



COMMENTARY



Generating translatable evidence to improve patient care: the contribution of human factors

Masoud A.M. Afnan^{1,*}, Khalid Saeed Khan², Ben Willem Mol³

ABSTRACT

Generating translatable evidence to improve patient care has proved challenging in reproductive medicine, with many 'add-on' treatments in routine assisted conception clinical practice that have not been reliably tested. This has consequences for patient care; specifically, IVF pregnancy rates have not improved. A change of culture is required in our profession, from indiscriminately applying the latest 'add-on' to large-scale participation in generating reliable translatable evidence.

GENERATING RELIABLE EVIDENCE FOR SMALL TREATMENT EFFECT SIZES

Editorials (*Macklon et al., 2019*) have bemoaned the undue reliance on data generated by randomized controlled trials (RCT) as a necessary validation for 'add-on' treatments in assisted conception (*Human Fertilisation and Embryology Authority, 2019*). This has had the effect of stimulating debate about evidence generation (*Wilkinson et al., 2019*).

Macklon et al. (2019) cite the practical difficulties and high costs of performing RCT in the current climate as obstacles to generating such data, and therefore, that we risk not offering 'add-ons' that might work. They advocate using 'big data' as a means of generating that data which appears so elusive. Whilst we agree that there is value to be gained from

big data, it tends to be without direct influence on decision-making because it merely observes what happened as a consequence of clinical decision-making, rather than the research question-determined effect. This will always make big data more prone to a biased observation of the truth (*Collins et al., 2020*).

This commentary aims to move the debate forward by addressing some of the human factors.

THE MODEL OF TRANSLATIONAL RESEARCH

A basic model of translational research that goes from basic to human clinical research ('bench to bedside') is familiar to health practitioners, but understanding its details is key to the complexity of implementation. Human clinical research is multiphase. Practice-based research

knowledge moves from early efficacy clinical studies to effectiveness RCT and then through to meta-analyses and guideline development, including tools for patients, clinicians and policy-makers. Confusion arises as translation is not made up of discreet steps, rather it is a continuum with feedback loops to inform research development along the way.

The role of human factors

As can be seen from the response of leaders to the current COVID-19 crisis, there is a strong desire to do something, anything, to help the patient. Doing something makes not only the patient feel better, but also the doctor (or the leader). This has led to advocating treatments based on data which are open to significant bias (*Bik, 2020; Gautret et al., 2020*).

Furthermore, much data are generated in ways which are not replicable or

KEYWORDS

Big data
Human factors
RCT
Translational evidence

¹ Department of Obstetrics and Gynaecology, Qingdao United Family Hospital, Qingdao, People's Republic of China

² Distinguished Investigator in Preventive Medicine and Public Health at the University of Granada funded by the Beatriz Galindo Program (Senior Modality) of the Spanish Ministry of Science, Innovation and Universities, Universidad de Granada, Granada, Spain

³ Department of Obstetrics and Gynaecology, Monash University, Clayton, Victoria, Australia

easily understandable, which limits their usefulness (*Chalmers and Glasziou, 2009*). Worse still, some data are used intentionally to obfuscate (*Ioannidis, 2016; Launer, 2020*). Obvious examples include 'nudging' and marketing – especially when done in the guise of 'educational meetings', or worse still, in the form of pseudo journal articles.

On the other hand, researchers and clinicians need to engage with society to make sure that their work is relevant to present-day problems, including those elements that make decisions, or fund research in a way that is transparent and minimizes the risk of bias (*Fauser and Macklon, 2019*).

Doctors care about their patients and want reasonable certainty that their treatments benefit the patient. Like all human beings, doctors are susceptible to influences that bias their judgement (*Fenton and Khan, 2011*). As Cochrane himself noted, we remember our successes more than our failures. Against this background, doctors are surrounded by 'evidence', some good, some not so good (*Moynihan et al., 2019*).

