PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Effects of basic carbohydrate counting vs. standard outpatient nutritional education (The BCC Study): study protocol for a randomised, parallel open-label, intervention study focusing on HbA1c and glucose variability in patients with type 2 diabetes
AUTHORS	Ewers, Bettina; Bruun, Jens; Vilsbøll, Tina

VERSION 1 - REVIEW

REVIEWER	Pamela Dyson University of Oxford UK
REVIEW RETURNED	12-Aug-2019

GENERAL COMMENTS	This manuscript describes a protocol for assessing the effects of carbohydrate counting on glycaemic control in people with type 2 diabetes. Despite almost universal acceptance for the role of carbohydrate counting (and insulin adjustment) in people with type 1 diabetes, this strategy had not been well studied in people with type 2 diabetes and the results of this study should prove informative. Major comments It now generally acknowledged that referring to people with diabetes as 'patients' is not acceptable and they should be described simply as 'people with diabetes' Aim: The aim of any study should be reflected in the primary outcome. Here the aims are stated as 'improving carbohydrate counting accuracy and day-to-day consistency of carbohydrate intake' and yet the primary outcome is not assessment of carbohydrate intake, but differences in glycaemic control as measured by HbA1c or MAGE Study duration: There appears to be some confusion here. The abstract mentions a 12-month study, but the methods and fig 3 state 48 weeks. In addition, the section detailing data collection on page 8 states that the end of the study is week 24. I would recommend using the description 48 weeks throughout. The authors state that 48 weeks is 'long-term follow-up'. This is debatable and requires supporting references. In general, dietary studies are considered long-term only if they extend beyond 1 year (52 weeks), and this definition has been applied in recent meta-analyses e.g. those reporting outcomes for low carbohydrate diets

	Outcomes: The power calculation states that a difference of 3.0 mmol/mol between the two groups is considered clinically meaningful. Although this is a subject of much controversy, most authorities agree that a difference of at least 5.0 mmol/mol is required in order to be clinically meaningful. The authors should justify why they have chosen 3.0 mmol/mol and provide supporting references. Minor comments Introduction Page 4, line 5. Replace 'calorie' with 'energy' – this applies throughout the manuscript Page 4, line 24/25. BCC aims to improve overall glycaemic control and this is not limited to plasma glucose Page 4, lines 27-34. Are there any references to support the barriers to ACC that are listed, if not, the authors need to make it clear that this is their opinion. When judgemental statements such as 'lack of motivation' are made, it is important to supply supporting references Page 5, line 7/8. References are required to support the statement that increased carbohydrate awareness may lead to reductions in carbohydrate intake Page 5, lines 10-13. References are required to support the statement about short-term effects of carbohydrate restriction and individualisation Methods Page 8. Intervention group. The final sentence mentions an app for carbohydrate estimation. Presumably this is only available to those with compatible devices, so should this be mentioned as part of the inclusion criteria? Page 8. Control group. One of the most common criticisms of dietary studies is the lack of detailed information about the interventions and this is especially true of control groups. I would recommend providing details of all interventions in a separate appendix including any dietary resources that are used in clinical practice. For the intervention group, the detailed curriculum of the education programme should be reported Page 9. Secondary outcomes: Clinical parameters include % time spent in range (3.9-10.0 mmol/l). An explanation of these limits is required, as is >10 mmol/l representing hyperglycaemia Page 9. Is the PCS a validated questionnaire? Supporting references are required Page 10. Sample size calculation. Supporting references for the calculation based on HbA1c should be supplied. The relationship between changes in MAGE and improved outcomes should be made explicit There are typos and errors with grammar and syntax throughout the manuscript and I would recommend editing by someone with English as their first language. Examples include (this is not an exhaustive list): Page 5, line 22. Delete the word 'in' between 'educated' and 'how' Page 5, line 25. 'ACC is targeted the patient' should read 'ACC is targeted at the patient' Page 6, line 4. 'Particularly' should be replaced with 'In particular' Page 6, line 4. The word 'high' is redundant, this should phrase should read 'energy-dense foods'

