## PEER REVIEW HISTORY



## ARTICLE DETAILS

| | |
|---|---|
| **TITLE (PROVISIONAL)** | |
| **AUTHORS** | |

## VERSION 1 – REVIEW

| | |
|---|---|
| **REVIEWER** | Noor, Nurulamin<br>University of Cambridge, Department of Medicine |
| **REVIEW RETURNED** | 10-Dec-2021 |

| | |
|---|---|
| **GENERAL COMMENTS** | Overall this is an interesting and well-written manuscript describing a protocol to conduct a double-blind, single centre, randomized controlled trial of low-dose naltrexone versus placebo for patients with mild to moderate Crohn's disease.<br><br>The findings from this trial would be of significant interest and novelty in the field given the relatively unexplored mechanism of action being assessed in this clinical trial.<br><br>I have provided several points, which are all minor, below. I hope these points help the authors with clarification or for addition of details, to help further understanding around the methodology of the trial for potential readers and to help understand some of the practicalities of the trial also:<br><br>1. The authors highlight results of a pilot study, and in particular draw attention to 74.5% clinical improvement and 25.5% clinical remission at week 12. Could the authors provide patient numbers for this pilot study as well as the percentages, as this would help to contextualize these results?<br><br>2. The authors highlight inclusion criterion of SES-CD score of 3-15 to define population of mild-moderate Crohn's disease. Given a patient with ileal disease and high SES-CD score would likely be considered to have severe Crohn's disease and not mild to moderate disease activity. Can the authors confirm whether they have considered stratification and use of different scores based on disease location e.g. for patients with ileal disease vs. those with ileocolonic or colonic disease?<br><br>3. Would the authors be able to clarify if stopping of oral aminosalicylate is being advised for patients who present with Crohn's disease and active inflammation at point of inclusion into the trial? Given the lack of sufficient evidence for aminosalicylates in Crohn's disease and particularly in the context of patients presenting with active disease whilst taking aminosalicylate as a concomitant medication. |

| | |
|---|---|
| | 4. With regards the naltrexone treatment intervention and dose. The authors provide citations which highlight use of low-dose naltrexone. Did the authors also consider assessment of multiple other doses? Could the authors include a sentence or two about the options for doses to select and the reasons for selecting the 4.5mg dose in this trial?<br><br>5. The authors list several secondary outcomes. Could the authors confirm if multiple correction will be employed for assessment of these secondary outcomes, and which statistical method will be used?<br><br>6. The authors provide a helpful overview and Figure of the scheduled trial visits. Given the association of increased placebo response for trials with high number of ad-hoc visits, can the authors also comment on the possibility for ad-hoc visits in this trial? Are ad-hoc visits allowed within the protocol and would any particular trial data or samples be collected at ad-hoc visits?<br><br>7. The authors state a secondary outcome measure of week 52 corticosteroid free clinical remission. Could the authors confirm the definition that will be used in this instance - how long would patients have to be “free” from corticosteroid medication? How would clinical remission be defined for this outcome measure?<br><br>8. The authors state a secondary outcome measure of week 52 endoscopic remission and response. Could the authors clarify how endoscopic remission and response would respectively be defined at this week 52 timepoint?<br><br>9. For sample size estimation. The authors state preliminary data of 25% mucosal healing rates at week 12 for use of low-dose naltrexone in Crohn’s disease. Could the authors clarify what is being referred to as “mucosal healing” here, is this endoscopic remission data? And could the authors clarify on what size of cohort this data is based. If published data, could the authors provide a reference to support this statement?<br><br>10. Could the authors clarify what software or process will be used to randomize patients into this trial? Will patients have randomization stratified on any elements which might affect their response to treatment, such as age or disease location of their Crohn’s disease?<br><br>11. Given the low and declining global recruitment to clinical trials in IBD, further details about recruitment would be helpful for potential readers. The authors state that this trial is being performed in academic and non-academic centers across the Netherlands. Can the authors state how many sites are participating across the Netherlands? When is the recruitment due to be completed?<br><br>12. The authors state that recruitment started on 14 January 2021. Could the authors then state how many participants have been recruited to date since 14 January 2021? Is the current recruitment timeline feasible to complete the trial based on recruitment data so far?<br><br>13. The authors describe that there will be the possibility of an open-label extension until week 52 for patients to participate in. |

