EBioMedicine 69 (2021) 103461

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Commentary

An ethically-motivated, Bayesian, adaptive design clinical trial bringing hope to women with menorrhagia. . .and warmth to statisticians' hearts

Chris R Palmer

Cambridge, UK

ARTICLE INFO

Article History:
Received 11 June 2021
Accepted 11 June 2021
Available online xxx

In this issue of *EBioMedicine* [1] Warner *et al.* report on a fascinating study, DexFEM, assessing dexamethasone for menorrhagia. Their dose-optimising research may help women enduring heavy menstrual bleeding and inform the doctors treating them. It will also be warmly welcomed by research methodologists in clinical trials by exemplifying a non-traditional *Bayesian, adaptive* design. Italicised terms are lesser-spotted and less commonly understood within the applied biomedical literature and hence will be the focus of this commentary.

First, Bayesianism and frequentism are contrasting schools of statistical thought, the latter being the classical norm. Both are perfectly valid but arrive at their statistical conclusions very differently. Frequentists define the probability a flipped coin lands heads 0.5 as the theoretical ratio of heads to tosses in a hypothetical, unending sequence. Bayesians, reckoning two sides of the coin are equally likely, just believe the same value, $^1/_2$. The Bayesian view of probability more naturally suits clinical trials, which are the cornerstone of evidence-based medicine, as one-off experiments. They also differ on which is fixed and which is a random variable—the population mean or the sample mean—with Bayesians invoking "prior distributions" on the former, updated by the evidence observed (i.e. data) into "posterior distributions" for making inferences from sample to population. In turn, Bayesian analyses are more intuitively interpreted, unlike p-values or confidence levels, which are convoluted for frequentists to explain, requiring considerable mental gymnastics. The DexFEM trial sensibly utilises a neutral, "uninformative prior" to become fully swamped by accruing data. The paper provides some justification and terminology, including a supplementary file on introduction to Bayesian concepts, but for further understanding of Bayesian thinking, see elsewhere, e.g. [2].

Secondly, adaptive design trials are those that plan pre-specified changes in the light of unfolding data collected during a study. DexFEM had built-in adaptations scheduled after 33 and 66 patients. It did not, but could have chosen to, adapt on overall sample size as well, as happens in sequential trials, another subtype of data-dependent design. Interestingly, with propitious timing in December 2019, US FDA released guidance [3] that should encourage more use of adaptive designs (speculatively more urgently post-pandemic and entirely coincidental that dexamethasone is prescribable for COVID-19). Today's trials trace their history to testing fertilisers in a 1926 English crop trial, when randomisation was pioneered. Human trials can be more sensible and flexible than agricultural experiments by exploiting obvious facts that patients enter trials and their outcome data arise in real time, without the need to await harvest time before learning from data. Also, in clinical trials, note experimental units are people, making ethical concerns paramount. Adaptive trials begin as sophisticated mathematical exercises, only feasible in the modern information age, with computer simulations conducted to explore design features, as was done at considerable effort (about a million program runs over six weeks of computer time alone) ahead of DexFEM [4]. One beauty of conducting simulations is that only hypothetical "imaginary patients" get harmed, whereas real patients can actually benefit from the intensive and extensive mathematical modelling. Women were preferentially randomised during DexFEM to receive the better performing dexamethasone dose levels to home in efficiently on optimal dosage. As a concomitant ethical advantage, participants were steered away from less promising dosages. Antithetical to the misleading practice of *post hoc* data-dredging, this sort of thinking and pre-planning exemplifies intelligent, learn-as-you-go, data-dependent design, representing the best that statistical theory can offer the world of medicine. Adaptive designs are abundant in theory yet remain scarcely implemented in actual clinical practice [5, 6]. Simulations beforehand and the complex analyses required to inform adaptation during a clinical trial's progress depend thoroughly on modern computing and communication technologies, which have arisen and advanced astonishingly over the past century. Some might argue it is scandalous that the vast majority of trials today are conducted largely oblivious to the opportunities offered by these amazing developments, a deprivation akin to taking telescopes away from astronomers. Just imagine the outcry if today's medical doctors were equally constrained to prescribe treatments only from among those already known back in the 1920's!

