Author's response to reviews

Title: The fitness consultation. A way to increase muscle strength and VO2max of patients with type 2 diabetes in general practice: an 18-month intervention study

Authors:

Henning Lohmann (Lohmann@dadlnet.dk)
Niels de Fine Olivarius (olivarius@sund.ku.dk)
Volkert Siersma (siersma@sund.ku.dk)

Version: 2 Date: 8 June 2010

Author's response to reviews: see over
Submission of a revised version of the manuscript “The fitness consultation. A way to increase muscle strength and \( \text{VO}_{2\text{max}} \) of patients with type 2 diabetes in general practice: an 18-month intervention study”

Thank you for inviting us to resubmit our manuscript (now with a new title) to BMC Family Practice. The two thorough reviewers have made it clear that several aspects of the methods need to be explained in more detail, which we have done. We have also clarified these aspects of the methodology in the manuscript. The reviewers have been so helpful that we thought it necessary to thank the reviewers in the Acknowledgements, if you and the reviewers find this appropriate.

Below is our answer to the critique (in New Times Roman and blue). Excerpts from the manuscript are in italics.

Please note: A comparison with “Word track changes” of the first submission with this last submission is uploaded as an additional material file.

We are looking forward to your reply.

Sincerely yours,

Henning Lohmann

Reviewer’s report (1)

Title: Motivational Interviewing and Fitness Tests used to increase Physical Fitness in Patients with Type 2 Diabetes in General Practice: an 18-month intervention study

Version: 1 Date: 25 November 2009

Reviewer: Susan Smith

Reviewer’s report:

Major compulsory revisions

Conclusions need to be more measured. The design does not support a statement that this intervention should be incorporated in routine service delivery.

We agree with the reviewer. Because the study is not an RCT, we have been very cautious about drawing conclusions on the basis of our findings. Our conclusion now states: Our results indicate that physical testing combined with motivational interviewing can be done in a primary health care setting. Here a fitness consultation tailored to the individual patient, his/her comorbidities and the conditions in the local area can be incorporated into the diabetes programme to improve patients’ muscle strength and cardio-respiratory fitness. The extra workload caused by fitness consultations in primary care could for instance be carried by practice nurses or physiotherapists who have received training in motivational interviewing and physical testing. Randomised trials are needed to confirm our findings and to optimise recommendations for the content and the frequency of the fitness consultations [1].

Needs to be tested in RCT with controls.

We agree, as we state in the conclusion cited above.

Intervention involved a high degree of clinical contact.

The usual diabetes control was only modestly prolonged. We did not measure exactly the extra time spent at the doctor’s consultation but our estimate is 10 minutes. In Results we have added the following sentence:

The extra time consumption attributed to the extension of the usual diabetes control to become a fitness consultation was estimated to 10 min.

The participation rate of patients was quite low and should be presented in terms of the acceptability of such an intervention for patients.

The idea of physical testing was completely new to the patients and this was probably one of the major reasons for non-participation. We have no information about reasons for non-participation. See our answer to reviewer 2, point 31. The patients had to sign a written consent which listed all the many uncommon side-effects or complications of physical testing, and this procedure in itself (especially the bicycle ergometer ride) caused uneasiness in some of the patients and discouraged them from participation. In Methods we state, as we did in the former manuscript, that
There was no statistically significant difference between the 227 non-responders and the 127 participants with regard to age (67.5/67.0 years, p=0.77), gender (women 42.3/42.5%, p=0.96) and HbA1c (6.8/6.9%, p=0.30) before the start of intervention.

In view of the median baseline age of 67 years it could even be considered positive that so many patients volunteered for this new treatment offer.

In the Discussion section we now write:

*The study was done in the setting of a primary health care unit using primary care practitioners to carry out the intervention. The whole city was aware of the project, which made it impossible to do a randomised controlled trial. Contrary to most other studies, patients with cardiovascular or musculoskeletal disease were not excluded [2]. The idea of physical testing was completely new to the patients and this in combination with the relatively high median age was probably the major reasons for the low participation rate.*

**Methods:** The setting and system of diabetes care needs to be described.

We have added a sentence about this subject to the Methods section (Study design and participants):

*In Denmark, routine care of patients with type 2 diabetes is usually given by GPs and practice nurses in primary care units in a national structured diabetes programme recommending control every 3 months.*

How were GPs selected and what proportion of those approached agreed to participate.

