PEER REVIEW HISTORY

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)	Sensing Interstitial Glucose to Nudge Active Lifestyles (SIGNAL): Feasibility of combining novel self-monitoring technologies for persuasive behaviour change.
AUTHORS	Whelan, Maxine; Kingsnorth, Andrew; Orme, Mark; Sherar, Lauren; Esliger, Dale

VERSION 1 – REVIEW

REVIEWER	Esther van Sluijs CEDAR, University of Cambridge, UK I am a collaborator with two authors on a different project, which has a different focus. I have no other conflicts of interests to declare.
REVIEW RETURNED	10-Jul-2017

GENERAL COMMENTS	This paper describes the protocol of a study to assess the feasibility of providing continuous (bio)behavioural feedback to participants at risk for diabetes. The ambitious project is generally well-described and the aims, study processes and data collection procedures align with the feasibility focus of the project. Below I provide some suggestions for consideration/revision. Major compulsory revisions 1. The project is predicated on the suggestion that there is a direct impact of small amounts of physical activity on glucose and that this is measurable with the device deployed. However, no evidence of this is provided in the introduction or methods. 2. The strengths and limitations section is not very strong and should be amended (i.e. the first bullet is not a strength of the study). 3. The authors suggest that they want to align the SIGNAL project with the NDPP, which is a strength. However, cost of the intervention, and cost-effectiveness will be important evidence to inform future implementation. The intervention appears relatively high cost (purchasing of equipment, maintenance, support), but no assessment of cost is included. Please amend or justify. 4. More detail on the effectiveness of (behavioural and outcome-based) self-monitoring interventions on behaviour should be provided. Is there any evidence that suggests that sustained monitoring is required for behaviour change maintenance (bottom p3)? How does this fit with behaviour change theory? 5. The amount, detail and potential identifiable nature of data collected/monitored and stored, also outside of the UK, could be a concern. I encourage the authors to provide more detail on this, in particular the relevant regulatory precautions and data management/storage procedures.

	6. The suggested use of default settings for step targets may be too ambitious for the target population, who are unlikely to be doing sufficient activity at the start of the study. Have the authors considered introducing graded targets and other motivational techniques to encourage behaviour change, continued engagement with the feedback and maintenance of behaviour change? Minor revisions (in order of manuscript) Abstract; 1. Introduction – please clarify what is meant with 'efforts are limited' – are there limited efforts or are the effects limited, or something else? 2. P3, 3rd para – the link between control theory, BCT and the current intervention is unclear. Please revise this paragraph to create clearer flow. Introduction 3. The focus of the introduction is on prevalence of known diabetes, but a large proportion of diabetes is undiagnosed. Please comment on this. 4. Clearer definitions of IGR and IFG should be provided. Methods 5. Please remove either investigate or assess from primary aim. 6. The ISRCTN registration does not specify HbA1c as an inclusion criterion, please clarify. 7. Are you able to provide examples of compatible Android phones – are these the major brands? What is the spread of these phones? 8. How/when will HbA1c be assessed for the inclusion criteria? 9. P7 (final appointment) – please specify what DKT refers to (I was not able to find it). 10. Device masking – to what extent will participants be able to change the settings or remove the tape during the data collection phase? How will this be prevented/monitored? Will devices also be masked at follow up? 11. P9 (BCT) – what is the purpose of the 250/hr step goal – this only adds up to less than 50% of the total goal if one would achieve this every waking hour of the day? 12. P9 (measures) – please provides a little bit more information on the proposed data processing decisions for the Actigraph (non-wear, cut points etc). 13. P11 (measures) – please provide a bit more detail on the protocol for the mCAFT, is this a treadmill- or bike-based test, maximal or submax, etc? 14. Quantitative data analyses – it may be useful to monitor through which route participants were recruited? Will this be monitored? Supplementary Table 1: Under 'why', this is referred to as a method to 'nudge' people, but this is not discussed in the introduction. To what extent can this approach be considered a 'nudge' intervention? If it is, please justify in the introduction.

