Author's response to reviews

Title: Lower leg muscle strengthening does not redistribute plantar load in diabetic polyneuropathy. A randomized controlled trial.

Authors:

Tom Melai (Tom.Melai@MaastrichtUniversity.nl)
Nicolaas C Schaper (N.Schaper@MUMC.nl)
Herman IJzerman (Herman.IJzerman@MaastrichtUniversity.nl)
Ton LH de Lange (ALH.deLange@hetnet.nl)
Paul JB Willems (Paul.Willems@maastrichtuniversity.nl)
Valéria Lima Passos (Valeria.Limapassos@maastrichtuniversity.nl)
Aloysius G Lieverse (L.Lieverse@mmc.nl)
Kenneth Meijer (Kenneth.Meijer@MaastrichtUniversity.nl)
Hans HCM Savelberg (Hans.Savelberg@MaastrichtUniversity.nl)

Version: 2 Date: 3 June 2013

Author's response to reviews: see over
Dear dr. Menz/ dear editor,

We appreciate the constructive comments made by the reviewers and based on these comments we would like to resubmit the revised manuscript to the Journal of Foot and Ankle Research. Please find enclosed the revised manuscript with our revisions underlined and a document in which we answer (A) in detail each point raised by the reviewers (R) with the alterations we made in manuscript (M) based upon these suggestions. As suggested by one of the reviewers, we have also uploaded a COSORT checklist as an additional file.

We look forward to your response.

On behalf of all co-authors,

Yours sincerely,

Hans H.C.M. Savelberg PhD
Department of Human Movement Science,
MaastrichtUniversity
PO Box 616; NL-6200MD, Maastricht, TheNetherlands
Phone: +31.43.3881392
Email: Hans.Savelberg@MaastrichtUniversity.nl
Answer to the reviewers

Please find below the comment of the reviewer R, our answer (A) in which we have tried to describe our response in detail and the alterations made in manuscript (M) based upon the suggestions of the reviewer.

Statistical review:

R 1. I found the data used for the statistical analysis (random intercept model) very difficult to interpret. I thought they would have been better off using the actual pressure readings rather than the changes in one variable compared to the rate of change in the control. The actual pressure readings would have been more clinically relevant as for instance on page 15, they discuss how high plantar pressures are a risk factor for foot ulceration. The beta values used to analyse the determinants (Table 2a-1) were very hard to interpret. I think the use of a two-way mixed ANOVA would have been more useful. It could have been applied to do post-hoc tests at end points and also to calculate effect sizes with CI 95%. I think this was the most clinically relevant statistic produced by the RCT.

A 1. We understand the reviewer’s reservations with respect to the random intercept model. To facilitate the interpretation, we adjusted the manuscript accordingly. Briefly, we motivate our choice as follows: 1. The use of the random intercept model is more and more encouraged in current literature above the use of an ANOVA for analysis of longitudinal follow-up data of a clinical trial (Nich 1997). One of the disadvantages of an ANOVA approach is that when several measurements are made over time, loss of one measurement would mean the exclusion of the complete data set of the subjects with a missing value. The mixed model approach (with random intercept) allows for all patients to be included in the analysis, not only the complete cases. 2. Admittedly, the model and their parameters may pose some interpretational challenge. Please, note that alternatives to present these multivariable data are many. For instance, as the reviewer suggested, one can choose a time point (e.g. the last one), and based on the final model parameters, compare the estimated groups’ means (this will give only groups effect but no time effect). Alternatively, one could compare different time points with each other per group separately. If we were to select these options, we may only report that other covariates had a significant effect, without providing the proper quantification. Besides, for a few outcomes, groups’ differences could be observed, whereas for others it could not. Changing every time the way the data are reported, according to observed statistical significances, would in our belief have led to inconsistencies in the presentation. Hence, on the grounds of completeness of information, consistency and parsimony, we deemed it better to report the model parameters (inclusive the estimates of the between and within subject variability - random intercept and residual variances), adding an interpretational note in the figure caption. Please note that availability of these details will in our opinion also facilitate future comparisons with similar studies.

M 1. Please see table caption, page 22: “Note: Measurements were obtained for the intervention and control group (Group) at different time points (Time) and at two gait velocities, preferred and imposed (see text). Estimated fixed (and random) effect parameters of the random intercept models are shown. β (95 % CI) represent average changes in the outcome variables per unit increase for the continuous explanatory variables. For categorical variables, the estimates represent average changes with respect to the reference (control group, females, or follow-up measurement at t= 52 weeks). Effects of time are therefore negative if baseline measurements were lower than during follow-up. Age was determined at onset of the study. Abbreviations and symbols: IG= intervention group; CG= control group; PP= peak pressure; *= p≤ 0.05, †= p≤ 0.01, ‡= p≤ 0.001.”

