PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Flash Glucose Monitoring in gestational diabetes mellitus: study protocol for a randomised controlled trial
AUTHORS	Majewska, Agata; Stanirowski, Paweł; Wielgoś, Mirosław; Bomba-Opon, Dorota

VERSION 1 – REVIEW

REVIEWER	Lynnsay M Dickson University of the Witwatersrand, South Africa
REVIEW RETURNED	16-Aug-2020

GENERAL COMMENTS	1. Is the research question or study objective clearly defined? # The study title 'Flash glucose monitoring in gestational diabetes mellitus: study protocol for a randomized controlled trial' may be misleading as the flash glucose monitoring (FGM) is only used by half of the participants for only one-third of the trial. It is not a core component of this study. No rationale is provided for this brief exposure to FGM. # The study purpose is to evaluate the impact of FGM on the efficacy of treatment of gestational diabetes mellitus (GDM). The IMPACT study and others have demonstrated that FGM may reduce the time spent in hypoglycaemia, which is a frequent complication of insulin therapy. A motivation or justification for the use of expensive FGM in the management of GDM, a common intermediate hyperglycaemia with low hypoglycaemic risk, is not established in the introduction. 3. Is the study design appropriate to answer the research question? # A comparison of outcomes between groups may be difficult to interpret because, for two-thirds of the trial duration, the two groups are identical. All study participants (both groups) undergo a complex health intervention that includes self-monitoring of blood glucose, dietary intervention, physical exercise intervention, and if required, insulin therapy. The use of FGM (Freestyle Libre) or glucometers (iXell) for self-monitoring, which is the only difference between groups, is only used for the first four weeks of an estimated twelve-week trial. Thereafter, all participants use a glucometer. This is not a true parallel-group study. # The primary outcome is the ‘mean glycaemia results’ (fasting and one hour postprandial) in the FGM and glucometer groups during the first month following the diagnosis of GDM. This is inappropriate for the following reasons: - FGM measures interstitial fluid glucose concentration and the glucometer measures whole blood glucose concentration and these are not directly interchangeable.[1] - Clinically significant differences between glucometer and laboratory measurement of blood glucose are known to occur, even for those

	glucometers complaint with ISO 15197:2013 recommended accuracy.[2] The ISO compliance of the iXell glucometer is not disclosed. - The Freestyle Libre and the iXell glucometer are effectively index tests in this study. There is no evaluation of their accuracy against a known reference test. No calibration exercise is described. This would allow meaningful clinical interpretation of mean blood glucose concentrations by these different methods of measurement. - According to the 'Standards of Polish Society of Gynaecologists and Obstetricians in management of women with diabetes' (Reference 10), routine HbA1c monitoring in GDM is not recommended. This test may not be used as a reference test for glucose concentration. 4. Are the methods described sufficiently to allow the study to be repeated? # Whether a universal or a selective GDM screening strategy will be applied should be disclosed. # Finger-prick testing 9–10 times a day (fasting + one hour postprandial after three meals + after three small meals + weekly midnight), which is above the six tests daily recommended by the 'Standards of Polish Society of Gynaecologists and Obstetricians in management of women with diabetes', should be referenced or justified. # The correct reference of the '10 000 steps' as the recommended physical activity in pregnancy should be provided. Are physical activity recommendations individualized based on participant's level of fitness? # Details of the Eating Assessment Test should be provided. Participants are separated into four groups based on this assessment. These subgroups should be described to enable the interpretation of results and study duplication. # Please confirm if participants in the FGM group use their cell phones. Are participants compensated? # The first and second authors are responsible for recruitment of GDM positive patients. The person responsible for making the diagnosis of GDM should be identified. # Appropriate randomisation procedures are used. # Blinding is not possible because blood glucose monitoring devices used, are visibly different to participants and researchers. # The first and second authors are responsible for data collection. Persons who will be analyzing participants glycaemia levels, diet control, and physical activity in this study are not identified. Persons modifying the various components of this complex health intervention per individual, based on this analysis, are not identified. # Predetermined protocols for insulin dose adjustments should be included in the study design, to reduce the risk of researcher bias when treatment decisions are made in the FGM and glucometer groups. # Exclusion criteria - The exclusion of overt diabetes in pregnancy, which is a separate entity to GDM according to the World Health Organisation, is not explicitly stated. - Complications of the birth such as prematurity, may be used to exclude data at the end of the study rather than at recruitment. - Will each participant undergo routine screening for chronic renal or hepatic disease at recruitment, or will only those known to have disease be excluded? 6. Are the outcomes clearly defined?

