PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Evaluation of the Digital Diabetes Prevention Programme Pilot: Uncontrolled Mixed Methods Study Protocol.
AUTHORS	Murray, Elizabeth; Daff, Kerry; Lavida, Anthi; Henley, W; Irwin, Jenny; Valabhji, Jonathan

VERSION 1 - REVIEW

REVIEWER	Karla Galaviz Emory University, USA.
REVIEW RETURNED	30-Aug-2018

GENERAL COMMENTS	Thank you for the opportunity to review this interesting study. This study aims to assess the effectiveness and implementation of a digital diabetes prevention program. It has a strong theoretical base and proposes to provide valuable data to inform the wide scale implementation of a digital diabetes prevention program. Though it is a protocol for a study, many important details about the methods and interventions proposed are lacking. To better understand what authors are proposing, I recommend revising the aims as follows: 1) pilot test the effectiveness of the digital interventions and 2) evaluate the implementation process of such interventions. These 2 objectives would fit nicely within the theoretical frameworks authors propose to use: one pertaining to the effectiveness of the digital diabetes prevention intervention (Aim 1), and one pertaining to the implementation processes (Aim 2). My main concern about the methods is that this study appears to be a clustered study where sites are selected, then GPs then patients; yet, there are no methods employed to address neither the power inefficiency of such design nor the inflated type I error in effectiveness analyses. Details about how clusters, GPs and patients will be selected are lacking. Finally, 5 interventions will be tested but none of these are described. A protocol paper should describe these in detail. Attention to writing quality is also needed and bullet points should be avoided. Introduction: • The bullet points in the introduction are distracting and break up the flow of the text. It is better to have a paragraph describing these and making a strong case of why this evaluation is needed. • The rational of why this evaluation is needed is not clear. Authors need to report what other digital programs have found in terms of impact and what gaps remain. Then, authors should state how this evaluation will address those gaps. This sets the stage to introduce the Aims of the study. • Unless the journal requires the Aims be listed separately from the introduction, I would report them at the end of the introduction.

	Also the objectives could be revised as follows: the main objective (to pilot test the effectiveness of the intervention) and a secondary objective (to evaluate the implementation process). The 9 objectives subsidiary objectives can be merged into the process evaluation objective. These 2 objectives would fit nicely within the theoretical frameworks: one pertaining to the effectiveness of the digital diabetes prevention intervention (Aim 1), and one pertaining to the implementation processes (Aim 2). Methods: • Consider reorganizing this section – information from related parts is scattered throughout and makes it difficult to understand what the authors are proposing. Consider organizing methods by aim or by data collection type – quantitative and qualitative. • How will the settings be selected? By convenience? This should be specified. Also, this seems to be a clustered study and details about how participants in each of the 8 demonstrator sites (i.e. clusters) will be selected are lacking. Strategies to account for the cluster design effect should be applied to improve power (e.g. inflate sample by design effect). The current sample size and power calculation does not account for the design effect. • Details about how participants will be selected (convenience?), recruited (posters?) and linked to the digital program are needed. • Avoid using bullet points: it is better to describe each of the 5 interventions in a paragraph. Using bullet points reflects poor writing quality. • Need to describe how eligible patients will be identified by GPs for referral. And where will patients will be referred if these are digital interventions? Specify how patients will be linked to or given access to digital programs. • Authors state “The interventions will be described according to the TiDIER Framework for describing complex interventions (41) and the CALO-RE Behaviour Change Technique taxonomy (40)” but interventions are not described. What are the 5 interventions selected? Interventions should be described in detail based on these frameworks. • Need to report what is the legibility criteria for GPs and how they will be selected and recruited. • How will sites, GPs and patients be assigned to interventions? • The primary outcomes are changes in A1c and weight; yet, the objective of the study does not mention effectiveness. The aims need to be revised to better reflect this. • How will focus groups be conducted? How many? Who will moderate them? Will they be recorded? Which groups of stakeholders will go together and why? How will stakeholders be selected? The same goes for interviews. More details about these qualitative data collection methods are needed. • There is no adjustment for clustered data planned for analyses. This is a major flaw since type I error is inflated in clustered studies. This needs to be adjusted in effectiveness analyses. • How will implementation data be used? Will this be used to understand program effectiveness? Will this be used to inform program adaptation and rollout?