REVIEWER	Rousset INRA France
REVIEW RETURNED	11-Jul-2017

GENERAL COMMENTS	The manuscript aims to evaluate the user engagement with two applications : Fitbit and Librelink. Several points are not clear: Is the aim to estimate only the proportion of time that the applications are used and the number of scans and syncs? If yes : what do you deduce from these usages? Is the aim to compare Actigraph and Fitbit? Are Actigraph and Fitbit reference scientific devices for estimating accurately physical activity levels? if yes, please add references. Is the aim to compare the usage of the applications with and without feedback ? or to compare the information influence (blood glucose versus physical activity) on behavioral change? Is the aim to determine if a physical activity increase improve blood glucose level? There are too many possible objectives and the assumptions are not sufficiently explicit. What is the Ethica Data? What is Diasend? Could the authors draw a graph to explain the relationships between the applications, Ethica Data and Diasend ? and if the glucose and physical activity data are available to researchers ? How can researchers evaluate a change in physical activity? It is difficult to understand what behaviour change techniques are proposed to volunteers and how they are evaluated (efficacy on what? is target complete or incomplete?): BCT:1.1 (10000 steps and 10 floors climbed), BCT1.3 (4.0-5.9 nmol/L), BCT 2.3 recording of physical activity without fixed goal?, BCT 2.4 recording of glucose without fixed goal?, BCT : recording of glucose and physical activity without fixed goal and BCT 7.1 alert to walk 250 steps for an hour ? How the ANCOVAs will be used (what independent and dependent variables?) These points (aim, assumption, measured variables, analyses, software activity, data available to volunteers and researchers) have to be clarified in the revised manuscript : the readers need to understand easily the aim of the study, the used methods and the expected outputs. Please correct the reference n°10 (authors).

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

This paper describes the protocol of a study to assess the feasibility of providing continuous (bio)behavioural feedback to participants at risk for diabetes. The ambitious project is generally well-described and the aims, study processes and data collection procedures align with the feasibility focus of the project. Below I provide some suggestions for consideration/revision.

Major compulsory revisions

Comment 1: The project is predicated on the suggestion that there is a direct impact of small amounts of physical activity on glucose and that this is measurable with the device deployed. However, no evidence of this is provided in the introduction or methods.

Response: Thank you for raising this point. We have amended the introduction and updated the references pertaining to the acute relationship between physical activity and glucose. In addition, we have added references detailing the validity of the glucose device in the Methods.
Introduction: "For example, studies investigating the acute effects of brief physical activity bouts or interruptions to prolonged sedentary behaviour on glucose levels in controlled settings have found reductions in postprandial glucose as a result of increased movement (e.g. Peddier et al., 2013; DiPietro et al., 2013; Dunstan et al., 2012; Reynolds et al., 2016). As a result, the present study proposes that delivering behavioural and physiological feedback in parallel may be more persuasive rather than when delivered in isolation."
Introduction: "With ongoing developments, technologies such as flash glucose monitoring offer a wealth of information to users without the need for invasive fingerprick samples; offering a useful tool for non-diabetic individuals (who are not accustomed to regular fingerprick blood samples) (Bailey et al., 2015)."
Methods: "The Freestyle Libre demonstrates consistent accuracy throughout the 14 days with a mean absolute relative difference of 11.4% compared with capillary blood glucose, a lag time of 4.5-4.8 minutes and is not impacted by physical characteristics including age, BMI and HbA1c (Bailey et al., 2015)."

Comment 2: The strengths and limitations section is not very strong and should be amended (i.e. the first bullet is not a strength of the study).

Response: Thank you for highlighting this. We have amended the strength and limitations section as follows:
• The study will present real-time biological and behavioural information to participants using wearable technologies; a novel concept which has not been utilised in physical activity research.
• The study will be the first to deploy flash glucose monitors to people at risk of developing Type 2 diabetes.
• We will employ quantitative and qualitative methodologies to explore user engagement with the technology.
• We will use a validated survey to identify at moderate-to-high risk of developing Type 2 diabetes, among which we expect a proportion to have prediabetes.
• Whilst the duration of the intervention (six continuous weeks) will allow us to examine how engagement changes over time, the absence of additional follow-ups prevents the assessment of long-term use engagement and behaviour change maintenance.
• Recruitment of at-risk individuals from the community aligns with the National Diabetes Prevention Programme.
• Cost-effectiveness analysis will not be undertaken in this study.