	# The primary outcome is restricted to ‘mean glycaemia results’ (fasting and one-hour postprandial) between the two groups. ‘Time spent in hypoglycaemia’ is used to estimate sample size. The ‘number of hypoglycaemic episodes’, listed as one of the secondary outcomes, is not quite the same entity. The study objective, and primary and secondary outcomes should be aligned. 7. If statistics are used are they appropriate and described fully? # Sample size. The clinical assumptions used in the sample size estimation should be revised. The authors Agata and colleagues use a 38% reduction in the time spent in hypoglycaemia to estimate the sample size of 76 patients. This is not reasonable because there is no risk of hypoglycaemia in the first two weeks and there is a low risk of hypoglycaemia in the second two weeks post recruitment. The Freestyle Libre FGM is inserted at recruitment and day 14 post recruitment in one group. Only diet and physical activity are used in the first two weeks of the study and this has no risk of hypoglycaemia. An assessment of the need for insulin therapy is first made on day 14 post recruitment. According to the ‘Standards of Polish Society of Gynaecologists and Obstetricians in management of women with diabetes’, 10–40% of GDM affected women may need insulin therapy during their pregnancy. A significantly larger study sample is required to demonstrate a reduction in hypoglycaemia, within this very small window of two weeks of insulin therapy while FGM is in use. The historical local experience of insulin requirements in GDM at the Medical University of Warsaw (10 % ?, 40 % ?) may help refine this sample size calculation. # Data management. - Participant retention strategies are not disclosed. - Strategies for the management of missing data, particularly in the glucometer group, are not discussed. - Excessive data collection with the use of FGM is known to occur. In the IMPACT study, the mean (SD) daily sensor scanning frequency was 14.7 (10.7).[3] Strategies to manage excess data, particularly in the FGM group, should be discussed. - A reduction in the time spent in hypoglycaemia is mentioned in the sample size calculation, however, the method of evaluating this between the two groups is not described. FGM measures interstitial fluid glucose concentration every 15 minutes and the duration of hypoglycaemia is given as a percentage of the day on the phone app. Glucometer blood glucose measurements at intermittent, predetermined time intervals may, theoretically, underestimate the frequency of hypoglycaemic events. # Statistical methods. - Mean glycaemia (fasting and 1-hour postprandial) in each group is the primary outcome. However, the FGM and the glucometer results are different for biochemical and technical reasons, and this difference may not be clinically true. - Statistical methods useful for comparisons between methods of measurement include the Bland Altman method, Passing-Bablok Regression, and Deming regression analysis. - Reduction in hypoglycaemia is mentioned in the sample size calculation but this is not included as a primary outcome, nor is it mentioned in the statistical methods section. - The researchers should specify which software programs, if any, will be used in the data analysis. 8. Are the references up-to-date and appropriate?

	# The authors cite the IMPACT study [3] (Page 4 of 19, line 54; Page 4 of 19, line 57), which is not included in the list of references. A similar study by Bolinder, et al.,[4] is listed instead as Reference 6. # Physical activity recommended to study participants is 10 000 steps per day and this will be measured by a Xiaomi Mi Band. The source of this aggressive recommendation for exercise in pregnancy is not clear. This is not in the cited source: Evenson, et al. (Reference 9). It is also not in the 'Standards of Polish Society of Gynaecologists and Obstetricians in management of women with diabetes' (Reference10) 12. Are the study limitations discussed adequately? # The limitations of using Freestyle Libre for only a third of the duration of the study was not discussed. 14. To the best of your knowledge is the paper free from concerns over publication ethics (e.g. plagiarism, redundant publication, undeclared conflicts of interest)? # The involvement of the Abbott Diabetes Care (Freestyle Libre), Genexo sp (iXell meter) and the Xiaomi Corporation (Xiamomi Mi Band) or their representatives in the study design, as well as their access to the study data, is not disclosed. References 1. Cobelli C, Schiavon M, Dalla Man C, Basu A, Basu R. Interstitial Fluid Glucose Is Not Just a Shifted-in-Time but a Distorted Mirror of Blood Glucose: Insight from an In Silico Study. Diabetes Technology & Therapeutics. 2016; 18(8) DOI: 10.1089/dia.2016.0112 2. Freckmann G, Schmid C, Annette Baumstark A, Pleus S, Link M, Haug C. System Accuracy Evaluation of 43 Blood Glucose Monitoring Systems for Self-Monitoring of Blood Glucose according to DIN EN ISO 15197. J. Diabetes Sci. Technol. 6, 1060–1075 (2012). 3. Oskarsson P, Antuna R, Geelhoed-Duijvestijn P, Kröger J, Weitgasser R, Bolinder J. Impact of flash glucose monitoring on hypoglycaemia in adults with type 1 diabetes managed with multiple daily injection therapy: a pre-specified subgroup analysis of the IMPACT randomized controlled trial. Diabetologia. 2018; 61:539–550. https://doi.org/10.1007/s00125-017-4527-5 4. Bolinder J, Antuna R, Geelhoed-Duijvestijn P, Kröger J, Weitgasser R. Novel glucose-sensing technology and hypoglycaemia in type 1 diabetes: a multicentre, non-masked, randomised controlled trial. Lancet. 2016; 388 (10057):2254–63 http://dx.doi.org/10.1016/S0140-6736(16)31535-5