REVIEWER	Gladys Block USA. Turnaround Health Univ Calif Berkeley (retired) We have developed a digital diabetes prevention program like those to be tested.
REVIEW RETURNED	01-Sep-2018

GENERAL COMMENTS	This is a well-written paper describing a protocol for a planned pre-post study of several digital diabetes prevention programs. It appears that invitations to digital programs were primarily focused on British organizations. Regarding statistical analysis plans, I suggest that the authors plan to include an intention-to-treat analysis. The numbers completing a one-year program will be relatively small, as the authors recognize. But some programs will be better than others at retaining subjects, and an ITT analysis might shed some further light on the relative merits of the programs.

VERSION 1 – AUTHOR RESPONSE

Reviewer Comment	Response	Page number
Reviewer: 1		
Thank you for the opportunity to review this interesting study. This study aims to assess the effectiveness and implementation of a digital diabetes prevention program. It has a strong theoretical base and proposes to provide valuable data to inform the wide scale implementation of a digital diabetes prevention program.	Thank you.	
I recommend revising the aims as follows: 1) pilot test the effectiveness of the digital interventions and 2) evaluate the implementation process of such interventions. These 2 objectives would fit nicely within the theoretical frameworks authors propose to use: one pertaining to the effectiveness of the digital diabetes prevention intervention (Aim 1), and one pertaining to the implementation processes (Aim 2).	Please see below	5
My main concern about the methods is that this study appears to be a clustered study where sites are selected, then GPs then patients; yet, there are no	We agree with the reviewer that this study can be considered as a clustered study because participants are naturally clustered within GP practices within 8 demonstrator sites. There is also the potential for clustering effects due to the allocation of five digital diabetes prevention interventions across the different	12 - 13

methods employed to address neither the power inefficiency of such design nor the inflated type I error in effectiveness analyses.	demonstrator sites. A key issue to be addressed is whether the form of clustering in this study is ignorable (Kahan and Morris BMC Medical Research Methodology, 2013, 13:58: https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-13-58). As the reviewer highlights, clustering can cause inflation of type 1 errors in some situations, such as where treatment is assigned randomly to clusters (as in a cluster randomised trial). Clustering is non-ignorable when there is both correlation between patient outcomes within clusters, and correlation between treatment assignments within clusters. Where clustering is non-ignorable, it should be accounted for in the sample size calculations and analysis. However, there are some situations where an analysis that does not account for clustering may be preferable. For example, when there are multiple layers of clustering (e.g. patients within therapists within hospitals within countries), attempting to control for all levels of clustering can lead to an overly complex analysis that may not work well in practice (Kahan and Morris, 2013). The clustering in this evaluation is complex involving multiple components of clustering. Furthermore, the Digital Diabetes Prevention Programme is based on a pretest – posttest design with each participant serving as their own control. Importantly in terms of considering whether clustering is ignorable, treatment assignment is not associated with the clustering under the assumption that the different digital interventions are interchangeable. In the terminology of Kahan and Morris, this would imply that V(E)=0 and the clustering does not need to be included in the analysis in order to obtain valid type I error rates. Given the complexity of the clustering and the potential for the clustering to be ignorable, we decided not to account for it in the original protocol. However, we have now extended our power calculations and analysis plan to allow for the possibility of non-ignorable clustering as follows: a) Power analysis: we considered the effect of clustering by demonstrator site (ignoring the further impact of clustering at the GP practice level and clustering by digital intervention / intervention component) and made the conservative assumption that assignment of the digital diabetes prevention interventions was highly correlated within sites. We then estimated minimum detectable effect sizes at 90% power and a 5% significance level for the NDH and overweight/obese groups, allowing for clustering using	