Please see section 2.6, page 10: “All data are expressed as mean with 95% confidence intervals (CI). The effect of the training program on the primary outcome was determined on an intention to treat analysis. As subjects participated in a training program with two components and were subsequently followed for 6 months, multiple measurements had to be performed over time. Given the hierarchical structure of the data, with repeated measurements nested within subjects, a random intercept model was fitted to evaluate average changes in plantar pressure variables over time for the intervention and control group, adjusting simultaneously for covariates. This model has the advantage of using all available observations for model estimation, inclusively those of subjects entered in the model until drop-out, and not only the complete cases like the classical repeated measures ANOVA. ”

R 2. The use of the ‘interaction’ value was not helpful as they excluded it from the model because the time/group was not significant. The real meaning of this non-significant interaction effect is apparent in figure...
where the graph shows absolutely no interaction effects. This of course means that there are no improvements over time in the treated group in comparison to the control.

A 2. The reviewer is correct that no interaction effects were observed for group and time in our analyses. The interaction effect was therefore excluded from the final model and we fully agree with the reviewer that in comparison to the control group no improvements over time were observed. Note however, that there may be changes over time, but this was the same for both groups, or that there were groups' differences that were the same for all times. We realize that the text may have been unclear and have changed it in the revised manuscript.

M 2. Please see section 2.6, page 11: “The interaction between group and time was considered in order to test whether temporal changes of the outcomes over time depended on the group, with a significance level of 5%. Estimates were determined by a restricted maximum likelihood method. For interaction terms, whose p-values were < 0.1, additional Likelihood Ratio (LR) tests were conducted (at 5% significance level). If non-significant, they were left out of the final model. All other fixed factors remained in the model, irrespective of significance.”

Please see section 3.2, page 12: “No significant effect of the intervention on the plantar pressures characteristics (peak pressures, PTI, tCOP and F/R-ratio) was observed (LR p-values >0.1). As the interaction effect (time*group) was for none of the variables significant it was left out of subsequent models.”

R 3. The descriptive statistics is very useful (fig 7.2 a1), however they are a little misleading as they exaggerate the difference and changes due to time because they didn’t use a zero point for the outcome variables but started with very high values. Nevertheless, the graphs clearly indicate that apart from initial differences between the controls and treated (e.g. graph ‘a’), and to some degree of deterioration in both groups (e.g. ?c?), there were absolutely no therapeutic benefits for the intervention.

A 3. We thank the reviewer for this comment and we agree that the used figures can be somehow misleading. We chose not to use a zero point as the figures would become unreadable (e.g. the intervention group could not be discerned from the control group). However, to avoid misleading we have adjusted the figures, adding a break in the y-axis and clarification in the legend.

M 3. Example of figure 2:
Please see the legend of figure 2: “Graphs have broken y-axis to improve readability.”

R 4. There is a basic question unanswered by the researchers: ‘did the exercise program result in increased muscle strength in the patients?’ They state on page 6 that ‘measures of muscle strength and mobility will be reported elsewhere’. This is odd because their models suggest that treatment leads to increased strength and to decreased plantar pressure. They should have reported muscle strength changes in the same paper.

R 4. Our dataset is very elaborate as the research project included among others measurements of isometric and isokinetic strength of the knee and ankle flexors and extensors, a 6 minute walk test, a ‘get up and go’ test, joint moments during gait and two questionnaires on mobility and quality of life. The large quantity of the data obtained from these measurements cannot be described in one article. Therefore the data on changes in muscle strength are submitted elsewhere and cannot be included in the current manuscript. We observed an increase of muscle plantar flexion strength of 44% (p=0.024). To aid the interpretation of the results of the current study we have described the general effects of the intervention on muscle strength in the discussion section.

M 4. Please see section 4, page 14: “The high drop-out rate might have influenced the outcome, but to be reported elsewhere, the intervention did increase maximal plantar flexion strength, although maximal dorsal flexion strength was not affected (IJzerman et al. Unpublished data).”

R 5. The failure of the treatment is indicated by the high percentage of dropout (41.7% - page 11) in the treatment group; almost twice that of the control group (22%). Perhaps a pilot study would have been relevant as it is evident that this intervention has been found unsuitable by the obese elderly participants with diabetes.
A 5. We did perform a pilot/feasibility study in 10 diabetic subjects with neuropathy, of which 9 completed the study with a training compliance of 88% (Duijkers et al. 2006, unpublished data). We feel that this marked difference with the current trial may be explained by differences in patient characteristics due to a more selective inclusion of subjects in the pilot study. The high drop-out in the trial stresses the importance of publication of these results to aid future research.

R 6. I am not sure from table 2a ‘1 which are the clinically meaningful moderator variables for identifying the patients who might benefit from the treatment. The researchers have some ideas on possible improvements for future research (pages 14-16), but I am ignorant of the underlying theory; as far as I can see, the best research would be to start from the beginning with interventions that are acceptable to the patients being helped.