| | |
|---|---|
| | Could the authors describe if any other longer term follow-up is planned to take place thereafter and if for example this were to be through the use of electronic health records, a brief explanation of how this would be done in the Netherlands and whether patient consent for longer term follow-up has been or will be considered?<br><br>14. Can the authors clarify if there is an independent data monitoring committee overseeing unblinded trial outcome and safety data? Can the authors confirm if there are planned formal interim analyses to assess safety, efficacy or even futility?<br><br>15. Can you elaborate on data sharing plans and whether a data sharing protocol has been developed or is due to be developed?<br><br>16. As a very minor point. The authors alternate between using “naltrexone” and “naltrexon” in the manuscript. In the English language section, the authors could perhaps stick to using “naltrexone”. |

| | |
|---|---|
| **REVIEWER** | Raknes, Guttorm<br>Raknes Research |
| **REVIEW RETURNED** | 11-Dec-2021 |

| | |
|---|---|
| **GENERAL COMMENTS** | I would like to thank for the opportunity to review this manuscript which contains a protocol for a randomized, double-blinded, placebo-controlled multicentre study investigating the effects of low dose naltrexone (LDN) in mild to moderate Crohn’s disease.<br><br>The study has already been initiated, and it is almost a year since the first patient was enrolled.<br><br>High-quality RCTs like this on LDN in various diseases are very welcome. For many conditions, including inflammatory bowel disease, LDN seems to have ended up in a limbo between being seen as an obscure alternative therapy and as a serious treatment alternative. For Crohn's disease, there are no studies that completely refute that LDN is efficacious, but the studies that have been published so far are small and can at best be called promising. The study outlined in the protocol here will bring the knowledge about LDN in Crohn's disease important steps forward. It will in many ways try to reproduce the results of Smith et al’s RCT from 2011, but with a significantly more solid study design. It will be very interesting to see the results of this study!<br><br>The manuscript is well written in a clear and concise language. Most of the items of the SPIRIT checklist have been adequately addressed. I do not have any major objections, but some issues need to be addressed:<br><br>P4, L38: The pilot study was without a control group. There is no sure basis for establishing a causal relationship, ie say that LDN _induced_ remission. Perhaps say something like: “among 47 patients that started LDN, 74.5% experienced clinical improvement and----“.<br><br>P5, L12-13: Are really ALL gastro/hepato centres (=hospitals?) participating in the study? How will the clinics be recruited/invited to take part in the study? |

| | |
|---|---|
| | P5, L26: More details on the recruitment of patients are needed. Will all attending patients Crohn's patients to all clinics in the Netherlands be asked to participate in standard consultations? Will they be asked by their gastroenterologist, or will they be invited by letter? Will there be announcements made through patient organisations, or in the press? Is there any chance of recruitment bias?<br><br>P5, L53-57: What does the contents of the placebo capsules consist of? Lactose? Ascorbic acid?<br><br>P6, L15-23: What measures were made to ensure that the endoscopic assessments are performed in a standardized way, there were multiple centres and investigators involved? Any education or algorithms to be followed? Were there any guidelines on the scope of the examination, including the number of biopsies, time spent etc? Wil the examinations only be performed by specialists (gastroenterologists), or will also junior doctors or endoscopy nurses also do the assessments?? I am no gastroenterologist, so for that matter the eexamination may be sufficiently standardized per se. Maybe this should have been described to readers not familiar to endoscopic procedures.<br><br>P4, L31: naltrexonE<br><br>P6, L27: Baseline blood samples should be specified in the main text, not only in the Table 1 note. Why should these blood samples be taken of all included patients? Any exclusion criteria based on blood samples, or outcomes other than CRP?<br><br>Will the study participants be asked whether they believe they received LDN or placebo at each visit or at the end of the study to assess the quality of blinding?<br><br>There are many outcomes in this study. Will there be any adjustments made for the interpretation of multiple tests?<br><br>How will participants withdrawing from the study for different reasons be handled? Both intention-to-treat and per-protocol analyses?<br><br>In such a large study it is likely that some data will be missing. What strategies are in place to handle missing data in the analyses? Are there any particular measures in place to avoid missing data?<br><br>It would be interesting to see the supplementary Dutch documents translated to English. |