The 127 patients with type 2 diabetes were included from a Danish general practice with 6 GPs (among them HL) and 9521 patients. The first author of this paper (HL) is one of the GPs, and he applied for and received financial support from the National Board of Health to develop methods to increase physical activity in patients with type 2 diabetes. In the Methods section we now write:

*The study was designed as an uncontrolled study in a primary health care unit with six GPs supplying health services to approx. half of the about 20,000 citizens in a town in Denmark. All health care providers in the unit participated.[...] Of 399 patients with known type 2 diabetes, 127 were included (Fig. 1). All were Caucasians. During the study, participants were considered lost to follow-up after non-response to one reminder. The reasons for exclusion or loss to follow-up appear from Fig. 1. The ethics committee of West Zealand approved the protocol. Patients gave written informed consent.*

**Results:** need to include details of the response rates etc and the information currently presented in the method section. Also refer to flow chart of participation rates in the text

We have now moved the following slightly edited paragraph from the Methods section to the Results section, added a sentence clarifying the response rate, and we now refer to Fig. 1 in this paragraph:

*Of 354 eligible patients, 127 (35.9%) participated in the study. There was no statistically significant difference between the 227 non-responders and the 127 participants with regard to age (67.5/67.0 years, p=0.77), gender (women 42.3/42.5%, p=0.96) and HbA1c (6.8/6.9%, p=0.30) before the start of intervention (Fig. 1).*

**Conclusions:** More specific recommendations for future research would be appropriate, given your experience conducting this study rather than the statement that more RCTs are needed.

We have added the following sentences to Conclusions:

*The extra workload caused by fitness consultations in primary care could for instance be carried by practice nurses or physiotherapists who have received training in motivational interviewing and physical testing.*

**Minor**

*Background: Grammer correction: A substantial proportion of patients with type 2 diabetes have low levels of physical fitness and do not engage in the recommended level of physical activity.*

*Done*

The hypothesis is difficult to understand and I suspect this is due to poor translation as it does not make sense as it is written.

This is true. Please read our answer below to a similar comment from reviewer 2, point 8. We have changed the wording as follows:
We assumed that by providing patients with knowledge of their own muscle strength and cardio-respiratory fitness, they would become aware of discrepancies between their current physical fitness and personal goals for future health, and this realisation could induce behavioural changes [3,4].

Methods: begins with reporting actual results relating to participants. This section should report the inclusion and exclusion criteria of those eligible to participate rather than actual details of those who finally did participate. We have moved the results relating to participants to the Results section. The inclusion and exclusion criteria are now all included in Fig. 1. This is why we have not repeated this information in the text.

Discussion: the summary should also report the outcomes that were unchanged. Before the study started, we chose 5 primary endpoints (Table 2, first row), and for these variables the p-values may be interpreted at the 5%-level. We also selected 8 secondary outcomes for which the p-values can be said to indicate trends (Table 3). Among our primary outcomes, improvements in 4 were statistically significant. We agree that the negative findings should be openly reported, and we have therefore included all the 13 outcomes in the summary at the beginning of the Discussion section:

In this 18-month uncontrolled intervention study, repeated fitness consultations including fitness testing and motivational interviewing resulted in increased muscle strength and VO$_{2\text{max}}$ and the lipid profile improved, while HbA$_1$c remained unchanged on a low level. Among the secondary outcome measures, waist circumference, BMI and fasting plasma glucose were unchanged, blood pressures increased slightly, whereas total cholesterol and LDL-cholesterol decreased.

Given the lack of change in weight, it is hard to know whether this intervention really has clinical significance for people with type 2 diabetes as there were no significant improvements in weight and HbA1c though the authors do acknowledge that the latter related to good HbA1c control at baseline. The clinical significance of increased fitness is not necessarily dependent on a weight loss. We discuss this briefly immediately prior to Conclusion in the following new sentences: The intervention had no impact on BMI and waist circumference. Nevertheless the intervention may have had a clinically significant effect on health as the inverse gradient between fitness and mortality in men with documented diabetes mentioned above is independent of BMI [5,6]. The improvement in glycaemic control following endurance and strength training may also be observed with unchanged BMI [7].

Level of interest: An article whose findings are important to those with closely related research interests
Quality of written English: Needs some language corrections before being Published
Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.
Declaration of competing interests: I declare that I have no competing interests

Reviewer's report (2)
Title: Motivational Interviewing and Fitness Tests used to increase Physical Fitness in Patients with Type 2 Diabetes in General Practice: an 18-month intervention study
Version: 1 Date: 18 February 2010
Reviewer: Paul Poirier
Reviewer's report:
TITLE
Motivational Interviewing and Fitness Tests used to increase Physical Fitness in Patients with Type 2 Diabetes in General Practice: an 18-month intervention study.
AIM OF THE STUDY
To investigate whether the introduction of motivational interviewing combined with fitness tests in the type 2 diabetes care program resulted in a change in cardio-respiratory fitness expressed by VO2max, muscle strength of upper and lower extremities, HbA1c and HDL-cholesterol.