Evidence-based medicine requires that three elements be taken into account when caring for a patient; the evidence, the patient perspective and the doctor's experience (*Sackett, 1997*). All have to be considered, and all have to be aligned. Understanding the patient perspective requires an appreciation of the outcome that is important to the patient, as opposed to a pandering to whatever the patient articulates. This means applying the best evidence available to achieve that which is important, in partnership with the patient. Likewise, 'the doctor's experience' is often misused to justify whatever care the doctor intuitively feels to be right, independent of the evidence. The use of high-quality evidence is therefore critical to prevent yielding to a patient's unjustified wants, or a doctor's biased inclinations. Understanding data is not intuitive and requires training in evidence and critical appraisal, which can go some way to mitigating this bias (*Kulier et al., 2012*).

The doctor has to work through all of these influences to provide real care to patients, and will end up providing care in line with his or her own 'understanding of the evidence', which necessarily will have some bias. The problem arises when

doctors either succumb to commercial influences, or become wrongly convinced that they are right.

The good doctor therefore requires, in addition to empathy with the patient, an understanding of the evidence, a posture of humility and unshakable integrity (*McIntyre, 2019*).

Effect size and certainty

Sometimes, knowledge is readily apparent when an incidental treatment is successful. This is the case when the treatment (the signal) indicates an enormous improvement as compared with the natural course of the disease (the noise) (*Glasziou et al., 2007*): for example when Lesley Brown conceived after IVF and gave birth to Louise, this proved the treatment to be effective, as Mrs Brown had had a double-sided tubectomy several months earlier (*Steptoe and Edwards, 1992*). Effectiveness was confirmed, based on an 'all or nothing' rule, when more women without any natural conception chances had a baby after IVF.

In reality, however, such large signal-to-noise ratios are rare. The large majority of medical treatments generate, if effective, much smaller treatment benefits, and the success rates in the control groups are usually not zero. We then need additional tools to know the truth, including RCT and large cohort studies (of which RCT are a subtype establishing a cohort for follow-up by randomization at inception). Below, we discuss the arguments for both RCT and big data.

The features of RCT

First, *Macklon et al. (2019)* argue that RCT take a long time, and technology and patient needs have evolved by the time the results of an RCT become available. While that may be true, it need not necessarily be so. The reason that RCT take a long time and are expensive can be found in human failure. Recent RCT in reproductive medicine from China and Vietnam (*Chen et al., 2016; Shi et al., 2018; Vuong et al., 2018*) have been executed in short time periods, thus proving that the long time and high costs are not intrinsic to RCT and are caused by external factors.

Second, *Macklon et al. (2019)* argue that RCT primarily serve to remove treatment options from the clinician

and their patient. This of course is not true. The fact that many RCT report negative findings is not the fault of the RCT, but is because the therapies they test apparently do not work. This is to be expected, as treatment effects, as argued above, are overestimated (*Fenton and Khan, 2011*). In any event, knowing that an intervention does not work is just as important. Giving a treatment that does not work to a patient is harmful, either because it prevents an effective treatment being given, or it causes actual physical or financial harm.

Another criticism of RCT is that the patient population is so tightly defined that the results are often not applicable to the patient in the clinic, and that we need to consider not only the 'evidence', but the specific circumstances or 'context' of the patient (*Macklon and Fauser, 2020*). This is of course true, and as stated, the doctor's judgement in applying the evidence is a necessary prerequisite for the practice of evidence-based medicine. The issue, however, is how to change uncertainty to certainty. This is best done through RCT. A good example is that, despite extensive data gathering over decades as to whether progesterone treatment reduces the rate of miscarriage in patients with threatened miscarriage, there was still uncertainty. It was only when a well-conceived and well-conducted RCT was carried out, with pre-specified subgroups across a whole range of patient characteristics, that the answer was reached (*Coomarasamy et al., 2019a, 2019b*).

Aside from the excessive regulation around RCT, the most significant obstacle to carrying out RCT is the doctor. It seems that doctors, either consciously or subconsciously, do not want to contribute to trials. RCT provide a clear view on truth; this, for a variety of reasons, may be uncomfortable (*Nietzsche, 2015*). This comes down to culture. Paediatric oncology has long had a culture of embracing uncertainty and entering patients into well-run randomized controlled studies to bring about certainty (*De Vries et al., 2011*). This is likely to have contributed to the remarkable improvement in survival of childhood cancer, something that unfortunately cannot be said for IVF success rates, which have not improved over the past 5 years. Performing an RCT is like doing the dishes: if everyone does their small part, it is quickly achieved.

