Author's response to reviews

Title: Autonomous exercise game use improves metabolic control and quality of life in type 2 diabetes patients - a randomized controlled trial

Authors:

   Kerstin Kempf (kerstin.kempf@wdgz.de)
   Stephan Martin (martin@wdgz.de)

Version: 2 Date: 29 August 2013

Author's response to reviews: see over
1. I am very impressed with the significant HbA1c improvement shown in this apparently "well-controlled" group, given their extremely low baseline levels. You may want to emphasize this in the DISCUSSION section.

Reply to 1. Thank you for that comment. We emphasized this point now in the discussion on page 8, lines 206-212: "A meta-analysis investigating the efficacy of pharmacological therapies demonstrated that the amount of HbA1c reduction essentially depends on baseline HbA1c. [12] HbA1c reduction seems to be greater the higher the baseline levels had been previously. With a mean baseline HbA1c of 7.0-7.9%, therapies with oral antidiabetic medication achieved an HbA1c reduction of 0.1%, while in our study the baseline HbA1c in the intervention group was 7.1 ± 1.3%, and a reduction of 0.3% was achieved."

2. I suggest that you move all description of sample size (page 3, METHODS/Study Subjects) starting at “Sample size had been determined ....” To “were planned to be recruited” to the Statistical Analysis section. I suggest that you put the first four sentences dealing with patient disposition (RESULTS) before the heading “Significant improvement ...”.

Reply to 2. According to the referees’ suggestion we moved the sample size calculation to the statistics section (page 4, lines 89ff) and the description of baseline parameters and completers before the heading (page 5, lines 109-113).

3. I suggest that you revise the RESULTS section subheadings to be neutral vs. declarative. For example: “Changes in glucometabolic control and weight” “Changes in physical activity” “Changes in diabetes-dependent …” “Changes in depression” The actual results in each of these sections tell the story.


4. I think once you have finalized your manuscript for resubmission, it would be helpful to have a native-English language speaker review and copyedit.

Reply to 4. The manuscript now had been proofread by a native speaker.
Reply to referee 2

1. Please state the exact primary outcome in your hypothesis. According to the ClinicalTrials.gov registration this is HbA1c. Other outcomes should be secondary outcomes.

Reply to 1. According to the referees’ comment we stated in the abstract in lines 3-7: “Therefore, in a randomized-controlled trial we tested the hypothesis that the autonomous use of the interactive exercise game Wii Fit Plus over a period of 12 weeks improves metabolic control, with HbA1c reduction as the primary outcome, and weight loss, reduction of cardiometabolic risk factors, physical activity and quality of life (secondary outcomes) in T2DM patients.”, in the background section on page 3, lines 45-48: “Therefore, in a randomized-controlled trial we investigated the hypothesis that autonomous use of the interactive exercise game Wii Fit Plus over a period of 12 weeks is able to improve HbA1c (primary outcome) as well as weight, cardiometabolic risk factors, physical activity and quality of life (secondary outcome) in T2DM patients.”, and in the methods section on page 4, lines 81-83: “Primary endpoint was reduction of HbA1c after 12 weeks of intervention; secondary endpoints were differences in weight, cardiometabolic risk factors, physical activity and quality of life measurements.”

2. Please state the differences between intervention (Ix) and control (Cx) instead the within group changes.

Reply to 2. We inserted a new Fig.1 and stated the comparison of both, the between group as well as the within group differences on page 5, lines 115ff: “Changes in HbA1c. The complete case analysis demonstrated that, during the 12 weeks of intervention, patients in the intervention group were able to significantly reduce their HbA1c from 7.1 ± 1.3% to 6.8 ± 1.0%; p=0.0002; Fig.2A) while no significant reduction (from 6.8 ± 0.9% to 6.7 ± 0.7%) was seen in the control group during the waiting phase. In detail, the reduction in HbA1c differed significantly from zero in the intervention group only (-0.3 ± 1.1%; Fig.2B), but not in the control group (-0.1 ± 0.5%). HbA1c values were not significantly different between groups at baseline or after the 12 week study period. Nevertheless, the within group differences reached statistical significance (p=0.01), and these differences remained statistically significant after adjustment to baseline HbA1c, age, sex, and weight difference (Table 2). The percentage of participants who reached the treatment goal of HbA1c <7.0% increased in the intervention group from 56 to 65% (+9%) and decreased in the control group from 68 to 67% (-1%). When the control group finished the waiting phase and also started intervention, a significant reduction of -0.2 ± 0.5% of HbA1c (p=0.003) was observed, leading to 78% reaching the treatment goal. If missing values were filled in according to the ‘Last-observation-carried-forward’ principle results stayed the same with a significant reduction of HbA1c of -0.24 ± 0.96% (p=0.0002) in the
intervention vs. -0.08 ± 0.47% in the control group. The results even remain robust after Bonferroni correction for multiple testing.

