ELSEVIER

Journal of Clinical Epidemiology 113 (2019) vi−viii

Journal of Clinical Epidemiology

EDITORIAL

Trial recruitment a continuing challenge

Over half of all controlled trials either take much longer to reach their sample size or fail to achieve them at all. Small but insufficient improvements of recruitment have been shown for a] telling people what they are receiving in the trial rather than not telling them and b] phoning people who do not respond to a postal invitation [1]. It is encouraging that a number of initiatives are addressing this challenge. Delaney et al report a systematic review of the effectiveness of education and training interventions for recruiters to trials. The review included randomized and non-randomized controlled trials of any type of education and training intervention for recruiters to trials, within any health care field, reporting recruitment rates, quality of informed consent, recruiter self-confidence, understanding/knowledge of trial information, numbers of potential trial participants approached, satisfaction with training, and retention rates. But the results were disappointing with only six studies meeting the eligibility criteria. The authors call for priority to be given to establish the evidence base around the effectiveness of education and training interventions for recruiters to trials and the establishment of a standardized set of core outcome measures in studies of trial recruitment education and training interventions. A second paper studied factors affecting trial recruitment in over 2,500 men with clinically localized prostate cancer offered randomisation for a trial of four options: monitoring, surgery, radiotherapy, and brachytherapy. Sixty two percent agreed to be randomised, while 38% rejected randomization primarily because they had a preference for one of the intervention arms. Another 22% refused the allocation after randomisation. In view of these disappointing results, the authors conclude that interventions need to be developed or extended to support recruiters to assess equipoise and patient preferences. A third paper reports on the use of a trial design to improve trial recruitment, namely the ‘Trials within Cohort’ design that employs a ‘staged-informed consent’ where over 5,000 patients randomized to the intervention were offered the intervention, and the same number of controls were not informed. This design improved logistics in recruitment and prevented contamination, but the rate of refusal of the intervention was still high.

Many disciplines have shown improvement in the methods and results in increasing sample sizes to reduce type 2 errors [2]. This has been thought to be in part due to the increased focus on methodological rigor and implementation of research guidelines such as the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) [15] and the Consolidated Standards of Reporting Trials (CONSORT) statements. Anthon et al report that surprisingly this has not occurred in ICU trials between 1977 and 2018 with the exception of better reporting of allocation concealment and an increase in the size and number of the largest RCTs, especially in recent years.

Subgroup analysis of interactions (addressing e.g., whether the effect of an intervention varies depending on patient characteristics such as age or on intervention characteristics such as dose) in large randomized trials are often highlighted in the conclusions and recommendations for clinical practice. However, some of these effect modifications are not credible and indeed may be spurious, and, if acted on, can be detrimental to patient care [3]. Schandelmaier et al found 36 different methods for effect modification. The most frequent were tests for interaction, a priori hypotheses, providing a causal explanation, accounting for multiplicity, and testing a small number of effect modifiers. Further work is needed to review the rationale for their credibility before recommendations can be made on a minimum set of criteria for reporting guidelines. A second paper on effect modification by Lopez et al suggests that one good option for demonstrating the rationale for effect modification is using Selection Diagrams. These are a form of Directed Acyclic Graphs that specifically include selection factors, that is, background factors or attributes that vary by population (or population subgroup), by which structural discrepancies between the two populations are suspected to take place. The authors demonstrate this with examples such as the mediating effect of melatonin in the skin, impacting on the reduction in the risk of skin cancer from using sunscreen.

Controlled trials of surgical interventions lag well behind pharmaceutical and behavioral interventions, regarding both frequency and methodology. One of the major challenges in conducting trials of surgical interventions is the representativeness and generalisability of the level of expertise of the surgical team and how to handle clustering effects of groups of patients treated by the same surgeon. Conroy et al report that a previously underappreciated quality assurance mechanism is the role of funders. In a sample of 50 trials of surgical interventions they document how UK funders have played a major role in ensuring that these issues are addressed.