• Due to the nature of feasibility studies, this study will not be powered for effectiveness.

Comment 3: The authors suggest that they want to align the SIGNAL project with the NDPP, which is a strength. However, cost of the intervention, and cost-effectiveness will be important evidence to inform future implementation. The intervention appears relatively high cost (purchasing of equipment, maintenance, support), but no assessment of cost is included. Please amend or justify.

Response: We thank the reviewer for this comment. We will not be conducting cost-effectiveness analysis because we believe it is too premature to do so at this stage. This is partly because the technologies we are deploying are not optimal yet (i.e. physical activity and interstitial glucose levels are recorded using two independent devices and their feedback presented on two independent applications rather than one). In addition, data collected from Ethica Data and the other platforms described will investigate how participants adhere to use and engage with the technologies; therefore, costing this study a priori will not be possible. If the results of this study support the potential for wearable technology to play a role in helping at-risk individuals engage with their health, we will incorporate a cost-effectiveness analysis in the next iteration. We have added this comment as a limitation of the study (please see response to Reviewer 1 Comment 2).

Introduction: “The present study intends to implement a community screening approach, monitor participant retention and to investigate whether self-monitoring technologies providing feedback about physical activity and interstitial glucose levels play a role in the prevention pathway (which may be amenable to the NDPP framework).”

Comment 4: More detail on the effectiveness of (behavioural and outcome-based) self-monitoring interventions on behaviour should be provided. Is there any evidence that suggests that sustained monitoring is required for behaviour change maintenance (bottom p3)? How does this fit with behaviour change theory?

Response: We recognise that further details should be included in the introduction. However, we would like to emphasise that this is not a long-term intervention due to an unpowered projected sample size and a feasibility focus. Therefore, we will not be assessing the effectiveness of the technologies on behaviour. We have subsequently amended the introduction, as follows:
Introduction: “Given recent consumer interest (Ferguson et al., 2015), wearable technologies permit people to self-monitor behaviour and health. Gardner and colleagues (2016) reviewed behavioural interventions and identified self-monitoring of behaviour as a particularly promising behaviour change technique. Similarly, continuous glucose monitoring technology has shown promise for longer term physiological outcomes (including glycated haemoglobin (HbA1c)) (Vigersky et al., 2012); supporting the suggestion that more frequent engagement leads to better health outcomes (Fonda et al., 2013). Self-monitoring of both behaviour and outcomes are listed within the taxonomy alongside 91 other ingredients (i.e. feedback and goal-setting) in behavioural interventions (Michie et al., 2013).”

Comment 5: The amount, detail and potential identifiable nature of data collected/monitored and stored, also outside of the UK, could be a concern. I encourage the authors to provide more detail on this, in particular the relevant regulatory precautions and data management/storage procedures.
Response: We acknowledge concerns directed toward the sensitive nature of the data that will be collected in this study. We will have safeguards to minimise the collection of identifiable information. Specifically, no accounts will use participant’s personal email address or name. Instead, we will assign ‘dummy’ email accounts and use the study ID to set up the technologies. All data sources monitored via the various platforms will be de-identified using study ID. GPS (or location services), despite being de-identified by Ethica, offers data that can theoretically be ‘reverse-engineered’ to re-identify an individual. All participants will be explicitly informed about all information monitored as part of the study. For individuals who do not wish to have their location services monitored, we will set up

a ‘reduced access’ version of Ethica (app use, screen state and survey responses only). All data collected as part of the study will be stored on password protected computers on password protected documents. A section called the “Data management and storage procedures” has been added, as follows:
“All data collected will be anonymised by assigning a participant ID. Accounts with the three applications (Fitbit, LibreLink and Ethica Health) will be setup using study-specific (‘dummy’) email addresses and passwords (accessible only to the research team) to minimise use of personalised information. All data will be stored securely on the Loughborough University server, as password protected, encrypted documents and original paperwork kept in locked storage. No directly personally identifiable information will be collected through these platforms. GPS (global positioning system) will be collected via Ethica Data which could theoretically be ‘reverse-engineered’ to re-identify individuals; however, all participants will be explicitly informed about all information monitored as part of the study. For individuals who do not wish to have their location services monitored, we will set up a ‘reduced access’ version of Ethica Data (application usage, screen state and survey responses only).”