The features of 'big data'

Modern electronic medical records with structured questionnaires are an ideal source for generating large amounts of data. There are ongoing governance issues that need to be addressed, such as where the data are stored, how to keep it safe, who owns the data and who is allowed to access the data. These problems are solvable but, in practice, are open to vested interests.

The emphasis is usually on the 'big' rather than the 'data'. Accuracy of data input is variable, depending on the setting and the particular variable in question, and so interpreting the data and therefore any putative associations must be treated with caution. We have a problem with getting accurate 'small data', let alone big data.

Because the sample size with big data is large, it runs the risk of presenting a finding that is precise, but incorrect. This 'precision' can be manipulated to obfuscate reality, and subsequently exploited by commercial interests.

The cornerstone of interventional clinical trials is to have two equal groups of patients and then apply an intervention to one group and not to the other, without letting the patient switch to the other group. One problem with retrospective analysis such as big data is that it is not known if the two groups are equal. Another problem is that it often does not determine the point at which a patient started a treatment, i.e. intrauterine insemination and IVF might be started after different periods of infertility, which hampers comparison.

Big data is of course useful to detect associations, which can then be tested. There are many existing quality control databases that have validated and well-defined datasets which could be useful for generating hypotheses. Furthermore, when there is a change in policy, such as moving to single embryo transfer, then the effect can be easily and quickly seen through registries.

What are we trying to achieve?

We should distinguish two broad categories of clinical research: (i) research that observes what happened (observational studies, including those using routinely collected data) and (ii) research that determines what we should do (e.g. RCT or other types of health

technology assessments). The former tends to produce publications but usually has little or no translational value. The latter is the only type of research that from the outset aims to produce translational value. The challenge is to change the use of big data from (i) to (ii). If big data is used only to solve problems of precision, then let's not get fooled – when small effects are immaterial clinically, big data (by producing statistical significance) is the enemy, not a friend. However, when we are able to let the research question determine what to do, for example in cluster RCT, then big data could be highly valuable.

RCT and big data complement each other

RCT primarily serve to generate actionable evidence. Big data serves to generate real-world historical data. Both are important. There are many situations where conclusions from big data differ from conclusions from RCT on the same topic, but we will only know the reality when both studies are available (*Hemkens et al., 2016*). Just as results from big data need confirmation in smaller RCT, the impact of clinical actions informed by evidence generated by RCT need to be monitored in clinical practice, preferably through big data.

SUMMARY

Both big data and RCT are helpful. Big data is useful for looking at associations and generating testable hypotheses. It can, in some cases of large treatment effect, provide compelling evidence. However, big data also runs the significant risk of obfuscating the truth. RCT, on the other hand, reduce uncertainty to levels that justify the use of treatments in clinical practice. To date, reproductive medicine and its add-ons have largely developed independent from RCT (*Farquhar, 2019*). The logistics of carrying out RCT need not be prohibitive, and rather than discard them, the solution is to make them more part of everyday medical care.

In our opinion, clinicians working in reproductive medicine have been too ready to embrace the latest innovations without properly testing them. Consequently, we continue to have uncertainty, we have fewer patients available for clinical studies, and less time as clinicians to participate in them. That is the fundamental problem.