3. Please state that these are intention to treat analyses.

Reply to 3. The results reported in the original manuscript were NO intention to treat analyses but **complete case analyses**. The number of completers had been stated in the results section as well as in the figure legends. The Last-observation-carried forward method so far only had been used for missing values in secondary outcomes from completers. HbA1c values have been available from all participants who completed the 12 week study. To make this clearer we added this information in the abstract on page 2, line 15 as well as in the methods section on page 4, lines 96-98: “Complete case analyses were performed. Missing values from completers concerning secondary outcomes were substituted by ‘last-observation-carried-forward’ principle.” and page 5, line 115ff: “The complete case analysis demonstrated that ...”.

4. Please conclude on the primary outcome.

Reply to 4. We concluded about HbA1c reduction reached on page 7-8, lines 196-212: “Also a recent meta-analysis of community-based physical activity for adults with type 2 diabetes revealed a significant lowering of HbA1c levels by -0.32% [95% CI -0.65, 0.01].[10] However, it needs to be emphasized that such interventions required an enormous amount of personal and economical effort. In contrast, in our approach, mentoring and costs were low, just mailing the *Wii* console with balance board, and the exercise game *Wii Fit Plus*. For the point of view of the participants, the use of the game was a low-threshold proposal for lifestyle intervention. In detail, we did not test the effect of physical activity on glucometabolic control, but just the effect of providing a device for interactive exercise in order to encourage self-motivation in the participants. Nevertheless, a comparable HbA1c reduction of 0.3% was reached and the percentage of participants reaching the HbA1c goal of <7.0% [11] increased by 9%. A meta-analysis investigating the efficacy of pharmacological therapies demonstrated that the amount of HbA1c reduction essentially depends on baseline HbA1c. [12] HbA1c reduction seems to be greater the higher the baseline levels had been previously. With a mean baseline HbA1c of 7.0-7.9%, therapies with oral antidiabetic medication achieved an HbA1c reduction of 0.1%, while in our study the baseline HbA1c in the intervention group was 7.1 ± 1.3%, and a reduction of 0.3% was achieved.”

5. The introduction is relatively brief. A mentioning of key studies of PA interventions (+ drawbacks) of the interventions could improve this section markedly. There are some community based studies on PA and not lifestyle intervention (which assumable also contains a diet intervention).
Reply to 5. We did not refer to physical activity studies because we did not test the effect of physical activity on glucometabolic control but just the effect of providing a device for interactive exercise playing in order to encourage self-motivation of participants. Nevertheless, we compared our results with those of a recent Meta-analysis of community-based physical activity in type 2 diabetes patients on page 7, lines 196-206: “Also a recent meta-analysis of community-based physical activity for adults with type 2 diabetes revealed a significant lowering of HbA1c levels by -0.32% [95% CI -0.65, 0.01].[10] However, it needs to be emphasized that such interventions required an enormous amount of personal and economical effort. In contrast, in our approach, mentoring and costs were low, just mailing the
\textit{Wii} console with balance board, and the exercise game \textit{Wii Fit Plus}. For the point of view of the participants, the use of the game was a low-threshold proposal for lifestyle intervention. In detail, we did not test the effect of physical activity on glucometabolic control, but just the effect of providing a device for interactive exercise in order to encourage self-motivation in the participants. Nevertheless, a comparable HbA1c reduction of 0.3% was reached and the percentage of participants reaching the HbA1c goal of <7.0% [11] increased by 9%.”

6. \textit{Again please clearly state primary, secondary and tertiary outcomes in this section (with aims and hypothesis).}

Reply to 6. Please see reply to 1.

