PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Randomized control trial for the feasibility of continuous glucose monitoring in patients with type 1 diabetes at two district hospitals in Neno, Malawi.
AUTHORS	Gomber, Apoorva; Valeta, Francis; Coates, Matthew M.; Trujillo, Celina; Ferrari, Gina; Boti, Medson; Kumwenda, Kenwood; Mailosi, Bright; Nakotwa, Dester; Drown, Laura; Wroe, Emily B.; Thapa, Ada; Mithi, Victor; Matanje, Beatrice; Msekandiana, Amos; Park, Paul H; Kachimanga, Chiyembekezo; Bukhman, Gene; Ruderman, Todd; Adler, Alma

VERSION 1 – REVIEW

REVIEWER	Moran, Antoinette University of Minnesota System
REVIEW RETURNED	17-Jul-2023

GENERAL COMMENTS	Summary Forty-three people with T1D living in rural Malawi were treated for three months with Dexcom G6 (n=29) or usual care (n=13) with SMBG 1-2x/day in a randomized feasibility study. The CGM devices were well tolerated and well accepted by patients. Only one subject was able to change their own sensor, and as subjects came to clinic every 14 days and the sensor only last 10 days there was 67% sensor wear. Patients with CGM had greater numbers of dose adjustments and lifestyle change suggestions, and there was a non-significant trend towards HbA1c improvement. These positive outcomes may have been related to the CGM, or to the fact that the CGM group had 8.3 interactions with the healthcare team vs 1.3 in the usual care group, or a combination of the two factors. Critique 1. There is a bit of confusion in the primary and secondary aims as stated in the abstract. This was a feasibility not an effectiveness study---the study was not powered for effectiveness. That is not to say that feasibility is not an important goal in and of itself, there are very real reasons why sensors present challenges in low income, low education, hot climate regions. 2. As a feasibility study there are some practical issues that it would be nice to see fleshed out a little more. Rashes are only briefly mentioned but this is a common problem with CGM and one which one might expect to be worse in a hot climate. In addition, keeping the sensor attached can be a problem. Whether their subjects did or did not experience these problems, a little more detail would be welcome, and if problems existed, some detail on how they were managed.

	3. An important part of feasibility is patient experience and acceptance. The issues raised by the qualitative interviews belong in this paper, not as a separate manuscript. 4. In table one, duration of diabetes would be more useful than year of diagnosis. I have a hard time believing some of these older patients (46, 51) have type 1 diabetes given that autoantibodies are not likely to have been done in the Malawi setting. However, it doesn't matter for the purposes of this paper if they have other forms of diabetes that are insulin requiring. 5. I am surprised by how good the baseline A1c levels are, given the infrequency of glucose monitoring and given the amount of time spent >250 as shown in figure 2. There are many factors that could affect hemoglobin and thus HbA1c, perhaps sensor GMI values would be more informative. 6. I am not concerned by only getting data 10/14 days---this is still substantially more data than those patients doing a fingerpoke once or twice a day have available. It is curious to me that even people with limited numeracy skills could not be taught to copy numbers into the Dexcom, but as the authors point out, this is a moot point with newer CGM systems.

REVIEWER	Dicembrini, Ilaria University of Florence
REVIEW RETURNED	18-Sep-2023

GENERAL COMMENTS	Dear Authors, your manuscript has the relevant aim to investigate the efficacy of continuous glucose monitoring in type 1 diabetic subjects living in rural area without electricity and having low literacy. T A randomized clinical trial was then conducted in 45 patients randomized to standard of therapy or CGM. However no power calculation has been performed in order to investigate significant differences between the two study groups arms on HbA1c change or other endopoints. Moreover differences on baseline characteristics seem to be of importance between the two study groups as HbA1c value. In order to detect an HbA1c reduction of at least 15% greater in CGM randomized patients in compariso to control individuals you need to enroll 92 subjects. Patients randomized to CGM have significant higher glycemic values than those randomized to standard of therapy (HbA1c 8.6 versus 8.1%).This difference could explain at least in part the greater efficacy observed with CGM versus standard of therapy in terms of HbA1c reduction. Further observation: diagnosis of Type 1 diabetes was not confirmed: are enrolled patients diabetic individuals in insulin intensive regimen? Likely this study has the importance to investigate how the future technology will be able to become easier to address the needs of type 1 diabetic people living in rural area and/or with lower literacy.

