PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	PROGRESSIVE RESISTANCE TRAINING IN PARKINSON'S DISEASE. A SYSTEMATIC REVIEW AND META-ANALYSIS
AUTHORS	Saltychev, M; Bärlund, Esa; Paltamaa, Jaana; Katajapuu, Niina; Laimi, Katri

VERSION 1 - REVIEW

REVIEWER	Barbara Strasser Institute for Nutritional Sciences and Physiology, UMIT, Austria
REVIEW RETURNED	05-Jul-2015

GENERAL COMMENTS	The present manuscript has several problems: Introduction: The authors should point out the need of a meta-analysis since most of the included studies are of week quality with insufficient data. Furthermore, the authors should address the pathogenesis of disease and why resistance training might have a positive effect on disease outcomes. Methods: More extensive information on inclusion and exclusion criteria, assessed outcomes, data extraction, and statistical analysis. Results: More extensive information on the interventions including the intensity (%1RM, or repetition maximum), dose (sets per muscle group per week) and frequency (sessions per week) of resistance training, and the duration (minutes per session). The authors should also include muscle strength as an outcome. Discussion: The authors should describe shortly possible mechanism why RT does not work in PD. Further, describe limitations of the included studies and of the present systematic review.

REVIEWER	Serene Paul The George Institute for Global Health Australia
REVIEW RETURNED	14-Jul-2015

GENERAL COMMENTS	This prospectively registered systematic review with meta-analysis investigates the effectiveness of progressive resistance training (PRT) for people with Parkinson's disease (PD). This is an area of interest for clinicians working with people with PD.

	Major comments: 1. Introduction, 2nd paragraph. Please distinguish the systematic reviews (Brienesse & Emerson 2013, Lima et al 2013) from the narrative reviews (Falvo et al 2008, David et al 2012) given that the former use methods to minimise bias and therefore provide more robust evidence. 2. Introduction. Please clarify what constitutes evidence of effectiveness versus evidence for ineffectiveness. This would assist the reader to interpret the results and follow the discussion. 3. Introduction and Methods, pp6-7. The PICO criteria are repeated in consecutive paragraphs; please remove one. 4. Methods. The definition of the intervention and comparison outlined in the PICO criteria on pp6-7 do not appear to have been correctly applied to two of the included studies: Shulman et al 2013 and Combs et al 2013. a. For Shulman's study, the two other groups participated in aerobic training; one arm a low-intensity intervention that progressed duration, the other arm a high-intensity intervention that progressed duration and intensity. Since neither of these groups participated in PRT, the results of both these control groups should have been included in the meta-analysis. b. For Comb's study, both groups included a strengthening (PRT) component. Based on descriptions of the boxing intervention as summarised by these authors and described by Combs et al 2011 it appears that there is less of a PRT component in the boxing intervention than the control intervention. Should this study not have been excluded given that the comparison does not fit the listed criteria? 5. Results, p10, 3rd paragraph. Line 50 of p10 states "of 12 studies, 11 reported positive effect…" yet Table 1 shows that only 10 studies reported a positive effect. Please correct this discrepancy. 6. Results, p11, lines 8-12. While it is true that the inability to blind participants and/or therapists is a common problem in RCTs investigating PRT, it would be helpful to discuss why this is the case in the Discussion. 7. Results, p12, line 22. The Timed Up and Go test is measured in s, not m/s. Please therefore clarify what is the previously published minimal detectable change for the TUG. 8. Discussion, p13 1st paragraph. Please clarify what is meant by "without reasonable follow-ups after the end of treatment". 9. Discussion, p14 lines 22-24 and 29. Please discuss the relevance and importance of following up on the effects of PRT after cessation of training. Do the authors expect that the effects of PRT will persist following cessation of such training? If so, for how long? 10. Discussion, p13, lines 34-38. This sentence is confusing. Please also clarify how this conclusion was reached. 11. Discussion. Please discuss the clinical implications of the current findings on the management and rehabilitation of people with PD. Should rehabilitation providers continue to use PRT in the management of their PD clients? Why or why not? Minor comments: 1. Methods, study selection. It would be helpful to list the initials of the two independent reviewers and that of the third reviewer who was called to resolve disagreements. 2. Results, p10, line 52. Please list the references for the six studies which compared PRT with 'usual activities'. 3. Please correct spelling of 'gate' throughout the manuscript. It should be 'gait'. 4. Abstract: Please clarify that the meta-analyses were conducted on