Clinical importance is the complement of statistical significance when estimating sample size for clinical trials and

https://doi.org/10.1016/j.jclinepi.2019.07.018
0895-4356

for interpreting their results as to efficacy and effectiveness. Derva et al report on a systematic review to provide an inventory of the magnitude of minimal clinically important differences of 16 hip, knee, foot and ankle patient reported outcome questionnaires, assessed by an anchor–method. This evidence base will provide useful guidance, but there is an urgent need to establish consensus, as for several of the questionnaires substantial differences in the magnitude of the MCID estimates were found. In addition, there is the challenge that these methods focus on estimates of the magnitude that can be detected rather than the actual importance to the patient. Thus it is encouraging to see the paper by Bennett et al, who describe a method for incorporating patient preference for diabetic patients deciding on whether to add second agent of either insulin or a sulphonyl urea when metformin alone is not controlling their diabetes: this takes into account risks of 6 outcomes (death, myocardial infarction, stroke, severe hypoglycemia, diarrhea, nausea). Insulin provides more benefits overall but when patient preferences are added many patients prefer the oral sulphonyl urea. This shows that this decision is indeed sensitive to treatment preferences.

Support by journals for reporting guidelines along with trial registration are arguably two of the most important advances in encouraging research transparency. The CONSORT guidelines have formed the foundation for these reporting guidelines. However these do not address the many variants of trial design, and so have needed to be supplemented by 14 ‘extensions’. On the Equator Website Ghosn et al found that these extensions do successfully address the objectives of over 80% of 1000 RCTs in Pubmed so they are obviously value-added. However this is inefficient with many extensions not having been updated for the current version of CONSORT, and with up to 67 additional pages of instructions needing to be followed. The authors call for a new model to address this.

Individual patient data of controlled trials is the ideal dataset for assessing the benefits and most common harms of interventions in systematic reviews. However this is extremely challenging for many reasons. Areti et al are to be commended for using a controlled trial design to assess the usefulness of a small financial incentive to request IPD datasets from authors and sponsors of 137 studies (29 RCTs for type I diabetes mellitus and 108 RCTs for Alzheimer’s dementia). The datasets were often obtained only after long delays and there was no benefit of the financial incentive found. The investigators found that authors of publicly sponsored studies with medium-large treatment effects of above 0.2 on the standardized mean difference scale, tended to respond positively to IPD requests, especially for large studies with a low risk of bias.

NMAs are becoming increasingly popular in he systematic review community as the preferred method for assessing comparative effectiveness. However, the statistics are less transparent and they are substantially more resource–intensive so it is even more important that they are of high quality. Gao et al reviewed the quality of the 42 NMAs published in the Cochrane Library before April 2018 and compared them with over 500 other published NMAs. Overall, the quality was adequate but more attention is needed for the assessment of publication bias, the geometry of the network, and assessment and exploration of inconsistency.

There is ongoing controversy over whether and how to combine the results of controlled trials and nonrandomized evidence in network meta-analyses. The Cochrane Handbook advises against this, but proponents of Real World Evidence synthesis argue that integration of both is needed for pragmatic estimates on benefit and harm [4]. Zhang et al report in this issue on a scoping review of the published literature on methods used to combine randomized and nonrandomized evidence in network meta-analyses. Twenty three NMAs met the predefined inclusion criteria, of which 74% used natıve pooling. Most NMAs were supplemented with additional analyses to investigate the effect estimates when only randomized evidence was included. The rationales for combining included generalizability, the limited amount of direct evidence from RCTs, evaluation of rare outcomes, to allow for a wider range of evidence, and to evaluate the robustness of the results.

Selective outcome reporting [sometimes called ‘cherry picking the ripest cherries but ignoring those less ripe’] is an ongoing problem in the reporting of benefits and harms of clinical trials. This is now a separate category in the most recent Cochrane Handbook Risk of Bias Instrument. Core outcome sets are an important means of preventing this selective reporting and these are increasingly used for assessing benefits [5,6]. However, as Mayo Wilson et al describe in a pair of articles in this issue, this is still a major problem for reporting of harms. They classified the adverse events (AEs) as either ‘systematic’ (proactively measured in every patient when AEs are anticipated because of preclinical studies or evidence from related clinical trials) or ‘nonsystematic’ (passively collected when AEs are not assessed for every participant’). The first paper reviewed systemic adverse event reporting in 159 analyses from 52 datasources of trials of quetiapine for bipolar depression; they identified major problems with how systematic AEs were defined and combined. Most systematic AEs, even though they were collected, were not even mentioned in public sources, and of those that were, over 40% did not provide adequate information on how the estimates were calculated. The authors call for core outcome sets requiring reporting of systematic adverse events that are collected in every patient. The second paper examined nonsystematic adverse events (i.e., adverse events that were not assessed for every participant) from 52 sources of evidence from 21 trials of gabapentin for neuropathic pain, and 80 sources of evidence from seven trials of quetiapine for bipolar depression. Most nonsystematic AEs, including serious AEs, were not reported in public sources; this suggests that nonsystematic AEs might be “cherry-picked” for reporting. The authors call for all AEs that occur in clinical trials to be reported publicly (e.g., on www.ClinicalTrials.gov).