REVIEWER	Aruna Nigam Hamdard Institute of medical sciences and Research, Jamia Hamdard, India
REVIEW RETURNED	12-Oct-2020

GENERAL COMMENTS	very good protocol. there are only few suggestions 1. author is required to justify the sample size 2. appropriate statistical analysis should be mentioned. author can take help of article Nigam a Et al. Comparative analysis of 2-week glycaemic profile of healthy versus mild gestational diabetic pregnant women using flash glucose monitoring system: an observational study 2019;126 3. mention , how frequently author is going to get the SMBG done during the study period and also how frequently the author is going

	to send the data of flash monitoring through mobile.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
1. Is the research question or study objective clearly defined?

Assuming, that the first month after diagnosis of GDM is essential for the proper implementation of dietary and physical activity recommendations by the patients, the favorable cost-effective strategy will be to apply FGM device only during that period of pregnancy. By analyzing results of this study, such as fasting and postprandial glycemia levels, number of nocturnal hypoglycemic episodes, number of women requiring insulin therapy, daily dosage of insulin and maternal-fetal perinatal outcomes, we will provide a scientific basis for more common use of FGM in the population of pregnant women affected by GDM.
The IMPACT study revealed that FGM reduces the time spent in hypoglycemia. We citied this study because although hyperglycemia is the most common alteration occurring in GDM patients, there is also an increased risk of masked hypoglycemia. It was showed that almost one third of GDM patients experienced hypoglycemic events during the course pregnancy that could have been easily detected using methods for continuous glycemia control (1). Diagnosis of these hypoglycemic episodes may be of particular importance in GDM patients before qualifying to insulin therapy.
The above sentences were added into the Introduction section.

2. Is the study design appropriate to answer the research question?

a. A comparison of outcomes between groups may be difficult to interpret because, for two-thirds of the trial duration, the two groups are identical. (…). This is not a true parallel-group study.
FLAMINGO is a single-center, non-blinded, randomized, crossover study with a nested qualitative evaluation and 1:1 allocation ratio.
The above sentences were added into the Study design section.

b. FGM measures interstitial fluid glucose concentration and the glucometer measures whole blood glucose concentration and these are not directly interchangeable.
FGM is a factory-calibrated sensor measuring glucose concentrations in the interstitial fluid. Although it represents different measurement technique than SMBG, previous studies have shown, that glycemia levels obtained by both methods are comparable. (2) (3) Apart from continuous glucose concentration measurements the system provides additional data by creating a 24-hour glycemic profile. In comparison, SMBG provides only single, intermittent measurements, that limits detection of glycemic variability or nocturnal hypoglycaemic events (2). Based on the results of studies comparing FGM and SMBG we assumed that the methods are comparable and could be interchangeable.
The above sentences were added into the Introduction section.

c. Clinically significant differences between glucometer and laboratory measurement of blood glucose are known to occur, even for those glucometers complaint with ISO 15197:2013 recommended accuracy. The ISO compliance of the iXell glucometer is not disclosed.
We added additional piece of information about ISO for iXell®; Genexo sp; Warsaw, Poland - ISO 15197:2015 in the Participant selection and recruitment section. As above mentioned it is different measurement method but results from SMBG and FGM, as it was shown in studies are comparable.
The above information was added into the Participant selection and recruitment section.

d. The Freestyle Libre and the iXell glucometer are effectively index tests in this study. There is no evaluation of their accuracy against a known reference test. No calibration exercise is described. This would allow meaningful clinical interpretation of mean blood glucose concentrations by these different methods of measurement.
FGM is a factory-calibrated sensor measuring glucose concentrations in the interstitial fluid. Although it represents different measurement technique than SMBG, previous studies have shown, that glycemia levels obtained by both methods are comparable (2) (3). Since the methods are factory-calibrated, there is no need for calibration in our study.
The above sentences were added into the Introduction section.