**VERSION 1 – AUTHOR RESPONSE**

Reviewer: 1

1. The authors highlight results of a pilot study, and in particular draw attention to 74.5% clinical improvement and 25.5% clinical remission at week 12. Could the authors provide patient numbers for this pilot study as well as the percentages, as this would help to contextualize these results?

Yes, the following citation is from the study of dr. M. Lie: 'Of the 47 patients, 35 (74.5%) achieved a clinical response. Of those 35 patients, 12 patients had a response of at least 3 months (25.5% of total cohort, 8 CD, 4 UC), whereas a short-lived (between 4 and 12 weeks) improvement was seen in the remaining 23 patients (48.9% of total cohort, 13 CD, 10 UC).'
- Lie, M. R., van der Giessen, J., Fuhler, G. M., de Lima, A., Peppelenbosch, M. P., van der Ent, C., & van der Woude, C. J. (2018). Low dose Naltrexone for induction of remission in inflammatory bowel disease patients. Journal of translational medicine, 16(1), 1-11.

2. The authors highlight inclusion criterion of SES-CD score of 3-15 to define population of mild-moderate Crohn's disease. Given a patient with ileal disease and high SES-CD score would likely be considered to have severe Crohn's disease and not mild to moderate disease activity. Can the authors confirm whether they have considered stratification and use of different scores based on disease location e.g. for patients with ileal disease vs. those with ileocolonic or colonic disease?
A SES-CD score of 3-15 will mostly correspond with mild to moderate disease as was validated in multiple studies. Because the scoring is composed of different criteria that need to be evaluated in 5 different regions of the bowel, it can be that severe but isolated disease scores between 3 and 15. Because periphery centers are not used to score colonoscopies, using multiple scoring systems will increase the risk for mistakes and will lead to lower inclusion rates. When analyzing the results after finishing inclusion, we can take disease location and behavior into account.

3. Would the authors be able to clarify if stopping of oral aminosalicylate is being advised for patients who present with Crohn's disease and active inflammation at point of inclusion into the trial? Given the lack of sufficient evidence for aminosalicylates in Crohn's disease and particularly in the context of patients presenting with active disease whilst taking aminosalicylate as a concomitant medication.
The protocol advices to continue oral 5-ASA compounds provided the dose prescribed. The dose must be stable for at least 4 weeks prior to randomization and 10 weeks after randomization. In The Netherlands, oral aminosalicylates are not regularly prescribed in Crohn's disease, consistent with the ECCO guidelines, so we don't expect this to be an issue.

4. With regards the naltrexone treatment intervention and dose. The authors provide citations which highlight use of low-dose naltrexone. Did the authors also consider assessment of multiple other doses? Could the authors include a sentence or two about the options for doses to select and the reasons for selecting the 4.5mg dose in this trial?
The dose is based on the agonistic effect of naltrexone on the mu opioid receptor when naltrexone is administered at a low dose. In previous animal studies and clinical trials in different immune mediated inflammatory disease the dose of 4.5 mg per day is investigated. We therefore choose to use this dose, because data on efficacy and safety is already available. If we wanted to test different doses in this trial, the number of participants would have increased which is not feasible.