REVIEWER	Gardiner, Joseph Michigan State University Department of Epidemiology, Biostatistics
REVIEW RETURNED	19-Nov-2023

GENERAL COMMENTS	Nov 20, 2023

	Review of Manuscript ID bmjopen-2023-075554, titled "Randomized control trial for the feasibility of continuous glucose monitoring in patients with type 1 diabetes at two district hospitals in Neno, Malawi.." Gomber et al PAGES 1-34 Additional reports: A qualitative study --- Thapa et al, PAGES 35-62 Clinical insights---Ferrari et al, PAGES 63-91 Overall assessment: The studies were conducted by an enthusiastic group of investigators engaging local personnel at two sites (hospitals) in a rural area of Malawi. The lessons learnt will be valuable in carrying out studies of this type in low resource settings. Although results from the CGM vs UC appear to favor the intervention CGM, the small sample size /lack of statistical power does not make for stronger conclusions. Nevertheless, the groundwork has been set up for future studies. With further training the feasibility of using CGM in the T1D patient population seems promising. General Comments Main manuscript Basic and obligatory features of the RCT are provided--- including a CONSORT statement, and Flow Diagram (Figure 1). Unsure what the page numbers refer to. At the upper right (or left)-hand corner of page numbers are overwritten and illegible---see after page 34. Details include challenges to conducting the trial in a LIC—and also in a remote location with few resources. Study required training of personnel, evaluators. Participants were given support on use of monitoring equipment. Line 166-171: Only 45 participants met the criteria for participation. Randomized 2:1 to CGM (continuous glucose monitoring with Dexcom G6) or UC (blood glucose meter, Safe-Accu). However, missing data reduced slightly the sample size: 29 in CGM, 13 in UC (Table 1). Both groups received diabetes education support. Table 1: Range of age of participants suggest some middle-aged subjects (>40 years). Suggest adding the median age of participants. Line 154: ---"diagnosis of T1D from any age group". Please explain. Implementation endpoints ---fidelity, acceptability, appropriateness Main clinical outcome: Change in HbA1c from baseline to 90 day follow-up. Line 131/Line 271: ---"standard deviation of HbA1C" --meaning estimated from all participants , at baseline, and subsequently at follow-up? Line 278: The power calculation needs some explanation. Presume the (unstated) significance level is alpha=.05. For HbA1c, a two-sided test (t-test) is conducted to test for no difference between the CGM and UC arms. The SD=2.05 (pooled

	SD). With group assignment CGM vs UC as 2:1, the total sample size needed to detect a mean difference of 1.05 in HbA1c between arms with 80% power, will require a much larger sample size than 45. The sentence –"we included the entire population---so no sample size calculations were conducted" needs explication. Line 280: add a description of the error term in the model Line 284: HbA1c Analysis. With a sample size of only 29+13=42 (minus 3)=39, estimating 8 parameters in model (line 290) could be problematic. Were any model diagnostics assessed? Line 413: Clinical outcomes (also Table 3). CGM arm did a little better—reduction in HbA1c, but UC arm did worse—increase in HbA1c. Perhaps none of these are significant, but the reduction seen in CGM is (clinically) promising. Line 441:---"average standard deviation of HbA1C across both arms" . Does with mean pooled standard deviation from estimates from each arm. In Table 1 it is from the baseline values? Line 445: Table 3. Second entry—crude change from baseline. The SD here is for the difference, follow-up minus baseline. Results for model 1 and model 2 are differences in least-squares means CGM vs UC) Line 452: Qol analyses—in Supplementary Table 1. This is a pro-forma analysis. Inference is limited by the small sample (n=38). Also see comment above for HbA1c. 1

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Antoinette Moran, University of Minnesota System
Comments to the Author:
Summary
Forty-three people with T1D living in rural Malawi were treated for three months with Dexcom G6 (n=29) or usual care (n=13) with SMBG 1-2x/day in a randomized feasibility study. The CGM devices were well tolerated and well accepted by patients. Only one subject was able to change their own sensor, and as subjects came to clinic every 14 days and the sensor only last 10 days there was 67% sensor wear. Patients with CGM had greater numbers of dose adjustments and lifestyle change suggestions, and there was a non-significant trend towards HbA1c improvement. These positive outcomes may have been related to the CGM, or to the fact that the CGM group had 8.3 interactions with the healthcare team vs 1.3 in the usual care group, or a combination of the two factors.