A 6. We agree with the reviewer that it is probably not possible to select the patients that might benefit from the treatment based on the variables or factors used in current manuscript. However, the manuscript does contain information that is helpful for future studies monitoring the effect of any intervention on plantar pressures of patients with diabetic polyneuropathy or studies exploring the effects of strength training on other outcome measures in this patient group. We thank the reviewer for this comment and have clarified the discussion section in the revised manuscript.

M 6. Please see section 4, page 15: “Our study resulted in some important findings that should be taken into account in future research exploring the effects of other or similar interventions on plantar pressures in patients with DPN. First of all, it is possible that we did not observe an effect of our intervention, due to a large between-subjects variability. In our research we included patients with different types of diabetic polyneuropathy, several mild comorbidities and several levels of motivation. To see which intervention is most suitable for the patients being helped, future research needs to focus on strictly defined specific patient groups before looking at general benefits of an intervention in a RCT.”

Reviewer’s report: Lower leg muscle strengthening does not redistribute plantar load in diabetic polyneuropathy. A randomized controlled trial.

Title: Lower leg muscle strengthening does not redistribute plantar load in diabetic polyneuropathy. A randomised controlled trial.

Version: 1 Date: 16 January 2013

Reviewer: Joanne S Paton

Reviewer’s report:

Lower leg muscle strengthening does not redistribute plantar load in diabetic polyneuropathy. A randomised controlled trial.

Major Revisions

R 1. Data should be converted to kPa, the accepted System International unit of pressure.

A 1. We thank the reviewer for the comment and have converted the units in the revised manuscript.

R 2. It would be useful for more detail to be provided with regard to exclusion criteria. For example was the diagnosis of severe cardiac disease etc patient reported or from hospital notes? What was the definition of severe cardiac disease/osteoarthritis. How did the researchers define foot deformities and did this include toe deformities? Did they include participants with a history of foot ulceration? How did they assess risk of falling.

A 2. We thank the reviewer for the comment and have specified the exclusion criteria.

M 2. Please see section 2.2, page 7: “Subjects were excluded if diagnosed with severe cardiac disease (NYHA≥ 3), renal dysfunction (creatinin> 180 μmol/l), intermittent claudication, neurological disorders other than DPN, rheumatoid arthritis, amputations, prior or current foot ulceration or were thought to be at risk of falling as assessed by the research staff. Subjects were also excluded if the presence of severe osteoarthritis or foot deformities interfered with the gait pattern.”

R 3. Authors should consider and describe how using merged data from two different pressure analysis systems could affect the validity of the results and specifically how they overcame this problem. For example it would be helpful if they could present data from a pilot study to confirm that the two systems are comparable or alternatively discuss with reference to their findings that the pressure values did not differ between sites.

A 3. The two pressure analysis systems used were not expected to lead to differences in outcome as our primary outcome measure was the change in plantar pressure. In addition, the two systems are of the same manufacturer and provide comparable results. Unfortunately we do not have other (pilot) data on differences
between these analysis systems. We agree with the reviewer that this information should be presented more clearly and have added this to the limitation section of the discussion.

M 3. Please see section 4, page 18: “Finally, it could be considered a limitation that we did not use identical pressure platforms at the two study sites, which in theory could have affected the outcome. Nevertheless we expect the difference in platforms to be negligible, as the platforms are of the same manufacturer, provide comparable data and in particular because our primary outcome was the change in plantar pressure.”

R 4. Authors should adjust p values to correct for multiple testing (Using for example the Bonferroni method). It would also help readers if the authors could expand on the implications of multiple testing and the risk of finding a significant outcome by chance.

A 4. We did not observe a significant effect of the intervention on our main outcome. Adjusting for multiple testing would not affect this outcome, so we opted for not correcting for the type I inflation. In the other analyses we observed trends over time in our mixed models, adjustment for multiple testing is in our opinion not relevant. We do agree that when several different parameters are tested in a group of subjects (each peak pressures, PTI and tCOP) one could argue that for each test the p-value should be corrected. This approach however can markedly increase the risk of a type II error.

R 5. Authors should be encouraged to include and comply with the CONSORT statement checklist for reporting RCT.

A 5. We agree with the reviewer and have added the checklist as supporting document. The revised manuscript complies with the CONSORT checklist.

R 6. I apologise that I am unfamiliar with the statistical tests used but I am confused by the results. I wonder if it would be helpful to clarify a couple of points. Whilst the results section explains that the intervention had no effect on peak pressure, table 2 Group (IG vs CG) suggests the difference between groups was significant. Does this actually mean that groups were not comparable for PP at baseline?

A 6. The reviewer is correct in the argumentation. Without a significant interaction term, the groups differences are the same for all time points. There are indeed differences between groups for the peak pressure under the heel and forefoot as has been described in section 3.4 and table 2. Both section and table have been changed in the revised manuscript.

