BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

**ARTICLE DETAILS**

| TITLE (PROVISIONAL) | Acceptability and Efficacy of the Zemedy App versus a Relaxation Training and Meditation App for IBS: Protocol for a randomized controlled trial |
|---------------------|----------------------------------------------------------------------------------------------------------------------------------|
| AUTHORS             | Hunt, Melissa; Dalvie, Anika; Ipek, Simay; Wasman, Ben                                                                            |

**GENERAL COMMENTS**

In the present manuscript "Acceptability and Efficacy of the Zemedy App versus a Relaxation Training and Meditation App for IBS: Protocol for a randomized controlled trial" the authors describe a study protocol used to test a novel CBT-based smartphone app. Although the protocol describes the planned procedures in detail, the following points should be addressed by the authors:

1. The authors thoroughly present the outcomes of their investigations of a previous version of the app. They describe significant improvement of primary and secondary measures; however, they miss to elaborate on the clinical significance of the outcomes.

2. The authors explain there is an intended sequence of modules that are supposed to be completed followed by additional 2 weeks after the 6 modules. The authors nonetheless do not mention whether there is a schedule/check list tool for patients to keep an overview over their progress.

3. The authors mention in the present manuscript that the app has some tools that need unlocking. They do not explain the required actions to unlock a tool and the motive behind locking certain tools. Locking tools in general appears inappropriate for an app supposed to reduce the obstacles for people suffering from IBS to receive therapy.

4. The authors decided to include recommendations for food diaries and exclusion diets in their control app, mentioning it stands in contrast to recommendations of CBT. Is there data showing inferiority of diet and diaries compared to CBT, if this is the case, recommending it to the control cohort could be interpreted as unethical.
5. The authors did not include enough information about the inclusion procedures. Who selects patients, how is the correctness of diagnosis controlled, etc.?

6. The authors list several exclusion criteria, but they do not say if an existing or past eating disorder is also regarded as an exclusion criteria.

7. The authors intend use questionnaires several times throughout the study; however, the description of the time points is not detailed enough, e.g. they miss to mention when they plan to introduce the baseline questionnaire to the participants. The authors could include a timeline for better understanding.

8. A quick online research shows that there is a lot information available about zemedy explaining that it is based on CBT. How do the authors intend to ensure study participants do not learn of their group allocation before the end of the trial?

9. The authors describe a compensation of 20$ for participants completing the study. They do not say when the participants learn about this compensation? In addition, one could ask if the authors expect that it could encourage inappropriate usage of the app.

10. The authors plan to conduct multiple comparisons (anxiety, catastrophizing, fear of food, GI symptoms, QoL); thus, they should consider the need for conducting a Bonferroni correction.

---

**REVIEWER**

Raman, Maitreyi  
University of Calgary, Department of Medicine, Division of Gastroenterology

**REVIEW RETURNED**  
28-Jul-2021

**GENERAL COMMENTS**

Thank you for the opportunity to review the manuscript by Hunt et al, titled Acceptability and Efficacy of the Zemedy App versus a Relaxation and Medication App for IBS: Protocol for a randomized controlled trial. The aim of this study is to examine the acceptability and efficacy of Zemedy 2.0, a digital health app that provides CBT-based treatment for IBS. Overall, this article is well written, but several details surrounding the methodology are lacking and need to be included. This review was completed by two individuals, a practicing gastroenterologist with expertise in lifestyle medicine, and a PhD student.

1. Abstract – Study duration and time point for primary outcomes are needed. Primary endpoints / outcomes need to be described. What changes in outcome measures will be considered significant?

2. Check prevalence of 6-7% in the intro. This seems very low. IBS prevalence rates range between 10-30%

3. Please provide a rationale for each module of the Zemedy 2.0 application. Describe motivational interviewing for those who may not be familiar with this technique.

4. Undoubtedly in some patients diet has a role to play in IBS management. Is guidance provided about food reintroduction?
5. Change the study design to describe it as a superiority, non-blinded trial and include the location of where they study is taking place (e.g., research hospital) and which countries participants are recruited from.

6. Who will enrol and screen the participants?

7. Will patients be sub-categorized based on IBS type? Constipation, diarrhea or mixed?

8. Under Interventions and Assessments, the authors describe the participants begin working through the modules as soon as they download the app. Is there a module schedule (e.g., one module per week), or is the intervention group self-directed?

9. Describe the primary outcome measure when justifying the sample size. The authors state that benefit for HRQOL was observed in the previous study and used this to power the current study. However, it is not clear whether HRQOL is the primary outcome measure for this study until much later.