In a third article on harms, Golder et al point out the difficulties in searching for adverse event data and provide evidence- based guidance in identifying adverse events data in the literature. They highlight that adverse events searching should be tailored to suit the intervention of interest, and recommend that adverse effects search strategies should aim to maximize sensitivity over precision to retrieve the totality of the evidence for appraisal.

As already mentioned trial registration is arguably one of the most important advances in encouraging research transparency. This is pivotal in reducing the likelihood of negative trials being underreported. It is therefore encouraging that Berber et al report that over 90% of the 5 most recent Cochrane intervention reviews published by each of the Cochrane review groups indeed searched trial registers. As the authors point out, details about completed studies obtained from trial register records can help to inform judgments about a study's risk of bias and the overall strength of evidence in a systematic review. Information about ongoing studies obtained from trial registers can help to inform judgments about the remaining uncertainties on a topic, inform research priorities and plans for updating systematic reviews including living systematic reviews.

However, a paper by Tan et al demonstrates that trial registration is not sufficiently complete for all systematic reviews to be restricted to those on trial registers, as has been proposed [7]. Based on a review of 28 high and low impact medical journals, although overall trials that comply with the prospective registration requirement tend to display lower risk of bias, trial registration is an imperfect proxy for risk of bias since this excludes some important low risk of bias studies. Thus systematic reviews only including trials that have been registered may not adequately represent the body of evidence for assessing benefit and harms of interventions, which may result in other biases [such as increasing the proportion of industry funded studies].

Barcot et al demonstrate that there is a continuing problem in assessing blinding in RCTs included in Cochrane systematic reviews. In over 10,000 trials in over 700 Cochrane reviews, a quarter did not accurately assess blinding. Some improvement was noted when the personnel and participants components of blinding were explicitly disaggregated.

Ioannidis has highlighted the problem of multiple systematic reviews and meta-analyses of the same trials being an example of research waste [8]. The paper by Kamis et al adds to this evidence base by reviewing the results of 22 meta-analyses based on 6 trials addressing the same question of use of implantable cardiac defibrillators.

(ICD) in patients with non-ischemic cardiomyopathy. Replication is a core element of the scientific method but this redundancy could be avoided by systematic review registration with independently developed detailed, explicit, and preplanned meta-analytical approaches.

Guidelines are another example of research waste unless they are used. Gagliardi et al found 178 articles discussing different aspects of guideline implementation. Together with a six country multidisciplinary team they have developed a 33 item Clinician Guideline Determinants Questionnaire that can be used at multiple time points in the guideline development cycle to assess determinants of the use of new, updated, or adapted guidelines, and before and after implementation to assess intervention impact.

Peter Tugwell
Andre Knottnerus
E-mail address: Tugwell.BB@uOttawa.ca (P. Tugwell)

References

[1] Treweek S, Pitkethly M, Cook J, Fraser C, Mitchell E, Sullivan F, et al. Strategies to improve recruitment to randomised trials. Cochrane Database Syst Rev 2018;2:MR000013.

[2] Reveiz L, Chapman E, Asial S, Munoz S, Bonfill X, Alonso-Coello P. Risk of bias of randomized trials over time. J Clin Epidemiol 2015;68:1036−45.

[3] Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. Lancet 2005;365:176−86.

[4] Available at. https://www.imi-getreal.eu/Portals/1/Documents/Publications%20website/Including%20Real%20World%20Evidence%20in%20network%20meta-analysis.pdf.

[5] Gargon E, Williamson PR, Altman DG, Blazeby JM, Clarke M. The COMET initiative database: progress and activities from 2011 to 2013. Trials 2014;15(1):279.

[6] Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. Trials 2007;8:38.

[7] Roberts I, Ker K, Edwards P, Beecher D, Manno D, Sydenham E. The knowledge system underpinning healthcare is not fit for purpose and must change. BMJ 2015;350:h2463.

[8] Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 2016;94:485−514.