In the future simulations will replace clinical trials

Martin A. Ebert[1,2] · Val Gebski[3] · Clive Baldock[4]

Introduction and overview: Clive Baldock, moderator

Since the coronavirus surfaced in Wuhan in 2019, there has been much progress in the diagnosis and treatment of the COVID-19 disease with significant developments of new vaccines [1]. An important component of these developments has been the need to undertake clinical trials with many simultaneously being undertaken globally. This has presented challenges with the need for the rapid development of COVID vaccines at an unprecedented speed and in unprecedented numbers [2].

In this very topical and timely debate Martin Ebert and Val Gebski debate whether in the future simulations will replace clinical trials.[1]

Arguing for the proposition is Prof. Martin Ebert. After failing to gain entry into medicine, Prof. Ebert decided to bide his time and started studying physics only to discover that, not only was physics an extremely enjoyable and rewarding subject, but that it was also a back door into medicine. He has subsequently enjoyed nearly thirty years working both as a clinical and research medical physicist, with most of that time spent at the Sir Charles Gairdner Hospital and University of Western Australia. One focus of Prof. Ebert's research has been collection and analysis of clinical trial data and he actively works with TROG Cancer Research and trial investigators to collect data that is invaluable for subsequent translation. However, he can taste the revolutions coming in computing

and data analytics and is actively advocating for physicists to apply their unique perspectives and creativity to rethinking medicine. Prof. Ebert finds the best place to contemplate the future of humanity is when out in nature and spends quality thinking-time hiking, cycling and surfing. Since the start of 2020, this has largely been restricted to the unique and pristine settings of Western Australia.

Martin Ebert

Arguing against the proposition is Prof. Val Gebski who is an international leader in clinical trials design, implementation and interpretation. He is an exceptional clinical biostatistician with a thorough clinical understanding of research problems together with the methodological and statistical underpinning of the underlying science. As a biostatistician for over 35 years, he is regarded as a key resource and authority in research design by national and international clinical trials groups.

A chief investigator on five National Health and Medical Research Council (NHMRC) Program Grants, Prof. Gebski works closely with collaborative clinical trials groups and individual investigators. He plays a significant role with respect to both research ideas and concepts, and established projects. He has attracted multiple research grants being co-awarded over $60 M.

✉ Clive Baldock
  C.Baldock@westernsydney.edu.au

1 Department of Radiation Oncology, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia

2 School of Physics, Mathematics and Computing, University of Western Australia, Crawley, WA 6009, Australia

3 NHMRC Clinical Trials Centre, University of Sydney, Camperdown, NSW 2050, Australia

4 Graduate Research School, Western Sydney University, Penrith, NSW 2747, Australia

[1] Contributors to Topical Debates are selected for their knowledge and expertise. Their position for or against a proposition may or may not reflect their personal opinions.

Considered an expert in the area of competing risk and methods relating to problems with time to event outcomes (survival analysis), Prof. Gebski's work has resulted in these methods being adopted as standard analysis approaches. He has led the methodological aspects of non-inferiority studies evaluating laparoscopic surgery (rectal, endometrial, cervix), the results being practice changing. His work and expertise in the meta-analysis of esophageal and breast cancer trials has resulted in global change in clinical management of these diseases. He leads in methodological areas related to data maturity, and recurrent time to event models with ordinal and multinomial responses and computational efficiency of models in Genome Wide Association Studies (GWAS).

Prof Gebski is the Group statistician for three collaborative cancer trials groups (Breast Cancer Trials, ANZ Gynaecological Oncology Group, Australasian Gastro-Intestinal Trials Group) as well as a key methodological leader in trials lead by international groups (US, Europe, Asia–Pacific). He is the statistical editor of the ANZ Journal of Surgery, an associated editor of the Journal of Pharmaceutical Statistics and a member of the editorial board of the Journal of Clinical Oncology (JCO). He has an honorary fellowship from Royal Australian and New Zealand College of Radiologists (RANZCR).and is a member of the Australasian Kidney Trials Network Advisory Board and the NHMRC Grants Management Solution Working Group.

