

Article Type: Commentary

Reproductive medicine: still more ART than science?

J Wilkinson¹, S Bhattacharya², JMN Duffy^{3,4}, MS Kamath⁵, J Marjoribanks⁶, S Repping⁷, A Vail¹, M van Wely⁷, CM Farquhar⁶

¹ Centre for Biostatistics, University of Manchester, Manchester, United Kingdom.

² The Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, United Kingdom.

³ Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom.

⁴ Balliol College, University of Oxford, Oxford, United Kingdom.

⁵ Reproductive Medicine Unit, Christian Medical College, Vellore, India.

⁶ Cochrane Gynecology and Fertility Group, University of Auckland, Auckland, New Zealand.

⁷ Centre for Reproductive Medicine, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1471-0528.15409

This article is protected by copyright. All rights reserved.

Accepted Article

Corresponding author:

J Wilkinson

Centre for Biostatistics, University of Manchester, Manchester, United Kingdom.

jack.wilkinson@manchester.ac.uk

Running title:

Reproductive medicine: still more ART than science?

The history of obstetrics and gynaecology is not a tale of evidence-based practice. Tradition, expert opinion, and the lure of new technology have frequently superseded evidence as the primary driver for clinical decision making¹. The proof can be found in a litany of dubious interventions which have gained widespread popularity despite an absence of high quality data attesting to their effectiveness and, in some cases, ample credible evidence demonstrating harm. As a specialty, we have relied on investigations including X-ray pelvimetry and antenatal stress tests, subjected innumerable women to stilboestrol and thalidomide, and have performed routine episiotomy in all primigravid women¹. It is no surprise that, in 1979, Archie Cochrane famously awarded obstetrics the 'wooden spoon' for being the least evidence-based specialty.

Reproductive medicine has come a long way since then and there have been dramatic advances in the field such that biological parenthood is a reality for many couples, for whom adoption was the only available route in the past. The gold standard in reproductive medicine is now "to provide childless couples with the best possible management of their fertility problems, while at the same time ensuring that they are not exposed to unnecessary

This article is protected by copyright. All rights reserved.

Accepted Article

risks or ineffective treatments³. In order to achieve this standard, we need reliable evidence which can be used to develop clinical practice guidelines. Unfortunately, even where we have good evidence, its uptake is far from universal¹. This is particularly true in the field of assisted reproduction, which has become increasingly commercialised with a pronounced predilection for overdiagnosis, overuse, and overtreatment.

Too much treatment?

One in seven couples will struggle to conceive within a year, and will be labelled as infertile. Yet, population based data show that half of this group will go on to conceive in the next 12 months⁵. Meanwhile, success rates for fertility treatments remain modest, with live birth rates for in vitro fertilisation (IVF) estimated to be in the region of 27% per treatment cycle (<https://www.hfea.gov.uk/media/2563/hfea-fertility-trends-and-figures-2017-v2.pdf>). Yet, media fuelled unrealistic expectations on the part of couples coupled with skilful marketing by commercial fertility clinics has caused a global explosion in the use of assisted reproductive technology. As few governments fund infertility treatment the private sector is buoyant. Contrary to guidelines, assisted reproductive technology is frequently initiated early-on in couples who still have a reasonable chance of conception, promoting overtreatment, increasing avoidable harm, and increasing healthcare costs⁴. A plethora of essentially unproven add-on interventions are offered to couples in a technological arms race between fertility clinics⁷. Since financial considerations often influence treatment policies, the novelty and apparent sophistication of these new technologies take precedence over proof of effectiveness and safety. For many couples, the additional expense of add-on therapies is itself interpreted as a sign of quality. It is the perfect storm for exploitation.

The gap between clinical practice and evidence

While the absence of evidence in many areas of reproductive medicine remains a problem, blatant disregard for evidence which is available is an inexcusable but all too common phenomenon. One of the most notable and regrettable examples of ignoring evidence in assisted reproductive technology is the persistent practice of transferring two or more embryos at a time. This option appears attractive to many couples, who see it as a means to maximise their chances while obviating the need for further treatment to have a second child. Some IVF clinics continue to support this practice on grounds of patient autonomy, whilst downplaying the considerable maternal, fetal, neonatal, and childhood risks associated with multiple pregnancy. Although multiple pregnancy is almost completely preventable by the practice of single embryo transfer, many healthcare providers across the United Kingdom, Europe, and North America continue to transfer multiple embryos. One explanation is that IVF is often self-funded whilst maternity and neonatal care are often covered by the state or by insurance companies. If the same funder paid not only for IVF but also for maternity and neonatal care, then there would be a strong incentive to reduce multiple pregnancies in order to avoid increases in fetal, neonatal, and childhood morbidity and the attendant costs. In 2005, New Zealand increased funding for IVF from one cycle to two provided that a single embryo transfer strategy was used. As a result, New Zealand now has the lowest rate of IVF-related multiple pregnancies (4%) in the world (<https://npesu.unsw.edu.au/surveillance/assisted-reproductive-technology-australia-and-new-zealand-2015>).