	Page 6, line 14. Delete ‘the’ between ‘managing’ and ‘diet’ and change ‘diet’ to ‘dietary intake’ Page 7, line 36. Delete ‘of’ between ‘use’ and ‘an’ Page 7, line 37. Delete ‘the’ between ‘affecting’ and ‘dietary’ Page 8, line 8. Insert ‘will’ between ‘participants’ and ‘receive’

REVIEWER	David Kerr Sansum Diabetes Research Institute United States
REVIEW RETURNED	23-Aug-2019

GENERAL COMMENTS	This study is important and the researchers are from an organization with a long and impressive track record in clinical research. It is very likely to be successful. However, the proposed approach has important methodological limitations in its current form. Specifically, (a) there is no measure of food security which is an established risk factor in type 2 diabetes (b) social determinants of health are recognized increasingly as important considerations in type 2 diabetes self-management. What steps are the authors taking to make sure that the materials are culturally appropriate and not lead to selection bias? (b) the inclusion criteria allow any glucose-lowering is allowed including insulin and other injectables and the dose and timing may change - this provides an opportunity for a health-economic comparison which would be of additional value. More importantly, there may be different magnitudes of effect with different sub-groups of pharmacological therapies which may make the findings difficult to interpret and perhaps the sample size calculation invalid. Would it make more sense to limit to therapies involved? Also, an additional statistical (i.e. actuarial) approach using "response" to interventions as the outcome and assessing the different weighted factors influencing the responses may add value. (c) the HCP to participant interactions between groups is unbalanced in terms of time - this is a frequent challenge in this type of research. By this, I mean that the content of the proposed education program may be less important than the simple fact that a professional is interacting (i.e. caring) more. Is there some way to control for this with another interaction unrelated to food? (d) the intervention has a number of facets and if successful it may be difficult to tease out which ones are most important. Why not offer the app to this group without any instruction? (e) why not use a wearable device to assess physical activity? I assume this is for economic reasons.

VERSION 1 – AUTHOR RESPONSE

Reviewers' Comments to Author:

Reviewer: 1

Reviewer Name: Pamela Dyson

Institution and Country: University of Oxford, UK Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below This manuscript describes a protocol for assessing the effects of carbohydrate counting on glycaemic control in people with type 2 diabetes. Despite almost universal acceptance for the role of carbohydrate counting (and insulin adjustment) in people with type 1 diabetes, this strategy had not been well studied in people with type 2 diabetes and the results of this study should prove informative.

Major comments R1: It now generally acknowledged that referring to people with diabetes as 'patients' is not acceptable, and they should be described simply as 'people with diabetes'. A: We have changed patients to people throughout the manuscript; however, the title cannot be changed since the study has already been registered at clinicaltrials.gov
Major comments R1: Aim: The aim of any study should be reflected in the primary outcome. Here the aims are stated as 'improving carbohydrate counting accuracy and day-to-day consistency of carbohydrate intake' and yet the primary outcome is not assessment of carbohydrate intake, but differences in glycaemic control as measured by HbA1c or MAGE A: We agree, and we have deleted the sentence "the BCC intervention aims at improving carbohydrate counting accuracy and day-to-day consistency of carbohydrate intake."
Major comments R1: Study duration: There appears to be some confusion here. The abstract mentions a 12-month study, but the methods and fig 3 state 48 weeks. In addition, the section detailing data collection on page 8 states that the end of the study is week 24. I would recommend using the description 48 weeks throughout. A: Thank you for pointing out this mistake. We have corrected it to 12 months throughout the manuscript as well as figures.
Major comments R1: The authors state that 48 weeks is 'long-term follow-up'. This is debatable and requires supporting references. In general, dietary studies are considered long-term only if they extend beyond 1 year (52 weeks), and this definition has been applied in recent meta-analyses e.g. those reporting outcomes for low carbohydrate diets A: A very relevant point. Indeed, no clear definition exists concerning the use of "long-term" in dietary intervention studies. A recent meta-analysis (Snorregaard O, Poulsen GM, Andersen HK, Astrup A. Systematic review and meta-analysis of carbohydrate restriction in patients with type 2 diabetes. BMJ Open Diabetes Research & Care 2017, 5) defined "long-term studies" as those collecting clinical important outcomes "at one year or later" which is in line with our study collecting outcomes at 12 months/1 year. We therefore consider our study as long-term.
Major comments R1: Outcomes: The power calculation states that a difference of 3.0 mmol/mol between the two groups is considered clinically meaningful. Although this is a subject of much controversy, most authorities agree that a difference of at least 5.0 mmol/mol is required in order to be clinically meaningful. The authors should justify why they have chosen 3.0 mmol/mol and provide supporting references.