	6 of the 12 included studies. References Brienesse LA and Emerson MN. Effects of resistance training for people with Parkinson's disease: a systematic review. J Am Med Dir Assoc 2013;14(4): 236-241. Combs SA, Diehl MD, Chrzastowski C, et al. Community-based group exercise for persons with Parkinson disease: a randomized controlled trial. NeuroRehabilitation 2013;32(1): 117-124. Combs SA, Diehl MD, Staples WH, et al. Boxing training for patients with Parkinson disease: a case series. Phys Ther 2011;91(1): 132-142. David FJ, Rafferty MR, Robichaud JA, et al. Progressive resistance exercise and Parkinson's disease: a review of potential mechanisms. Parkinson's Disease 2012;2012: 124527. Falvo MJ, Schilling BK and Earhart GM. Parkinson's disease and resistive exercise: rationale, review, and recommendations. Movement Disorders 2008;23(1): 1-11. Lima LO, Scianni A and Rodrigues de Paula F. Progressive resistance exercise improves strength and physical performance in people with mild to moderate Parkinson's disease: a systematic review. Journal of Physiotherapy 2013;59(1): 7-13. Shulman LM, Katzel LI, Ivey FM, et al. Randomized clinical trial of 3 types of physical exercise for patients with Parkinson disease. JAMA Neurol 2013;70(2): 183-190.

REVIEWER	Melissa Raymond Caulfield Hospital, Alfred Health Melbourne Australia
REVIEW RETURNED	24-Jul-2015

GENERAL COMMENTS	Please clarify how your systematic review is different to the other recently completed reviews that you have referenced? I am aware that all aren't systematic reviews, however this is not evident to the reader without further accessing the referenced papers. Please discuss in context as the reader may question why you chose to do another systematic review when other recent very reviews have been published. Why did you feel that you found more results than previous recent studies ? Out of interest, why were some studies not included in your review included in the other review studies ? Exclusions of 6 papers – please explain 'outcome diversity' further – meta-analysis may be difficult with few data on outcomes, however may highlight areas for further research. What outcomes did these papers include that you excluded? Page 5 – Please include references in evidence for lines 15-20 Page 7 line 10 – 'small number of repetitions' - please specify. 'Small' number is subjective and non-specific, 20 repetitions may be considered not small but if progressed to maintain intensity may still be considered low intensity progressive resistance strength training. Page 10 – line 50 perhaps the word you may be meaning is frequency not intensity.

	Page 10 - line 57- appears to be word missing after treadmill e.g. walking Page 12 – line 22 –the units for the previously suggested level of minimal detectable change are different to the pooled effect size units, thus not enabling the reader to draw any conclusions from this. Looking directly at the Steffan study (ref 23 cited) perhaps this should be 11 seconds? Discussion – page 13 – line 15 ‘gate’ – perhaps 'gait' or ‘measures of gait’? Page 13 – line 24 – what were the limits? ‘almost without limits’? Page 14 – line 3 – consider ‘disease severity’ versus ‘disease’s severity’ Table 3- please check raw mean difference data – e.g Combs 2013 walking speed. Some values may be missing the negative signs or be calculated incorrectly. Table 3 – figure inserted into table – is this accidental? Otherwise unreadable at A4 page size. PRST and vitamins versus vitamins – in your PICO statement, you stated that it your Intervention would be PRST versus nothing/placebo/intervention, but not PRST PLUS another intervention versus another intervention. Can you please make this clearer or state why you made exception to your initial statement. 2010 Allen paper (as per summary table)– please comment on the potential for delay in progression for exercises may not have been not progressed for 4 weeks? Discussion – you may consider commenting on importance of intensity of strength training as this appears to have varied between studies. Page 32- comfortable walking speed effect size tables – consider showing all of confidence intervals in graph as you have for the large intervals of 6 minute walk. Consider altering graph for fast walking speed to same x axis units if you are wanting to keep the same units on the x-axis. Please note if you have/have not pooled the two results from the same study (Ref 17) in your effect size graph – this may need to be reviewed by statistician as to whether this is statistically appropriate (i.e. if you have pooled 2 samples from the same study in your meta-analysis).

REVIEWER	Rod Taylor University of Exeter UK
REVIEW RETURNED	29-Aug-2015

GENERAL COMMENTS	The paper as written has two major problems 1. The author's definition of resistance training is poor and need better precision. As a consequence it is not clear that the studies included in this systematic review appropriately capture the full body of RCT evidence in this area 2. scope of outcomes considered - the review focuses on measures of exercise capacity and misses what are key outcomes for patients - quality of life, activities of daily living and policy maker - healthcare utilisation The write up has a number of areas that require improvement

	- abstract: no mention of RCTs, dates of search cut off, effect sizes not defined - introduction: pico should be methods not hear - methods: clarify study design inclusion, why random effects meta-analysis?, need to systematic review of within person correlation coeff,

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 (Dr. Barbara Strasser)

Comment 1

Introduction: The authors should point out the need of a meta-analysis since most of the included studies are of week quality with insufficient data. Furthermore, the authors should address the pathogenesis of disease and why resistance training might have a positive effect on disease outcomes.