7. \textit{What does the routine care consist of?}

Reply to 7. We described the meaning of “routine care” on page 3, lines 62-64: “(i.e. quarterly visits to the attending physician according to the DMP T2DM with check of HbA1c, weight, BMI, blood pressure, and blood lipids and therapy adjustment if necessary).”

8. \textit{Where was the blood samples drawn (veins or..) and how where the parameters assessed (assay or calculation)?}

Reply to 8. Handling of blood collection now is described on page 4, line 69-72: “Venous blood was collected after an overnight fast of 8 h by inserting an intravenous cannula into the forearm vein, and laboratory parameters (i.e. HbA1c, levels of fasting blood glucose (FBG), triglycerides, total-, HDL- and LDL cholesterol) were analyzed at local laboratories.”

9. \textit{Which procedures where used to derive BP and anthropometrics?}
Reply to 9. These methods are now described on page 3-4, lines 67-69: “Body weight was measured in light clothing to the closest 0.1 kg and height to the closest 0.5 cm. Blood pressure was measured after a five-minute rest in a sitting position at both arms.”

10. Please move the statement regarding primary and secondary outcomes to the introduction.

Reply to 10. We have done this. Please see reply to 1.

11. LOCF is a simplified method to assess outcomes in drop-outs and it increases the risk of over reporting efficacy as TD2 often progresses over time (meaning that a 0 effect might be overoptimistic). At least perform a sensitivity analysis were the worst and the best observations (in terms of change in Hba1c) to check the robustness of your findings. Another possibility is to impute (e.g. using boot strapping or multiple imputation maneuvers) the missing values. Before doing so, I advise you to consult a statistician to check assumptions about imputing values.

Reply to 11. As suggested by the referee we consulted a statistician but he was wondering about the complexity of the recommended analysis concerning missing values. Maybe, there might have been a misunderstanding. The primary endpoint of our analysis had been reduction of HbA1c after 12 weeks compared between the intervention and the control group. HbA1c reduction was only analyzed for those participants who completed the 12 week study, i.e. 93 out of 120 in the intervention group and 83 out of 100 in the control group. In the methods section we described that missing values had been filled in by the Last-observation-carried-forward (LOCF) principle but this was only used for secondary endpoints because HbA1c had been available from all 93 and 83 completers after 12 weeks.

Even if we would use LOCF for the drop outs (not described in the original manuscript), the results would stay the same. We added this information on page 5, lines 128-131: “If missing values were filled in according to the ‘Last-observation-carried-forward’ principle results stayed the same with a significant reduction of HbA1c of -0.24 ± 0.96% (p=0.0002) in the intervention group vs. -0.08 ± 0.47% in the control group.”

As described by Hollis and Campbell (BMJ 1999; 319: 670) LOCF is a commonly used method for handling of missing values if there are not sufficient data to allow good estimation. By assuming that all missing responses were constant, the estimation would rather be conservative because any effect of intervention or even study effects were negotiated. Extreme case analysis were reported to be unlikely to yield a conclusive answer in practice (Meyer K, Windeler J, 19th International Society for Clinical Biostatistics meeting, Dundee 1998)

12. Please include a flowchart (of participants).
Reply to 12. A flowchart according to the CONSORT guidelines had been included in the original manuscript. We now added the flowchart as new Fig.1.

13. You state that no statistical differences (in the results section) were observed between Ix and Cx. However, there is a 0.3% difference in Hba1c (lower in Ix) and the variation is larger. This could potentially affect the differences between groups.

Therefore I suggest that you analyze the changes (instead of the follow up differences) between groups and adjust the analysis for the baseline values.

Reply to 13. We used the Mann-Whitney test to compare the baseline HbA1c values between groups and found no statistically significant difference ($p=0.1690$). Mean + SD in the intervention group had been 7.1 ± 1.3% ranging from 5.5 to 12.8% and 6.8 ± 0.9% in the control group with a range of 5.5-10.5%. This broad range caused the standard deviations. Nevertheless, the differences of baseline HbA1c were not statistically significant.

Unfortunately, I did not fully understand the intention of the second part of your comment. What do you mean by “changes between groups”? We analyzed the differences in HbA1c within groups, i.e. ‘values at the end of intervention’ minus ‘baseline values’ for each patient. These differences were -0.3 ± 1.1% HbA1c for the intervention group and -0.1 ± 0.5% for the control group. Additionally, we now performed a regression analysis in order to demonstrate the effect of intervention on the difference in HbA1c and adjusted for baseline HbA1c values and other parameters. The results are shown in the new Table 2 and described in the results section on page 5, lines 122-124: “... these difference remained statistically significant after adjustment to baseline HbA1c, age, sex, and weight difference (Table 2).”