e. According to the 'Standards of Polish Society of Gynaecologists and Obstetricians in management of women with diabetes' (Reference 10), routine HbA1c monitoring in GDM is not recommended. This test may not be used as a reference test for glucose concentration.
The secondary objectives will be to compare the two groups for the number of patients requiring insulin therapy, dosage of insulin, number of hypoglycemic episodes, as well as to compare blood glycated hemoglobin (HbA1c) and fructosamine serum concentrations as potential markers of long-term glycemic control and predictors of perinatal complications, based on previous studies (4), (5).
We agree with the reviewer that this is not a reference test. We would like to check the markers in the study as potential predictors in the future if the results will show their role of predicting perinatal complications.
The above sentences were added into the Aim of the study and objectives section.

4. Are the methods described sufficiently to allow the study to be repeated?

a. Whether a universal or a selective GDM screening strategy will be applied should be disclosed.
We aim to recruit 100 pregnant women diagnosed with GDM based on the results of 75g oral glucose tolerance test (OGTT), performed between 24-28 gestational weeks, in accordance with the universal criteria defined by World Health Organization (6).
The above sentence was added into the Study population and eligibility criteria section.

b. Finger-prick testing 9–10 times a day (fasting + one hour postprandial after three meals + after three small meals + weekly midnight), which is above the six tests daily recommended by the 'Standards of Polish Society of Gynaecologists and Obstetricians in management of women with diabetes', should be referenced or justified.
According to PSOG recommendations all participants will be obliged to measure fasting and 1-h postprandial glucose concentrations in a daily manner, together with once per week midnight measurement (15). Postprandial glucose measurements will be performed after three main meals (breakfast, dinner, supper). To avoid excessive data collection, we do not aim to analyze all the results obtained with the FGM. Summarizing we will recommend 4 tests daily plus once per week one additional midnight measurement.
The above sentences were added into the Methods and analysis: Visit 1 section.

c. The correct reference of the '10 000 steps' as the recommended physical activity in pregnancy should be provided. Are physical activity recommendations individualized based on participant's level of fitness?
According to the criteria proposed by Tudor-Locke we will divide patients into four groups of physical activity based on a number of daily steps collected by the wristband: sedentary (< 5000 daily steps), low active (5000~7500 daily steps), somewhat active (7500~10000 daily steps) and active (≥ 10000 daily steps) (7). Physical activity recommendations will not be individualized, unless the patient will have a contraindication for above mentioned level of physical activity.
The above sentence was added into the Methods and analysis: Visit 2 section.

d. Details of the Eating Assessment Test should be provided. Participants are separated into four groups based on this assessment. These subgroups should be described to enable the interpretation of results and study duplication.
To evaluate participants' dietary habits, we will use Eating Assessment Test prepared by the Polish National Institute of Public Health – National Institute of Hygiene. This is a short questionnaire (20 items for diet) intended to evaluate dietary habits in patients diagnosed with GDM. The summary points base on a number of meals per day, length of breaks between meals, daily portions of fruits and vegetables, frequency of fried food and sweets consumption per week (questionnaire and list of points for each element provided in Supplementary files). The maximum test result is 42 points, the minimum 4 points. Based on the points obtained patients will be assigned to one of the four diet groups (good: 39-42, satisfactory: 30-38, demanding diet modification 12-29, and not satisfactory < 12 points).
The above sentences were added into the Methods and analysis: Visit 1 section.

e. Please confirm if participants in the FGM group use their cell phones. Are participants compensated?
Participants in the FGM group will use their cell phones. Patients without mobile phone will obtain Freestyle Libre Reader and instructions for using the device. Participants will not be compensated.
The above sentence was added into the Methods and analysis: Visit 1 section.

f. The first and second authors are responsible for recruitment of GDM positive patients. The person responsible for making the diagnosis of GDM should be identified.
Pregnant women diagnosed with GDM by the trial research staff (AM, PS) who will meet study inclusion criteria will be invited to participate in the project (see Figure 1).
The above sentence was added into the Methods and analysis: Participant selection and recruitment section.

g. Appropriate randomisation procedures are used. Blinding is not possible because blood glucose monitoring devices used, are visibly different to participants and researchers.

h. The first and second authors are responsible for data collection. Persons who will be analyzing participant's glycaemia levels, diet control, and physical activity in this study are not identified. Persons modifying the various components of this complex health intervention per individual, based on this analysis, are not identified.
The trial research staff (AM, PS) will be responsible for analyzing participants glycemia results, diet control and physical activity as well as for the modifications of health interventions during the follow-up visits.
The above sentence was added into the Methods and analysis: Study procedures section.