REFERENCES

- Bik, E., 2020. Thoughts on the Gautret et al. paper about Hydroxychloroquine and Azithromycin treatment of COVID-19 infections [WWW Document]. *Sci. Integr. Dig.* URL <https://scienceintegritydigest.com/2020/03/24/thoughts-on-the-gautret-et-al-paper-about-hydroxychloroquine-and-azithromycin-treatment-of-covid-19-infections/> (accessed 4.9.20).
- Chalmers, I., Glasziou, P. **Avoidable waste in the production and reporting of research evidence.** *Lancet* (London, England) 2009; 374: 86–89. doi:10.1016/S0140-6736(09)60329-9
- Chen, Z.J., Shi, Y., Sun, Y., Zhang, B., Liang, X., Cao, Y., Yang, J., Liu, J., Wei, D., Weng, N., Tian, L., Hao, C., Yang, D., Zhou, F., Shi, J., Xu, Y., Li, J., Yan, J., Qin, Y., Zhao, H., Zhang, H., Legro, R.S. **Fresh versus frozen embryos for infertility in the polycystic ovary syndrome.** *N. Engl. J. Med.* 2016; 375: 523–533. doi:10.1056/NEJMoa1513873
- Collins, R., Bowman, L., Landray, M., Peto, R. **The Magic of Randomization versus the Myth of Real-World Evidence.** *N. Engl. J. Med.* 2020; 382: 674–678
- Coomarasamy, A., Devall, A.J., Brosens, J.J., Quenby, S., Stephenson, M.D., Sierra, S., Christiansen, O.B., Small, R., Brewin, J., Roberts, T.E., Dhillon-Smith, R., Harb, H., Noordali, H., Papadopoulou, A., Eapen, A., Prior, M., Carlo Di Renzo, G., Hinshaw, K., Mol, B.W., Ann Lumsden, M., Khalaf, Y., Shennan, A., Goddijn, M., van Wely, M., Al-Memar, M., Bennett, P., Bourne, T., Rai, R., Regan, L., Gallos, I.D., Tommy, F., 2019a. Micronized vaginal progesterone to prevent miscarriage: a critical evaluation of randomized evidence. <https://doi.org/10.1016/j.ajog.2019.12.006>
- Coomarasamy, A., Devall, A.J., Cheed, V., Harb, H., Middleton, L.J., Gallos, I.D., Williams, H., Eapen, A.K., Roberts, T., Ogwulu, C.C., Goranitis, I., Daniels, J.P., Ahmed, A., Bender-Atik, R., Bhatia, K., Bottomley, C., Brewin, J., Choudhary, M., Crosfill, F., Deb, S., Duncan, W.C., Ewer, A., Hinshaw, K., Holland, T., Izzat, F., Johns, J., Kriedt, K., Lumsden, M.A., Manda, P., Norman, J.E., Nunes, N., Overton, C.E., Quenby, S., Rao, S., Ross, J., Shahid, A., Underwood, M., Vaithilingam, N., Watkins, L., Wykes, C., Horne, A., Jurkovic, D. **A randomized trial of progesterone in women with bleeding in early pregnancy.** *N. Engl. J. Med.* 2019; 380: 1815–1824. doi:10.1056/NEJMoa1813730
- De Vries, M.C., Houtlosser, M., Wit, J.M., Engberts, D.P., Bresters, D., Kaspers, G.J., Van Leeuwen, E. **Ethical issues at the interface of clinical care and research practice in pediatric oncology: A narrative review of parents' and physicians' experiences.** *BMC Med. Ethics* 2011; 12: 18. doi:10.1186/1472-6939-12-18
- Farquhar, C. **Introduction: Add-ons for assisted reproductive technology: can we be honest here?** *Fertil. Steril.* 2019; 112: 971–972. doi:10.1016/j.fertnstert.2019.10.010
- Fausser, B.C.J.M., Macklon, N.S. **May the colleague who truly has no conflict of interest now please stand up!** *Reprod. Biomed. Online* 2019; 39: 541–544. doi:10.1016/j.rbmo.2019.09.001
- Fenton, J.J., Khan, K.S. **The intuitive appeal of case series thinking: a challenge for evidence-**