14. Due to the large amount of outcomes I suggest that the p-values are corrected for multiple testing.

Reply to 14. We added Bonferroni correction for multiple testing and described this in the methods section on page 5, lines 103-104: “The Bonferroni correction was used for multiple testing ($n=15$) resulting in $p=0.05/15=0.033$. “ and for the main outcome in the results section on page 5-6, line 135-137: “In the intervention group also FBG levels improved significantly although the reduction no longer remained statistically significant after correction for multiple testing.”

15. It is unclear whether the Ix-waiting group was included in the primary analysis. If so the analysis should be a repeated measures analysis.

Reply to 15. As defined the primary analysis concerned the difference in HbA1c reduction during the first 12 weeks of study, i.e. exercise game intervention for the intervention group and waiting phase for the control group. After these 12 weeks the control group also were
sent the equipment for exercise gaming and we documented their health parameters after additional 12 weeks. These results are descriptive but confirm the effects that already had been observed in the intervention group. We now performed Bonferroni’s Multiple Comparison Test for the control group and described it in the methods section on page 4, lines 100-101: “... Bonferroni’s Multiple Comparison Test for analyses of repeated measurements within groups ...” but results remained significant.

16. Please clearly state whether or not an effort of allocation concealment was made. If an effort was made, how was the maneuver performed?

Reply to 16. We now described the details of allocation concealment on page 3, lines 52-59: „Eligible patients … were randomized according to an electronically generated randomization list (created by trial statistician) into parallel groups (assigned by study nurse). In detail, each participant was assigned a serial study ID. For each ID there was a closed envelope with the group assignment. Both, the participants and the study nurse were blinded for sequence of allocation concealment. “

17. When was the baseline measurements performed (before or after randomization)?

Reply to 17. The baseline measurement had been performed before randomization.

18. Please justify why three different scores of depression were used. It seems to me (from figure 4) that all outcomes are depression (yes/no).

Reply to 18. The questionnaire ‘Problem Areas in Diabetes’ (PAID) was developed for determination of diabetes-dependent impairment; with the SF-12 questionnaire physical and mental health can be assessed; the WHO-5 questionnaire is thought for estimation of subjective wellbeing, and the ‘General Depression Scale’ (CES-D) of which we used the German version ADS-L (Allgemeine Depressionsskala) is for assessment of quality of life. We described these details on page 4, lines 77-81. But you are right; all questionnaires include estimation for the presence of depression.

19. First header in results should be entitled “Sample characteristics”

Reply to 19. According to the suggestion of referee 1 we moved the description of baseline parameters and completers before the heading end entitled it ‘Sample characteristics’ according to referee’s 2 suggestion (page 5, line 109ff).
20. There is a lot of focus on the within-group differences. I urge the authors to describe the between-group changes (as stated above) as we are looking at alternative treatment of T2D.

Reply to 20. Please see reply to 2.

21. In regard to changes in medication use; who prescribed the medication and changes herein? Please state this in the methods section.

Reply to 21. We now described this on page 3, lines 62-64: “(i.e. quarterly visits to the attending physician according to the DMP T2DM with check of HbA1c, weight, BMI, blood pressure, and blood lipids and therapy adjustment if necessary)”. 

22. The results section is hard to read. I suggest that the numbers are presented in a table instead (for the within- and between-group differences) for all the secondary outcomes and leave in a figure of glycemic control (Hb1ac and potentially glucose). This would increase the flow of this section. In doing so, figure 2, 3 and 4 would be redundant. This applies to the following sections as well.

Reply to 22. According to the referees’ suggestion only the data for HbA1c were presented in the new Fig.2 (the new Fig.1 includes the flowchart) and all other parameters from the original Fig.1 and Fig.2 were included into Table 1. Nevertheless, we felt, that we would lose information if the quality of life data would only be presented as numbers in the table. In our opinion, it would be easier for the reader if those secondary outcomes were presented as figures because they also include the descriptive results of the control group during intervention phase.

