Reviewer #1:
Scientific Quality: Grade B (Very good)
Language Quality: Grade A (Priority publishing)
Conclusion: Accept (General priority)
Specific Comments to Authors: The manuscript suggested by Morris Gordon et al focused on a very important methodologic aspect of RCT: the minimum sample size estimate. The authors performed a systematic review of available RCTs in IBDs. On 105 trials, a third of intervention studies in IBD within the last 25 years are underpowered. Also, the authors present a sample size estimate resource for future researchers. This is a very interesting paper.

Thank you for your feedback.

To be more useful, the paper should be more practice with a clear framework to help reader for calculating appropriate sample size estimate for future studies. The title reflects the main subject of the manuscript The abstract correctly summarizes and reflects the work described in the manuscript The Key words reflect the focus of the manuscript The Background adequately describes the overview on the topic and correctly present the significance of the study The methods part is not very well distributed. A lot of details are given about the eligibility criteria for selection of papers included in systematic review but less details are given about calculation details for sample size estimation which is quite a pity because this is the purpose of the paper et the most interesting part.

Thank you for your comment. The parameters we used were: Two independent study groups, dichotomous endpoint, power 80%, type 1 error 0.05. In group 1 we have put the rate reported by the study of the intervention drug, and in group 2 we have put the rate of the placebo. We have now added it to the methods: page 9, lines 200-202: "The parameters we used were two independent groups, dichotomous outcomes. In group 1 we have put the rate reported by the study of the intervention drug, and in group 2 we have put the rate of the placebo."

Also, authors (line 186, page 8) told about inconsistencies. These inconsistencies could be listed in a supplementary table for reader information. Again, details to obtain the recalculate sample size should be introduced in the methodology part. This would represent a practice framework for readers

At first, we were recalculating the sample size for each study drug VS placebo irrespective of the achieved difference. This gave us massive figures of sample size for the studies with achieved difference of 10% and less. The decision was made for us to

recalculate the sample sizes only for the studies with achieved difference of 10% and above, as it was the smallest MCID used.
We don't think we need to produce the table for these inconsistences, however, we have added "regarding the use of sample size calculations for the studies with achieved difference of less than 10%," to methodology – line 193-194 p 8.

In the result part, le paragraph line 208 to 212 is not informative and not useful for the topic. To calculate a mean proportion of clinical remission by pooling is not useful at all as there is a huge heterogeneity in study populations and tested drugs.
There are obvious issues of heterogeneity that limit the utility or appropriateness of pooling for purposes such as meta-analysis. However, sample size calculation requires an MCID to be used and currently the only method to do this is to select in a subjective fashion individual past. Therefore, this forms a different source of information to inform this process. It is not to be used in a draconian fashion as there may be very many valid reasons to not use the specific result from these tables. Instead, it forms a further useful resource and highlights the need for such consideration at the beginning of trial design.
We added on page 13 lines 318 – 320 this as limitation. "There are obvious issues of heterogeneity limiting the appropriateness of pooling the data, however, the only way to obtain the previously used MCID was through looking at the past studies. "
The discussion is well constructed. In line 262 to 264, the authors write that the work present a resource for sample size estimation not just for future study authors, but for study peer reviewers but at this stage, the resource seems not practical enough to be a useful resource. Does the manuscript interpret the findings adequately and appropriately, highlighting the key points concisely, clearly and logically? Are the findings and their applicability/relevance to the literature stated in a clear and definite manner? Is the discussion accurate and does it discuss the paper's scientific significance and/or relevance to clinical practice sufficiently?
Thank you for your feedback. As mentioned in the discussion on page 13 lines 333-335 this resource will require regular updates. We also do not mandate the use of this resource, but rather presented it to be aware of the common trends happening with power calculation in the trials.
About illustrations and tables, it would be easier to read Table 2 if the table 2 is split in Table 2A UC and Table 2B CD No other comments on Biostatistics, Units. References, Quality of manuscript organization and presentation. Research methods and reporting and Ethics statements

Reviewer #2:
Scientific Quality: Grade C (Good)
Language Quality: Grade B (Minor language polishing)
Conclusion: Major revision
Specific Comments to Authors: The authors have performed a novel systematic review to evaluate the minimum sample size that should be used to achieve adequate sample size based on randomised controlled trials in IBD. They found that

approximately one third of interventional studies since 1996 were underpowered. The manuscript provides important food for thought in a potentially overlooked area of trial design by clinicians that is actually very important in IBD trials. Some details in the inclusion of trials needs to be clarified as low molecular weight heparin and cannabis were included in the study while ustekinumab and vedolizumab were considered exploratory interventions which does not reflect clinical practice.
I have the following suggestions: Page 7, line 150-53 eligibility of trials – the medications mentioned as exploratory therapies have received FDA/NHS/ approval for use in inflammatory bowel disease by necessary authorities in many developed nations. Can this statement be clarified or perhaps more clearly specify which therapies are considered established and only include those. The fact that LMWH and cannabis were included would not fit what is considered standard IBD therapy.