Comment 6: The suggested use of default settings for step targets may be too ambitious for the target population, who are unlikely to be doing sufficient activity at the start of the study. Have the authors considered introducing graded targets and other motivational techniques to encourage behaviour change, continued engagement with the feedback and maintenance of behaviour change?

Response: We acknowledge that the default settings may not be appropriate for all participants. However, by personalising participant goals for them, we will lose valuable insight into whether or not people decide to create their own goals. Participants will be informed that they can change the goals via their smartphone. As our primary outcome is engagement, we feel this will be one of the key pieces of information used to evaluate this objective. In addition, we have implemented efforts to encourage continued engagement for participants according to the feedback they receive. For instance, if participants receive physical activity feedback, they will be asked to sync the Fitbit at least once every five days (to obtain minute-level data) and to charge the device overnight. Participants receiving interstitial glucose feedback will be asked to scan the Freestyle Libre sensor at least once every seven hours. It will be made clear that participants can sync the Fitbit and scan the Freestyle Libre as much as they want to. The manuscript has been amended as follows:
Behaviour Change Techniques: “Participants will be fully informed that they can freely change the goals set for physical activity as preferred (i.e. should the default value be too easy/difficult) via the Fitbit application.”
Behaviour Change Techniques: “Participants will be asked to sync the Fitbit (at least once every five days) and scan the Freestyle Libre (at least once every seven hours) if they are in the respective group to receive feedback from these devices. This action has a dual purpose; to minimise data loss and to encourage continued engagement with the technologies.”

Comment 7: Introduction – please clarify what is meant with ‘efforts are limited’ – are there limited efforts or are the effects limited, or something else?
Response: We thank the reviewer for highlighting this ambiguity and the sentence has been rewritten to clarify that behaviour change efforts can be challenging as follows:
Abstract: “Changing lifestyle behaviours is difficult and is often predicated on the assumption that…”

Comment 8: P3, 3rd para – the link between control theory, BCT and the current intervention is unclear. Please revise this paragraph to create clearer flow.

Response: We thank the reviewer for the prompt to revise this paragraph. We have changed the introduction as follows:
"As well as delivering key behaviour change techniques, self-monitoring technologies also support Control Theory (Carver et al., 1982). More specifically, people are presented with information about a present state via feedback (e.g. 9,000 steps) and are often provided a set goal to achieve (i.e. 10,000 steps). Equipped with this information, people may make efforts to achieve the goal or desired outcome (i.e. ≥10,000 steps) because they have been informed how they are performing relative to it. The majority of research to date has focused on the deployment of technologies to self-monitor movement behaviours (e.g. (Cadmus-Bertram et al., 2015)) or specific health markers (e.g. (Polonsky et al., 2013)) in isolation. Although these approaches have shown to be beneficial to behaviour change in the short term, most user engagement is not sustained beyond six months (Ledger et al., 2014). Despite research conducted on short-term improvements, it is not yet clear whether results are sustained with prolonged use (Barwais et al., 2013; Tudor-Locke et al., 2009). However, the rationale is that when provided with information about their current levels of activity, people may feel motivated to improve their behaviour."

Comment 9: The focus of the introduction is on prevalence of known diabetes, but a large proportion of diabetes is undiagnosed. Please comment on this.

Response: In the opening paragraph we have added content that clarifies the vast prevalence of undiagnosed diabetes, as follows:
Introduction: "Another imposing challenge is the proportion of the population living with undiagnosed diabetes (current prevalence estimated at 45.8%) (Beagley et al., 2014); which is possibly, in part, attributable to its asymptomatic state prior to the presentation of complications."

Comment 10: Clearer definitions of IGR and IFG should be provided.

Response: Definitions offered by Gutham and colleagues have been provided to offer clarify within the introduction opening paragraph, as follows:
Introduction: "Prediabetes, categorised as either impaired fasting glucose or impaired glucose tolerance represents abnormal glucose homeostasis and is placed between diabetes and normal regulation. Impaired fasting glucose has been defined as elevated fasting plasma glucose (100-126 mg/dl) whilst impaired glucose tolerance is characterised by an elevated two hour plasma glucose concentration (140-199 mg/dl) following intake of a 75g glucose load (Genuth et al., 2003)."