5. The authors list several secondary outcomes. Could the authors confirm if multiple correction will be employed for assessment of these secondary outcomes, and which statistical method will be used?
In the study design, we identified one single outcome as the primary outcome (endoscopic remission) and treated the remaining outcomes as secondary. According to the literature, there is no need to adjust for multiplicity because of our single primary outcome investigating endoscopic remission; the findings for secondary outcomes are considered subsidiary and exploratory, and not confirmatory. The questionnaires that measure fatigue, sleep and quality of life for example, do not say something about the effectiveness of low dose naltrexone for the induction of endoscopic remission (the gold standard).
- Li, G., Taljaard, M., Van den Heuvel, E. R., Levine, M. A., Cook, D. J., Wells, G. A., ... & Thabane, L. (2017). An introduction to multiplicity issues in clinical trials: the what, why, when and how. International journal of epidemiology.

6. The authors provide a helpful overview and Figure of the scheduled trial visits. Given the association of increased placebo response for trials with high number of ad-hoc visits, can the authors also comment on the possibility for ad-hoc visits in this trial? Are ad-hoc visits allowed within the protocol and would any particular trial data or samples be collected at ad-hoc visits?
Ad-hoc visits are allowed if the disease worsens and a patient needs to withdraw blood because of that, or escape medication is started. Especially during the induction phase, all patients are regularly seen at the outpatient clinic or contacted by phone, so in our experience no extra visits are requested for.

7. The authors state a secondary outcome measure of week 52 corticosteroid free clinical remission. Could the authors confirm the definition that will be used in this instance - how long would patients have to be "free" from corticosteroid medication? How would clinical remission be defined for this outcome measure?
Thank you for this comment. Most important is that patients are steroid free at the time of the endoscopic evaluation. If patients are tapered off corticosteroids just before the endoscopy, this needs to be taken into account and will be described in the results. Clinical remission is defined as an HBI score of ≤4.

8. The authors state a secondary outcome measure of week 52 endoscopic remission and response. Could the authors clarify how endoscopic remission and response would respectively be defined at this week 52 timepoint?
The endoscopic evaluation at week 52 will be the same as for week 12, so endoscopic remission is defined as a SES-CD ≤2 and ulcerated surface subscore ≤1 in all five segments. Endoscopic response is defined as a reduction of the SES-CD score by ≥50% compared to the colonoscopy at week 12.

9. For sample size estimation. The authors state preliminary data of 25% mucosal healing rates at week 12 for use of low-dose naltrexone in Crohn's disease. Could the authors clarify what is being referred to as "mucosal healing" here, is this endoscopic remission data? And could the authors clarify on what size of cohort this data is based. If published data, could the authors provide a reference to support this statement?
Mucosal healing means the absence of inflammation during endoscopic evaluation of the bowel. For the sample size calculation, data of the study of M. Lie was used. The only other data available was from the group of Smith et al (2011), but with the 33% endoscopic response of LDN and 8% response of placebo less patients were needed. We were unsure if this patient population was representative of our population, risking a sample size that is too small which would have reduced the power of the study and would have increased the margin of error. A few examples: inclusion of patients between 2006 – 2009 when less therapeutics options were available so patients started earlier with LDN in their disease process, only 40 patients were included, cross-over was possible, high response rate LDN and placebo. Therefore we based our sample size calculations on Lie et al, 2018.

10. Could the authors clarify what software or process will be used to randomize patients into this trial? Will patients have randomization stratified on any elements which might affect their response to treatment, such as age or disease location of their Crohn's disease?
Randomization was done by the pharmacy that produces the study medication. After the production of the placebo and naltrexone capsules, they have numbered all the bottles that contain either placebo capsules or naltrexone in a random order. The trial pharmacy of the Erasmus MC received the bottles and they deliver the medication to the participating centers in the order of inclusion of the individual patients. No patient characteristics are taken into account.

11. Given the low and declining global recruitment to clinical trials in IBD, further details about recruitment would be helpful for potential readers. The authors state that this trial is being performed in academic and non-academic centers across the Netherlands. Can the authors state how many sites are participating across the Netherlands? When is the recruitment due to be completed?
At this moment, one academic (Erasmus MC, Rotterdam) and 6 non-academic centers (Maasstad hospital, Deventer hospital, Bernhoven hospital, Albert Schweitzer hospital, Sint Franciscus Gasthuis, Bravis hospital) are participating, and we are busy with recruiting extra centers. Hospitals in the region are free to refer patients to our center for the course of the study. We have added a sentence to the study setting paragraph. We estimate that the inclusion will be finished in three years.