Another example where common practice defies the evidence is the expansion of intracytoplasmic sperm injection (ICSI), originally designed for men with very poor semen quality, to include couples where the semen parameters are normal. Indeed, it has now become the dominant form of in vitro fertilisation in many countries. The Cochrane review of in vitro

fertilisation versus intra-cytoplasmic sperm injection for non-male infertility includes a single randomised trial involving 415 couples which reported 70/213 ongoing pregnancies following in vitro fertilisation compared to 51/202 in the intra-cytoplasmic sperm injection group (Odds ratio 1.44, 95% confidence interval 0.95 to 2.21) ¹¹. Data from a subsequent Canadian study which randomised 60 women with unexplained infertility to in vitro fertilisation or intra-cytoplasmic sperm injection failed to show a statistically significant difference in live birth rate ¹². In this context it is the focus on a surrogate outcome (fertilisation rate) rather than the clinical outcome (live birth) which has presumably influenced clinical decision making; trials of randomised oocytes suggest that a strategy of ICSI in non-male infertility could reduce the number of cases of total fertilisation failure, but an improvement in live births has not been demonstrated. Similarly, preimplantation genetic testing and time lapse systems for embryo incubation are becoming routine practice in many clinics, despite a lack of robust evidence of clinical benefit ^{13, 14}.

Marketing trumps informed patient choice

The primary criterion used by many people with infertility to select an IVF clinic is its success rate. In order to be competitive, fertility clinics must convince couples that the path to parenthood leads through their doors. But there is a clear tension between direct to consumer advertising of assisted reproductive technology and informed patient choice. Clinics frequently wield emotive language and creative statistics to persuade rather than to inform ^{16, 17}. Websites speak of 'dreams' and 'miracles'¹⁶, while bespoke success rates can be constructed from a large array of statistics in order to cast a clinic's performance in a favourable light ¹⁷. Due to a lack of binding standards on assisted reproductive technology outcome reporting, couples may be misled about the likelihood of success and may compare clinics on the basis of incomparable figures. Self-regulation does not appear to furnish

couples with impartial information regarding the effectiveness of different treatment interventions.

Challenges for evidence-based assisted reproductive technology

Before being implemented as routine, any assisted reproductive technology intervention should undergo preclinical and clinical trials, followed by randomised controlled trials (RCTs) to prove efficacy and safety, with long-term follow up to evaluate ongoing effectiveness and safety⁷. However, in assisted reproductive technology, eagerness to identify attractive new interventions has led to a quantity over quality approach to RCTs. Consequently, the clinical effectiveness of interventions is often not robustly assessed. A pivotal limitation is the size of the trials undertaken. In order to achieve 80% power to detect an improvement in birth rate from 27% to 32% at a 5% significance threshold, a trial size of 2,610 women is required. This far exceeds the numbers of women we actually see in the vast majority of in vitro fertilisation trials. Routinely collected data and large electronic databases have been touted as a solution but the absence of random allocation introduces doubt regarding the accuracy of any conclusions drawn. These are essentially large-scale observational studies, and do not offer a free pass to assisted reproductive technology treatment evaluation.

In principle, the machinery of meta-analysis offers a means to overcome the limitations of small individual trials, but this is predicated on sufficient cumulative sample sizes after pooling studies. The trend to 'test' many interventions in small trials does not guarantee this will hold. Moreover, if trials do not report outcomes in a consistent fashion, meta-analysis is precluded. Recent reviews of reporting standards in this field suggest that they do not¹⁸. Moreover, many trials measure outcomes in such a way that the benefits of randomisation are lost, and many fail to report cumulative birth rates per couple, which is probably the most relevant outcome for patients. The need to establish statistically valid, patient-centred

outcome measures is the motivation behind the Core Outcome Measures for Infertility Trials (COMMIT) project, which will establish core outcomes to be reported in all assisted reproductive technology trials

(<https://www.phc.ox.ac.uk/research/hypertension/pregnancy/commit>). Fewer, larger RCTs developed through priority-sharing partnerships and enabled by clinical trial networks committed to the use of common reporting standards will ensure high quality evidence with maximum clinical utility. A new priority setting partnership has been established for infertility aiming to bring together professionals, researchers, and people with infertility to prioritise unanswered research questions (www.phc.ox.ac.uk/infertility). In Table 1, we present our top ten research gaps in reproductive medicine.

Conclusions:

How to introduce new technologies in reproductive medicine responsibly?

With respect to new assisted reproductive technologies, our responsibility extends beyond the couples we treat to the wellbeing of future generations. Under the status quo however, robust evidence of effectiveness and safety is not a prerequisite for adoption of novel interventions[added Harper 2017 here]. One solution to the trend for treatment to be delivered in defiance of evidence would be stronger regulatory standards governing the introduction of new therapies in assisted reproductive technology. A study on the acceptability of regulating such techniques found broad support from clinicians, patients and the general public for regulation of new reproductive technologies by a national bioethics committee, which would consult with advisors from various backgrounds ²⁴. We are not sympathetic to the argument that a requirement for rigorous prospective treatment evaluation would impede progress in the field, since it is not at all clear that a 'right to try' philosophy, where treatments are sold to vulnerable people on a speculative basis, benefits anyone other than the people making the sale, and may well cause harm. Until the time comes when we insist that generation of evidence precedes implementation of new

technologies rather than follows it, reproductive medical research will be largely preoccupied with reversing perverse practice. We believe that this ongoing need for medical reversal in reproductive medicine is the real enemy of progress.