Comment 11: Please remove either investigate or assess from primary aim.

Response: We thank the reviewer for spotting this and we have made the necessary change.

Comment 12: The ISRCTN registration does not specify HbA1c as an inclusion criterion, please clarify.

Response: Thank you for highlighting this. We have moved the HbA1c criteria to the exclusion criteria to align with our ISRCTN registration, as follows:
Exclusion criteria: "Individuals with a clinical diagnosis of type 1 or type 2 diabetes, a HbA1c of ≥6.5%, or have suspected/confirmed pregnancy will be excluded. Participants unable/unwilling to provide informed consent, cannot/unwilling to adhere to the study protocol or cannot read/write English will also be excluded."

Comment 13: Are you able to provide examples of compatible Android phones – are these the major brands? What is the spread of these phones?

Response: We agree with the reviewer that this would help the reader. We have added details of Android requirements to the inclusion criteria, as follows:
Inclusion criteria: “Participants will be aged at least 40 years old, have a moderate-to-high risk of developing type 2 diabetes and use a compatible Android smartphone.”
Inclusion criteria: “Compatible smartphones at the time of the study will be defined as having the following characteristics: An Android operating system of 4.0 or higher, Near Field Communication (NFC), a screen resolution of 480x800 to 1080x1920 and a screen size of 8.9-14.5cm. Exceptions at the time of the study are the Samsung Galaxy 7, Samsung S8, Nexus 5X and Nexus 6P which cannot install the LibreLink application.”

Comment 14: How/when will HbA1c be assessed for the inclusion criteria?

Response: HbA1c will be assessed at appointment 1 (baseline appointment). We have restructured the “Health, physical functioning and fitness” section, as follows:
“HbA1c will be assessed at the first appointment using a point-of-care system, (Afinion AS100 Analyser, Alere Inc., Waltham, MA). Results will be processed immediately following collection. Participants receiving a result ≥6.5% will be ineligible, readings of 5.7-6.4% classified as pre-diabetic (American Diabetes Association, 2014) and readings of <5.7% classified as euglycemia.”

Comment 15: P7 (final appointment) – please specify what DKT refers to (I was not able to find it).

Response: Thank you for spotting this. We have expanded DKT to Diabetes Knowledge Test throughout the manuscript.

Comment 16: Device masking – to what extent will participants be able to change the settings or remove the tape during the data collection phase? How will this be prevented/monitored? Will devices also be masked at follow up?

Response: We hope that we have addressed the participants’ ability to change settings in our response to Reviewer 1 Comment 6. We have added more detail about the masking procedures, as follows:
Device masking: “All email accounts and password combinations will be manually generated and managed by the research team to prevent use of identifiable information. During baseline wear, the activity tracker will be physically masked using black tape applied to the screen; leaving only time and date viewable. Participants will be asked not to tamper with the screen; however, if they do manipulate the masking, it should be readily apparent to the research team. Settings on the application will also be adjusted to remove physical activity metrics from the device screen and notifications fully restricted on their phone and activity tracker. However, participants will not be locked out of the application due to the requirement to sync the device. Time spent on the Fitbit application will be inspected using Ethica Data (Kitchener, Ontario, Canada) to identify potential unauthorised use.”

Comment 17: P9 (BCT) – what is the purpose of the 250/hr step goal – this only adds up to less than 50% of the total goal if one would achieve this every waking hour of the day?

Response: We have amended the “Behaviour Change Techniques” paragraph to offer further clarification about the 250/hr step goal and how this contributes to the study. This characteristic of the Fitbit aligns with many physical activity guidelines which recommend minimising sedentary time. If the wearer has not taken at least 250 steps before ten-to the hour, they will be prompted to move.

Therefore, this prompt should be considered independent to the goal of achieving 10,000 steps each day. We have amended the text, as follows:
“Participants will also receive haptic feedback (BCT 7.1: Prompts/cues; i.e. a gentle vibration) as a reminder to move by the Fitbit 10 minutes prior to the end of each hour (default 09:00-18:00) if 250 steps have not been taken. The reminder to move prompt aims to encourage interruptions in prolonged sedentary bouts as is recommended by the UK Physical Activity Guidelines (UK Department of Health, 2011).”