A: We used the Diabetes Control and Complications Trial (DCCT) to define a clinically meaningful cut-off point for HbA1c reducing the risk of microvascular complications (Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. N Engl J Med. 1993;329(14):977-986).

Minor comments
R1: Page 4, line 5. Replace 'calorie' with 'energy' – this applies throughout the manuscript
A: Thanks for pointing this out. As suggested, we have changed this throughout the manuscript.

Introduction
R1: Page 4, line 24/25. BCC aims to improve overall glycaemic control and this is not limited to plasma glucose
A: We agree – this has been corrected.

Introduction
R1: Page 4, lines 27-34. Are there any references to support the barriers to ACC that are listed, if not, the authors need to make it clear that this is their opinion?
A: Our reference for the listed barriers to ACC is based on a qualitative evaluation of the Dose Adjustment for Normal Eating (DAFNE) programme from the UK, since most research in ACC has been conducted in patients with type 1 diabetes. Reference: Chapter 3 Qualitative and quantitative evaluation of the DAFNE intervention: the psychosocial study. Improving management of type 1 diabetes in the UK: The Dose Adjustment for Normal Eating (DAFNE) programme as a research test-bed. A mixed-method analysis of the barriers to and facilitators of successful diabetes self-management, a health economic analysis, a cluster randomised controlled trial of different models of delivery of an educational intervention and the potential of insulin pumps and additional educator input to improve outcomes. National Institute for Health Research. Programme Grants for Applied Research 2014; Vol. 2: No. 5, DOI: 10.3310/pgfar02050

Introduction
R1: When judgemental statements such as 'lack of motivation' are made, it is important to supply supporting references
A: We believe that we have made it clear that it is the authors' opinion since we write "potential patient barriers" and then list the possible barriers including a possible lack of motivation without reference.

Introduction
R1: Page 5, line 7/8. References are required to support the statement that increased carbohydrate awareness may lead to reductions in carbohydrate intake
A: The statement is based on our clinical experience and academic assessments why no reference is available.

Introduction
R1: Page 5, lines 10-13. References are required to support the statement about short-term effects of carbohydrate restriction and individualisation
A: We wrote "The short-term effects of low-carbohydrate diets may be due to a decline in dietary adherence over time indicating that the recommended intake of carbohydrates should be individualised and based on an assessment of the patient's current eating patterns and preferences as practised in the BCC concept." This is the authors' personal opinions based on data from carbohydrate restriction studies including the previously mentioned meta-analysis (Snorregaard O, Poulsen GM, Andersen HK, Astrup A. Systematic review and meta-analysis of carbohydrate restriction in patients with type 2 diabetes. BMJ Open Diabetes Research & Care 2017, 5). This meta-analysis found that dropout rates tended to be larger in the low-carbohydrate groups in trials with long follow-up indicating that it may be difficult to adhere to a low-carbohydrate diet in the long