Val Gebski

Opening statement: Martin Ebert

Practicality will drive virtual trials[2]

Virtual trials avoid the practical problems that plague in vivo trials, which must be powered to answer very specific questions. A need for equipoise must be balanced with cost and

---

[2] The term *virtual trial* is being used rather than **in silico** *trial*, often found in the literature, to accommodate future computational approaches that are not based on silicon.

the availability of participants. They must navigate complex and potentially-biasing ethics, patient selection and consent processes [3]. Analysis and interpretation of in vivo trials are hampered by losses in fidelity and quality that result from manual reporting and data collation. In vivo trials are dependent on maintenance of social and economic stability over lengthy periods. The detrimental consequences have been highlighted in the context of COVID-19; in the months required for assessment of vaccines, the world saw increasing rates of suffering, death, societal and economic disaster.

Demand will drive virtual trials

It is broadly acknowledged that COVID-19 has provided just a sampling of future pandemics with existential implications and has accelerated antimicrobial multidrug resistance through the over-use of antibiotics [4]. The urgency to rapidly assess vaccines and drugs in an ongoing manner demands revision of the trial process. Computational models are being applied now to accelerate the process of drug discovery [5, 6]. The demand for rapid drug development parallels rapid technological evolution, and virtual trials have already been employed for the purpose of regulation [7].

Feasibility will drive virtual trials

It may be difficult to conceive of simulations which could accommodate the complexity of physiology, biochemistry, multi-omics and cognition at multiple scales. Yet efforts are underway to make such simulations viable and accurate. 130,000 processor cores have recently been used to simulate a single gene locus at 1 physical nanosecond per day of simulation [8]. Extensible, multi-scale and novel computational models will enable upscaling to a full individual. Such solutions are being pursued by consortia such as the Virtual Physiological Human project [9], and commercialised by emerging companies such as ELEM (https://elem.bio).

Fecundity will drive virtual trials

In contrast to in vivo trials, the precision of virtual trials can be scaled as required. Accuracy can be continuously improved via feedback and learning processes. Additionally, simulations can be weighted according to primary intention, allowing, for example, a bias towards identification of potential adverse events. Simulations could be performed to accommodate diversity of almost unlimited populations or made specific to an individual. Simulations enable exploration of arbitrary adaptations to trial arms, adjuvant treatments, comorbidities and patient phenotype, impossible with in vivo trials. They provide multi-scale information on processes that can be used to design new targeted treatments. Simulations enable arbitrary variation of participant

environment (could we ever power a physical trial for assessment of drug efficacy in space?), including socio-economic context.

Given the inexorable march of science towards processes based on computation, the replacement of in vivo trials with virtual ones is essentially a fait accompli. In vivo trials will become unethical when replaceable with a virtual equivalent.

Opening statement: Val Gebski

Clinical trials are essential in forging advances in modern medicine. They cannot be replaced by methods offering convenience over rigour, expedience over substance.

In the physical sciences, particularly engineering, mathematics and physics, computer simulations have been used to obtain and understand solutions to analytically intractable problems. Outside very precise boundaries, how successful have simulations been to advancing knowledge? They have yielded notably disastrous results in economics, finance and climatology, being less informative than asking a starfish for directions. However, I like starfish, they are quite good looking and aim to be helpful.

In the clinical setting, simulations and modelling have been the basis of predictions of the impact of COVID-19 on spread, mortality and medical equipment resources such as ICU respirator usage. Even from some of the most expert applied mathematical modeler's, the modelling results have been disastrous [10], with predictions of up to 150,000 deaths and between 5 M-15 M (20%-60%) of the population being infected [11]. As of 9/06/2021 Australia has 30,210 cases and 910 deaths [12] with many of the deaths being as a result of policy decisions rather than the medical management assumed in the simulated models [13, 14].

In biostatistical methodology, resorting to simulation to understand behaviour under certain assumptions is now commonplace although we need to distinguish between simulated data on hypothetical populations as opposed to methods applied to observed data. The key here is generalizability. Simulated models follow defined processes heavily dependent on assumptions, regardless of how well they represent actual populations. Patients entering clinical trials however, are not. *No* amount of modelling could predict the result of the Laparoscopic Approach to Cervical Carcinoma (LACC) trial, that laparoscopic surgery demonstrated a survival detriment [15] or ILLUMINATE heart failure trial showing a 60% mortality rate increase, despite a reduction of 13% in LDL and 12% in triglycerides for Torcetrapib [16].