FINDING
The authors concluded that fitness consultations combined with motivational interviewing resulted in increased muscle strength and VO2max, and the lipid profile improved.

GENERAL COMMENTS
The purpose of this study is relevant but we have identified major errors in the methodology. The study aims to look whether the intervention (motivational interviewing and fitness tests) resulted in a change in VO2max, muscle strength of upper and lower extremities, HbA1c and HDL-cholesterol but the research design does not include a control group. Medical practice could not be influenced by the results of this study because we don’t know if changes were related to the experimental intervention or other factors.

We acknowledge this important limitation (see above) which is discussed thoroughly in the manuscript.

The use of the term “motivational interviewing” (MI) is not appropriate.
We acknowledge that we cannot document that the consultations included all the principles from MI. The 6 GPs, however, had a 3-hour session with an introduction to MI led by an experienced MI-researcher. At every fitness consultation the GPs had to judge and note in the patient record where the patient was in his or her decision circle with its six stages: pre-contemplation, contemplation, preparation, action, maintenance, and relapse. We agree, however, that it is inappropriate to mention MI in the title of the paper as the focus of the paper is rather on the fitness tests, and we have changed the title to The fitness consultation: A way to increase muscle strength and VO2max of patients with type 2 diabetes in General Practice: an 18-month intervention study.

I suggest reviewing the structure of the manuscript. Furthermore, I would like to address some issues to the authors.

The manuscript has been rewritten to a great extent (enclosed please find a comparison of the old and the new manuscript)

SPECIFIC COMMENTS
Abstract section:
1. Could the authors define GP?
   Done.
2. Could the authors add the p-value for the increments in VO2max among patients without pain from function limitation?
   We have added this and the similar information about cardiovascular disease to the abstract and to Results: Patients without cardiovascular disease or pain from function limitation increased their VO2max by 6.3% (p<0.0001) and 8.3% (p<0.0001), respectively.

Background session
3. I suggest rewriting the first sentence of this section because the literature cited (Wei & al., Hu & al., Church & al. and Blair & al.) did not demonstrate the association between increment of cardiovascular fitness and decrement of mortality but demonstrates the association between level (low, moderate or high) of cardiovascular fitness and mortality.
   We have added this reference by S.N. Blair: “Changes in physical fitness and all-cause mortality. A prospective study of healthy and unhealthy men.” After this addition, we believe that the statement is documented: Increases in cardio-respiratory fitness, muscle strength and level of physical activity are associated with decreased mortality and protect against age-related disabilities [5,6,8-12].

4. The primary outcome of this study is change in VO2max, muscle strength, HbA1c and HDL-cholesterol. Could the authors add references regarding the association between those variables and physical activity (PA) behavior? Indeed, the aim of the intervention is to motivate people to increase and maintain their...
muscle strength and cardio-respiratory fitness by self-managed PA, but the authors did not present the scientific evidence that PA behavior is associated with improvement of the study’s variables (VO2max, muscle strength, HbA1c and HDL-cholesterol) among people with type 2 diabetes.

It is true that we failed to discuss this important point in the Background section. We have now added these sentences with relevant references:

Regular exercise in type 2 diabetic individuals may have a significant effect on VO2max and may result in decreased HbA1c [2,13]. Similarly, HDL-cholesterol increases with cardio-respiratory fitness [14,15]. It has also been demonstrated that progressive resistance training increases muscle strength in type 2 diabetic patients [16].

5. The authors justify the utilization of MI based on one study that concluded that positive attitude toward exercise and sense of control over it was associated with adherence to the exercise regimen. However, this explanation is not sufficient to justify the use of MI. Indeed, MI is based on a variety of theories and concepts such as the Transtheoric Model of Prochaska and Diclemente (1982), the Protection Motivation Theory of Rogers (1975), the Decisional Balance of Jamis and Man (1977), the Reactance Theory of Brehm (1981), the Self-Perception Theory of Bem (1967), the readiness to change (Rollnick, Mason and Butler, 1999) and the ambivalence exploration (Miller and Rollnick, 1991). Those theories and concepts contain much more behavioral determinants than “positive attitude toward the behavior” and “sense of control”. Initially, showing that other behavioral determinants used in MI are present within the study population would be necessary in order to explain the relevance of this technique. Furthermore, the literature possesses a range of researches regarding behavioral determinants conducted among people with type 2 diabetes. It would be more appropriate to refer to those studies.