	a design effect. We used an ICC (intra-cluster correlation) of 0.02 based on a median estimate of 0.0185 in a study of intra-cluster correlation coefficients in adults with diabetes in primary care practices (Littenberg and MacLean. BMC Med Res Methodol. 2006; 6: 20; doi: 10.1186/1471-2288-6-20). Assuming equal numbers of patients recruited per site, this gave minimum detectable effect sizes of d=0.18 and 0.22, assuming a 25% completion rate at 12 months. Both effect sizes are consistent with the weighted mean effect size of d=0.22 estimated in a meta-analysis by Johnson et al (ref 48 in manuscript) for behaviour change interventions targeting eating and physical activity, suggesting the study is adequately powered to detect changes in the primary outcomes after allowing for clustering by healthcare site. The sample size section of the manuscript has been updated accordingly. b) Analysis plan: The proposed generalised linear modelling framework for analysing changes in outcomes will be extended by inclusion of random effects for the demonstrator site to account for potential clustering effects. The possibility of fitting three level models that account for clustering at the GP practice level will also be explored. The Data Analysis section of the manuscript has been updated accordingly.	
Details about how clusters, GPs and patients will be selected are lacking.	This is the responsibility of the demonstrator sites. This has been clarified in the section on populations and participants	6.
Finally, 5 interventions will be tested but none of these are described. A protocol paper should describe these in detail.	We agree that it is important that interventions are well described, and are following current advice on best practice by writing a parallel paper describing the 5 interventions, including characterising their active components according to the CALO-RE taxonomy. We have also summarised the important features of each intervention in this paper.	7 - 8
Introduction		
The bullet points in the introduction are distracting and break up the flow of the text. It is better to have a paragraph describing these and making a strong case of why this evaluation is needed. The rational of why this evaluation is needed is not clear. Authors need to report what other digital programs	This has been rewritten without the use of bullet points. There is a clear statement of the need for an evaluation, in terms of the potential benefits, reference to systematic review evidence of potential effectiveness, and clear statement of challenges. This leads into the aims and objectives.	4 - 5

have found in terms of impact and what gaps remain. Then, authors should state how this evaluation will address those gaps. This sets the stage to introduce the Aims of the study.		
Unless the journal requires the Aims be listed separately from the introduction, I would report them at the end of the introduction.	This has been done.	5
Also the objectives could be revised as follows: the main objective (to pilot test the effectiveness of the intervention) and a secondary objective (to evaluate the implementation process). The 9 objectives subsidiary objectives can be merged into the process evaluation objective. These 2 objectives would fit nicely within the theoretical frameworks: one pertaining to the effectiveness of the digital diabetes prevention intervention (Aim 1), and one pertaining to the implementation processes (Aim 2).	This is a good suggestion, but does not accurately reflect either the commissioning brief or the research questions that the evaluation will address. The commissioning brief (tender) clearly specified what was required and what was out of scope. We have rewritten the section on aims and objectives to make it clearer that these were largely determined by the commissioning brief (tender).	5
Methods		
Consider reorganizing this section – information from related parts is scattered throughout and makes it difficult to understand what the authors are proposing. Consider organizing methods by aim or by data collection type – quantitative and qualitative.	We agree that this is a complex protocol to report, as there are 2 different populations, 8 different settings, 5 different interventions and both quantitative and qualitative data. We have reorganised the methods sections to make it clearer. We have used the standard subheadings of Design Setting Population and participants Interventions Outcomes Data collection Data analysis Ethics Dissemination.	5 - 13