Comment 18: P9 (measures) – please provides a little bit more information on the proposed data processing decisions for the Actigraph (non-wear, cut points etc).

Response: Thank you for spotting this omission from the ActiGraph paragraph. We have added the following text to this section, as follows:
ActiGraph: “Non-wear will be defined as 60 minutes of consecutive zeros (allowing for up to two minutes of interruptions) with a minimum wear of 600 waking minutes used to define a valid day (Troiano et al., 2008). A minimum of 4 valid days will be used to define a valid file with sedentary time classified as <100cpm, light activity as 100-2019cpm and MVPA as >= 2020cpm (Troiano et al., 2008).”

Comment 19: P11 (measures) – please provide a bit more detail on the protocol for the mCAFT, is this a treadmill- or bike-based test, maximal or submax, etc?

Response: The paragraph has been amended to include the following text:
Health, physical functioning and fitness: “The mCAFT is a sub-maximal step-test protocol with participants instructed to complete ≥1 three-minute stages of stepping at a speed dictated by an audio track. Heart rate will be monitored throughout with the stepping stages continued until heart rate ≥85% of age-predicted maximal heart rate.”

Comment 20: Quantitative data analyses – it may be useful to monitor through which route participants were recruited? Will this be monitored?

Response: Thank you for highlighting this point. We have added the following sentence to clarify this within the “Analysis of Secondary Outcomes” paragraph:
“In addition, the screening survey will also identify recruitment sources.”

Comment 21: Under ‘why’, this is referred to as a method to ‘nudge’ people, but this is not discussed in the introduction. To what extent can this approach be considered a ‘nudge’ intervention? If it is, please justify in the introduction.

Response: We thank the reviewer for highlighting this. We have removed “nudge” from the “Why” section because the study incorporates a number of behaviour change techniques (including feedback, self-monitoring and goal setting) as well as prompt/cues (nudges).

Reviewer 2

Comment 1: The manuscript aims to evaluate the user engagement with two applications: Fitbit and Librelink. Several points are not clear: Is the aim to estimate only the proportion of time that the applications are used and the number of scans and syncs? If yes: what do you deduce from these usages?

Response: Thank you for requesting clarification on the user engagement. Engagement using the two applications (LibreLink and Fitbit) and devices (Freestyle Libre and Fitbit) is the primary aim of the study and will be quantified using those key metrics (syncs, scans and application usage). Together, these metrics will offer data into how often and when participants used the devices and applications. We will also be able to compare their use with existing mobile applications on participants' phones. This will offer unique insight into 'competition between apps' and may offer insight into the importance people place on mobile health applications. Moreover, we will employ semi-structured qualitative interviews which will look at engagement in the context of participants' daily lives. These interviews will help us to understand the context and reasons behind why people engaged (or did not engage) with the technology, whether they understood what was being shown to them and whether this information empowered them to change their behaviour. Together, the quantitative and qualitative insights will offer a comprehensive assessment of engagement and the use of wearable technologies and mobile applications for better health. We have added the following text:
"Number of times the activity tracker syncs (occurs when the application is opened, assumed to see feedback about physical activity) and scans of the glucose sensor (occurs when the participant scans and to see feedback about interstitial glucose levels) will also be recorded."

Comment 2: Is the aim to compare Actigraph and Fitbit? Are Actigraph and Fitbit reference scientific devices for estimating accurately physical activity levels? if yes, please add references.

Response: Thank you for your comment. We will collect ActiGraph and Fitbit data in parallel at baseline (7 days) which will allow us to compare physical activity data measured by the Fitbit to data collected by the ActiGraph. In this study, the ActiGraph is the reference scientific device for estimating physical activity levels but previous models of the Fitbit have been shown to be valid for counting steps. We will conduct free-living validation assessment as part of this study. Results from these analyses will be reported with the results of the trial. As a result, we have added the following sentences:
"ActiGraph accelerometers were employed because they have been shown to offer high validity and reliability in free-living settings (Aadland et al., 2015)."
"Previous models of the Fitbit have been validated for step count (Lee et al., 2014). Free-living concurrent validation of the Fitbit Charge 2 will be conducted and reported with the results of the trial"

Comment 3: Is the aim to compare the usage of the applications with and without feedback? or to compare the information influence (blood glucose versus physical activity) on behavioral change? Is the aim to determine if a physical activity increase improve blood glucose level? There are too many possible objectives and the assumptions are not sufficiently explicit.