M 6. Please see table 2 and section 3.4, page 13: “At all times of measurement, there were no differences between the intervention and control group for PTI, tCOP and F/R-ratio, but the peak pressures under heel and forefoot were significantly higher for the intervention group for both the preferred and imposed gait velocity (p≤ 0.01, table and figure 2a-d).”

R 7. I am not clear if the analysis actually compares PP before and after intervention between groups? Is that not your primary outcome given that you have undertaken a RCT? If it is rather to build a model to determine the association of a number of predictor variables on plantar pressure then why use an RCT design when a observational cohort study with an appropriate sample size would be more appropriate. It is normal practice to use intention to treat analysis when applying an RCT. I have serious concerns given the high attrition rate that this approach was not adhered particularly given that the sample size calculation was not based on the primary outcome measure but I do not have the expertise on how to give advice on how to proceed. In summary my primary concern is that the research question and analysis is not a comparison of intervention effectiveness and therefore not aligned to the RCT research design.

A 7. The analysis does compare the before and after intervention between groups. The regression coefficients (betas) associated with the dummy variables for time are indicative, if significant, of average differences between the first three time points and the last one, taken as a reference. Time point 0 is the pre measurement.

The reviewer addresses several pertinent issues. First of all, the primary goal of our research was not to build a model of predictive values for plantar pressures, but rather to study the effect of the intervention on plantar pressures. Predictive models are usually less concerned with potential explanations of the observed associations, but rather with the precision of the prediction, irrespective of the theoretical underpinnings. We hypothesized that treatment may improve plantar loading, whereas no improvement was expected for the control patients (this differential effect would require the interaction term between time and groups to be significant). This, however, was not observed, suggesting no effect of treatment.

Second, we fully agree that the drop-out rate was high and may thus introduce biases to our inferences, if not pondered upon carefully. Because the pattern of missingness (of drop-out) may not be at random, readers
should be reminded to interpret the results carefully, as the issue of selective drop-out cannot be solved by using the mixed model. In addition, it should be noted that high drop-out rates in long term exercise programs for patients with diabetes are not exceptional, as discussed in our manuscript on page 17.

As for the research design: The applied mixed model approach can be used in both follow-up settings, either experimental (RCT) or observational (cohort). Either way, this technique accounts for the correlated nature of the measurements. The use of this technique warrants not only the statistical validity of the results but also the internal validity of the conclusions, but only when the missingness is at random. (Please see the discussion section on page 16)

At last, conducting this study in an observational setting, for the sake of sample size, would make causal explanations even more difficult to attain. One may end up with non-equivalent groups, and as such prone to a series of other confounding bias, additional to the attrition problem.

To aid to the interpretation of our method and results we have added extra information to the results section.

M 7. Please see section 3.2, page 12: “No significant effect of the intervention on the plantar pressures characteristics (peak pressures, PTI, tCOP and F/R-ratio) was observed (LR p-values >0.1). As the interaction effect (time*group) was for none of the variables significant it was left out of subsequent models.”

Minor Essential Revisions

R 1. Abstract results section: 3rd line change the word increased with to increased by.
A 1. We thank the reviewer for the suggested improvement and have changed this in the revised manuscript.

M 1. Please see the abstract results section, page 3: “In both the intervention and control groups the peak pressure and the pressure-time-integral under the forefoot increased by 55.7 kPa (CI: 14.7, 96.8) and 2.0 kPa.s (0.9, 3.2) over 52 weeks, respectively”

R 2. Authors need to acknowledge within the design overview section the limitations and implications of using a sample size calculation based upon an un-reported primary outcome measure.
A 2. We thank the reviewer for this comment and have explicitly described that we used a convenience sample in the design overview. In addition, we have added the lack of sample size calculation for the primary outcome as one of the limitations of the current study.

M 2. Please see section 2.1, page 6: “Based on scarce literature [14, 15] the sample size of the trial was determined at 50 subjects with DPN for both the intervention and control group. The sample size of current study was based on this sample size determination and therefore can be considered a convenience sample.”

Please see section 4, page 17: “Another limitation of the current study is that sample size was determined based on the expected distribution and effects of exercise on muscle function, not on plantar pressures. Also since no correction for multiple testing was applied, significant findings should be interpreted with caution.”

R 3. It would be useful to have more details on the recruitment settings. For example how many recruitment centres were participants enrolled from and if they community or hospital based.
A 3. We thank the reviewer for the suggested improvement and have added the suggested information in the methods section.

M 3. Please see section 2.2, page 7: “Subjects were recruited from five diabetes outpatient clinics (one university and four regional hospitals) at the southern part of the Netherlands between December 2006 and November 2010.”

R 4. Typo page 7. Should gnostic read diagnostic?
A 4. We agree with the reviewer that the term gnostic might be confusing and have changed the text of the manuscript by specifying the tests we performed.