It was not our intention to justify the use of MI with the reference concerning adherence and participation and we have now changed the sequence of the sentences to remove this misunderstanding.

In the Background section, we have added some references:

[...]It seems relevant to use motivational interviewing and include testing of muscle strength, cardio-respiratory fitness and exercise prescription in the motivational armamentarium of a diabetes care programme [3,17-21].

6. References are required for the sentence beginning by “These patients have many comorbidities…”

We have added a relevant reference [22].

and for the sentence beginning by “There are many laboratory studies with supervised exercise…”

We have added a relevant reference to a metaanalysis [2].

Furthermore, could the authors add information regarding the cardiovascular mortality among persons with type 2 diabetes?

We have changed the last sentence in the first paragraph of the background section as follows:

These patients have increased cardiovascular mortality [23] and many comorbidities [22], e. g. hypertension, cardiovascular disease and arthritis, which may preclude some physical activities or require evaluation by a physician before the activities can be undertaken.

7. The sentence: “Feedback was given in fitness consultations every three
months with simple measurements of muscle strength and cardio-respiratory fitness followed by motivational interview” seems to be more relevant in method section since this sentence explains the experimental intervention. This is correct. We have actually been able to delete this sentence from the manuscript since its content is already included in the Methods section.

8. The hypothesis proposed by authors is not related with dependents variables (eg. VO2max, muscle strength, HbA1c and HDL-cholesterol). Indeed, authors’ hypothesis is that patient’s knowledge of their own muscle strength and cardio-respiratory fitness (...) may induce behavioral changes. However, they did not evaluate patient’s knowledge of their own muscle strength and cardio-respiratory fitness neither the relation between their knowledge and behavioral changes.

It is quite correct that we were not precise enough in our use of common concepts. When we wrote “Our hypothesis was...”, we had no intention of presenting a hypothesis to be tested in the analyses. We have now changed the wording, removed the word hypothesis and added the concepts “over-all purpose” and “specific aim”:

The over-all purpose of the present study was to see whether general practitioners (GPs) and their staff could motivate people with type 2 diabetes to increase and maintain their muscle strength and cardio-respiratory fitness by self-managed physical activities during an 18-month intervention period. We assumed that by providing patients with knowledge of their own muscle strength and cardio-respiratory fitness, they would become aware of discrepancies between their current physical fitness and personal goals for future health, and this realisation could induce behavioural changes [3,4]. The specific aim of the study was to evaluate the impact of expanding the regular diabetes control consultation to include fitness tests and motivational interviewing. The primary outcomes that were measured were VO2max (maximal oxygen uptake, ml O2 kg\(^{-1}\) min\(^{-1}\)), muscle strength of upper and lower extremities, haemoglobin A1c (HbA1c), and HDL-cholesterol during the 18-month intervention period. Secondary outcomes were waist circumference, body mass index (BMI), systolic and diastolic blood pressure, fasting plasma glucose, total cholesterol, LDL-cholesterol, and triglycerides.

9. The last sentence of the background section could be written as: “The aim of this study was to evaluate the impact of the motivational interviewing and fitness tests in a change in...”

Done, as indicated above.

Methods section:

10. Uncontrolled study was used to evaluate the impact of the motivational interviewing and fitness tests in a change in VO2max, muscle strength, HbA1c and HDL-cholesterol. However, this study design is not appropriate to achieve the aim of this study since the absence of a control group does not allow knowing if the changes were related to the intervention or to other factors. Could the authors justify why they chose this study design?

We have added the following sentence to the beginning of the Discussion section:

The study was done in the setting of a primary health care unit using primary care practitioners to carry out the intervention. The whole city was aware of the project, which made it impossible to do a randomised controlled trial.

11. Could the authors add information regarding the inclusion and exclusion criteria, the recruitment and the consent? Two inclusion criteria are listed in table 1, but not in the text.

From the practice list of patients we included all patients who were diagnosed with type 2 diabetes according to national guidelines, i.e. WHO criteria (n=399). We excluded patients who were unable to perform at least one of the physical tests (n=45). These were mainly patients who were confined to bed or unable to understand instructions. All the remaining patients (n=354) received a letter inviting them to participate. All those invited were offered verbal information from doctors and staff. To become included patients had to sign a written consent which listed the theoretical risks of physical testing and this procedure made some of the patients uneasy and discouraged them from participation. We did not register how many there were.