Discussion: The authors should describe shortly possible mechanism why RT does not work in PD. Further, describe limitations of the included studies and of the present systematic review.

Response 1

We think that when conducting a systematic review, it is important to distinguish study questions that belong to an original research from those of a review. In a systematic review, we are not seeking for an answer to the question “does particular intervention work and, if so, then why?” but rather to the question “what evidence exists on the effects of a particular intervention and what is the quality of that evidence?” We think that explanation on potential reasons of effectiveness or ineffectiveness of an intervention as well as the pathogenesis of a disease may be usually irrelevant for the purpose of a systematic review. They should be barely touched focusing on seeking for evidence existing, analysing its possible “black spots”, and suggesting explanations for the strengths and weaknesses of that evidence.

We have now modified the “Weaknesses” section of the Discussion as follows:

“The case and control treatments, as well as intensity, duration, and frequency of PRT, employed in the selected studies were diverse and sometimes hardly comparable. The included studies have been conducted on relatively small samples and the effects were followed up for only a few months at most. In this review ‘small number of repetitions’ was defined according to the classic work of DeLorme and Watkins in 1948. The use of more precise definition given by the American College of Sports Medicine, defining ‘small number of repetitions’ as <12 repetitions, might alter our results. Due to uncertain definition of PRT, it is possible that some relevant studies remained undetected.”

Comment 2

Methods: More extensive information on inclusion and exclusion criteria, assessed outcomes, data extraction, and statistical analysis.

Response 2

While we have tried to keep our manuscript balanced, even in its present form, the Methods section is turned out to be almost as long as the Results. Before we consider extending the text, would it be possible to get more specified suggestions what kind of additional information may be needed in the Methods section?

Comment 3

Results:

a) More extensive information on the interventions including the intensity (%1RM, or repetition maximum), dose (sets per muscle group per week) and frequency (sessions per week) of resistance training, and the duration (minutes per session).
b) The authors should also include muscle strength as an outcome.

Response 3

a) All these data are presented in Table 1. We have tried to avoid repeating the presentation of the results in both text and tables.
b) We agree that such analysis might be of interest. Unfortunately, we had to leave some potentially important outcome measures outside the present review. Even in its present form, our manuscript contents five tables and six graphs. However, we have now added a sentence to the end of the Discussion section as follows:
"Further studies may also reveal the effects of resistance training on such important outcome measures as quality of life, activities of daily living, cost-effectiveness, and muscle strength left out of the scope of this review."

Reviewer 2 (Dr. Paul Serene)

Comment 1

Introduction, 2nd paragraph. Please distinguish the systematic reviews (Brienesse & Emerson 2013, Lima et al 2013) from the narrative reviews (Falvo et al 2008, David et al 2012) given that the former use methods to minimise bias and therefore provide more robust evidence.

Response 1

We have now changed the order of reviews' presentations in the text as follows:

"In a recent systematic review by Brienesse et al. of five randomized controlled trials (RCT), PRT was found to have a positive effect on muscle strength, mobility, endurance, fat free mass, and performance in functional tasks. Another recent systematic review by Lima et al. of four controlled trials suggested that PRT could be effective in increasing walking capacity in Parkinson's disease. Narrative review by David et al. reported favorable effect of PRT on muscle strength and function and non-motor symptoms of Parkinson's disease. Also narrative review by Falvo et al. emphasized the lack of robust data on the topic."

Comment 2

Introduction. Please clarify what constitutes evidence of effectiveness versus evidence for ineffectiveness. This would assist the reader to interpret the results and follow the discussion.

Response 2

We have now modified the text of Introduction as follows:

"While Lima et al. and David et al. ended up with a strong conclusion that there is evidence that progressive resistance training should be implemented in Parkinson's disease rehabilitation, the conclusions of Brienesse et al. and Falvo et al. indicated more cautiously that data are insufficient to make robust recommendations and further research is needed."

Comment 3

Introduction and Methods, pp6-7. The PICO criteria are repeated in consecutive paragraphs; please remove one.

Response

The typo has now been corrected and PICO criteria are presented in the text of Methods only.

Comment 4

Methods. The definition of the intervention and comparison outlined in the PICO criteria on pp6-7 do not appear to have been correctly applied to two of the included studies: Shulman et al 2013 and Combs et al 2013.

a. For Shulman's study, the two other groups participated in aerobic training; one arm a low-intensity intervention that progressed duration, the other arm a high-intensity intervention that progressed duration and intensity. Since neither of these groups participated in PRT, the results of both these control groups should have been included in the meta-analysis.

b. For Comb's study, both groups included a strengthening (PRT) component. Based on descriptions of the boxing intervention as summarised by these authors and described by Combs et al 2011 it appears that there is less of a PRT component in the boxing intervention than the control intervention. Should this study not have been excluded given that the comparison does not fit the listed criteria?