Critique
1. There is a bit of confusion in the primary and secondary aims as stated in the abstract. This was a

feasibility not an effectiveness study---the study was not powered for effectiveness. That is not to say that feasibility is not an important goal in and of itself, there are very real reasons why sensors present challenges in low income, low education, hot climate regions.

Thank you for this comment-we have tried to rework this in the abstract and throughout the document to make it clearer that this was a feasibility study with some information on clinical outcomes. We do believe that it is vital to report on some clinical outcomes- while CGM is standard of care in high resource settings it is relatively untested in low resource settings. Given the challenges of use, we need to see if it has the ability to influence clinical outcomes-and it is important to ensure it does not make people worse.
2. As a feasibility study there are some practical issues that it would be nice to see fleshed out a little more. Rashes are only briefly mentioned but this is a common problem with CGM and one which one might expect to be worse in a hot climate. In addition, keeping the sensor attached can be a problem. Whether their subjects did or did not experience these problems, a little more detail would be welcome, and if problems existed, some detail on how they were managed.

We have now included the following lines: “Rashes and skin irritation were not a commonly encountered complaint in the CGM arm. The hot weather caused a few participants difficulty with keeping the sensor attached. We overcame this using skin Tac adhesive and cut-out tape overpatches to secure the sensors in place and prevent removal. No sensor related bleeding or potential skin reaction around or under the sensor was observed..”

3. An important part of feasibility is patient experience and acceptance. The issues raised by the qualitative interviews belong in this paper, not as a separate manuscript.

Thank you, we agree that the qualitative portion is of the utmost importance, which is one of the reasons that we chose to have it as a separate paper. This allowed us to go more in depth on themes than we would have done otherwise.

The second reason that we opted to have the qualitative paper as a separate paper was due to the author list. In this study, like many done in low resource settings, many of the researchers involved were also the people providing the intervention. We cannot have a first author who is also an informant in qualitative analysis. If we included some of the important qualitative work in this paper, we would not have been able to include five of our authors who were vital in the carrying out and writing of the manuscript. For these reasons we respectfully ask that the structure be preserved.

4. In table one, duration of diabetes would be more useful than year of diagnosis. I have a hard time believing some of these older patients (46, 51) have type 1 diabetes given that autoantibodies are not likely to have been done in the Malawi setting. However, it doesn’t matter for the purposes of this paper if they have other forms of diabetes that are insulin requiring.

We have made this change in the tables, we have also added in age at diagnosis. As a note to the reviewer, we recognize the unusual age make-up but note that this is a common but not well understood phenomenon in Sub Saharan Africa.

In addition, we also would like to note that some authors have argued that there may be differences in genetic and autoimmune underpinnings between type 1 diabetes in patients of African and European, Asian and Arab descent. This is beyond the scope of our current study findings and we believe there is a growing interest in African phenotype understanding the etio-pathogeny for t1d. (if you are interested in this here are some articles that may be of interest) 1. Katte JC, McDonald TJ, Sobngwi E, Jones AG. The phenotype of type 1 diabetes in sub-Saharan Africa. Front Public Health. 2023 Jan 27;11:1014626. doi: 10.3389/fpubh.2023.1014626. PMID: 36778553; PMCID: PMC9912986.

2. Gill GV, Mbanya JC, Ramaiya KL, Tesfaye S. A sub-Saharan African perspective of diabetes. Diabetologia. 2009 Jan;52(1):8-16. doi: 10.1007/s00125-008-1167-9. Epub 2008 Oct 10. PMID: 18846363. 3. Osei K, Schuster DP, Amoah AG, Owusu SK. Diabetes in Africa. Pathogenesis of type 1 and type 2 diabetes mellitus in sub-Saharan Africa: implications for transitional populations. J Cardiovasc Risk. 2003 Apr;10(2):85-96. doi: 10.1097/01.hjr.0000060841.48106.a3. PMID: 12668905.