i. Predetermined protocols for insulin dose adjustments should be included in the study design, to reduce the risk of researcher bias when treatment decisions are made in the FGM and glucometer groups.
Initial predetermined insulin dose planned for the study is four units for long-acting insulin and three units per meal for short-acting insulin. The final insulin dosage will be individualized based on the glycemic results during the follow-up visits (8).
The above sentence was added into the Methods and analysis: Visit 2 section.

j. Exclusion criteria
- The exclusion of overt diabetes in pregnancy, which is a separate entity to GDM according to the World Health Organisation, is not explicitly stated.
Multiple pregnancy, fetal malformations, pre-gestational diabetes mellitus (overt diabetes in

pregnancy) (…).
The above information was added into the Methods and analysis: Exclusion criteria section.

- Complications of the birth such as prematurity, may be used to exclude data at the end of the study rather than at recruitment.
Above mentioned complications were removed from the protocol.

- Will each participant undergo routine screening for chronic renal or hepatic disease at recruitment, or will only those known to have disease be excluded?
(…) chronic renal or hepatic disease diagnosed prior to study entry (…).
Only the patients known to have disease will be excluded at recruitment.
The above information was added into the Methods and analysis: Exclusion criteria section.

6. Are the outcomes clearly defined?

a. The primary outcome is restricted to 'mean glycaemia results' (fasting and one-hour postprandial) between the two groups. 'Time spent in hypoglycaemia' is used to estimate sample size. The 'number of hypoglycaemic episodes', listed as one of the secondary outcomes, is not quite the same entity. The study objective, and primary and secondary outcomes should be aligned.
We agree with the reviewer. Based on the suggestions we clarified primary outcome "mean glycemia results" as mean results of fasting and 1-h postprandial glucose concentratios between the groups in conjuction with the percentage of results in the target glycemic range.
We decided to unify secondary outcome relative to hypoglycemia. The secondary outcome will be the number of hypoglycemic episodes. We changed also outcome used for estimation of sample size. Newly calculated sample size is based on "percentage of results in the target glycemic range", that is the primary outcome in our study.
Our primary outcome is mean glycemia results (fasting and 1-h postprandial glucose concentrations) in each group (FGM/SMBG) during the first month following GDM diagnosis in conjuction with the percentage of results in the target glycemic range. The secondary objectives will be to compare the two groups for the number of patients requiring insulin therapy, dosage of insulin, number of hypoglycemic episodes (…).
The above sentences were added into the Methods and analysis: Aim of the study and objectives section.

7. If statistics are used are they appropriate and described fully?

a. Sample size
The clinical assumptions used in the sample size estimation should be revised. The authors Agata and colleagues use a 38% reduction in the time spent in hypoglycaemia to estimate the sample size of 76 patients. This is not reasonable because there is no risk of hypoglycaemia in the first two weeks and there is a low risk of hypoglycaemia in the second two weeks post recruitment. The Freestyle Libre FGM is inserted at recruitment and day 14 post recruitment in one group. Only diet and physical activity are used in the first two weeks of the study and this has no risk of hypoglycaemia. An assessment of the need for insulin therapy is first made on day 14 post recruitment. According to the 'Standards of Polish Society of Gynaecologists and Obstetricians in management of women with diabetes', 10–40% of GDM affected women may need insulin therapy during their pregnancy. A significantly larger study sample is required to demonstrate a reduction in hypoglycaemia, within this very small window of two weeks of insulin therapy while FGM is in use. The historical local experience of insulin requirements in GDM at the Medical University of Warsaw (10 % ?, 40 % ?) may help refine this sample size calculation.
We agree with the reviewer. After analysis of primary and secondary outcome we decided to calculate

the sample size based on the results of the study that analyzed the difference between FGM and SMBG group in above others the percentage of results in the target glycemic range.
The performed power analysis (power of 80%, significance level of 5%, two-sided) estimated a required sample size of a total of 80 patients (40 patients in each group). The analysis was based on the results of a previous report comparing FGM with SMBG and estimation to detect a difference in the percentage of results in the target glycemic range between study and control groups(9). Sample size is further increased to 100 patients to account for a potential exclusion and drop out of approximately 10%.
The above sentences were added into the Sample size calculation and statistical analysis section.

b. Data management.
- Participant retention strategies are not disclosed.
Patients recruited to the study will obtain an email with the schedule of the follow up visits. They will also obtain regular phone contact before each visit.
The above sentences were added into the Trial monitoring and management section.

- Strategies for the management of missing data, particularly in the glucometer group, are not discussed.
For missing glycemic data, in particular in the SMBG group mean glycemia levels from the past 5 days will be calculated, in accordance with the time of measurement (fasting, postprandial).
he above sentences were added into the Trial monitoring and management section.