	For outcomes, data collection and data analysis we have presented first the quantitative methods and then the qualitative methods.	
How will the settings be selected? By convenience? This should be specified. Also, this seems to be a clustered study and details about how participants in each of the 8 demonstrator sites (i.e. clusters) will be selected are lacking. Strategies to account for the cluster design effect should be applied to improve power (e.g. inflate sample by design effect). The current sample size and power calculation does not account for the design effect.	These were selected by NHSE and not the evaluation team. This has been clarified. See above	6 12 - 13
Details about how participants will be selected (convenience?), recruited (posters?) and linked to the digital program are needed.	Describing and understanding how patients are identified and referred are part of the research questions. It is important to realise that the implementation of the clinical service is not under the control of the evaluation team, and the evaluation team are not responsible for identifying or referring patients. It is an error to conceive of the patients as research participants who are selected and recruited – the participants are patients identified through standard clinical workflows and referred following standard clinical practice. Each local health economy has determined its own workflows / pathways for identifying the target population and referring them to the available intervention. Part of the evaluation involves studying these different pathways. This has been clarified in the section on participants and population.	6
Avoid using bullet points: it is better to describe each of the 5 interventions in a paragraph. Using bullet points reflects poor writing quality.	All bullet points have been removed and each intervention described in a paragraph	7-8
Need to describe how eligible patients will be identified by GPs for referral. And where will patients will be referred if these are digital	Please see response above. The text on page 6 has been rewritten and clarified to make it clearer that each demonstrator site will determine their own methods for identification and referral of patients. How patients are onboarded to the digital interventions is now described	6

interventions? Specify how patients will be linked to or given access to digital programs.	in the intervention section. In addition, as the mechanisms are the same as for the face-to-face programme we have referenced that.	
Authors state "The interventions will be described according to the TiDIER Framework for describing complex interventions (41) and the CALO-RE Behaviour Change Technique taxonomy (40)" but interventions are not described. What are the 5 interventions selected? Interventions should be described in detail based on these frameworks.	They have now been described. The selection of the interventions was undertaken by NHSE (as described in the paper), not by the evaluation team. We are in the process of familiarising ourselves with the interventions, and at present, are heavily reliant on information provided by the commercial companies. We are therefore reluctant to put too much detail in, as we have not been able to verify the detail yet. Doing this forms part of the evaluation, and hence the reference to the TiDIER and CALO-RE taxonomies, so readers can understand that this work is still to be done.	6 - 7
Need to report what is the eligibility criteria for GPs and how they will be selected and recruited.	Again, this is not under the evaluation teams control, and is the same as for the face-toface programme. This has been stated on page 6 and a reference provided	6
How will sites, GPs and patients be assigned to interventions?	This is the responsibility of the demonstrator sites. This has been clarified on page 6.	6
The primary outcomes are changes in A1c and weight; yet, the objective of the study does not mention effectiveness. The aims need to be revised to better reflect this.	The objectives mention impact. We have carefully avoided using the word "effectiveness" as effectiveness is best determined through a trial. Given the design of this study, any changes in HbA1c or weight cannot be causally ascribed to the digital diabetes prevention intervention. This is stated in the "strengths and limitations" section.	5
How will focus groups be conducted? How many? Who will moderate them? Will they be recorded? Which groups of stakeholders will go together and why? How will stakeholders be selected? The same goes for interviews. More details about these qualitative data collection methods are needed.	Since submitting the protocol we have decided not to undertake focus groups, and reference to this has been removed. More detail has been provided about the interviews.	9
There is no adjustment for clustered data planned for analyses. This is a major flaw	Please see response above.	12 - 13