We have clarified these inclusion and exclusion criteria in Figure 1, which is the only place in the manuscript where we document these criteria in the new version if the manuscript.

12. The first paragraph in methods section presents the sample size and the
characteristics at baseline. The description of sample’s characteristics seems to be more relevant in the results section.

We have followed this suggestion as indicated above (please see our answer to the first reviewer).

13. The paragraph beginning by “The six GPs involved were …” and the section entitled “The fitness consultation” could be gathered.
We have followed this suggestion.

14. The skills to use MI required a specific training; could the authors give more details about the 3-hour session followed by the GPs?
We have changed the wording slightly and added the following description of the 3-hour session:

*The six GPs involved were trained in the principles of aerobic and resistance exercise testing and training in a 3-hour session, and they were introduced to strategies and techniques of motivational interviewing in another 3-hour session conducted by a researcher with wide experience from motivational interviewing.*

The content of the training is described in the manuscript as mentioned under point 5 above.

15. (4th page, 2nd paragraph)
Could the authors review the section about the use of the principles of MI? I suggest describing how the GPs applied the 4 principles of MI during their consultations: the expression of empathy, the divergence development, working with resistance and increasing self-efficacy.

They could also add the 6 essential ingredients to brief interventions according to Miller and Rollnick (1991). The presence of those principles within the interventions is necessary in order to be sure that GPs effectively carried out a MI and did not only listen and encourage the participants.

We have changed the text as described under point 5 above and added relevant references.

16. Could the authors include the information about validity and reliability of the tests (fitness and muscle strength) in the subsection entitled “Tests of fitness and muscle strength”?
We have added the following sentences to the first paragraph of “Tests of fitness and muscle strength” of the Methods section:

*The chair stand test is a measure of lower body strength and has a moderately high correlation to leg press scores (R=0.78 for men and R=0.71 for women). The arm curl test is a measure of upper body strength and has a moderately high correlation to combined 1-RM (repetition max) biceps, chest, and upper back (R=0.84 for men and R=0.79 for women). The test-retest reliability (95% confidence interval) for the chair stand test is 0.89 (0.79-0.93) and for the arm curl test 0.81 (0.72-0.88) [9].*

17. Could the authors add information about the validation of the protocol used to measure the cardio-respiratory fitness?

We have added this description to the Methods section:

*This method of measuring cardio-respiratory fitness has a high correlation, R=0.97, with measurements of pulmonary ventilation and gas exchange [24].*

18. Could the authors explain why they chose HDL-C as a primary outcome when LDL-C is a primary goal regarding dyslipidemia/lipid management for persons with type 2 diabetes (see: Diabetes Care, 2010, 33 (suppl. 1), s11-s61).
Furthermore, no specific targets for HDL-C level have been determined in clinical trials (Can J Cardiol, 2009, 25 (10): 567-579)

The primary outcomes measures were all chosen well ahead of the beginning of the data collection. We decided to limit the number of outcomes to five in order to be able to interpret the p-values at the 5%-level. HDL-C was chosen as a primary outcome because we considered that increases in HDL-C lead to reduced cardiovascular risk.

19. Why blood samples were drawn after 8-hour fast only while the standard procedure to lipid profile is 12-hour fast without consumption of alcohol within 72-hour? If those recommendations were not respected, the results of the lipid profile could be biased
With regard to blood sampling we followed the standard procedure in our national guidelines for diabetes and cardiovascular disease. If this procedure has affected the level of the lipids, there is no reason to believe that this misclassification has changed over time and thereby affected the results.

20. (Page 6, line 3). Authors used function limitation and cardiovascular fitness to identify barriers to increased fitness. However, I think that the term “identify” is not properly used since “to identify barriers to increased fitness” they should select those variables that have a significant negative impact on fitness. To do that, they should consider all variables in a multivariable model and select those significantly associated with the outcomes. Furthermore, could authors justify why they chose only function limitation and cardiovascular fitness to identify barriers to increased fitness?

We admit that the wording is unfortunate. In addition, the subgroup analyses are not limited to function limitation and CVD, and these subgroup analyses are better described in the last paragraph of the Statistical analysis sub-section. Accordingly, we have changed the text as follows:

Pain with function limitation was defined as pain from joints and/or muscles in arms, shoulders, legs and/or back which reduced the performance at the physical tests at the 9- and/or 18-month follow-up as indicated by the patient. Cardiovascular disease (CVD) was defined as history of myocardial infarction and/or verified stenosis of coronary arteries and/or stroke and/or arteriosclerosis of the lower extremities verified by distal pressure measurement recorded at baseline and/or after 9 and/or 18 months.