Response: We thank the reviewer for asking for clarification on the aims of the study. We have adjusted the structure of the manuscript which now has major headings (primary and secondary outcomes) and subheadings (engagement, feasibility, physical activity, interstitial glucose and questionnaire items) to help clarify the aims of the study.

Comment 4: What is the Ethica Data? What is Diasend? Could the authors draw a graph to explain the relationships between the applications, Ethica Data and Diasend ? and if the glucose and physical activity data are available to researchers? How can researchers evaluate a change in physical activity?

Response: We thank the reviewer for these questions. We have amended Figure 2 to highlight what the participants and researchers will have access to and the subsequent data to be analysed. In addition, we have amended the sections referring to Ethica Data and Diasend, as follows:
Analysis of primary outcomes: “Ethica Data is a fee-for-service platform that will be used to provide time-stamped data relating to application usage. This is an application installed on the participants phone and sits idle during the study period.”
Analysis of primary outcomes: “Fitabase is a fee-for-service platform that permits access to download 60-s epoch Fitbit data (i.e. levels of physical activity) and remote monitoring of Fitbit devices (e.g. battery level and time since last sync event) via Bluetooth and Wi-Fi.”
User engagement: “Diasend will connect with the Freestyle Libre via the LibreLink application and data will be recorded and accessed through this software.”
Analysis of secondary outcomes: “Diasend is a fee-for-service platform that permits access to download 15 minute epoch Freestyle Libre data and remote monitoring of multiple LibreLink accounts.” Regarding evaluating changes in physical activity, participants will continuously wear the Fitbit for the full study duration of seven weeks (1 week baseline and 6 week intervention regardless of group allocation). Therefore, we will be able to identify changes in activity levels per day for the full study duration. This has been clarified in the manuscript, as follows:
Fitbit: “To examine changes in physical activity over the study duration, participants will be requested to wear the device for the full seven weeks and…”

Comment 5: It is difficult to understand what behaviour change techniques are proposed to volunteers and how they are evaluated (efficacy on what? is target complete or incomplete?): BCT:1.1 (10000 steps and 10 floors climbed), BCT1.3 (4.0-5.9 nmol/L), BCT 2.3 recording of physical activity without fixed goal?, BCT 2.4 recording of glucose without fixed goal?, BCT : recording of glucose and physical activity without fixed goal and BCT 7.1 alert to walk 250 steps for an hour ?

Response: We thank the reviewer for highlighting this ambiguity. The techniques will be evaluated based on whether the target is complete or incomplete. For example, BCT 1.3 of 4.0-5.9 mmol/L will be assessed by time spent in this target range as well as time spent above and below this range. The section has been amended with the following text:
“Attainment of a goal will be assessed as either complete or incomplete.”

Comment 6: How the ANCOVAs will be used (what independent and dependent variables?)

Response: We have amended the “Quantitative Data Analysis” section, as follows:
“Descriptive statistics of the sample will be conducted. In addition, repeated ANCOVAs will be conducted to assess change in engagement (dependent) according to group (independent) having adjusted for participant characteristics. Similarly, ANCOVAs will be conducted to assess changes in physical activity (dependent) according to group (independent) having adjusted for baseline physical activity and Fitbit wear time. All data will be analysed using Statistical Package for Social Sciences (SPSS Inc. Chicago, IL).”

Comment 7: These points (aim, assumption, measured variables, analyses, software activity, data available to volunteers and researchers) have to be clarified in the revised manuscript: the readers need to understand easily the aim of the study, the used methods and the expected outputs.

Response: We thank the reviewer for highlighting that clarifications of the study aim, methods and expected outputs are required. We believe we have responded to this comment by addressing Reviewer 2 Comments 1 to 6).