run. Thus, as already recommend in the recent guidelines by the American Diabetes Association, the diet should be individualized and based on preferences and current eating patterns.
Methods R1: Page 8. Intervention group. The final sentence mentions an app for carbohydrate estimation. Presumably this is only available to those with compatible devices, so should this be mentioned as part of the inclusion criteria? A: No, it is not a criterion for participating. The people with diabetes participating were offered individualised methods.
Methods R1: Page 8. Control group. One of the most common criticisms of dietary studies is the lack of detailed information about the interventions and this is especially true of control groups. I would recommend providing details of all interventions in a separate appendix including any dietary resources that are used in clinical practice. For the intervention group, the detailed curriculum of the education programme should be reported A: We will provide details of the curriculum in an appendix when the study has been conducted, so the study can be replicated by others.
Methods R1: Page 9. Secondary outcomes: Clinical parameters include % time spent in range (3.9-10.0 mmol/l). An explanation of these limits is required, as is >10 mmol/l representing hyperglycaemia A: Time in range (TIR) (3.9-10.0 mmol/l) and hyperglycaemia (>10 mmol/l) has been defined by a large expert group in International consensus on use of continuous glucose monitoring. Diabetes Care 2017;40:1631-40.
Methods R1: Page 9. Is the PCS a validated questionnaire? Supporting references are required A: Yes, it is a validated questionnaire. The following reference has been placed in the manuscript: Williams GC, McGregor HA, Zeldman A, Freedman ZR, Deci EL. Testing a self-determination theory process model for promoting glycemic control through diabetes self-management. Health Psychol 2004; 23: 58-66.
Methods R1: Page 10. Sample size calculation. Supporting references for the calculation based on HbA1c should be supplied. A: As described in the text the SD and dropout rate used for sample size calculation was based on what we have found when evaluating previous BCC courses at our clinic after 6 months in completers with T2D. BMJ Open do not allow for citations using “unpublished data”, and the data were not published, so we cannot change this.
Methods R1: The relationship between changes in MAGE and improved outcomes should be made explicit A: MAGE is used to capture mealtime-related glucose excursions. MAGE has been associated with coronary artery disease independent of HbA1c (Su G, Mi S, Tao H, et al. Association of glycemic variability and the presence and severity of coronary artery disease in patients with type 2 diabetes. Cardiovasc Diabetol 2011;10:19–9), and in a recent meta-analysis and systematic review MAGE and other measurements of glycaemic variability have been found to reduce carotid intima-media thickness and insulin resistance indicating a casual effect (Liang S, Yin H, Wei C, Xie L, He H, Liu X. Glucose variability for cardiovascular risk factors in type 2 diabetes: a meta-analysis Journal of Diabetes & Metabolic Disorders 2017; 16:45:1-9). The relationship between MAGE and improved

outcomes is now mentioned together with the references in the section on Data Collection and primary outcomes.
R1: There are typos and errors with grammar and syntax throughout the manuscript and I would recommend editing by someone with English as their first language. A: We have revised our manuscript for typos and errors with grammar and syntax.
Typos and errors with grammar and syntax: R1: Page 5, line 22. Delete the word 'in' between 'educated' and 'how' A: We are very sorry, but all three co-authors are unsure what is meant by the reviewer. The sentence is therefore kept unchanged and is as follows:" People with diabetes are educated in how to manage a consistent carbohydrate intake with respect to time and amount, which foods are rich in carbohydrates, how to read food labels and estimate carbohydrate portion sizes accurately."
Typos and errors with grammar and syntax: R1: Page 5, line 25. 'ACC is targeted the patient' should read 'ACC is targeted at the patient' A: Done
Typos and errors with grammar and syntax: R1: Page 6, line 4. 'Particularly' should be replaced with 'In particular' A: Done
Typos and errors with grammar and syntax: R1: Page 6, line 4. The word 'high' is redundant, this should phrase should read 'energy-dense foods' A: Done
Typos and errors with grammar and syntax: R1: Page 6, line 14. Delete 'the' between 'managing' and 'diet' and change 'diet' to 'dietary intake' A: Done
Typos and errors with grammar and syntax: R1: Page 7, line 36. Delete 'of' between 'use' and 'an' A: Done
Typos and errors with grammar and syntax: R1: Page 7, line 37. Delete 'the' between 'affecting' and 'dietary' A: Done
Typos and errors with grammar and syntax: R1: Page 8, line 8. Insert 'will' between 'participants' and 'receive' A: Done

Reviewer: 2

Reviewer Name: David Kerr

Institution and Country: Sansum Diabetes Research Institute, United States Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below This study is important and the researchers are from an organization with a long and impressive track record in clinical research. It is very likely to be successful.