12. The authors state that recruitment started on 14 January 2021. Could the authors then state how many participants have been recruited to date since 14 January 2021? Is the current recruitment timeline feasible to complete the trial based on recruitment data so far?
Unfortunately, because of COVID, the start of the study in our and our collaborating centers was delayed. The first patient was included in our center, the Erasmus MC, January 2021. Over the course of the year the study was initiated in the collaborating centers and to date 10 participants have started the trial. This number will go up fast, because the trial has just been enrolled in the other centers and they are ready to start including patients. Therefore, we hope and believe this timeline is feasible. Every six months we discuss with our sponsor (ZonMw) how we are doing and what our recruitment planning is.

13. The authors describe that there will be the possibility of an open-label extension until week 52 for patients to participate in. Could the authors describe if any other longer term follow-up is planned to take place thereafter and if for example this were to be through the use of electronic health records, a brief explanation of how this would be done in the Netherlands and whether patient consent for longer term follow-up has been or will be considered?
Unfortunately, no longer follow-up is possible because the study medication that is produced is for one year maximum. Patients are allowed to continue treatment with low dose naltrexone afterwards, but this is not always reimbursed so a contribution is necessary. If we would like to extend the follow-up for patients that continue treatment, an amendment needs to be made to the protocol and extra study medication needs to be produced. In addition, the focus of our study is induction of remission, so the first three months are most important.

14. Can the authors clarify if there is an independent data monitoring committee overseeing unblinded trial outcome and safety data? Can the authors confirm if there are planned formal interim analyses to assess safety, efficacy or even futility?
The local ethics committee agreed on no data monitoring committee being necessary. No interim analysis are planned, because unblinding will take place after the last patients finishes the study. Adverse events are reported according to the protocol, so adverse events are reported in the online questionnaires, and serious adverse events are reported to the study coordinator and ethics committee .

15. Can you elaborate on data sharing plans and whether a data sharing protocol has been developed or is due to be developed?
We have created a data management plan together with a data expert of the Erasmus MC, and this plan is reviewed and approved by ZonMw, the organization that funds our research. We intend to archive our data at an external data company. The database can then be requested by submitting an application to principle investigator professor dr C.J. van der Woude.

16. As a very minor point. The authors alternate between using “naltrexone” and “naltrexon” in the manuscript. In the English language section, the authors could perhaps stick to using “naltrexone”.
Thank you for your alertness. We have corrected this.

Reviewer 2

P4, L38: The pilot study was without a control group. There is no sure basis for establishing a causal relationship, ie say that LDN _induced_ remission. Perhaps say something like: “among 47 patients that started LDN, 74.5% experienced clinical improvement and----“.
Thank you for the suggestion. We have changed the wording.

P5, L12-13: Are really ALL gastro/hepato centres (=hospitals?) participating in the study? How will the clinics be recruited/invited to take part in the study?
At this moment, one academic (Erasmus MC, Rotterdam) and 6 non-academic centers (Maasstad hospital, Deventer hospital, Bernhoven hospital, Albert Schweitzer hospital, Sint Franciscus Gasthuis, Bravis hospital) are participating, and we are busy with recruiting additional hospitals. Most of the participating hospitals were contacted by email during the grant proposal phase in 2018/2019. One hospital was interested in participation in the trial after a protocol presentation on a symposium. Additional hospitals are and will be approached by email. Patients that are treated in a hospital that is not participating in the trial, can be referred to our hospital.