So, will simulation studies demonstrate such monumental treatment failures as well as the potential benefits? Absolutely *not*, unless the mechanism of action and the clinical process of benefit are thoroughly understood. In these trials, reasons for failure are still not well understood to this day, and these are results from actual patients! It is trial data, from actual patients, which continually improves this understanding (albeit sometimes only by minute amounts) processes and mechanisms of action. In clinical study design a key component in sample size determination is the success rate estimates in the standard arm. However, from completed studies, there is a huge variability in this estimate to that originally assumed. Given the difficulty of accurately determining estimates when actual data exists, validity of simulations and their result, never confirmed or tested on actual patients would forever go unquestioned.

Implicitly trusting simulations, lulled into a realm of dynamic inactivity, putting trust in the clinical wisdom of computer, mathematical, and statistical modellers while ignoring the critical component of scientific inquiry – real world observations A program of mass psycho-immunization for researchers and clinicians into acceptance of this new religion would need to instigated. The biggest losers in such a system will be patients!

Rebuttal: Martin Ebert

Prof Gebski's perspective has been born out of his tireless efforts to better our lives. For both his perspective and efforts I am extremely grateful.

Prof Gebski's reference to inaccurate model predictions for COVID-19 is a questionable exemplar of poor model performance. In this context, can we distinguish between a model that is inaccurate from an accurate one that has succeeded in its job of subsequently impacting outcomes?

As with many past models, many clinical trials have failed to translate. The artificial contexts in which trials must operate and their frequent use of biased outcomes measures have led to substantial translation gaps [17, 18]. In addition, trials depend on the weak law of large numbers [19] to predict an average outcome for a population. But I want to know what *my* outcome will be. Like my esteemed colleague, I too am fond of starfish though highlight that their lack of a brain (the starfish that is) means they cannot even pretend to want to be helpful. Relative to the clumsy models applied in the past, future models will appear as agile dolphins. They shall adapt via constant real-world feedback processes and their accuracy for predicting individual outcomes progressively increase.

My arguments have particularly diverged from those of Prof Gebski in terms of *how* trials might be simulated. Such simulations will not appear as models as we now know them, but develop in response to revolutions in computing. Exascale or quantum computing will enable algorithms that we are yet to conceive of, or that we have conceived but currently cannot execute (such as complete neural cortex

simulation, not achievable with current computing capability [20]). Such simulations will focus on the two extremes that in vivo trials have no hope of approaching: $N \to \infty$ (for discovery) and $N = 1$ (for personalisation).

Rebuttal: Val Gebski

The argument for replacing *actual* clinical trials with simulated ones suggest the most important aspects for advancement of medicine are (a) practicality; (b) demand for virtual models (by whom?); (c) the proliferation/speed with which these simulated reports can choke up medical journals. Issues such as scientific enquiry; patient benefit; actual applicability/generalizability of results do not appear to matter. The exposé by Prof Ebert centres around computational models which explain *observed* data and attempts to predict outcomes for future patients. While this is not unreasonable, it is somewhat presumptuous to now suggest that simulated trials, based on artificial data (the title of this debate), have complete relevance in predicting outcomes of future patients.

A plenary session of the American Society of Clinical Oncology Annual Meeting (1999), presented four trials evaluating high-dose chemotherapy (HDC) for breast cancer. Clinical benefit was only observed in one study but with further scrutiny, 100% of patients in the control and 23% in the HDC groups were found to be 'simulated' [21, 22]. Had not the other studies with 100% *actual* patients been available for comparison, ineffective/toxic HDC would now be the standard of care. Artificial data however is not just confined to oncology with COVID-19 treatments not escaping the eagle eye of data trawlists to create hybrid models and present these as being accurate [23, 24]. The only consolation with simulated trials is that without question, *they are all fake*! Unfortunately, their results would be then be represented as *truth* by both politicians and a hysterical media pack as we have currently witnessed in many other aspects which impact on our lives—lockdowns, curfews, climate etc., etc. Are we now about to enter an era of "the emperor's new clothes"?

Advancement in medical research is also driven by study failures. Medicine has made major leaps through learning from promising ideas failing to demonstrate benefit. The STEADFAST phase III trial of *azeliragon* in 475 patients with mild Alzheimer's disease, failed to show any benefit for major clinical endpoints. A post-hoc subgroup analysis (those with type II diabetes) did show some benefit for azeliragon—this now requiring evaluation in a future study (underway) on actual patients. A simulated trial would completely mask any potential benefit in this subgroup.

Simulations can be useful to gain insights into therapies only when we know/understand the mechanisms of action.