R2: (a) there is no measure of food security which is an established risk factor in type 2 diabetes
A: To our best knowledge we have not stated this in the manuscript and therefore the question is a bit confusing for us. No changes have been conducted in the resubmitted version.

R2: (b) social determinants of health are recognized increasingly as important considerations in type 2 diabetes self-management. What steps are the authors taking to make sure that the materials are culturally appropriate and not lead to selection bias?

A: In Denmark we do not have that many people with type 2 diabetes from Ethnic minorities as in the US. We cannot avoid the risk of selection bias, since we can only include people who understanding and talk sufficiently in Danish to participate.

R2: (b) the inclusion criteria allow any glucose-lowering is allowed including insulin and other injectables and the dose and timing may change - this provides an opportunity for a health-economic comparison which would be of additional value. More importantly, there may be different magnitudes of effect with different sub-groups of pharmacological therapies which may make the findings difficult to interpret and perhaps the sample size calculation invalid. Would it make more sense to limit to therapies involved? Also, an additional statistical (i.e. actuarial) approach using "response" to interventions as the outcome and assessing the different weighted factors influencing the responses may add value.
A: We thank the reviewer for this very important point. We are aware of the impact of the different pharmacological therapies and we will adjust for this accordingly. The study is ongoing. It is not possible to change the inclusion criteria now and we are interested in examining the effect of BCC in a large population of patients with type 2 diabetes.

R2: (c) the HCP to participant interactions between groups is unbalanced in terms of time - this is a frequent challenge in this type of research. By this, I mean that the content of the proposed education program may be less important than the simple fact that a professional is interacting (i.e. caring) more. Is there some way to control for this with another interaction unrelated to food?
A: We share this concern which is why this is mentioned as a study limitation in this paper, however, our study is being carried out to examine if it has any value to offer a dietetic course in carbohydrate counting as an add-on to standard treatment for patients with type 2 diabetes and therefore cannot be conducted in any other way. In real life most patients only receive a mean of up to three visits with a dietitian, so we need to compare with this standard treatment to be able to examine an additional effect of adding and offering a 3-day course in carb. counting.

R2: (d) the intervention has a number of facets and if successful it may be difficult to tease out which ones are most important. Why not offer the app to this group without any instruction?
A: Because our aim isn't to examine the effect of an intervention compared to using an app without further instruction. Some participants may already know about and use this app before talking to a dietitian. Both participants in the control and the intervention group can use this app and may or may not be instructed in how to use it depending on needs and motivation.

R2: (e) why not use a wearable device to assess physical activity? I assume this is for economic reasons.
A: It is to avoid burdening the participants more the necessary. We already ask the participants to collect urine and record their dietary intake for 4 days and wear a CGM and measure their blood glucose levels four times daily for 6 days. Physical activity is not a primary endpoint. We only assess their level of physical of activity to check if the participants change their level of physical active during and after the intervention period. Additionally, it is more costly and time-consuming for the researches to use wearable devices to asses physical activity in the study.

VERSION 2 – REVIEW

REVIEWER	Pamela Dyson University of Oxford UK
REVIEW RETURNED	08-Oct-2019

GENERAL COMMENTS	There are a couple of comments which I made in my first reveiw and which do not seem to have been addressed, please see detals below 1. The power calculation states that a difference in HbA1c of 3.0 mmol/mol between the two groups is considered clinically meaningful. Although this is a subject of much controversy, most authorities agree that a difference of at least 5.0 mmol/mol is required in order to be clinically meaningful. The authors should justify why they have chosen 3.0 mmol/mol and provide supporting references. 2. Secondary outcomes: Clinical parameters include % time spent in range (3.9-10.0 mmol/l). An explanation of these limits is required, as is >10 mmol/l representing hyperglycaemia