- Excessive data collection with the use of FGM is known to occur. In the IMPACT study, the mean (SD) daily sensor scanning frequency was 14.7 (10.7).[3] Strategies to manage excess data, particularly in the FGM group, should be discussed.
We will analyze fasting, postprandial and once per week midnight glycemia results between the groups (4 results per day plus one per week). We are aware that in FGM group we will obtain additional data due to the daily sensor scanning frequency about 14.7. We will analyze this superiority of FGM over SMBG only as the argument for explanation if FGM results will be statistically different better from SMBG results.
To avoid excessive data collection, we do not aim to analyze all the results obtained with the FGM.
The above sentence was added into the Methods and analysis: Visit 1 section.

- A reduction in the time spent in hypoglycaemia is mentioned in the sample size calculation, however, the method of evaluating this between the two groups is not described. FGM measures interstitial fluid glucose concentration every 15 minutes and the duration of hypoglycaemia is given as a percentage of the day on the phone app. Glucometer blood glucose measurements at intermittent, predetermined time intervals may, theoretically, underestimate the frequency of hypoglycaemic events.
We agree with the reviewer. We clarified in the manuscript that as secondary outcome instead of time spent in hypoglycemia will be the number of hypoglycemic episodes. The analyzed results will be: fasting and postprandial glucose measurements plus once per week midnight measurement. This will unify measurements and avoid underestimation in the SMBG group. Furthermore, the sample size in the revised manuscript is calculated based on difference in percentage of results in the target glycemic range between FGM and SMBG group (the primary outcome in the study).
According to PSOG recommendations all participants will be obliged to measure fasting and 1-h postprandial glucose concentrations in a daily manner, together with once per week midnight measurement. Postprandial glucose measurements will be performed after three main meals (breakfast, dinner, supper). To avoid excessive data collection, we do not aim to analyze all the results obtained with the FGM.
The above sentences were added into the Methods and analysis: Visit 1 section.

c. Statistical methods.
- Mean glycaemia (fasting and 1-hour postprandial) in each group is the primary outcome. However, the FGM and the glucometer results are different for biochemical and technical reasons, and this difference may not be clinically true.
Based on the previous studies that analyzed above mentioned measurement methods we assume that FGM and SMBG results are comparable, although the technique of glycemia measurement is different.
FGM is a factory-calibrated sensor measuring glucose concentrations in the interstitial fluid. Although it represents different measurement technique than SMBG, previous studies have shown, that glycemia levels obtained by both methods are comparable. (2) (3)
The above sentence was added into the Introduction section.

- Statistical methods useful for comparisons between methods of measurement include the Bland Altman method, Passing-Bablok Regression, and Deming regression analysis.
We agree with the reviewer. Based on the suggestions we included the Bland Altman method, Passing-Bablok Regression and Deming regression analysis. We agree that above mentioned statistical methods will be useful for comparison FGM with SMBG glucose level results.

- Reduction in hypoglycaemia is mentioned in the sample size calculation but this is not included as a primary outcome, nor is it mentioned in the statistical methods section.
We mentioned as secondary outcome number of hypoglycemic episodes.
• Number of hypoglycemic episodes (glucose concentration <70 mg/dl) during one month analysis (episodes per day in 0-4 weeks after the recruitment visit)
The above sentence was added into the Study outcome section.

As above written we changed the sample size calculation (in revised manuscript it is based on the difference between FGM and SMBG group in above others the percentage of results in the target glycemic range).

- The researchers should specify which software programs, if any, will be used in the data analysis.
Statistical analyses will be performed using SAS software, version 9.2 or later (SAS Institute, Cary, NC).
The above sentence was added into the Sample size calculation and statistical analysis section.

8. Are the references up-to-date and appropriate?

a. The authors cite the IMPACT study [3] (Page 4 of 19, line 54; Page 4 of 19, line 57), which is not included in the list of references. A similar study by Bolinder, et al.,[4] is listed instead as Reference 6.
We agree with the reviewer and changed the listed references. The IMPACT study was registered with ClinicalTrials.gov, number NCT02232698. The results of the study were published in Lancet by Bolinder, et al, listed as Reference 6(3). We changed the matching reference to the citation.

b. Physical activity recommended to study participants is 10 000 steps per day and this will be measured by a Xiaomi Mi Band. The source of this aggressive recommendation for exercise in pregnancy is not clear. This is not in the cited source: Evenson, et al. (Reference 9). It is also not in the ‘Standards of Polish Society of Gynaecologists and Obstetricians in management of women with diabetes’ (Reference10)
We added additional references including recommendations of Polish Society of Gynaecologists and Obstetricians (14). We found in previous studies that walking is completely safe and has protective effect on adverse outcomes for pregnant women.
According to Polish Society of Obstetricians and Gynecologists (PSOG) the recommended number of

footsteps in pregnancy is 10.000 per day. As previously demonstrated mild physical activity, such as walking has protective effect on excessive gestational weight gain and decreases the risk of preterm birth and fetal macrosomia ((10) (11, 12) (13)).
The above sentence was added into the Methods and analysis: Visit 1 section.