10. Defend why there are two primary outcomes in your study since in your first trial using Zemedy 1.0 only HRQL was evaluated. Do you have enough power to examine both?

11. Please provide more information on the IBS-QOL measure similarly to the GSRS-IBS.

12. For each outcome measure, please include the cut-offs that will demonstrate change, either statistically and/or clinically significant.

13. Why is the Modified Rome IV Questionnaire used as an outcome measure to determine whether participants meet current Rome IV diagnostic criteria for IBS? What is the goal of having this information? Shouldn’t this tool rather be used to ensure participants have IBS in the eligibility screening process?

14. For acceptability, please consider measuring attrition rates and defining what an acceptable withdrawal rate would be for this trial to demonstrate acceptability.

15. The data analysis section described that the results will be compared at the 3-month follow-up mark. What is the authors’ plan for the data collected at 6 and 12 months?

16. What confounding variables will the authors control for in the analysis? Depression, Anxiety, IBS sub-type, physical activity, medications, etc

17. A discussion section would be interesting. While the authors discuss their previous Zemedy application, it would be interesting to understand if there is literature that discusses experience with other IBS mobile apps, and how Zemedy is complementary.

REVIEWER          Fikree, Asma
                   The Royal London Hospital
REVIEW RETURNED   29-Jul-2021
**GENERAL COMMENTS**

Please can you include more information about ethical approval and how you will consent patients.

In Figure 1 - you have not accounted for attrition rate of active controls - there is likely to be drop out at that stage so you will have fewer than 75 who cross over to treatment group. Will you need to adjust your numbers accordingly?

How do you ensure engagement with active control group? If participants know that they will cross over into remedy group 8 weeks after active control group, perhaps that will not engage with active control and simply wait for enrolment with Zemedy. what measures do you take to prevent that?

You send patients a reminder email 4 weeks after the start of the active control or treatment group - is this enough notice? If you are sending them feedback questionnaires at 8 weeks may it be a good idea to email them 2 weeks after the start to ensure engagement, and then again at 4 and 6 weeks?

With active control group - they cross over to Remedy after 8 weeks, and are given a questionnaire 8 weeks following that. Is this enough time or should you build some slack into the system? If cross over at 8 and then further post Zemedy questionnaire at 8 weeks following that you risk not getting accurate feedback from those that might be a few weeks behind in group which crossed over from active control to remedy.

data analysis: you mention that you will be assessing catastrophising as well as visceral anxiety (VSI) and fear of food (FFQ) - how will you measure catastrophising? You have not included a questionnaire for this - please specify.

---

**VERSION 1 – AUTHOR RESPONSE**

Response to Reviewer 1

1. We have noted that participants in the RCT of V 1.0 of the app achieved clinically significant improvement, as well as statistically significant improvement.

2. We have clarified that the Zemedy app does track the user’s progress, but that users work through the modules at their own pace.

3. We agree that locked tools might be frustrating to users. We have clarified that this is a standard approach to “gamifying” apps that is supposed to increase engagement, but that we will be surveying users about the acceptability of this feature of the app at the end of the trial.

4. We have clarified that the low FODMAP diet is an efficacious treatment for IBS, particularly with regard to GI symptom severity, and that it is non-inferior to self-help CBT with respect to HRQL at least in the short term.

5. The reviewer requested clarification of our inclusion procedures. We have added text to that section providing further explanation of our procedures.
6. Existing or past eating disorders are not among the exclusion criteria. Many patients with IBS will meet criteria for fear based ARFID, but the CBT protocol actually addresses fear and avoidance of food. We have clarified this in the text.

7. We have added a sentence to the paragraph on Intervention and Assessments to clarify that the baseline questionnaires are administered during screening of participants prior to enrollment and randomization.

8. As specified in the paragraph on blinding, because of the nature of the trial participants are not blinded and are aware of their group allocation upon randomization.

9. We have added a paragraph on Consent and have clarified that participants learn about the compensation for completing study questionnaires at several follow-up timepoints from the consent form. The compensation to try to limit loss of follow-up data. Participants are not compensated for using the app itself.

10. The reviewer suggested that we consider the need for Bonferroni correction to compensate for multiple tests. We only have two primary outcome measures which we will examine at two main time points. Moreover, we view Bonferroni correction as overly conservative (e.g. see VanderWeele & Mathur, 2019. https://doi.org/10.1093/aje/kwy250). Indeed, we would argue (with others) that statistical significance is overrated (see Amrhein, Greenland, McShane et al. (2019). Scientists rise up against statistical significance. Nature, 567, 305-307). Rather, we think it is important to examine the clinical significance of change and fully intend to do so.