23. Were the changes in PA (other than Wii usage) formally tested. Please state the stats. In this regard, do the authors have any information on the Wii usage? It would be nice to include a measure of compliance or a per protocol analysis.

Reply to 23. Physical activity has been estimated by questionnaires. We added this information in the methods section on page 4, lines 72-76: “Self-reported physical activity was assessed according to a 6-point questionnaire asking for intensity of physical activity during working and leisure time, frequency of physical activity in summer and winter, duration of walking and bicycling time per day. Single values were summed up with the maximum reachable value set to 100%.”

We agree with the referee that a measurement of frequency intensity and compliance would have been very helpful. For this reason participants were sent a sheet where they had to fill the values for exercise duration and energy expenditure the Wii systems offers after the end of each training session. Unfortunately, this sheet had not filled in from the majority of participants so that a meaningful analysis had not been possible.
24. “Thus, an exercise game seems to be reasonable and adequate to motivate T2DM patients to increase their physical activity leading to improved glucometabolic control and quality of life.” I do not feel that the authors can demonstrate this in the current design. PA is subjectively measured and the exposure (Wii) or PA for that instance is not recorded through the trial. The introduction of the Wii leads to an improvement (assuming that the trial stats are correct), but the data does not say anything about this affect the outcome through increased PA. It could as well affect a generally improved lifestyle (such a diet behavior).

Reply to 24. We agree with the referee that the sentence might be too provocative and therefore changed it into “Therefore, exercise games may potentially be used in a home setting as a tool to reduce sedentary behavior in T2DM.” (Page 7, lines 187-188)

25. I urge the authors to compare their results with the larger exercise trial to elaborate on the effects of their trial. Furthermore, there are multiple community based PA/exercise trials they need to address.

Reply to 25. Please see reply to 5.

26. What is the clinical meaning of the effects sizes (are the changes clinically meaningful)? Did they actually get the patients within the therapeutic target?

Reply to 26. The percentage of patients reaching the therapeutic target increased. These results are described on page 5, lines 124-126: “The percentage of participants who reached the treatment goal of HbA1c <7.0% increased in the intervention group from 56 to 65% (+9%) and decreased in the control group from 68 to 67% (-1%). When the control group finished waiting phase and also started intervention significant reduction of -0.2 ± 0.5% of HbA1c (p=0.003) was observed leading to 78% reaching the treatment goal.”

An HbA1c reduction of 0.3% is clinically relevant and we discussed this point on page 7-8, lines 196-206: “Also a recent meta-analysis of community-based physical activity for adults with type 2 diabetes revealed significant lowering of HbA1c levels by -0.32% [95% CI -0.65, 0.01],[10] … Nevertheless, a comparable HbA1c reduction of 0.3% was reached and the percentage of participants reaching the HbA1c goal of <7.0% [11] increased by 9%.”

27. “Since a key obstacle to physical activity is lack of motivation exercise games motivate players to exercise, and some games take advantage of group dynamics to motivate players’ duration of exercise.” - Please provide a reference.

Reply to 27. We rewrote the sentence into a more positive statement and provided a reference. It now reads on page 8, lines 226-229: “Since key motivators to physical activity were weight management, feelings of physical and mental well being as well as social relationships
associated with physical activity, exercise games should help to motivate players to exercise, and could take advantage of group dynamics to motivate players in terms of the duration of the exercise period.[17]"

28. *Furthermore, there is something going on grammatically??*

Reply to 28. The manuscript now had been proofread by a native speaker.

29. *Please check if the drop-out is differential and state so.*

Reply to 29. We analyzed that point and stated the result on page 5, line 113: “The dropout rate did not differ significantly between both groups.”

30. *In limitations, some mentioning of the generalizability of your findings is warranted. Is this group different from the general group of T2D patients if so how.*

Reply to 30. We discussed this point on page 9-10, lines 269-275: “It might be speculated that the patients participating in this study might have been more strongly motivated than the general T2DM population. Nevertheless, that might be true for all patients participating in clinical studies. Perhaps those patients who had heard about the Wii and exercise games from their children or grandchildren might have been more willing to participate, but generally, for all T2DM patients who are physically able, exercise games might offer an alternative form of home exercise.”

31. *Furthermore, I do not agree on the use of proof of principle. This is a pragmatic/effectiveness trial rather than an efficacy trial.*

Reply to 31. We deleted “proof-of-principle” throughout the manuscript.