12. Are the study limitations discussed adequately?
a. The limitations of using Freestyle Libre for only a third of the duration of the study was not discussed.
We added the limitations of the study to revised protocol.
• To reduce costs FGM device will be applied only for the one third of the trial which might have an impact on the results of the study.
The above limitations were added into the Limitations of the study section.

14. To the best of your knowledge is the paper free from concerns over publication ethics (e.g. plagiarism, redundant publication, undeclared conflicts of interest)?
The involvement of the Abbott Diabetes Care (Freestyle Libre), Genexo sp (iXell meter) and the Xiaomi Corporation (Xiaomi Mi Band) or their representatives in the study design, as well as their access to the study data, is not disclosed.
Patients and the manufacturers (Abbott Diabetes Care, Genexo sp, Xiaomi Corporation) will not be involved in the process of the design, conduct, reporting or dissemination plans of the study.
The above sentence was added into the Patient and public involvement section.

Reviewer: 2
Comments to the Author
very good protocol. there are only few suggestions

1. author is required to justify the sample size:
After analysis of primary and secondary outcome we decided to calculate the sample size based on the results of the study that analyzed the difference between FGM and SMBG group in above others the percentage of results in the target glycemic range.
The performed power analysis (power of 80%, significance level of 5%, two-sided) estimated a required sample size of a total of 80 patients (40 patients in each group). The analysis was based on the results of a previous report comparing FGM with SMBG and estimation to detect a difference in the percentage of results in the target glycemic range between study and control groups(9). Sample size is further increased to 100 patients to account for a potential exclusion and drop out of approximately 10%.
The above sentences were added into the Sample size calculation and statistical analysis section.

2. Appropriate statistical analysis should be mentioned. author can take help of article Nigam a Et al. Comparative analysis of 2-week glycaemic profile of healthy versus mild gestational diabetic pregnant women using flash glucose monitoring system: an observational study 2019;126
We agree with the reviewer. Based on the suggestions we included the Bland Altman method, Passing-Bablok Regression and Deming regression analysis. We agree that above mentioned statistical methods will be useful for comparison FGM with SMBG glucose level results.

3. Mention, how frequently author is going to get the SMBG done during the study period and also how frequently the author is going to send the data of flash monitoring through mobile.
According to PSOG recommendations all participants will be obliged to measure fasting and 1-h postprandial glucose concentrations in a daily manner, together with once per week midnight measurement. Postprandial glucose measurements will be performed after three main meals (breakfast, dinner, supper). To avoid excessive data collection, we do not aim to analyze all the results obtained with the FGM.

The above sentences were added into the Methods and analysis: Visit 1 section.

The author is going to send the data of FGM through mobile to the research staff during the follow-up visits (once per 2 weeks).
The trial research staff (AM, PS) will be responsible for analyzing participants glycemia results (…) during the follow-up visits.
The above sentence was added into the Methods and analysis: Study procedures section.