Response to Reviewer 2
1. We have edited the abstract to clarify the study duration and time points at which we are administering primary outcome measures. Additionally, the primary outcomes and clinically significant change have been defined.

2. As described in Palsson (2016), the Rome IV diagnostic criteria for IBS are more stringent than previous versions, and thus the prevalence is lower. However, we have added an additional reference (Van den Houte et al., 2019) which details the much higher prevalence of self-reported IBS that fails to meet strict diagnostic criteria, but still results in considerable distress and disability.

3. We have added text providing a brief rationale for every module and explaining the nature of motivational interviewing.

4. We have clarified the nature of the intervention in the module on food.

5. We have changed the description of the study design to a “randomized, superiority, non-blinded cross-over trial”. We have also added a sentence specifying that participants are recruited from the United States, but because all of the study is conducted remotely, there is no physical location where the study is taking place.

6. We have added a sentence explaining that participants will be screened and enrolled by the study coordinator.

7. We have added a sentence explaining that participants have the opportunity to report their IBS subtype in the modified Rome IV measure completed at baseline.
8. The intervention group is self-directed, and we have clarified that.

9. We have added discussions of the primary outcome measures in the Power Analysis section. We have also described in the abstract what our primary outcome measures are.

10. On page 5 the text states that the RCT of Zemedy 1.0 had two primary outcome measures - GI symptom severity and HRQL. This study uses the same two measures. We have clarified that in the power analysis and added the effect size of Zemedy 1.0 on GI symptom severity.

11. We have added information to the description of the IBS-QOL including qualitative score ranges and sample items. Please provide more information on the IBS-QOL measure similarly to the GSRS-IBS.

12. We have added qualitative score ranges and cut-offs for clinically significant change.

13. As described in the Modified Rome IV Questionnaire section, we will use the questionnaire both at eligibility screening and at post-treatment and follow-up timepoints. This is to determine if participants still meet Rome IV criteria following treatment with the Zemedy app.

14. We will certainly measure and report attrition. As described in our power analysis section, we expect an attrition rate of around 50%, which is typical for behavioral health studies using online recruitment and low intensity, distance interventions. We have also added a sentence regarding attrition to the intervention and assessment section.

15. We have added a sentence in the Data Analysis section explaining that we will also compare results at 3, 6, and 12 month follow-up.

16. We have added a section in the Data Analysis section specifying that baseline symptom severity, depression and IBS subtype will be examined as potential moderators of treatment efficacy.

17. Protocol papers submitted to BMJ Open do not typically include a discussion section.

Response to Reviewer 3

1. We have added information about the ethical approval and the consent process.

2. The reviewer is correct that Figure 1 did not document the full level of expected attrition in the active control group over the full length of the study. We have amended the figure to be more accurate.

3. The reviewer inquired as to how we plan to ensure engagement with the active control group. The sham app still contains a good deal of informative text and a number of links to relaxation videos. Individuals in the active control group will receive an email at 4 weeks encouraging them to stick with the app. IBS has a high placebo response rate, but beyond providing a credible sham app, we are not doing anything specific to ensure engagement.

4. The reviewer suggested increasing the frequency of emailing participants. We are balancing engagement and potential attrition against ecological validity. In the real world when an IBS patient
downloads an app, they don’t get any emails from study personnel. We are hoping that our protocol strikes a balance between the oversight of a clinical trial and the real world use of the app.

5. The reviewer raised a concern that 8 weeks post-crossover was too fast for obtaining accurate follow-up data. We do indeed have “slack” in the system, insofar as the actual treatment protocol is about 6 weeks, with two weeks of extra time built in.

6. We have added a phrase explaining that catastrophizing will be measured by the GI-Cog.

We hope that with these revisions, the paper is now acceptable.

| REVIEWER       | stengel, Andreas  |
|----------------|--------------------|
|                | University Hospital Tübingen Department of Psychiatry and Psychotherapy |

**GENERAL COMMENTS**

In the present manuscript "Acceptability and Efficacy of the Zemedy App versus a Relaxation Training and Meditation App for IBS: Protocol for a randomized controlled trial" the authors describe a study protocol used to test a novel CBT-based smartphone app. Although the authors have revised the manuscript, the following points should still be addressed:

1. The authors decided to include recommendations for food diaries in their control app, mentioning it is contraindicated in CBT; if this is indeed the case, recommending it to the control cohort can be interpreted as unethical.