5. I am surprised by how good the baseline A1c levels are, given the infrequency of glucose monitoring and given the amount of time spent >250 as shown in figure 2. There are many factors that could affect hemoglobin and thus HbA1c, perhaps sensor GMI values would be more informative.

Thank you for this comment-we had done an exploration of this separately, but agree that this is important to address in this manuscript. We have now added a section on exploring the HbA1c in the methods, results, and discussion.

We now say in the results: “The endline point-of-care HbA1c values were compared to the estimated HbA1c based on 90-day CGM values using a paired t-test. Mean endline point-of-care HbA1c was 7.4% (95% CI 6.6%, 8.1%). Mean estimated HbA1c was significantly higher, at 10.1% (95% CI 9.3%, 10.8%) and mean difference of 2.7% (95% CI 2.2%, 3.2%; $p < 0.05$). Supplementary Figure 1 shows point-of-care HbA1c and estimated HbA1c for each participant.”

In the discussion we say:

“Comparison of endline point-of-care HbA1c to estimated HbA1c based on CGM values showed that point-of-care HbA1c may be overestimating glycemic control. A few theories for the discrepancy between HbA1c and mean blood glucose levels have been proposed, including the presence of hemoglobinopathies, individual variations in the lifespan of red blood cells, renal impairment, and nutritional deficiencies (e.g., iron-deficiency anemia, Kwashiorkor, Marasmus). No hemoglobinopathies were present in this population. Additionally, numerous assays for point-of-care HbA1c testing have become available over the last decade of possibly varying quality. These findings reinforce that HbA1c alone may not be adequate to evaluate glycemic control in PLWT1D, adding to current literature highlighting the importance of availability for additional ways to evaluate glycemic control, such as SMBG or CGM.”

6. I am not concerned by only getting data 10/14 days---this is still substantially more data than those patients doing a fingerpoke once or twice a day have available. It is curious to me that even people with limited numeracy skills could not be taught to copy numbers into the Dexcom, but as the authors point out, this is a moot point with newer CGM systems.

The purpose of this study was not to focus on the ability of participants to be able to change sensors, we were focusing on the feasibility of wear and the use of information.

As a note to the reviewer, some of the people in this study have never seen a number before. The entire concept was new to them. With time, we agree that they can probably be taught as part of diabetes self-management education (DSME) and strengthening diabetes care in low-resource settings, but that was not the focus of this study.

Reviewer: 2
Dr. Ilaria Dicembrini, University of Florence

Comments to the Author:
Dear Authors, your manuscript has the relevant aim to investigate the efficacy of continuous glucose monitoring in type 1 diabetic subjects living in rural area without electricity and having low literacy. T A randomized clinical trial was then conducted in 45 patients randomized to standard of therapy or CGM. However no power calculation has been performed in order to investigate significant differences between the two study groups arms on HbA1c change or other endopoints. In order to detect an HbA1c reduction of at least 15% greater in CGM randomized patients in comparison to control individuals you need to enroll 92 subjects.

Thank you, we did report power calculations for HbA1c under particular assumptions about the HbA1c standard deviation (see statistical methods section, line 280 in original submission). These power calculations were performed for the published protocol for this study (reference cited in the manuscript) and can be seen there in Figure 1 for a range of different assumptions about HbA1c standard deviations and effect sizes. However, these power calculations were for the original sample size that we expected to be able to obtain (50) and for a one-tailed test. We thank the reviewer for noting this issue. We enrolled all possible patients at the clinics, so we would not have been able to enroll 92 subjects. We have updated the sentence in the manuscript containing the power calculation to reflect the actual sample size rather than the original anticipated sample size and for a two-tailed test instead of a one-tailed test (below).

“We did not conduct sample size calculations because we recruited all PLWT1D receiving care at two PIH assisted hospitals where this study was being conducted. Rather, we calculated power to detect the difference in HbA1c with the number of patients who participated (29 in the CGM arm and 13 in the usual care arm). Given a pooled standard deviation of 2.05 and an alpha level of 0.05, we had 80% power to detect a 1.96 percentage point difference in HbA1c between the two study arms in a two-sample t-test. Initial power calculations relied on expected larger numbers of participants(1).”