References:
1. Naik D, Hesarghatta Shyamasunder A, Doddabelavangala Mruthyunjaya M, Gupta Patil R, Paul TV, Christina F, et al. Masked hypoglycemia in pregnancy. J Diabetes. 2017;9(8):778-86.
2. Bailey T, Bode BW, Christiansen MP, Klaff LJ, Alva S. The Performance and Usability of a Factory-Calibrated Flash Glucose Monitoring System. Diabetes Technol Ther. 2015;17(11):787-94.
3. Bolinder J, Antuna R, Geelhoed-Duijvestijn P, Kroger J, Weitgasser R. Novel glucose-sensing technology and hypoglycaemia in type 1 diabetes: a multicentre, non-masked, randomised controlled trial. Lancet. 2016;388(10057):2254-63.
4. Ayyappan S, Philips S, Kumar CK, Vaithiyanandane V, Sasikala C. Serum fructosamine a better indicator than glycated hemoglobin for monitoring gestational diabetes mellitus. J Pharm Bioallied Sci. 2015;7(Suppl 1):S32-4.
5. Mendes N, Alves M, Andrade R, Ribeiro RT, Papoila AL, Serrano F. Association between glycated albumin, fructosamine, and HbA1c with neonatal outcomes in a prospective cohort of women with gestational diabetes mellitus. Int J Gynaecol Obstet. 2019;146(3):326-32.
6. Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy: a World Health Organization Guideline. Diabetes Res Clin Pract. 2014;103(3):341-63.
7. Tudor-Locke C, Bassett DR, Jr. How many steps/day are enough? Preliminary pedometer indices for public health. Sports Med. 2004;34(1):1-8.
8. American Diabetes A. Insulin administration. Diabetes Care. 2003;26 Suppl 1:S121-4.
9. Costa D, Lourenco J, Monteiro AM, Castro B, Oliveira P, Tinoco MC, et al. Clinical Performance of Flash Glucose Monitoring System in Patients with Liver Cirrhosis and Diabetes Mellitus. Sci Rep. 2020;10(1):7460.
10. Jiang H, Qian X, Li M, Lynn H, Fan Y, Jiang H, et al. Can physical activity reduce excessive gestational weight gain? Findings from a Chinese urban pregnant women cohort study. Int J Behav Nutr Phys Act. 2012;9:12.
11. Takito MY, Benicio MH. Physical activity during pregnancy and fetal outcomes: a case-control study. Rev Saude Publica. 2010;44(1):90-101.
12. Marshall MR, Montoye AHK, George AJ. Pregnancy walking cadence does not vary by trimester. Gait Posture. 2018;65:81-5.
13. Connolly CP, Conger SA, Montoye AHK, Marshall MR, Schlaff RA, Badon SE, et al. Walking for health during pregnancy: A literature review and considerations for future research. J Sport Health Sci. 2019;8(5):401-11.
14. Standardy Polskiego Towarzystwa Ginekologicznego. Opieka położnicza nad ciężarną otyłą. Ginekol Pol. 2012, 83, 795-780.
15. Wender-Ożegowska, E.; Bomba-Opoń, D.; Brązert, J.; Towarzystwa, S.P. Ginekologicznego: Postępowanie u kobiet z cukrzycą. Ginekol. Pol. 2017, 2, 215–229.

VERSION 2 – REVIEW

REVIEWER	Lynnsay M Dickson University of the Witwatersrand, South Africa
REVIEW RETURNED	23-Dec-2020

GENERAL COMMENTS	Within-group comparison using Bland-Altman, Passing-Bablok and

	Deming regression analysis may be difficult to execute and/or interpret. Consider omitting. (Lines 33-34, page 8 of 36) Consider including the English translation of the following attachments: 1. Questionnaire for qualitative assessment of menus of patients (pages 14-16 of 36) 2. Informed consent (pages 24-25 of 36)

REVIEWER	Aruna Nigam Hamdard Institute of Medical Sciences and Research, Jamia Hamdard, New Delhi, India
REVIEW RETURNED	16-Dec-2020

GENERAL COMMENTS	Reviewer submitted the review without further comments.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1:
1. Within-group comparison using Bland-Altman, Passing-Bablok and Deming regression analysis may be difficult to execute and/or interpret. Consider omitting. (Lines 33-34, page 8 of 36).

As suggested by the Reviewer we removed Deming regression analysis from statistical analysis plan.

We would like to use below mentioned statistical methods to compare the groups. We agree that within the groups the analysis would be difficult to execute (we changed in the study protocol “within” for “between”).

We would like to use Bland-Altman plot to compare two methods (FGM and SMBG). It will plot the difference between the two measurements on the Y axis, and the average of the two measurements on the X axis from the whole period of the study (each result: three postprandial and fasting glucose level will have own plot).

We would like to use Passing-Bablok regression for glucose measurement method comparison as it is nonparametric method for fitting a straight line to two-dimensional data where both variables, X and Y are measured. By using this method, we would like to compare FGM with SMBG.

2. Consider including the English translation of the following attachments:
1. Questionnaire for qualitative assessment of menus of patients (pages 14-16 of 36)
2. Informed consent (pages 24-25 of 36)

We agree that the English translation of the attachments mentioned above would be useful for the study, if we would have foreign patients. On the other hand, in our outpatient’s clinic, where we are recruiting participants we do not have any foreign patients and this is the reason why the informed consent and questionnaire for qualitative assessment of menus of patients are submitted only in Polish.

VERSION 3 – REVIEW

REVIEWER	Lynnsay M Dickson University of the Witwatersrand. South Africa.

REVIEW RETURNED	25-Jan-2021

GENERAL COMMENTS	No additional comments