P5, L26: More details on the recruitment of patients are needed. Will all attending patients Crohn’s patients to all clinics in the Netherlands be asked to participate in standard consultations? Will they be asked by their gastroenterologist, or will they be invited by letter? Will there be announcements made through patient organisations, or in the press? Is there any chance of recruitment bias?
Thank you for the suggestion. We have added information to the study setting paragraph.
In most hospitals, patients that undergo an colonoscopy that shows mild to moderate disease activity will be invited to receive information about the study. Patients will then be seen at the clinic by the researcher/doctor, or they receive information by phone. After reading the patient information and signing the consent form, they can continue the process of screening for eligibility and start the trial. The Crohns&Colitis Netherlands patient organization has published an article about the study in their newsletter 3 years ago, but at that time the study was in a preparatory phase. Therefore, recruitment bias is not likely to occur.

P5, L53-57: What does the contents of the placebo capsules consist of? Lactose? Ascorbic acid?
Please see the following information on the placebo capsules that is described in our IMPD: ‘The placebo IMP for this trial has the same qualitative composition, except naltrexone hydrochloride. Because of the fact that the appearance of the drug product is a white powder and the share of naltrexone hydrochloride will be about 2 per cent of the total amount of powder in a capsule, leaving out the naltrexone (for the placebo product) will not change the appearance of the capsules and therefore will not cause any blinding issues.’. The remaining ingredients are: Microcrystalline cellulose PH-102, Colloidal anhydrous silica, Magnesium stearate and Lactose monohydrate 100 mesh.

P6, L15-23: What measures were made to ensure that the endoscopic assessments are performed in a standardized way, there were multiple centres and investigators involved? Any education or algorithms to be followed? Were there any guidelines on the scope of the examination, including the number of biopsies, time spent etc? Wil the examinations only be performed by specialists (gastroenterologists), or will also junior doctors or endoscopy nurses also do the assessments?? I am no gastroenterologist, so for that matter the eexamination may be sufficiently standardized per se. Maybe this should have been described to readers not familiar to endoscopic procedures.
Thank you for the comments. Colonoscopies in the participating centers are performed by gastroenterologist or gastroenterologist in training under supervision of a gastroenterologist. Because the indication of the colonoscopies concerning our study is to evaluate inflammation resulting from Crohn’s disease, biopsies are not mandatory and time spent depends, among others, on the extent of the inflammation. The colonoscopies are sufficiently standardized indeed. We assume that, overall,

the readers of our protocol will have some knowledge of gastroenterology, internal medicine or surgery.

P4, L31: naltrexone
Thank you for your alertness, we have corrected this.

P6, L27: Baseline blood samples should be specified in the main text, not only in the Table 1 note. Why should these blood samples be taken of all included patients? Any exclusion criteria based on blood samples, or outcomes other than CRP?
Thank you, we added the information on laboratory tests in the main text. Blood samples are taken beforehand to compare laboratory measures during the study with baseline, to evaluate possible adverse events of naltrexone, for example the impact on the liver. In the protocol, no exclusion criteria are specified with regards to the laboratory results.

Will the study participants be asked whether they believe they received LDN or placebo at each visit or at the end of the study to assess the quality of blinding?
This is not a standard question that is included in the online questionnaires, but this will often be discussed and registered during the study visits indeed. Therefore, it might be possible to assess the quality of blinding at the end of the study. Because the capsules of the study medication are identical, percentage of side effects are low and both patients and investigators/doctors are blinded, we assume the quality of blinding is good.

There are many outcomes in this study. Will there be any adjustments made for the interpretation of multiple tests?
Thank you for addressing this. In the study design, we identified one single outcome as the primary outcome (endoscopic remission) and treated the remaining outcomes as secondary. According to the literature, there is no need to adjust for multiplicity because of our single primary outcome investigating endoscopic remission, the main conclusion of our study; the findings for secondary outcomes are considered subsidiary and exploratory, and not confirmatory. The questionnaires that measure fatigue, sleep and quality of life for example do not say something about the effectiveness of low dose naltrexone for the induction of endoscopic remission (the gold standard).
Li, G., Taljaard, M., Van den Heuvel, E. R., Levine, M. A., Cook, D. J., Wells, G. A., ... & Thabane, L. (2017). An introduction to multiplicity issues in clinical trials: the what, why, when and how. International journal of epidemiology

How will participants withdrawing from the study for different reasons be handled? Both intention-to-treat and per-protocol analyses?
When subjects are withdrawn from the study, they will not be replaced. When missings are random, imputation of missing data can be carried out. Patients that are lost to follow up will be incorporated until their lost to follow up date only, as it is a per protocol approach. Missing and lost to follow up patients will be described.