- based teaching and practice.** *Evid. Based. Med.* 2011. doi:10.1136/ebm-2011-100073
- Gautret, P., Lagier, J.-C., Parola, P., Hoang, V.T., Meddeb, L., Mailhe, M., Doudier, B., Courjon, J., Erie Giordanengo, V., Vieira, V.E., Hervé, H., Dupont, H.T., Sté, S., Honoré, S., Honoré, H., Colson, P., Chabri, E., Scola, B.L., Rolain, J.-M., Brouqui, P., Raoult, D. **Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial.** *Int. J. Antimicrob. Agents.* 2020. doi:10.1016/j.ijantimicag.2020.105949
- Glasziou, P., Chalmers, I., Rawlins, M., McCulloch, P. **When are randomized trials unnecessary? Picking signal from noise.** *BMJ* 2007; 334: 349–351. doi:10.1136/bmj.39070.527986.68
- Hemkens, L.G., Contopoulos-Ioannidis, D.G., Ioannidis, J.P.A. **Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: Meta-epidemiological survey.** *BMJ* 2016; 352. doi:10.1136/bmj.i493
- Human Fertilisation & Embryology Authority, 2019. Treatment add-ons [WWW Document]. URL <https://www.hfea.gov.uk/treatments/explore-all-treatments/treatment-add-ons/> (accessed 1.4.20).
- Ioannidis, J.P.A. **The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses.** *Millbank Q* 2016; 94: 485–514. doi:10.1111/1468-0009.12210
- Kulier, R., Gülmezoglu, A.M., Zamora, J., Plana, M.N., Carroli, G., Cecatti, J.G., Germar, M.J., Pisake, L., Mittal, S., Pattinson, R., Wolomy-Molondo, J.J., Bergh, A.M., May, W., Souza, J.P., Koppenhoefer, S., Khan, K.S. **Effectiveness of a clinically integrated e-learning course in evidence-based medicine for reproductive health training: A randomized trial.** *JAMA–J. Am. Med. Assoc.* 2012; 308: 2218–2225. doi:10.1001/jama.2012.33640
- Launer, J. **The production of ignorance.** *Postgrad. Med. J.* 2020. doi:10.1136/postgradmedj-2020-137494
- Macklon, N., Fauser, B. **Context-based infertility care.** *Reproductive BioMedicine Online* 2020. doi:10.1016/j.rbmo.2019.12.001
- Macklon, N.S., Ahuja, K.K., Fauser, B.C.J.M. **Building an evidence base for IVF 'add-ons.'** *Reprod. Biomed. Online* 2019; 38: 853–856. doi:10.1016/j.rbmo.2019.04.005
- McIntyre, L.C. 2019 **The scientific attitude: defending science from denial, fraud and pseudoscience.** *The MIT Press Cambridge MA*
- Moynihan, R., Bero, L., Hill, S., Johansson, M., Lexchin, J., MacDonald, H., Mintzes, B., Pearson, C., Rodwin, M.A., Stavdal, A., Stegenga, J., Thombs, B.D., Thornton, H., Vandvik, P.O., Wieseler, B., Godlee, F. **Pathways to independence: Towards producing and using trustworthy evidence.** *BMJ* 2019; 367: 1–5. doi:10.1136/bmj.l6576
- Nietzsche, F., 2015. *Thoughts Out of Season, Part Two, The Use and Abuse of History.* Complete Works of Friedrich Nietzsche., Kindle. ed. Delphi Classics.
- Sackett, D.L. 1997 **Evidence-based medicine.** *Seminars in Perinatology* W.B. Saunders: 3–5. doi:10.1016/S0146-0005(97)80013-4
- Shi, Y., Sun, Y., Hao, C., Zhang, H., Wei, D., Zhang, Y., Zhu, Y., Deng, X., Qi, X., Li, H., Ma, X., Ren, H., Wang, Y., Zhang, D., Wang, B., Liu, F., Wu, Q., Wang, Z., Bai, H., Li, Y., Zhou, Y., Sun, M., Liu, H., Li, J., Zhang, L., Chen, X., Zhang, S., Sun, X., Legro, R.S., Chen, Z.J. **Transfer of fresh versus frozen embryos in ovulatory women.** *N. Engl. J. Med.* 2018; 378: 126–136. doi:10.1056/NEJMoa1705334
- Steptoe, P.C., Edwards, R.G. **Birth after the reimplantation of a human embryo.** *Arch. Pathol. Lab. Med.* 1992. doi:10.1016/s0140-6736(78)92957-4
- Vuong, L.N., Dang, V.Q., Ho, T.M., Huynh, B.G., Ha, D.T., Pham, T.D., Nguyen, L.K., Norman, R.J., Mol, B.W. **IVF transfer of fresh or frozen embryos in women without polycystic ovaries.** *N. Engl. J. Med.* 2018; 378: 137–147. doi:10.1056/NEJMoa1703768
- Wilkinson, J., Brison, D.R., Duffy, J.M.N., Farquhar, C.M., Lensen, S., Mastenbroek, S., van Wely, M., Vail, A. **Don't abandon RCTs in IVF. We don't even understand them.** *Hum. Reprod.* 2019: 1–6. doi:10.1093/humrep/dez199

Received 24 February 2020; received in revised form 15 April 2020; accepted 25 April 2020.