REVIEWER	David Kerr Sansum Diabetes Research Institute, California, USA
REVIEW RETURNED	03-Oct-2019

GENERAL COMMENTS	Thank you for asking me to review the revised submission. I must admit to being surprised at some of the responses to my comments. Specifically: The lack of appreciation that in any study of nutrition it may be useful to assess food security (there are standard measures to do this) as this is recognized increasingly to be an important factor in the self-management of diabetes. I cannot believe that there are no food insecure individuals in Denmark! The statement " we do not have that many people with type 2 diabetes from Ethnic minorities as in the US" is somewhat surprising given the published Danish data showing that compared with native-born Danes, the incidence of diabetes is about 2.5 times higher among migrants from Africa, Asia, and the Middle East, and these migrant groups also showed significantly higher prevalence (https://www.ncbi.nlm.nih.gov/pubmed/27750090). The reason this matters is because the authors state in the section Strengths and Limitations that "the results obtained have applicability beyond Denmark and has the potential to be included in the recommendations in future T2D guidelines". I would suggest revising this statement. I am still concerned about the potential for bias based on the fact that there may be different glucose-lowering treatments in one arm especially for injectable therapies and I note with interest that the study is "ongoing". I would suggest the authors consider how they plan to deal with this. Physical activity is important in terms of the potential to impact glucose control and I would suggest adding an additional assessment at 3 months.

	On re-reading I do think the authors need to justify the value of 3 mmol/mol on which they have based the power calculation without reference to other work.

VERSION 2 – AUTHOR RESPONSE

Reviewers' Comments to Author:

Minor comments:

Reviewer: 1

Reviewer Name: Pamela Dyson

Institution and Country: University of Oxford, UK Please state any competing interests or state 'None declared': None declared

R1: The power calculation states that a difference in HbA1c of 3.0 mmol/mol between the two groups is considered clinically meaningful. Although this is a subject of much controversy, most authorities agree that a difference of at least 5.0 mmol/mol is required in order to be clinically meaningful. The authors should justify why they have chosen 3.0 mmol/mol and provide supporting references.

A: Our power calculations are based on the conservative assumption that we expect to see a smaller difference in HbA1c means of 3 mmol/mol between the two groups. If we had based our power calculations on detecting a 5 mmol/mol difference in HbA1c between the two study groups, we would only need to include a total of 84 participants allowing for a dropout of 30% compared with a total 226 participants, which we are including in our study to detect a difference of 3 mmol/mol (according to table 1). Thus, we have enough power to detect a difference of 5 mmol/mol between the two groups. Secondly, we have several important secondary outcomes and we may have enough power to detect differences in means between the two groups in secondary outcomes by including this large a sample size. We have deleted the wording "clinically meaningful change" from the manuscript.

Table 1. Sample size calculation

Difference in HbA1c, mmol/mol (SD 7.0)	Difference in MAGE, mmol/l (SD 0.7)	Net no of participants in the study (BCC vs control)	No of participants given a dropout rate of 30%
3.0	0.30	174	226
5.0	0.50	64	84

R1: Secondary outcomes: Clinical parameters include % time spent in range (3.9-10.0 mmol/l). An explanation of these limits is required, as is >10 mmol/l representing hyperglycaemia

A: Time in range (TIR) (3.9-10.0 mmol/L) and hyperglycaemia (>10 mmol/L) has been defined as cut-off points by a large expert group in the International consensus on use of continuous glucose monitoring (Diabetes Care 2017;40:1631-40). A TIR (3.9-10.0 mmol/l) of 70 % corresponds to a HbA1c level of 53 mmol/mol which is why these are clinically meaningful cut-off points.