M 4. Please see section 2.2, page 7: “The diagnosis of DPN was made in 94 subjects, scoring at least 4 points out of the maximal 33 during a standardized Clinical Neurological Examination (CNE) which included reflexes, light touch, vibration, position sense, pinprick and lower extremity muscle strength testing [16].”

R 5. In the same sentence I don’t really understand the word ‘sensibility’ may be better to rephrase or define.
A 5. We thank the reviewer for this comment, the term sensibility is confusing and we have changed the text of the manuscript by specifying the tests we performed.
Please see section 2.2, page 7: “The diagnosis of DPN was made in 94 subjects, scoring at least 4 points out of the maximal 33 during a standardized Clinical Neurological Examination (CNE) which included reflexes, light touch, vibration, position sense, pinprick and lower extremity muscle strength testing [16].”

When describing the randomisation procedure. Details about how the randomisation allocation sequence was generated and whether it was stratified to centre should be included. Also confirmation about when the group allocation occurred could be mentioned for improved clarity. It appears from Fig 1 that it was after baseline data collection?

The reviewer is correct that we performed the randomization after baseline measurement. We thank the reviewer for the suggested improvement and have changed the text in the revised manuscript.

Please see section 2.3 on page 8: “After baseline measurement, using the envelope method stratified over 5 groups (2 at the first institute and 3 the other), the participants were randomized by research staff, to either the intervention group receiving physical therapeutic training or to the control group receiving no training (figure1).”

Whilst authors describe a high dropout rate in terms of attendance to exercise training sessions there is no mention of home exercise compliance rates. Did authors record home compliance to the exercise programme? This would be important additional information.

We agree with the reviewer that this is important information and have added that home compliance was not controlled during our experiments.

Please see section 2.3, page 8: “In addition to these plenary sessions, subjects were provided with an exercise manual and were asked to carry on the exercises of the second section, two times a week at home, without supervision or monitoring.”

When describing the data collection procedure it would be useful to know how long the test track was at each site and if the test surface was the same length at both centres.

We thank the reviewer for this comment and have changed the text in the revised manuscript.

In addition to these plenary sessions, subjects were provided with an exercise manual and were asked to carry on the exercises of the second section, two times a week at home, without supervision or monitoring.”

Authors should give more detail with regard to the data collection protocol. For example which foot was chosen for pressure analysis, how the decision was made and was it consistent within individuals across time.

We thank the reviewer for this comment and have changed the text in the revised manuscript.

Please see section 2.4, page 9: “For practical reasons and comparability over time only data was collected of the right foot.”

Also was the peak pressure region selected consistent across time or did the site of peak pressure change?

We appreciate the suggestion of the reviewer, but we did not analyze the shift of peak pressure location. There is large variety of parameters to choose from when analyzing plantar pressures and we considered the location of peak pressure of secondary importance. If the peak pressure would have changed due to the intervention, then peak pressure region would have been more relevant.

It is not until the discussion section that it becomes clear that patients preferred velocity was inconsistent over time. This should be detailed in the methods. Also the method of determining gait velocity should be described.

We thank the reviewer for this comment and have changed the text in the revised manuscript.

Please see section 2.4, page 9: “At first the subjects were asked to walk over the test track at their own preferred gait velocity, which was allowed to vary each trial.”

Page 11. The phrase ‘more on that later’ should be revised to sound less chatty.
We thank the reviewer for this comment and have changed the text in the revised manuscript.

Please see section 2.6, page 11: “Therefore, results should be interpreted with caution (as will be discussed further on).”

Table 1. Length is more commonly referred to as height.

We thank the reviewer for this comment and have changed the text in the revised manuscript.

Please see table 1

There should be a section entitled ‘Blinding’ to describe and define for the reader who was blind to the treatment allocation and when.

We thank the reviewer for this comment and have added information on blinding to the revised manuscript.

Please see section 2.1, page 6: “Due to practical reasons it was not possible for research staff and subjects to be blinded for treatment allocation.”

The results text does not seem to accurately describe table 2. When tabulating the effect of time on plantar loading, it would improve clarity for readers to add that negative values represent an increase in peak pressure and positive a decrease (as this seems counter intuitive). 1) Change in PP at the heel over 52 weeks using the imposed walking velocity was not significant if the table is correct. 2) Likewise change in PTI at the heel over 52 weeks using the imposed walking velocity was not significant if the table is correct.