The last question can only be answered satisfactorily with a supplementary multivariate analysis (we assume that the reviewer meant to write “cardiovascular disease” not “cardiovascular fitness” as fitness is the outcome). This analysis - which confirms our univariate findings - is now mentioned in the Results section (sub-section of Subgroup analyses) as follows:

In a full multivariate model including all the baseline variables listed in Table 2 as predictors and VO\textsubscript{2\text{max}} as outcome, only CVD (p=0.001) and pain with function limitation (p=0.023) are statistically significant.

Statistical analysis:
21. Authors should present a power calculation to be sure that the non significant differences observed in this study are not due to a lack of power.

We have added the following sentence at the end of the first paragraph in Statistical methods:

A power calculation shows that the study has a power of 80% to detect a difference in change in VO\textsubscript{2\text{max}} from 0 (no change) to 1.3 mg/kg/min during 18 months when n=127.

22. What are the beta and alpha errors?

The beta error is now defined above. Alpha is already defined (0.05) in the second-last sentence of Statistical methods.

23. (Page 6, line 17). The sentence beginning by “Thirteen patients had atypical courses due to severe disease…” could be placed in the results section with description of the sample.

It is correct that this information is primarily of clinical importance, but it cannot be separated from the information about how patients with atypical courses were treated in the analysis. Therefore, we have moved this small paragraph to the first paragraph of the Results section:

Thirteen patients had atypical courses due to severe disease (myocardial infarction, stroke, accidents, and cancer) and one patient started participation in a placebo trial after inclusion. Test results from these patients were excluded from analysis from the date of the event.

24. Could the authors add references regarding the procedure used to evaluate clinical variables (eg. Waist circumference, blood pressure)

Body weight, waist circumference and blood pressures were measured with common routine methods, which are accepted by consensus in both clinical work and research. We are not aware of any general reference that describes these pieces of clinical information, but we will be happy to add one, if the reviewer could inform us about such a reference. The three variables are thoroughly described in the manuscript in a way that enables readers to repeat the procedures, if necessary:

Body weight and height were measured without shoes and outer garments on the same scales throughout the study. BMI was calculated as \text{weight in kg}/\text{height in metres}^2. Waist circumference was measured to the nearest cm in the mid-horizontal plane between lowest rib and iliac crest. Blood pressure was measured after 10 min. in the seated position as the lowest of three values using a mercury sphygmomanometer.
We realize that we have not described the laboratory analyses precisely. Therefore, we have added the following description to “Biochemical and clinical variables” after “Slagelse Hospital”:

All blood samples were taken in the morning after an eight-hour overnight fast and a resting period of at least 15 min and no hard physical activity within the foregoing two hours. Samples were analysed at Slagelse Hospital. Fraction of HbA$_1c$ was measured by a high performance liquid chromatography method (a Tosoh Automated Glycohaemoglobin Analyzer HLC-723 G. Reference interval: 0.042-0.063). Serum total cholesterol concentration was measured enzymatically with cholesterol esterase-cholesterol oxidase-peroxidase reagent. Serum triglyceride concentrations were determined enzymatically with a lipase-glycerokinase-glycerol-3-phosphate oxidase-peroxidase reagent. HDL-cholesterol was determined by a homogeneous enzymatic colorimetric method. Plasma glucose was measured by a hexokinase method. In freshly voided morning urine, creatinine was determined by a Jaffé reaction and albumin by an immunoturbidimetric method.

Results section:

25. Seven (7) participants were withdrawn during the first 9 months follow-up because they had severe disease and those subjects were not included in the analysis at 9 and 18 months follow-up. How those withdrawals influenced the results? Why the authors kept those subjects in the baseline analysis? (Figure 2: the greater improvements of muscle strength and VO2max were in first 9 months).

The subjects who were lost to follow-up contribute to the analyses with data as long as the event prompting their discontinuation has not occurred, i.e. they contribute not only to the baseline analysis but with up to 6 measuring points in Figure 2 and Table 2. These patients probably made up a group of patients with relatively poor health. In that case the estimates of the effect of the intervention would increase if we deleted information from these patients from the analysis. By using all available information in a mixed model we make much better use of the prospective data in an open, straightforward and honest way compared to a situation where these subjects were to be excluded.