Otherwise, they yield results based on non-existent information having no/limited interpretation and no generalizability.

References

1. Davis JS, Ferreira D, Denholm JT et al (2020) Clinical trials for the prevention and treatment of COVID-19: current state of play. Med J Aust 213:86–93
2. Seidler AL, Aberoumand M, Williams JG et al (2021) The landscape of COVID-19 trials in Australia. Med J Aust 215:58-61.e1
3. Kho ME, Duffett M, Willison DJ et al (2009) Written informed consent and selection bias in observational studies using medical records: systematic review. BMJ 338:866
4. Afshinnekoo E, Bhattacharya C, Burguete-García A et al (2021) COVID-19 drug practices risk antimicrobial resistance evolution. Lancet Microbe 2:e135–e136
5. Zhavoronkov A, Ivanenkov YA, Aliper A et al (2019) Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nat Biotechnol 37:1038–1040
6. Mohapatra S, Nath P, Chatterjee M et al (2020) Repurposing therapeutics for COVID-19: rapid prediction of commercially available drugs through machine learning and docking. PLoS ONE 15:0241543
7. Badano A, Graff CG, Badal A et al (2018) Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an in silico imaging trial. JAMA Netw Open 1:e185474–e185474
8. Jung J, Nishima W, Daniels M et al (2019) Scaling molecular dynamics beyond 100,000 processor cores for large-scale biophysical simulations. J Comput Chem 40:1919–1930
9. Viceconti M, Henney A, Morley-Fletcher E (2016) In silico clinical trials: how computer simulation will transform the biomedical industry. Int J Clinical Trials 3:37–46
10. McGuinness P (2020) No model answers for this crisis. Australian Financial Review. Accessed 26 May 2020
11. McCauley D, Bagshaw E, Harris R (2020) Australia prepares for 50,000 to 150,000 coronavirus deaths. Sydney Morning Herald. Accessed 16 Mar 2020
12. Johns Hopkins Coronavirus Resource Centre. https://coronavirus.jhu.edu/map.html. Accessed 9 Jun 2021
13. McCauley D, Duke J (2020) Thousands of predicted COVID-19 deaths never eventuated - was it poor modelling or our response? Sydney Morning Herald. Accessed 27 May 2020
14. University of Sydney (2020) US COVID-19 deaths "poorly predicted" by IHME model. https://www.sydney.edu.au/news-opinion/news/2020/04/14/us-covid-19-deaths-poorly-predicted-by-ihme-model.html. Accessed 1 Oct 2021
15. Ramirez PT, Frumovitz M, Pareja R et al (2018) Minimally invasive versus abdominal radical hysterectomy for cervical cancer. N Engl J Med 379:1895–1904
16. Pearson H (2006) Why did a promising heart drug fail? Nature. https://doi.org/10.1038/news061204-8
17. Heneghan C, Goldacre B, Mahtani KR (2017) Why clinical trial outcomes fail to translate into benefits for patients. Trials 18:122
18. Stuart EA, Bradshaw CP, Leaf PJ (2015) Assessing the generalizability of randomized trial results to target populations. Prev Sci 16:475–485
19. Bernoulli J (1713) Ars conjectandi, opus posthumum: accedit tractatus de seriebus infinitis, et epistola Gallice scripta de ludo pilæ reticularis. Basileae: Impensis Thurnisiorum Fratrum
20. Jordan J, Ippen T, Helias M et al (2018) Extremely scalable spiking neuronal network simulation code: from laptops to exascale

computers. Front Neuroinform. https://doi.org/10.3389/fninf.2018.00002

21. Weiss R, Rifkin R, Stewart M et al (2000) High-dose chemotherapy for high-risk primary breast cancer: an on-site review of the Bezwoda study. Lancet 355:999–1003

22. Bezwoda 1985 Breast Cancer Transplant Study Fraudulent. Oncology NEWS International, Oncology NEWS International 10(6). https://www.cancernetwork.com/view/bezwoda-1985-breast-cancer-transplant-study-fraudulent. Accessed 1 Oct 2021

23. Mehra M, Desai S, Kuy S (2020) Cardiovascular disease, drug therapy, and mortality in Covid-[19]. N Engl J Med. https://doi.org/10.1056/NEJMoa2007621

24. Mehra M, Desai S, Ruschitzka F et al (2020) Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. Lancet. https://doi.org/10.1016/S0140-6736(20)31180-6