We thank the reviewer for these comments and have changed the result text

Please see section 3.3, page 13: “Time effects were present for almost all plantar pressure parameters, regardless of the group. At the preferred gait velocity, the peak pressure under the heel increased over 52 weeks with 45.8 kPa (95% CI: 25.3, 66.2; p ≤ 0.001; table and figure 2a). At the preferred and imposed gait velocity, the peak pressure under the forefoot increased in both groups with 52.8 (CI: 18.0, 87.5) and 55.7 (CI: 14.7, 96.8) kPa, respectively (p ≤ 0.01; table and figure 2c-d). The PTI of the heel and forefoot decreased in both groups with 4.5 (CI: 2.6, 6.7) and 2.0 (CI: 0.4, 3.5) kPa.s, respectively (p ≤ 0.01; table and figure 2e and g), at the preferred gait velocity. In contrast, at the imposed gait velocity the PTI increased under the forefoot in both groups with 2.0 kPa.s (CI: 0.9, 3.2; p ≤ 0.01; table and figure 2h). Moreover, tCOP decreased in both groups during the preferred gait velocity with 3.4% (CI: 1.92, 4.90; p ≤ 0.01) as shown in the table and figure 2i and j, indicating a faster forward transfer of COP. Also F/R-ratio increased over time for the preferred gait velocity with 0.11 (CI: 0.18, 0.38; table and figure 2k; p ≤ 0.01).”

Please see the legend of table 2: “Effects of time are therefore negative if baseline measurements were lower than during follow-up.”

Discussion section. Stating that the exercise program did not affect plantar pressure patterns does not from your results appear to mean you did not observe a difference between intervention group’s pre and post intervention which is what you claim in the opening of the discussion (although the figures seem to suggest that is likely to be the case). I am unable to locate significance test data comparing the two intervention groups before and after intervention. Given that this should be the primary outcome of the RCT described it is important for the reader to see this presented statistically.

We observed no effect of the intervention on plantar pressures, as there was no significant interaction between group and time. Therefore, the interaction effect was left out of the model. To add more information and to clarify, we re-wrote the result section.

Please see section 3.2, page 12: “No significant effect of the intervention on the plantar pressures characteristics (peak pressures, PTI, tCOP and F/R-ratio) was observed (LR p-values >0.1). As the interaction effect (time*group) was for none of the variables significant it was left out of subsequent models.”

The requirement to standardise gait velocity when evaluating an intervention over time using pressure analysis, whilst important, is not a new finding as implied and would be better discussed within study limitations.

Indeed, the finding that gait velocity is an important confounder of plantar pressures is not new. However, gait velocity is frequently not monitored or controlled in studies reporting on plantar pressures. In addition, the application of a standardized gait velocity is less described in literature, especially in subjects with diabetes and polyneuropathy. We consider the analysis of both a preferred and standardized gait velocity as a strength of the study and we think it would benefit the interpretation of pressure data if standardization is more frequently applied in future research.
R 18. The discussion appears unbalanced. Little attention or depth is given to addressing the main aim of the study. Instead much discussion is centred on 1) method limitations and 2) the natural PP progression of DPN. I am unclear from the current manuscript what it is within these two topic areas that the authors are suggesting is new information.

A 18. We have rewritten the discussion paragraph. In paragraphs 2-4 of the revised manuscript we discuss possible explanations of the lack of effect of our intervention, resulting in a more balanced discussion. In the other paragraphs we hope to give information that could aid further research in the design of experiments and interventions.

R 19. Conclusion section. The conclusion appears to focus on the finding that PP increases over time in patients with DPN. This is not a new finding or the primary research question being addressed. The conclusion should therefore be revised.

A 19. The conclusion firstly addresses that there is no effect of the intervention on plantar pressure. We agree that the increase of plantar pressure does not answer to the primary research question, however as previously addressed, gait velocity is an important confounder and to the best of our knowledge our study is the first to quantify this increase in pressure over time at a controlled gait velocity. Therefore, we choose to add one sentence to the conclusion section.

Discretionary Revisions

R 1. Abstract Background. 3rd line It would be advisable for authors to reduce the strength of supposition by replacing the words responsible for to associated with.

A 1. We thank the reviewer for this comment and have changed the text in the revised manuscript.

M 1. Please see section: abstract background, page 3: “Higher plantar pressures play an important role in the development of plantar foot ulceration in diabetic polyneuropathy and earlier studies suggested that the higher pressures under the forefoot may be related to a decrease in lower leg muscle strength.”

R 2. Table 1. The new IFCC measurement unit for HbA1C is mmols/mol. It would be preferable to convert from %.

A 2. We thank the reviewer for this comment and have changed the measurement units in the revised manuscript.

M 2. Please see table 1

Referee 2:
Title: Lower leg muscle strengthening does not redistribute plantar load in diabetic polyneuropathy. A randomized controlled trial.
Version: 1 Date: 22 January 2013
Reviewer: Anita Raspopovic
Reviewer’s report: This is an important and well written paper that warrants publication after the revisions below are addressed.

Major Compulsory Revisions

R 1. Methods: The suitability of determining sample size for a pressure study using data on muscle function is queried, even though an explanation is provided in the methods. Further justification of this approach is suggested. Is the assumption that muscle function is hypothesised to be related to plantar pressure, and thus a sample size adequate to detect a change in the former should be sufficient to detect changes in the later? Or perhaps the authors have another rationale for the approach taken? Please provide more explanation to the readers.