2. The authors state that they want to include participants that report having been diagnosed with IBS by a physician, but do not currently meet strict Rome IV diagnostic criteria on the questionnaire. Since the authors indicate in the summary that patients will not need to present any proof of diagnosis the risk of a heterogenous study population including patients that do not meet the necessary inclusion criteria and their effect on the results needs to be evaluated by the authors.

3. The authors should include a timeline depicting time of interventions and assessments.

4. The authors describe a compensation worth 20$ for participants completing the study. They do not say when the participants learn about this compensation? In addition, one could ask whether the authors expect that it could encourage inappropriate usage of the app.

5. The authors plan to conduct multiple comparisons (anxiety, catastrophizing, fear of food, GI symptoms, QoL); thus, they should consider the need for conducting Bonferroni correction, thus the need for adjustment of the significance levels.
Response to Reviewer 1

A number of Reviewer 1’s concerns are repeated verbatim from their first round review. We believe we have addressed all of them, including the one novel concern about our diagnostic and inclusion criteria.

1. The authors decided to include recommendations for food diaries in their control app, mentioning it is contraindicated in CBT; if this is indeed the case, recommending it to the control cohort can be interpreted as unethical.

We have already addressed this concern previously, but have added a bit more explanatory text on page 7 to clarify that including food diaries in the control app is NOT unethical: “because they work via opposing mechanisms. Nevertheless, restrictive diets are empirically supported,”

2. The authors state that they want to include participants that report having been diagnosed with IBS by a physician, but do not currently meet strict Rome IV diagnostic criteria on the questionnaire. Since the authors indicate in the summary that patients will not need to present any proof of diagnosis the risk of a heterogenous study population including patients that do not meet the necessary inclusion criteria and their effect on the results needs to be evaluated by the authors.

As described in Palsson (2016), the Rome IV diagnostic criteria for IBS are more stringent than previous iterations of diagnostic criteria for IBS. Many IBS patients were diagnosed prior to 2016 when the Rome IV criteria were promulgated. They would still have been told by a physician that they had IBS. In addition, we previously added an additional reference (Van den Houte et al., 2019) which details the much higher prevalence of self-reported IBS that fails to meet strict diagnostic criteria, but still results in considerable distress and disability. Thus, we are not concerned about the heterogeneity of the sample. To the contrary, our inclusion criteria are sampling from precisely the population of interest – people who believe that they have IBS, having been told so by a physician, and/or meeting strict Rome IV criteria, who are unhappy with their health status and are interested in a self-help app.

3. The authors should include a timeline depicting time of interventions and assessments.

On page 9 we have attempted to clarify the timeline of interventions and assessments further.

4. The authors describe a compensation worth 20$ for participants completing the study. They do not say when the participants learn about this compensation? In addition, one could ask whether the authors expect that it could encourage inappropriate usage of the app.
We have clarified that the information about compensation is included in the consent form (page 8). We are uncertain what sort of “inappropriate usage” the reviewer is concerned about. The biggest concern in trials of this type is attrition and lack of follow-up data completion. The compensation is intended to incentivize participants to complete the follow-up questionnaires, and has no bearing on their use of the app itself.

5. The authors plan to conduct multiple comparisons (anxiety, catastrophizing, fear of food, GI symptoms, QoL); thus, they should consider the need for conducting Bonferroni correction, thus the need for adjustment of the significance levels.

We have addressed this previously and refer again to our comments in the last letter. We stand by this position firmly.

“The reviewer suggested that we consider the need for Bonferroni correction to compensate for multiple tests. We only have two primary outcome measures which we will examine at two main time points. Moreover, we view Bonferroni correction as overly conservative (e.g. see VanderWeele & Mathur, 2019. https://doi.org/10.1093/aje/kwy250). Indeed, we would argue (with others) that statistical significance is overrated (see Amrhein, Geenland, McShane et al. (2019). Scientists rise up against statistical significance. Nature, 567, 305-307). Rather, we think it is important to examine the clinical significance of change and fully intend to do so.”

VERSION 3 – REVIEW

| REVIEWER     | stengel, Andreas  |
|--------------|-------------------|
|              | University Hospital Tübingen Department of Psychiatry and Psychotherapy |
| REVIEW RETURNED | 16-Nov-2021 |
| GENERAL COMMENTS | The authors responded sufficiently to all queries. |