Disclosure of interests

JW declares that publishing in peer reviewed journals benefits his career. SB reports grants from the Chief Scientist Office, Scotland, during the period of this Study. JMND reports grants from The Royal Society of New Zealand, during the conduct of the study. MSK, JM, SR, and MVW have nothing to disclose. AV has received grants from the Human Fertilisation and Embryology Authority, outside the submitted work. CMF declares that she is the coordinating editor of Cochrane Gynaecology and Fertility. Completed disclosure of interest forms are available to view online as supporting information.

Contributions of authors

All authors devised and wrote the manuscript, and gave approval for submission.

Details of ethical approval

No ethical approval was sought in writing this commentary.

Funding

No specific funding supported the writing of this commentary.

References

1. Grimes DA. Discovering the need for randomized controlled trials in obstetrics: a personal odyssey. *JLL Bulletin: Commentaries on the history of treatment evaluation*. 2007.
3. ESHRE. Good clinical treatment in assisted reproduction - An ESHRE position paper. 2008.
5. te Velde ER, Eijkemans R, Habbema HDF. Variation in couple fecundity and time to pregnancy, an essential concept in human reproduction. *Lancet*. 2000 Jun 3;355(9219):1928-9.
7. Harper J, Jackson E, Sermon K, Aitken RJ, Harbottle S, Mocanu E, et al. Adjuncts in the IVF laboratory: where is the evidence for 'add-on' interventions? *Human Reproduction*. 2017 Mar;32(3):485-91.
11. Bhattacharya S, Hamilton MPR, Shaaban M, Khalaf Y, Seddler M, Ghobara T, et al. Conventional in-vitro fertilisation versus intracytoplasmic sperm injection for the treatment of non-male-factor infertility: a randomised controlled trial. *Lancet*. 2001 Jun 30;357(9274):2075-9.
12. Foong SC, Fleetham JA, O'Keane JA, Scott SG, Tough SC, Greene CA. A prospective randomized trial of conventional in vitro fertilization versus intracytoplasmic sperm injection in unexplained infertility. *Journal of assisted reproduction and genetics*. 2006 Mar;23(3):137-40.
13. Franco JG. Seven Reasons To Be Concerned About the Use of the New Preimplantation Genetic Screening (PGS). *J Bras Reprod Assist*. 2015 Oct-Dec;19(4):189-91.
14. Chen MH, Wei SY, Hu JY, Yuan J, Liu FH. Does time-lapse imaging have favorable results for embryo incubation and selection compared with conventional methods in clinical in vitro fertilization? A meta-analysis and systematic review of randomized controlled trials. *PloS one*. 2017 Jun 1;12(6).
16. Hawkins J. Selling ART: An Empirical Assessment of Advertising on Fertility Clinics' Websites. *Indiana Law J*. 2013 Fal;88(4):1147-79.
17. Wilkinson J, Vail A, Roberts SA. Direct-to-consumer advertising of success rates for medically assisted reproduction: a review of national clinic websites. *BMJ open*. 2017 Jan 12;7(1):e012218.
18. Duffy JMN, Rolph R, Gale C, Hirsch M, Khan KS, Ziebland S, et al. Core outcome sets in women's and newborn health: a systematic review. *Bjog-Int J Obstet Gy*. 2017 Sep;124(10):1481-9.
24. Hendriks S, Vliegthart R, Repping S, Dancet EAF. Broad support for regulating the clinical implementation of future reproductive techniques. *Human Reproduction*. 2018 Jan;33(1):39-46.

Table 1: Ten research gaps in reproductive medicine

Population	Interventions
In women with unexplained subfertility	What is the effect of IVF versus expectant management
In men with azoospermia and severe oligospermia	What is the evidence for available screening tests (e.g. genetic testing, ultrasound, screening for cancer, others)
In women with repeated IVF failure	What is the effect of different treatment adjuncts such as endometrial scratch, hysteroscopy, immunoglobulin, G-CSF
In women having IVF who are undergoing frozen embryo cycles	What is the optimal regimen (e.g. endometrial preparation, triggering, luteal phase support)
In women having IVF and undergoing final oocyte maturation	What is the effect of Gn RhA trigger versus HcG trigger
In men with reduced semen quality	What is the most optimal treatment modality in relation to the semen quality (expectant management, IUI, IVF or ICSI)
In women having IVF	What is the effect of cleavage stage versus blastocyst transfer
In women having IVF	What is the optimal culture environment
In women having IVF and their children	What is the long term effect (over > 20 years)
In women having IVF	What is the optimal embryo selection method (morphology, genetic testing (PGS), time-lapse algorithm)?