Moreover differences on baseline characteristics seem to be of importance between the two study groups as HbA1c value. Patients randomized to CGM have significant higher glycemic values than those randomized to standard of therapy (HbA1c 8.6 versus 8.1%).This difference could explain at least in part the greater efficacy observed with CGM versus standard of therapy in terms of HbA1c reduction.

As the reviewer notes, although randomization removes confounding by baseline HbA1c in expectation, it is true that random differences in baseline values can be a factor in differential changes observed between groups if the baseline level is an important determinant of the change. However, we adjusted for baseline HbA1c in our analysis (regression equation in line 291), accounting for this. Results for Model 1 and Model 2 in Table 3 show slightly smaller adjusted mean difference point estimates compared to the crude difference (as the reviewer suggests may happen). Confidence intervals for all 3 approaches are compatible with a moderate to null reduction in the CGM arm compared to the usual care arm.

Further observation: diagnosis of Type 1 diabetes was not confirmed: are enrolled patients diabetic individuals in insulin intensive regimen? Likely this study has the importance to investigate how the future technology will be able to become easier to address the needs of type 1 diabetic people living in rural area and/or with lower literacy.

Type 1 diabetes is not confirmed in this setting due to limitations in resources available to make this determination. All people were diagnosed as having Type 1 diabetes based on local guidelines, and the clinicians who had been treating them all were confident in their diagnosis. The diagnostic criteria for all types of diabetes are based on symptoms, age, clinical signs of polyuria, polydipsia, etc., and available laboratory measurements of blood glucose levels. Measurement of islet autoantibodies, C-

peptide, or further genetic tests is not routinely recommended and practiced for the diagnosis of T1D in many low-resource settings like Malawi.

All patients were on an intensive insulin regimen We have now included the following sentence in the methods section: “PLWT1D typically use human insulin, intermediate-acting (NPH) two times daily and fast- acting (regular) two to three times daily.”

Reviewer: 3
Dr. Joseph Gardiner, Michigan State University Department of Epidemiology
Comments to the Author:
Difficult study in a remote location. Enthusiastic group of investigators--engaging local personnel and stakeholders.

Review of Manuscript ID bmjopen-2023-075554, titled "Randomized control trial for the feasibility of continuous glucose monitoring in patients with type 1 diabetes at two district hospitals in Neno, Malawi.." Gomber et al PAGES 1-34
Additional reports: A qualitative study --- Thapa et al, PAGES 35-62
Clinical insights---Ferrari et al, PAGES 63-91
Overall assessment: The studies were conducted by an enthusiastic group of investigators engaging local personnel at two sites (hospitals) in a rural area of Malawi. The lessons learnt will be valuable in carrying out studies of this type in low resource settings. Although results from the CGM vs UC appear to favor the intervention CGM, the small sample size /lack of statistical power does not make for stronger conclusions. Nevertheless, the groundwork has been set up for future studies. With further training the feasibility of using CGM in the T1D patient population seems promising.
General Comments Main manuscript
Basic and obligatory features of the RCT are provided--- including a CONSORT statement, and Flow Diagram (Figure 1). Unsure what the page numbers refer to. At the upper right (or left)-hand corner of page numbers are overwritten and illegible---see after page 34.

With apologies this appears to be a production issue. I see your point in the file generated by BMJ open, but in our documents this does not occur, and I don’t think we can fix this as this is something generated by the production process.

Details include challenges to conducting the trial in a LIC—and also in a remote location with few resources. Study required training of personnel, evaluators. Participants were given support on use of monitoring equipment.
Line 166-171: Only 45 participants met the criteria for participation. Randomized 2:1 to CGM (continuous glucose monitoring with Dexcom G6) or UC (blood glucose meter, Safe-Accu). However, missing data reduced slightly the sample size: 29 in CGM, 13 in UC (Table 1). Both groups received diabetes education support.

This is correct. We noted 1 patient from the CGM arm and 2 from the usual care arm did not attend on the day of the trial initiation in the first paragraph of the results section.