Thank you for clarifying this point – these should have been removed in earlier phases of the manuscript as they do not meet the inclusion criteria – this has been resolved. Page 7 lines 158-159 added "or not under the three core headings (biologic, immunomodulators or anti-inflammatories)."

The search strategy of the sustematic review primarily focuses on a Cochrane search which is not truely a systematic review - why weren't Medline, Scopus or Web of Science searched? An alternative would be to remove the term systematic review from the manuscript. PRISMA guidelines require the search terms on at least one search engine to be provided and the results of searches should be provided for full transparency. Please provide these.

Page 7, lines 147 – 149 mentions "We conducted a comprehensive search of the Cochrane IBD Specialized Trials Register, CENTRAL and hand searched within the Cochrane library of IBD reviews for further primary RCTs." These are 3 sources – as such, this review can be considered systematic. The search termes used are now included in the appendix – we have now added them to methodology page 7 lines 149. We have also included our choice of resources in the limitations – page 12, lines 309-312 "The search methods used limited the parameters of the search for pragmatic reasons. However, this does not represent any systematic bias, hence we do not believe it invalidates the findings, and in the future this resource can be updated prospectively."

Has this systematic review been registered on PROSPERO or through another review registry?
The systematic review identified 105 papers and the manuscript has only 6 papers referenced, none of which are interventional trials in inflammatory bowel disease. This does not provide transparency in which trials were used in the study. Are the authors able to provide a list of included studies and references?
Thank you for your comment. We didn't register this systematic review on Prospero but it is is on our online repository and as such freely available to ensure high quality and transparent systematic review practice (Please, see reference 8). Whilst Prospero is

useful in encouraging this, like many authors we elected to use an alternative a priori location to deposit the protocol.
We now have included the references of included studies in the Appendix.
Minor comments: Page 6, line 115 consider changing 'than' to 'then' Page 10, line 221 when referring to in clinical practice can this be further defined. Only randomized controlled trials were included so does this refer to investigator initiated trials?

Thank you for the comment. This has been corrected – sorry for the lack of clarity. By clinical practice we meant the studies that we used for this review. Page 6 line 119 "then" has been corrected. Page 10 line 235 has been corrected: " … reported rarely matched the actual differences achieved by these studies".

Reviewer #3:
Scientific Quality: Grade C (Good)
Language Quality: Grade A (Priority publishing)
Conclusion: Accept (General priority)
Specific Comments to Authors: In this study, the Authors aimed to review minimum sample size estimation for trials in Inflammatory Bowel Disease. The Authors demonstrated that there is no clear basis or accepted standard for current practice for minimal clinically important difference (MCID) estimation when producing a power calculation for a primary randomised controlled trials within IBD. The Authors showed that a third of intervention studies in Inflammatory Bowel Disease within the last 25 years are underpowered. This review gives us valuable information for the future studies.

Thank you for your feedback!

Science editor: 1 Scientific quality: The manuscript describes a Frontier of the estimates for trials in Inflammatory Bowel Disease. The topic is within the scope of the WJG. (1) Classification: Grade B, C and C; (2) Summary of the Peer-Review Report: The manuscript provides overlooked area of trial design that actually very important in IBD trials. Some details in the inclusion of trials needs to be clarified.

The questions raised by the reviewers should be answered; (3) Format: There are 2 tables and 1 figure; (4) References: A total of 8 references are cited, including 2 references published in the last 3 years; (5) Self-cited references: There is 1 self-cited references. The self-referencing rates should be less than 10%. Please keep the reasonable self-citations (i.e. those that are most closely related to the topic of the manuscript) and remove all other improper self-citations. If the authors fail to address the critical issue of self-citation, the editing process of this manuscript will be terminated; and (6) References recommendations (kindly remind): The authors have the right to refuse to cite improper references recommended by the peer reviewer(s),

especially references published by the peer reviewer(s) him/herself (themselves). If the authors find the peer reviewer(s) request for the authors to cite improper references published by him/herself (themselves), please send the peer reviewer's ID number to editorialoffice@wjgnet.com. The Editorial Office will close and remove the peer reviewer from the F6Publishing system immediately. 2 Language evaluation: Classification: Grade A, A and B. The authors are native English speakers. 3 Academic norms and rules: No academic misconduct was found in the Bing search. 4 Supplementary comments: This is an invited manuscript. No financial support was obtained for the study. The topic has not previously been published in the WJG. 5 Issues raised: The authors did not provide original pictures. Please provide the original figure documents. Please prepare and arrange the figures using PowerPoint to ensure that all graphs or arrows or text portions can be reprocessed by the editor. 6 Recommendation: Conditional acceptance.

(2) ***Company editor-in-chief:*** I have reviewed the Peer-Review Report, full text of the manuscript, and the relevant ethics documents, all of which have met the basic publishing requirements of the World Journal of Gastroenterology, and the manuscript is conditionally accepted. I have sent the manuscript to the author(s) for its revision according to the Peer-Review Report, Editorial Office's comments and the Criteria for Manuscript Revision by Authors.