Sample size estimation in inflammatory bowel disease trials included in Cochrane systematic reviews: a protocol

Zipporah Iheozor-Ejiofor[1], Morris Gordon[1,2], Anthony Akobeng[3]

[1]School of Medicine, University of Central Lancashire, Preston, UK
[2]Families Division, Blackpool Victoria Hospital, Blackpool, UK
[3]Sidra Medicine, Doha, Qatar

Background

Clinical trials with inadequate sample sizes are likely to produce erroneous results. Sample size estimations are necessary to determine the optimal information size required to detect a difference of effect between study interventions being compared. A study with a small sample size may be unable to detect a difference (type 2 error) while a sample size that is too large may falsely show a difference of effect (type 1 error) (Jones 2003). Systematic reviews aim to mitigate this problem by combining data from multiple studies in the hope of reaching the optimal information size. However, this is not being achieved as a recent study found imprecise results in 65% of Cochrane systematic reviews (Castellini 2018).

Apart from the risk of type 1 and type 2 error, inadequately powered trials are wasteful and unethical given the needless potential adverse events which patients are exposed to. This is of great concern in inflammatory bowel disease (IBD), a debilitating disease which is easily exacerbated by using the wrong medication. The problem of inadequate sample size is widespread in trials across different fields (Puffer 2003). However, there is no understanding of the scale of the problem in IBD trials nor the factors unique to the condition such as disease type or activity which prevent recruitment to trials. We need to understand the magnitude of the problem in IBD to improve future clinical trials.

Objective

We plan to examine how sample size estimation is conducted and reported in IBD trials.

Method

Types of study

Inclusion criteria*:* RCTs included in Cochrane IBD reviews involving the induction or maintenance of remission in IBD

Exclusion criteria*:* Reports, such as abstracts, lacking clear information on number of participants randomised and assessed, will be excluded. We do not anticipate finding cluster RCTs, however, if there happen to be any included in the systematic reviews, they will be excluded. We will also exclude pilot or feasibility studies, those with a mixed population of people with and without IBD and studies on secondary analyses of follow-up data collected after discontinuation of treatment. We will also exclude RCTs that were published before 1996 when the CONSORT statement was first published (CONSORT 1996) as most studies published before the CONSORT statement are unlikely to have included sample size calculation. This restriction ensures that we focus on studies which have been published since good reporting standards were established.

Literature search

We will search the Cochrane IBD specialised register and Cochrane database of systematic reviews for all IBD systematic reviews published in the last eight years. The restriction on publication date ensures that we avoid inactive or duplicates of old versions of reviews and include updates only. Using the reference list of included studies, we will search for full text articles that meet our inclusion criteria. We will indicate the reasons for excluding individual RCTs. Where possible, the results of the literature search and the study flow will be presented in an adapted PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) diagram (Moher 2009).

Data extraction

We will use a piloted data extraction form in collecting relevant information from the studies. This will be done independently by two reviewers. We will extract data on the following:

1. Bibliometric data
2. Study population
3. Study intervention and comparator
4. Study duration
5. Sample size information: we will indicate whether sample size calculation was reported or not. Further details of sample size estimation such as clinically important difference, power and significance level will be extracted.
6. Target sample size
7. Achieved sample size: This includes information on recruitment success, unsuccessful recruitment and reasons for unsuccessful recruitment.
8. Attrition rate: number of drop-outs at the end of study period
9. Number assessed at the end of study period
10. Outcome: Primary outcome(s), definition and follow-up time.
11. Funding source

Where study details are not clearly reported, we will attempt to contact authors for clarifications. Any study reaching its original or revised target sample size or terminated early for reaching significant results will be classed as a 'recruitment success'. Those terminated for other reasons or failing to meet the criteria for recruitment success will be regarded as unsuccessful. Information from follow-up periods which occurred after the discontinuation of treatment will be disregarded.

Data analysis

Phase 1: We will carry out a descriptive analysis of RCTs which report and do not report a sample size calculation. In studies which report sample size calculation, we will assess how well these are reported to enable its repetition based on the significance level, power and the minimal clinically relevant difference. Results will be summarised in percentages, means, median and interquartile range where appropriate.

Phase 2: Based on the reported parameters for sample size calculation, we will recalculate the sample sizes to assess consistency in the reported and recalculated values. This will be achieved using the *power onemean* command in STATA 15 following methods described in Sully 2013. We will then evaluate recruitment success in these trials by comparing the reported target sample size with achieved sample sizes. The number of study participants lost to follow-up will be used in assessing whether adequate sample size was maintained at the point of outcome reporting. The analysis in phase 2 will only be performed for the outcomes of clinical or endoscopic remission/relapse.

Phase 3: Data permitting we plan to match various study characteristics, with study sample sizes to determine the markers of adequate study power in IBD trials. Where possible, chi square tests will be used in testing differences between trials with different characteristics. Study characteristics tested will be limited to avoid false positive findings (Burke 2015). The following study characteristics will be evaluated:

Source of funding: industry funding versus other sources

Disease activity: active versus inactive disease

Disease type: Crohn's disease versus ulcerative colitis

Intervention type: pharmacological versus non-pharmacological

References

Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. Emerg Med J2003; 20:453–458

Castellini G, Bruschettini M, Gianola S, Gluud C, Moja L. Assessing imprecision in Cochrane systematic reviews: a comparison of GRADE and Trial Sequential Analysis. Systematic Reviews (2018) 7:110 https://doi.org/10.1186/s13643-018-0770-1

Puffer S, Torgerson DJ. Recruitment difficulties in randomised controlled trials. Controlled clinical trials. 2003 Jun; 24:214S-215S.

The CONSORT Group (1996). Improving the quality of reporting of randomized controlled trials. The CONSORT Statement. JAMA 276:637-639.

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal. pmed1000097

Sully BGO, Julious SA, Nicholl J. A reinvestigation of recruitment to randomised, controlled, multicenter trials: a review of trials funded by two UK funding agencies. Trials 2013, 14:166 Page 2 of 9 http://www.trialsjournal.com/content/14/1/166

Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. BMJ 2015; 351 doi: https://doi.org/10.1136/bmj.h5651