Table 1: Range of age of participants suggest some middle-aged subjects (>40 years). Suggest adding the median age of participants.

Thank you, median age was added to Table 1.

Line 154: ---"diagnosis of T1D from any age group”. Please explain.

We have now reworded this to say:

“Eligibility criteria for this study included a clinical diagnosis of T1D in PLWT1D, in diabetes care for at least one year, and seeking care at either of the PIH-supported MOH hospitals. We did not exclude anyone based on age.”

Implementation endpoints ---fidelity, acceptability, appropriateness
Main clinical outcome: Change in HbA1c from baseline to 90 day follow-up.

Line 131/Line 271: ---"standard deviation of HbA1C” --meaning estimated from all participants , at baseline, and subsequently at follow-up?

Yes, standard deviation of HbA1c at baseline may be of interest to others (e.g., for sample size/power calculations for future interventions). We have added “at baseline” in line 131. Baseline was already specified in line 271.

We report the standard deviations of HbA1c in both arms of the study separately at baseline and follow-up as well, but these specific lines of the manuscript refer to a secondary outcome specified in the protocol that might inform additional studies. We have also added a column in Table 1 for the total study population.

Line 278: The power calculation needs some explanation. Presume the (unstated) significance level is alpha=.05. For HbA1c, a two-sided test (t-test) is conducted to test for no difference between the CGM and UC arms. The SD=2.05 (pooled SD). With group assignment CGM vs UC as 2:1, the total sample size needed to detect a mean difference of 1.05 in HbA1c between arms with 80% power, will require a much larger sample size than 45. The sentence –“we included the entire population---so no sample size calculations were conducted” needs explication.

Thank you for this comment. This sentence in line 278 was based on an initial power calculation prior to implementation of the study using alpha=0.05 for a one-sided t-test for a mean difference of 1.5 for the number of subjects that we initially anticipated being able to have participate in the study (33 and 17 in the respective groups). We have corrected this, sought to clarify the sample size statement, and included the relevant details about the power calculation:

“We did not conduct sample size calculations because we recruited all PLWT1D receiving care at two PIH assisted hospitals where this study was being conducted. Rather, we calculated power to detect the difference in HbA1c with the number of patients who participated (29 in the CGM arm and 13 in the usual care arm). Given a pooled standard deviation of 2.05 and an alpha level of 0.05, we had 80% power to detect a 1.96 percentage point difference in HbA1c between the two study arms in a two-sample t-test. Initial power calculations relied on expected larger numbers of participants(1).”

Line 280: add a description of the error term in the model

We added the error term to the sentence describing the model.
“To test whether the change in HbA1c differed between the CGM and usual care arms, we used the linear regression model specified below, equivalent to longitudinal analysis of covariance, where HbA1c at follow-up (HbA1ct1) is predicted by study arm (SA), HbA1c at baseline (HbA1ct0), facility site (Site), age (Age), female gender (Fem), diagnosis year (DY), and body mass index (BMI), with an error term, ε, assumed normally distributed.”

Line 284: HbA1c Analysis. With a sample size of only 29+13=42 (minus 3)=39, estimating 8

parameters in model (line 290) could be problematic. Were any model diagnostics assessed?

We agree that including too many covariates can be problematic. On the other hand, including these parameters may also increase precision if they are helpful in predicting the outcome. We additionally included a minimally adjusted model that only adjusted for baseline HbA1c in estimating the effect of study arm (Table 3, Model 1). The model with 8 parameters gave a 95% confidence interval for our parameter of interest that was largely similar to that estimated by the model only adjusting for baseline HbA1c (-2.39 to 0.26 in the fully adjusted model versus -2.15 to 0.40 in the minimally adjusted model), suggesting that the inclusion of these additional parameters did not greatly inflate uncertainty. The adjusted R2 values for the two models were 0.21 (fully adjusted) and 0.16 (minimally adjusted), though of course metrics like adjusted R2 are not without limitations. We reported results for the crude change, the minimally adjusted model, and the adjusted model for transparency.