Reviewer: 2

Reviewer Name: David Kerr

Institution and Country: Sansum Diabetes Research Institute, United States Please state any competing interests or state ‘None declared’: None declared

R2: The lack of appreciation that in any study of nutrition it may be useful to assess food security (there are standard measures to do this) as this is recognized increasingly to be an important factor in the self-management of diabetes. I cannot believe that there are no food insecure individuals in Denmark! A: We are indeed sorry for this misunderstanding in relation to the reviewer’s question concerning food insecurity. We of course acknowledge that food insecurity is an important factor in relation to diabetes self-management. The diet-related quality of life questionnaire used in our study includes questions concerning whether the recommended diet is a financial or social burden for the individuals. The results on food insecurity will be an important part of the participants’ self-assessment of adherence and dietary self-management in the study. We thank the reviewer for pointing this out.
R2: The statement "we do not have that many people with type 2 diabetes from Ethnic minorities as in the US" is somewhat surprising given the published Danish data showing that compared with native-born Danes, the incidence of diabetes is about 2.5 times higher among migrants from Africa, Asia, and the Middle East, and these migrant groups also showed significantly higher prevalence (https://www.ncbi.nlm.nih.gov/pubmed/27750090). The reason this matters is because the authors state in the section Strengths and Limitations that "the results obtained have applicability beyond Denmark and has the potential to be included in the recommendations in future T2D guidelines". I would suggest revising this statement. A: We acknowledge that the incidence of diabetes in the Danish population is about 2.5 times higher among migrants from Africa, Asia, and the Middle East and that the prevalence of these migrant groups are higher. However, the number of Ethnic minorities in Denmark is rather low why the challenge numerically is limited. Although, we do indeed agree with the reviewer that applicability of our study results primarily will be to a Caucasian population which we have now added at a limitation in the discussion related to the applicability for future T2D guidelines.
R2: I am still concerned about the potential for bias based on the fact that there may be different glucose-lowering treatments in one arm especially for injectable therapies and I note with interest that the study is "ongoing". I would suggest the authors consider how they plan to deal with this. A: Once again, we thank the reviewer for this very important point. We are aware of the potential for bias, but we are recording all changes in glucose-lowering treatments for each study participant during the study period (at baseline, 6 months, and 12 months) and we will examine the possible impact of any changes in glucose-lowering by performing subgroup analyses according to glucose-lowering medication. This is also described in the section Statistical methods in the manuscript: “Heterogeneity in responsiveness to the interventions will be tested by dividing each intervention group into smaller groups based on data distribution (medians) or clinically meaningful cut-points”.
R2: Physical activity is important in terms of the potential to impact glucose control and I would suggest adding an additional assessment at 3 months. A: This is indeed a relevant point and we acknowledge the relationship between any change (increase/decrease) in the levels of physical activity and the potential impact on metabolic control.

However, according to our study design and to minimize patient burden it is not possible to assess physical activity after 3 months. We assess levels of physical activity at baseline, after 6 months, and after 12 months. HbA1c reflects the last 3 months and MAGE the last six days; in concordance with the IPAQ questionnaire which reflects the level of physical activity within the last seven days. Finally, since physical activity is not a part of our intervention, we trust that any major change in levels of physical activity will be evenly distributed between the groups.

R2: On re-reading I do think the authors need to justify the value of 3 mmol/mol on which they have based the power calculation without reference to other work.

A: Our power calculations are based on the conservative assumption that we expect to see a smaller difference in HbA1c means of 3 mmol/mol between the two groups. If we had based our power calculations on detecting a 5 mmol/mol difference in HbA1c between the two study groups, we would only need to include a total of 84 participants allowing for a dropout of 30% compared with a total 226 participants, which we are including in our study to detect a difference of 3 mmol/mol (according to table 1). Thus, we have enough power to detect a difference of 5 mmol/mol between the two groups. Secondly, we have several important secondary outcomes and we may have enough power to detect differences in means between the two groups in secondary outcomes by including this large a sample size. We have deleted the wording “clinically meaningful change” from the manuscript.

Table 1. Sample size calculation

Difference in HbA1c, mmol/mol (SD 7.0)	Difference in MAGE, mmol/l (SD 0.7)	Net no of participants in the study (BCC vs control)	No of participants given a dropout rate of 30%
3.0	0.30	174	226
5.0	0.50	64	84