notwithstanding the fact that these would only be graded B in the world of evidence-based medicine (15). It might then be possible to design an appropriately powered study to answer a tightly defined question. However, it is worth remembering that when the World Health Organization was carrying out its studies on infertility, well-formulated studies of sufficient power using multinational recruitment still entered patients extremely slowly. The patients ideal for those studies were few and far between.

In the meantime, we are left with comparisons between two variables that show no significant differences. However, type II errors remain in view of insufficient power (16). Is it worth doing these studies again? Probably only when there is adequate power and as many conditions as possible are standardized prospectively. Will such a study be funded? Almost certainly not. What then can we do? Authors should reanalyze recent data, taking the above features into account, and particularly assess the economic aspects in various ways. We will then have a few regimens that are affordable and worthy of more detailed, prospective study. There is no alternative if there is to be national funding. There is no alternative if the patients are paying and if we are to treat our patients optimally.

References