The question one must ask is, "Was it really worth the immense, extra statistical effort put into the design and conduct of DexFEM compared to simply keeping equal allocation of dose levels without adaptation?" Allow ethical thinking to help answer. Clinical trials are a delicate balance between individual ethics (putting foremost

DOI of original article: http://dx.doi.org/10.1016/j.ebiom.2021.103434.
E-mail address: crpalmer153@gmail.com

https://doi.org/10.1016/j.ebiom.2021.103461

interests of *current* patients within a trial) and collective ethics (considering primarily *future* patients standing to benefit from trial results) [7]. In research matters, the Declaration of Helsinki concurs with newly-qualified doctors' Hippocratic Oath pledging to prioritise current patients. Individual ethics' concerns are heightened in trials involving conditions that are relatively rare, serious or worse fatal and investigating new treatments, in contrast to more prevalent, less serious conditions and involving safety-profiled treatments. Frequentism (and large, fixed sample size trials) seems better matched to collective ethics, whereas Bayesianism (and adaptive designs) better suited to individual ethics [8]. Dexamethasone is far from novel, menorrhagia is moderately common worldwide and DexFEM did not have patient outcome data entries of '0's and '1's corresponding to 'alive' and 'dead'. This all suggests an honest answer to the question just posed is, "Not really, but it's a terrific example for even more relevant situations".

The multidisciplinary DexFEM study team is nonetheless to be congratulated for their triumphant efforts: methodological contributors for producing an excellent case study that will be gleefully received and much cited in the medical statistics literature; and clinical collaborators for successfully implementing the design after consulting the statisticians long before the first patient was randomised. This was so much better than deferring statistical input until after harvesting final data endpoints, a sadly common characteristic of poor medical research [9]. The DexFEM trial, even if not quite the perfect application of Bayesian, adaptive design theory, is commendably of the highest methodological quality.

About the Author

Chris Palmer MA(Oxon) MA(Cantab) MS PhD is an academic medical statistician (now early semi-retired) with a longstanding, personal desire to see best statistical thinking applied in clinical medicine and whose PhD dissertation developed a theoretical example of another ethically-motivated, adaptive, Bayesian clinical trials model (1988, UNC-Chapel Hill). He was post-doctoral Research Fellow at Harvard's Biostatistics Department (1988-89); has been Statistical Reviewer for *The Lancet* family of journals (1993-present) and for *Human Reproduction* (2006-present); and served as Deputy Editor, *Statistics in Medicine* (1996-2000). He was founding Director, Centre for Applied Medical Statistics, University of Cambridge (1996-2014) and was Co-editor of *Encyclopaedic Companion to Medical Statistics* [10], contributing thereto pertinent articles including: Consulting a statistician; Data-dependent designs; Ethics and clinical trials and Probability.

Declaration of Competing Interest

None to declare.

References

[1] Warner P, Whitaker LHR, Parker RA, Weir CJ, Douglas A, Hansen CH, et al. Low dose dexamethasone as treatment for women with heavy menstrual bleeding: a response-adaptive randomised placebo-controlled dose-finding parallel group trial (DexFEM). EBioMedicine 2021. doi: 10.1016/j.ebiom.2021.103434.
[2] Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. Chichester: Wiley; 2004.
[3] FDA guidance on adaptive designs. Center for Biologics Evaluation and Research, Center for Drug Evaluation and Research. 2019 Dec. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adaptive-design-clinical-trials-drugs-and-biologics-guidance-industry. Accessed 7 June 2021
[4] Hansen CH, Warner P, Walker A, Parker RA, Whitaker L, Critchley HOD, Weir CJ. A practical guide to pre-trial simulations for Bayesian adaptive trials using SAS and BUGS. Pharm Stat 2018;17(6):854–65 Nov.
[5] Palmer CR. Ethics, data-dependent designs, and the strategy of clinical trials: time to start learning-as-we-go? Stat Methods Med Res 2002; 11:381–402.
[6] Pallmann P, Bedding AW, Choodari-Oskooei B, Dimairo M, Flight L, Hampson LV, et al. Adaptive designs in clinical trials: why use them, and how to run and report them. BMC Med 2018 Feb;16(1):29.
[7] Pocock SJ. Clinical Trials: A Practical Approach. Chichester: Wiley; 1983.
[8] Palmer CR. Ethics and statistical methodology in clinical trials. J Med Ethics 1993;4(19):219–22.
[9] Altman DG. The scandal of poor medical research. BMJ 1994;308:283.
[10] Everitt BS, Palmer CR. (Eds). Encyclopaedic Companion to Medical Statistics. 2nd ed. Chichester: Wiley; 2011.