A 1. As the current study was embedded in a larger randomized clinical trial, the sample size was determined based on the expected distribution and effects of exercise on muscle function, not on plantar pressures. Therefore, the sample size of the current study was based on a convenience sample and significant findings should be interpreted with caution. We thank the reviewer for the comment and have added the lack of a sample size calculation to the limitations section.

M 1. Please see section 2.1, page 6: “Based on scarce literature [14, 15] the sample size of the trial was determined at 50 subjects with DPN for both the intervention and control group. The sample size of current study was based on this sample size determination and therefore can be considered a convenience sample.”
Please see section 4, page 18: “Another limitation of the current study is that sample size was determined based on the expected distribution and effects of exercise on muscle function, not on plantar pressures. Also since no correction for multiple testing was applied, significant findings should be interpreted with caution.”

R 2. Results, discussion, conclusions: The confounding factor of walking velocity changes on plantar pressure output, requires earlier identification, increased reporting and more consistent recognition in the discussion and conclusions of the paper. The mean difference and overall magnitude of change in walking velocity across measurement sessions should be reported (in terms of % change); alongside the overall magnitude of change in forefoot plantar pressure across measurement sessions (again % change would suffice). Currently this data did not appear to be reported but this would give the reader greater insight into the actual size of the changes being discussed.

The first paragraph of the discussion (second sentence) should acknowledge upfront that the increase in forefoot loading found over time may be due, at least in part, to commensurate changes in walking velocity. Similarly, the second last sentence of the abstract and the last statement in the conclusion do not acknowledge this significant confounding issue. Risk of foot ulceration may not progressively rise as stated in the conclusion, if the artefact of walking velocity changes were in fact responsible for the increased forefoot loading reported. The abstract and conclusion should be better contextualised to reflect this significant study limitation. Finally, could further statistical analysis or other research addressing this issue better identify what % of increased the forefoot loading might be attributable to walking velocity changes as compared to other events such as worsening impact of DPN?

A 2. We thank the reviewer for these comments and have emphasized that gait velocity may vary during preferred gait velocity. We also agree with the reviewer that gait velocity is an important confounding factor, which could affect plantar pressures during the preferred gait velocity, as we stated in the third paragraph of the discussion. However, we also observed an increase in plantar pressures over time at a standardized gait velocity. Therefore, the observed increase in plantar pressure during this condition could not be attributed to an increase in gait velocity. Even more, an increase in gait velocity would lead to an increase in peak pressure, but a decrease in pressure time integral. We observed an increase in both the peak pressure and pressure time integral, which further supports that gait velocity is not the factor responsible for the increased loading we observed during the standardized imposed gait velocity.

We agree that it would be very interesting to determine if the increase in forefoot loading can be attributed to the level of diabetic polyneuropathy. This is outside the scope of current study and would require extra measurements such as a neurological EMG. However, the results of the current study can provide a basis for such research.

We also agree that reporting the magnitude of change in the text would benefit the interpretation. Expressing change as a percentage is complicated, as the estimates of the statistical model are constructed based on the factors included. Leaving out one factor to obtain the values needed to calculate overall changes in percentages would therefore affect the estimates. For this reason we provided the absolute values, as well as reference data on the impact of these changes in peak pressure.

Based on the reviewers comment we have made several changes in the revised manuscript.

M 2. Please see section 2.4, page 9: “At first the subjects were asked to walk over the test track at their own preferred gait velocity, which was allowed to vary each trial.”

Please see section 3.3, page 13: “Time effects were present for almost all plantar pressure parameters, regardless of the group. At the preferred gait velocity, the peak pressure under the heel increased over 52 weeks with 45.8 kPa (95% CI: 25.3, 66.2; p ≤ 0.001; table and figure 2a). At the preferred and imposed gait velocity, the peak pressure under the forefoot increased in both groups with 52.8 (CI: 18.0, 87.5) and 55.7 (CI: 14.7, 96.8) kPa, respectively (p ≤ 0.01; table and figure 2c-d). The PTI of the heel and forefoot decreased in both groups with 4.5 (CI: 2.6, 6.7) and 2.0 (CI: 0.4, 3.5) kPa.s, respectively (p ≤ 0.01; table and figure 2e and g), at the preferred gait velocity. In contrast, at the imposed gait velocity the PTI increased under the forefoot in both groups with 2.0 kPa.s (CI: 0.9, 3.2; p ≤ 0.01; table and figure 2h). Moreover, tCOP decreased in both groups during the preferred gait velocity with 3.4% (CI: 1.92, 4.90; p ≤ 0.01) as shown in the table and figure 2i and j, indicating a faster forward transfer of COP. Also F/R-ratio increased over time for the preferred gait velocity with 0.11 (CI: 0.18, 0.38; table and figure 2k; p ≤ 0.01).”