Line 413: Clinical outcomes (also Table 3). CGM arm did a little better—reduction in HbA1c, but UC arm did worse—increase in HbA1c. Perhaps none of these are significant, but the reduction seen in CGM is (clinically) promising.

We agree with the reviewer. We have rewritten a sentence in the clinical outcomes to more explicitly interpret the finding.

“After adjusting for baseline HbA1c levels and other covariates, participation in CGM compared to usual care was associated with a 1.1 percentage point lower HbA1c; confidence intervals were compatible with a moderate to null reduction in the CGM arm relative to the usual care arm (95% CI: 2.4 percentage point reduction to 0.3 percentage point increase, Table 3).”

Line 441:---"average standard deviation of HbA1C across both arms” . Does with mean pooled standard deviation from estimates from each arm. In Table 1 it is from the baseline values?

Correct, this is the pooled baseline standard deviation. We have amended the phrasing and removed the reference to Table 3 to avoid confusion, adding a reference to Table 1, which now shows pooled characteristics as well. The reviewer is correct that this is the baseline values in Table 1.

“Pretrial, there was a standard deviation of 2.1 in HbA1c pooled across two arms (Table 1), although baseline HbA1c was low overall compared to what is generally expected in this type of setting(23-25).”

Line 445: Table 3. Second entry—crude change from baseline. The SD here is for the difference, follow-up minus baseline. Results for model 1 and model 2 are differences in least-squares means CGM vs UC)

Yes, this is correct

Line 452: Qol analyses—in Supplementary Table 1. This is a pro-forma analysis. Inference is limited by the small sample (n=38). Also see comment above for HbA1c.

Yes, we are limited by the sample size. To reduce the chance of overinterpretation of the QoL analysis results, we have rephrased the results section:

“Over the course of the study, QoL (N=28 in CGM and N=10 in UC) was assessed using WHO-BREF increased across all domains (Supplementary Table 1). Though unadjusted QoL increased slightly more in the UC arm (9.0) than the CGM arm (6.7), confidence intervals for differences in the change

in QoL between groups were large, and we did not find any strong evidence of differences."

In addition, we explicitly added HbA1c and QoL to the first sentence of the limitations section: "This was a feasibility trial with only 42 individuals, so may not have been powered for seeing differences between study arms in outcomes like HbA1c and QoL."

VERSION 2 – REVIEW

REVIEWER	Moran, Antoinette University of Minnesota System
REVIEW RETURNED	11-Feb-2024

GENERAL COMMENTS	This manuscript is improved. I do believe the results are highly important because of the paucity of sensor data in low income countries. There are two issues that were not addressed from my review, one of which I feel strongly about, and one which the editor can decide. Minor issue---the references do not match the references numbers, at least by the time they get to the 20's. 1. This is a feasibility study. That is good enough for publication. They are not powered for efficacy (that would require about 180 subjects). They took out the word "efficacy", but still call the change in HbA1c a primary outcome (abstract) and say in their objectives "to determine if CGM use can have an effect on diabetes clinical outcomes.." This is efficacy. Efficacy cannot be a primary outcome in a grossly underpowered study. Again, feasibility is enough. Also it is fine to use A1c for future power analysis. 2. As a feasibility study, I still believe the patient attitudes should go in this paper rather than a companion paper, but that is up to the editor.

VERSION 2 – AUTHOR RESPONSE

1. This is a feasibility study. That is good enough for publication. They are not powered for efficacy (that would require about 180 subjects). They took out the word "efficacy", but still call the change in HbA1c a primary outcome (abstract) and say in their objectives "to determine if CGM use can have an effect on diabetes clinical outcomes.." This is efficacy. Efficacy cannot be a primary outcome in a grossly underpowered study. Again, feasibility is enough. Also it is fine to use A1c for future power analysis.

We have now changed change in HbA1c to a secondary outcome. We have made changes in the abstract, methods, results and differences between the protocol and review sections.

2. As a feasibility study, I still believe the patient attitudes should go in this paper rather than a companion paper, but that is up to the editor.

We have now written a justification in the introduction.

VERSION 3 – REVIEW

REVIEWER	Moran, Antoinette University of Minnesota System
REVIEW RETURNED	07-Mar-2024

GENERAL COMMENTS	I am satisfied with their revisions.