Comment 8: Please correct the reference n°10 (authors).
Response: Thank you for spotting this mistake in the bibliography, we have amended the reference accordingly.

VERSION 2 – REVIEW

REVIEWER	Esther van Sluijs University of Cambridge,UK
REVIEW RETURNED	17-Aug-2017

GENERAL COMMENTS	The authors have addressed the reviewers' comments appropriately in a substantially revised manuscript. I wish them best of the luck with the study.

REVIEWER	Rousset INRA France
REVIEW RETURNED	29-Aug-2017

GENERAL COMMENTS	I appreciate the major modifications made to the text. The manuscript is now clearer. There is just one slight point of confusion: the use of Actigraph for the first week. The comparison between Actigraph wGT3x and Fitbit Charge 2 is not stated as the primary nor the secondary aim of the study. Moreover the article of Aadland et al. 2015 dealt with the reliabilty of two GT3x+accelerometers and not with the validity of this device (against methods of reference such as indirect calorimetry). The second article that you mentioned studied the validity of several devices in controlled conditions for a short period (69 minutes)(Lee et al, 2014) . In this article the authors found that Actigraph GT3x was less accurate that Fitbit one or Fitbit Zip. In these conditions, it seems problematic to consider Actigraph as a reference.

VERSION 2 – AUTHOR RESPONSE

Reviewer 2

Comment 1: I appreciate the major modifications made to the text. The manuscript is now clearer. There is just one slight point of confusion: the use of Actigraph for the first week. The comparison between Actigraph wGT3x and Fitbit Charge 2 is not stated as the primary nor the secondary aim of the study. Moreover the article of Aadland et al. 2015 dealt with the reliabilty of two GT3x+accelerometers and not with the validity of this device (against methods of reference such as indirect calorimetry).

Response: Thank you for requesting clarification on the use of ActiGraph for the first week. We have amended the manuscript to reflect that we will not conduct any validation assessment of the Fitbit Charge 2 in this study. The reason for deploying an ActiGraph is rather to offer an evaluation tool (independent of the intervention device, the Fitbit Charge 2) to capture habitual physical activity during the first week. Subsequently, we have reworded the ActiGraph paragraph to confirm these accelerometers are well-validated and deployed in several large-scale studies. We have changed the manuscript as follows:

First appointment and baseline: *deleted* "…measuring physical activity whilst also presenting an opportunity to validate the activity tracker."

ActiGraph: "In an effort to determine the physical activity levels of the participants relative to general population, participants will be asked to wear an ActiGraph wGT3X-BT (ActiGraph, Pensacola, FL, USA) accelerometer for seven days during waking hours and to remove for any water-based activities (e.g. showering and swimming). The waist-worn (i.e. over the right hip, mid-clavicular line) ActiGraph will quantify time spent sedentary, in light and moderate-to-vigorous physical activity (MVPA) as well as daily step counts and will function as a data logger (i.e. no feedback provided). ActiGraph accelerometers have been validated (Plasqui et al., 2007; Melanson et al., 1995) and extensively deployed (Hagstromer et al., 2007; Troiano et al., 2008; Chaudry and Esliger 2008) to measure physical activity under free-living conditions. Data from the ActiGraph…"

ActiGraph: *deleted* "…because they have been shown to offer high validity and reliability in free-living settings (Aadland et al.)."

Reviewer 2 comment 2: The second article that you mentioned studied the validity of several devices in controlled conditions for a short period (69 minutes) (Lee et al, 2014). In this article the authors found that Actigraph GT3x was less accurate that Fitbit one or Fitbit Zip. In these conditions, it seems problematic to consider Actigraph as a reference.

Response to comment 2: We agree with the reviewer. As we have clarified that we do not intend to do a validation of the Fitbit Charge 2, we believe this reference can remain. However, we have removed a sentence to help clarify that no validation will be conducted, as follows:

Fitbit: *deleted* "Free-living concurrent validation of the Fitbit Charge 2 will be conducted and reported with the results of the trial."

VERSION 3 – REVIEW

REVIEWER	Rousset INRA France
REVIEW RETURNED	01-Sep-2017

GENERAL COMMENTS	The last modifications are satisfactory.