R 3. Discussion, paragraph 2: A possible lack of study power is an important limitation to this study regarding the finding of no significance for the intervention trialled, particularly given that just over half of the
participants in the intervention group did not complete the study. This point should be made up-front in paragraph 2 of the discussion where possible reasons for the study results are being presented.

A 4. We thank the reviewer for this comment and have explicitly described that we used a convenience sample in the design overview. In addition, we have emphasized the high drop-out rate in paragraph 2 of the discussion and have added the lack of sample size calculation for the primary outcome as one of the limitations of the current study.

M 4. Please see section 2.1, page 6: “Based on scarce literature [14, 15] the sample size of the trial was determined at 50 subjects with DPN for both the intervention and control group. The sample size of current study was based on this sample size determination and therefore can be considered a convenience sample.”

Please see section 4, page 14: “The high drop-out rate might have influenced the outcome, but to be reported elsewhere, the intervention did increase maximal plantar flexion strength, although maximal dorsal flexion strength was not affected (IJzerman et al. Unpublished data).”

Please see section 4, page 17: “Another limitation of the current study is that sample size was determined based on the expected distribution and effects of exercise on muscle function, not on plantar pressures. Also since no correction for multiple testing was applied, significant findings should be interpreted with caution.”

Minor Essential Revisions.

R 1. Abstract, methods, sentence 2: Should read ‘were’ not ‘where’

A 1. We thank the reviewer for this correction and have changed it in the revised manuscript.

M 1. Please see the method section of the abstract on page 3.

R 2. Introduction: Clarification is required from the Editors of JFAR, as to the suitability of referencing unpublished work.

A 2. The work referred to is submitted elsewhere and other studies leading to similar information are not available.

R 3. Methods, 2.2 Settings and Participants: Please provide (somewhere in the paper) an indicative range of values for the CNE used to indicate lack of sensation (DPN), with cut-off scores for various levels of neuropathy severity.

A 3. We agree with the reviewer and have changed the manuscript as suggested.

M 3. Please see section 2.2, page 7: “The diagnosis of DPN was made in 94 subjects, scoring at least 4 points out of the maximal 33 during a standardized Clinical Neurological Examination (CNE), which included reflexes, light touch, vibration, position sense, pinprick and lower extremity muscle strength testing [16].”

R 4. Methods, 2.2 Settings and Participants: Was any pilot testing conducted to establish the consistency of pressure data output between the two different E-MED systems used?

A 4. The two pressure analysis systems used were not expected to lead to differences in outcome as our primary outcome measure was the change in plantar pressure. In addition, the two systems are of the same manufacturer and provide comparable results. Unfortunately we do not have other (pilot) data on differences between these analysis systems. We agree with the reviewer that this information should be presented more clearly and have added this to the limitation section of the discussion.

M 4. Please see section 4, page 18: “Finally, it could be considered a limitation that we did not use identical pressure platforms at the two study sites, which in theory could have affected the outcome. Nevertheless we expect the difference in platforms to be negligible, as the platforms are of the same manufacturer, provide comparable data and in particular because our primary outcome was the change in plantar pressure.”

R 5. Methods, 2.4 Outcomes and follow-up: How was standardised imposed gait velocity implemented and measured? What technique was utilised to check that participants did not alter their gait pattern to target the platform? Please state in the paper.

A 5. We thank the reviewer for these comments and have added this information to the methods section.

M 5. Please see section 2.2, page 7: “Gait velocity was measured using custom made infrared detection gates placed 2 meters apart over the platform.”

Please see section 2.4, page 9: “A trial was successful if the subjects did not alter their gait pattern to target the platform, which was assessed by the research staff.”
R 6. Methods section: Formatting requires cross checking to JFAR conventions through this section particularly where p values are reported (paper uses upper case P instead of lower case p) and referrals to tables require additional spaces.

A 6. We thank the reviewer for these suggestions and have adjusted the formatting of the manuscript as suggested.

R 7. Discussion, paragraph 6: There is a typographical error, should read ‘plantar’ not ‘planter’.

A 7. We thank the reviewer for this correction and have changed it in the revised manuscript.

M 7. Please see section 4, page 16.

Discretionary Revisions
R 1. Consider adding the following (or similar) as keywords: diabetic peripheral neuropathy and strength.

A 1. The suggested keywords indeed describe the manuscript, but as these are already in the title we opted not to add them to the keywords list.

R 2. The heading in the Methods section, 2.4 Outcomes and follow-up, could be more descriptive of the content covered under that heading. As a suggestion, this section could be broken into two sections. The first could be called 2.4 Measurement Protocol and the second might then be, 2.5 Outcome Measures.

A 2. We thank the reviewer for this advice and have changed section 2.4 outcome measures and follow-up into section 2.4 Measurement protocol and 2.5 Data analysis.

M 2. Please see section 2.4 and 2.5 on page 9

Level of interest: An article of importance in its field
Quality of written English: Acceptable
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 below