26. How many subjects were, or begun, insulin-dependedants? Does this condition influence the results?

In order to answer this question we have split up “pharmacological” into “oral agents” and “insulin” for the variable “Diabetes treatment” in Table 2. Please find the revised table as an attached file. As the number of insulin-treated diabetic patients is quite small, we are very cautious about drawing conclusions from the analyses, but the insulin-treated patients tend to show smaller improvements in fitness and muscle strength.

27. Regarding the change of the HDL-C, why the authors did not compare the group of patients without change of lipid-lowering medication (n=116) with patients who started treatment with simvastatin during the study (n=10) rather than the whole group? Furthermore, could the authors explain what is about one subject who did not change of lipid-lowering medication and did not start treatment with simvastatin?

Firstly, n=11, not 10 as indicated in the manuscript. We apologize for this error. Secondly, this analysis mentioned by the reviewer was primarily meant as a “control analysis” because we knew that some patients had started statin treatment during the study. We have now followed the suggestions of the reviewer and added statin treatment to Table 2 (please find attached). Again we are cautious not to over-interpret the findings.

28. (Page 7). In subgroup analyses section, I suggest to rewrite the sentence beginning by “The development of VO2max…” since it was observed a “change” in this variable instead of a “development”.

We believe that the correct word is “development” as the mixed model actually analyses the development rather than the “change”. In order to present the results, the changes in the primary outcomes are presented in Table 2, but the statistical analysis and its p-values relate to the development of these 5 variables over 18 months, as described in Statistical analysis.

29. The authors present secondary outcomes but they did not introduce those outcomes in the introduction or in the study’s question.

We admit that we overlooked this obvious flaw. The information fits into the manuscript several places, but we have chosen to place it at the end of the Background section:
The specific aim of the study was to evaluate the impact of extending the usual diabetes control consultation with fitness tests and motivational interviewing with the following primary outcomes: $V_{O_{2}max}$ (maximal oxygen uptake, ml O$_2$ kg$^{-1}$ min$^{-1}$), muscle strength of upper and lower extremities, haemoglobin A$_1c$ (HbA$_1c$), and HDL-cholesterol during the 18-month intervention period. Secondary outcomes were waist circumference, body mass index (BMI), blood pressures, fasting plasma glucose, total cholesterol, LDL-cholesterol, and triglycerides.

30. (Page 8), Why the authors include information regarding the course of the intervention in the results section?
We believe that some information from the process evaluation of the information is necessary so the reader can understand how the results were obtained. This includes information about the possible side effects of the maximal exercise test. Therefore, we would prefer to keep this information in the Results section.

31. Authors did not observe significant differences regarding HbA1c, gender and age between respondents and non respondents at baseline. However, it would be relevant to compare other characteristics that could influence the outcomes. For example, it would be interesting to compare comorbidities and fitness at baseline. If non respondents had more comorbidities and worse fitness than respondents, exclusion of such individuals could have overestimated the impact of the intervention on the outcomes.
Yes, it could have been interesting to compare comorbidities and fitness at baseline, but those data were not available from the non-participants. For ethical reasons these patients could not be approached to obtain supplementary information because they did not volunteer for the study.

Discussion
32. Authors should start this section by comparing the results with the scientific literature; strengths and limitations of the study should be placed at the end of the discussion section.
Done.

33. The explanation regarding the impact of the absence of control group is not sufficient.
We have changed the first part of this paragraph to comply with the reviewer’s critique. We haven’t mentioned the necessity of doing RCTs in this area of research as we have done so later in the Conclusion section:
The lack of a control group is a major limitation of the study which leaves the possibility that the improvements in outcomes could be due to the general development in the natural history of type 2 diabetes and, for the physical tests in particular, to some degree of habituation. However, it is unlikely that the observed improvements in cardio-respiratory fitness and muscle strength can be explained entirely by these effects. Firstly, control groups in previous randomised studies show a decrease in $V_{O_{2}max}$ of 1% over a period of 20 weeks and small non-significant increases of 1.5% and 5% in upper and lower body strength [2,25]. Secondly, the expected age-related decline in muscle strength and $V_{O_{2}max}$ over 18 months can be estimated to be 2.5% and 2.2% respectively in the population of the present study [9,26]. Thirdly, if habituation explained the improvement, the same development should be expected in the different subgroups, and this is not the case (Table 2).

34. (Page 9, 2nd paragraph). The references about the validity of the arm curl test and the chair stand test are more relevant in the method section. This is now briefly mentioned in the Discussion section and described more extensively in the Methods section.