In such a large study it is likely that some data will be missing. What strategies are in place to handle missing data in the analyses? Are there any particular measures in place to avoid missing data?
The questions in the online questionnaires for patients are mandatory, and you can only continue to the next question if you completed the current question. At every study visit, patients receive an email with a link to the questionnaire and they are reminded by the staff to fill them out. If they forget to, they receive reminder emails. During monitor visits, additional information and information that staff needs to collect during the study will be checked by the monitor. Gemstracker, the online questionnaire program, shows incomplete questionnaires to staff so they can act on it. When missings are random, imputation of missing data can be carried out.

It would be interesting to see the supplementary Dutch documents translated to English.
The patient information can be translated if necessary, but this will take a while. I have translated the headings. Hopefully you will extract meaningful information from it: Introduction, 1. General information, 2. Purpose of the study, 3. Background of the study. 4. What does it mean if you take part, 5. What will be expected of you, 6. Possible risks of taking part in the study, 7. Possible advantages and disadvantages, 8. If you don't want to take part in the study, or if you want to quit, 9. End of the study, 10. Storing of documents and samples, 11. Insurance of subjects, 12. Informing general practioner, treating physician and pharmacist, 13. Compensation for taking part., 14. Questions., 15. Signing the consent form, 16. Appendices, Appendix A. Contact information Erasmus MC, Appendix B. Study procedures, Appendix C. Adverse events and risks, Appendix D. Information about insurance, Appendix E. Consent form subject.

**VERSION 2 – REVIEW**

| | |
|---|---|
| **REVIEWER** | Noor, Nurulamin<br>University of Cambridge, Department of Medicine |
| **REVIEW RETURNED** | 09-Feb-2022 |

| | |
|---|---|
| **GENERAL COMMENTS** | All my queries have been answered and additional clarification provided to the manuscript in some sections, for which I thank the authors.<br><br>The trial and its eventual primary results should be of broad interest to the gastroenterology and IBD community. I wish the authors luck with the ongoing conduct and recruitment to this trial and have no further comments. |

| | |
|---|---|
| **REVIEWER** | Raknes, Guttorm<br>Raknes Research |
| **REVIEW RETURNED** | 14-Feb-2022 |

| | |
|---|---|
| **GENERAL COMMENTS** | After reading the revised version, I still think the study design seems sound, and that no major changes are needed. The authors have made some changes to respond to both mine and reviewer 1's comments. The manuscript has improved somewhat.<br><br>Many of the other inputs have also been discussed and answered in a good way, but without this having resulted in changes to the manuscript.<br><br>I think addidional information in the authors' response to reviewers should have been have been reproduced in the manuscript as well. Many of these objections will certainly come up when the results of the study are to be published.<br><br>Some examples:<br>-details on sample size calculation (ref. to Smith et al).<br>-Rationale for choice 4.5 mg naltrexone add as intervention dose<br>-Composition of placebo<br>-The handling of dropouts (estimated 5% in sample size calculation) and missing data, the use of per protocol analysis |

## VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

I think addidional information in the authors' response to reviewers should have been have been reproduced in the manuscript as well. Many of these objections will certainly come up when the results of the study are to be published.

Some examples:

-Rationale for choice 4.5 mg naltrexone add as intervention dose.

We described the rationale for the 4.5 mg dose in the manuscript, as was explained in our cover letter response.

-details on sample size calculation (ref. to Smith et al).

We added two sentences on the available data of Smith et al and added the reference.

-Composition of placebo

We have added the information about the placebo composition from the IMPD to the manuscript.

-The handling of dropouts (estimated 5% in sample size calculation) and missing data, the use of per protocol analysis.

Thank you. We have added our response from the cover letter to the manuscript.