since type I error is inflated in clustered studies. This needs to be adjusted in effectiveness analyses.		
How will implementation data be used? Will this be used to understand program effectiveness? Will this be used to inform program adaptation and rollout?	Yes, this is stated in the objectives. (objective 9; and overall aim).	5
Reviewer 2		
This is a well-written paper describing a protocol for a planned pre-post study of several digital diabetes prevention programs. It appears that invitations to digital programs were primarily focused on British organizations.	Thank you.	
Regarding statistical analysis plans, I suggest that the authors plan to include an intention-to-treat analysis. The numbers completing a one-year program will be relatively small, as the authors recognize. But some programs will be better than others at retaining subjects, and an ITT analysis might shed some further light on the relative merits of the programs.	We thank the reviewer for this comment. We agree that it would be desirable to conduct an intention-to-treat analysis so that estimates of effectiveness can be generalised to the target population. In order to conduct the ITT analysis we would need access to relevant outcomes at 6 and 12 months follow-up irrespective of whether the patients continued in the programme. It may be possible to achieve this for some patients by accessing routine medical records but for many patients this may not be possible. Instead, we will address the potential for bias due to non-random attrition by fitting a propensity score model to account for drop-out on the basis of baseline characteristics and then using inverse probability weighting (IPW) based on the propensity score to fit the treatment effectiveness model (Cole SR, Hernán MA. Am J Epidemiol. 2008 Sep 15;168(6):656-64. doi: 10.1093/aje/kwn164.).	12 - 13

VERSION 2 – REVIEW

REVIEWER	Gladys Block USA. Turnaround Health University of California Berkeley (retired) I am the owner and developer of a digital diabetes prevention program, Turnaround Health's Alive-PD.
REVIEW RETURNED	03-Jan-2019

GENERAL COMMENTS	This continues to be a well-written paper describing the protocol for a planned study of digital diabetes prevention programs in the UK. This will be very important research. I have minor suggestions below, but I don't think publication needs to be further delayed for extensive revision. Apparently the study was launched in 2017. The paper should include information about when it is expected to be completed. Is it already undergoing analysis? I believe the fact that analysis will be on participants with complete data should come earlier in the Data Analysis section. For data analysis, the authors assume 25% completion rate (sensible!). How will "completion" be defined? Provision of blood and weight data in 100% of sessions? Some proportion of that? Provision of blood in the final session? Regarding analysis, the authors will examine costs of the intervention. In addition to other planned analyses, I suggest the authors include an analysis of costs per successfully-treated persons — that is, who reach normal levels of HbA1c or fasting glucose. Personally, I think the "Ethics, research governance and data security" section is longer than it needs to be. PPI was not defined. Minor typos: p.8 line 21: need period at end. p.8, line 35: 'initiate' rather than 'initiative'

VERSION 2 – AUTHOR RESPONSE

Reviewer comment	Response	Page number
This continues to be a well-written paper describing the protocol for a planned study of digital diabetes prevention programs in the UK. This will be very important research. I have minor suggestions below, but I don't think publication needs to be further delayed for extensive revision.	Thank you.	
Apparently the study was launched in 2017. The paper should include information about when it is expected to be completed. Is it already undergoing analysis?	We have added the due date for the final report (2020)	5

I believe the fact that analysis will be on participants with complete data should come earlier in the Data Analysis section.	This has been moved up to the start of the section on analysis of outcomes.	P12
For data analysis, the authors assume 25% completion rate (sensible!). How will "completion" be defined? Provision of blood and weight data in 100% of sessions? Some proportion of that? Provision of blood in the final session?	The definition of completion has been added	P 13
Regarding analysis, the authors will examine costs of the intervention. In addition to other planned analyses, I suggest the authors include an analysis of costs per successfully-treated persons — that is, who reach normal levels of HbA1c or fasting glucose.	Thank you for this interesting suggestion. As mentioned, formal economic analysis is out of scope, but we will consider the practicability of this additional analysis. We have not changed the protocol as this would be an additional analysis beyond that agreed with NHSE.	
Personally, I think the "Ethics, research governance and data security" section is longer than it needs to be.	Thank you, yes, on review we agree. We have shortened it.	P13 -14
PPI was not defined.	This has been done	P6.
Minor typos: p.8 line 21: need period at end. p.8, line 35: 'initiate' rather than 'initiative'	Thank you, both these have been corrected	P8