35. The authors write that one third of the type 2 diabetic patients were unable to do the fitness test because of contraindications and co morbidity. However, this information was not presented in the results section. Furthermore, the authors did not discuss how this observation may have influenced the results.
The information about the missing fitness tests is in the second-last sentence in the Results section: Reasons for not performing all three tests at the final session were musculoskeletal disease (14/102), blood pressure $>180/110$ (7/102), heart disease (9/102), and acute illness (1/102). The patients could have missing test results at some follow-up consultations and not at others. The number (proportion, %) of patients who did not obtain a $V_{O_{2}max}$ estimation at baseline and after 3, 6, 9, 12, 15, and 18 months was 17 (13.7), 45 (36.3), 46 (37.1), 40 (32.3), 48 (40.0), 45 (39.1), and 50 (43.9). These missing values were treated appropriately in mixed models as explained above (point 25).
36. The authors report a significant increase in VO2max (+0.46 ml/kg/min or 2.5%). This increase is statistically significant (p=0.032). However, could they discuss about the clinical implication of this increase?

We have added the following new sentences to the Discussion section:

*There is a steep inverse relationship between cardio-respiratory fitness and mortality in men with documented diabetes [5]. This could mean that an improvement in fitness, like the one we have observed, is of clinical significance. In an observational prospective study of men, an increase of 7.0 ml/kg/min in VO2max over 4.9 years was associated with an estimated reduction of 30% in mortality risk during the following 5.1 years [12].*

37. According to the authors, the increase in HDL-c may be the result of increased muscle strength and increased VO2max. Indeed, those changes may improve metabolic profile and increase HDL-c but, could the authors explain why waist circumference and triglycerides did not change and the blood pressure increased slightly?

The intervention did not include diet counselling and the goal for physical activity in this ageing population was 2500 kcal/week. This may help to explain why waist circumference and triglycerides did not change. We have no obvious explanation for the increase in blood pressure, but it could be due to age-related increase in blood pressure in a population where 66 of 127 patients had clinical manifestations of cardiovascular disease.

38. 227/354 patients (~66%) not responding to the invitation to participate at this study. Could the authors discuss how this result influences their study’s results?

As stated above, we have access to very few data from the non-participants. Based on our clinical experience from treating these diabetic patients, we have no reason to believe that there were any major differences between participants and non-participants with regard to comorbidities. Almost all responders were deemed by their GPs to be in the preparation, action or maintenance stage. This cannot be expected to be the case in the group of non-responders where a greater number might be expected to be in the pre-contemplation stage. Furthermore, it could be hypothesised that there is a greater potential for increasing physical activity among the non-participants. All these considerations, however, are so speculative that we have not included them in the manuscript.

Tables and figures

39. In table 1, could the authors add samples characteristics regarding LDL-C and triglycerides?

Done

40. In table 2, could the authors add standard deviation (SD)?

The results in Table 2 are calculated from the parameter estimates in a mixed model. Hence, an SD is not well defined. It is, however, possible to calculate standard errors (SEs), but the addition of so many SEs to Table 2 would make this table very large, and we do not think that these SEs would give new information on the uncertainty of the estimates beyond that inherent in the p-values. If the editor finds it necessary, we will, of course, calculate the SEs and include them in Table 2.

41. In table 2, the values of the arm curl test and chair stand test could not have a decimal because this is the number of flexions or number of stand ups.

It is true that the results of these tests come in numbers without decimals. In Table 2, however, the mixed models produce results corresponding to average values of the test results, and here one decimal is appropriate.

42. In table 2, could the authors round off to one decimal the values of HbA1c and HDL-C to be in accordance with the clinical guidelines?

We agree that we have included one more decimal than we use when we normally report values of HbA1c and HDL-C. In Table 2, however, we report differences between the estimated value of these two variables at baseline and at 18-month follow-up. In this case it seems reasonable to list an extra decimal for two reasons: Firstly, it is customary to include “the uncertain decimal”. Secondly, the number of decimals should correspond to the percentage, which is rather dependent on the last decimal.

43. In table 2, could the authors use the same terms that in Table 1 (eg. Antihypertensive medicine vs. Antihypertensive treatment). Furthermore, could the authors add results for subgroups using or not the lipid-lowering agents?
We have changed the name of this variable to antihypertensive medication. As all patients in lipid-lowering medication receive statins, we have changed the name of this variable to “Statin treatment”. We have also added statin treatment in three groups to Table 2, as indicated above (point 27).

44. In table 3, could the authors add standard deviation (SD) for the difference (#) values? Please read our answer above (point 40). SD is not the appropriate measure, but we can calculate the SEs, if necessary.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**
'I declare that I have no competing interests'

---

**Reference List**