Response 4

a) In the trial by Shulman et al., there were three groups involved: 1) high-intensity treadmill; 2) low-intensity treadmill; and 3) stretching and resistance. In our opinion, the high-intensity treadmill training used in that study represented the different type of progressive training that is not consistent with our PICO criteria. Therefore, only group 2 and 3 were included in the analysis. We have referred to that consideration in the footnotes of Table 1.
b) The paper by Combs et al. mentioned the progression of exercise intensity only for boxing group. Therefore, we considered that study consistent with our PICO criteria.

Comment 5

Results, p10, 3rd paragraph. Line 50 of p10 states "of 12 studies, 11 reported positive effect…" yet Table 1 shows that only 10 studies reported a positive effect. Please correct this discrepancy.

Response 5

We have now corrected the sentence as follows:

"Of 12 studies, 10 reported positive effect of intervention."

Comment 6

Results, p11, lines 8-12. While it is true that the inability to blind participants and/or therapists is a

common problem in RCTs investigating PRT, it would be helpful to discuss why this is the case in the Discussion.

Response 6

We agree that this is a very important point. We have now added the following text to the Discussion along with a new reference:

"The most common source of potential systematic bias in the selected studies was the lack of blinding of participants and personnel. This source of bias is hardly avoidable when physical therapy is involved as the involvement bases on the close participation of both a patient and a therapist in the entire chain of planning, performing, and assessing the intervention. While it is barely preventable, it could be statistically controlled, for example, by using repeated measures of expectancy and beliefs about the demands of the research throughout the trial."

[Button KS, Munafo MR. Addressing risk of bias in trials of cognitive behavioral therapy. Shanghai Arch Psychiatry 2015;27(3):144-8.]

Comment 7

Results, p12, line 22. The Timed Up and Go test is measured in s, not m/s. Please therefore clarify what is the previously published minimal detectable change for the TUG.

Response 7

We have now corrected the "m/s" to "s". The minimal detectable change for the TUG has been corrected to 3.5 s and a new reference has been added

[Huang SL, Hsieh CL, Wu RM, et al. Minimal detectable change of the timed "up & go" test and the dynamic gait index in people with Parkinson disease. Physical therapy 2011;91(1):114-21.]

Comment 8

Discussion, p13 1st paragraph. Please clarify what is meant by "without reasonable follow-ups after the end of treatment".

Response 8

We have now modified the sentence as follows:

"Few studies conducted on small sample sizes with short periods of follow-up reported some positive

effects of PRT on freezing symptoms, gate, cognitive performance, and muscle strength."

Comment 9

Discussion, p14 lines 22-24 and 29. Please discuss the relevance and importance of following up on the effects of PRT after cessation of training. Do the authors expect that the effects of PRT will persist following cessation of such training? If so, for how long?

Response 9

We have now modified the text as follows:

"Unexpectedly, none of trials followed the effects of PRT more than 1 month after the end of a supervised training program. It has been previously reported that beneficial effects of training may persist for several months after the cessation of training."

[Ellis T, de Goede CJ, Feldman RG, et al. Efficacy of a physical therapy program in patients with Parkinson's disease: a randomized controlled trial. Archives of Physical Medicine and Rehabilitation 2005 Apr;86(4):626-632 2005.]

Comment 10

Discussion, p13, lines 34-38. This sentence is confusing. Please also clarify how this conclusion was reached.

Response 10

We have now modified the sentence as follows:

"We ended up, however, with more robust conclusion that, based on several small-sample good-quality RCTs, there is limited evidence on PRT being no more effective in Parkinson's disease than other physical training schemes."

Comment 11

Discussion. Please discuss the clinical implications of the current findings on the management and rehabilitation of people with PD. Should rehabilitation providers continue to use PRT in the management of their PD clients? Why or why not?

Response 11

We have now added a sentence as follows:

"As there is no evidence on the superiority or better safeness of one specific training scheme over another in patients with PD, rehabilitation providers may include or avoid PRT depending on the settled practice and costs of a particular rehabilitation program."

Comment 12

Methods, study selection. It would be helpful to list the initials of the two independent reviewers and that of the third reviewer who was called to resolve disagreements.

Response 12

The initials have now been added as suggested by the reviewer.

Comment 13

Results, p10, line 52. Please list the references for the six studies which compared PRT with 'usual activities'.

Response 13

The references have now been listed as suggested by the reviewer.

Comment 14

Please correct spelling of 'gate' throughout the manuscript. It should be 'gait'.

Response

The typos have now been corrected.

Comment 15

Abstract: Please clarify that the meta-analyses were conducted on 6 of the 12 included studies.

Response 15

The text has now been modified as suggested by the reviewer:

"Results: Of 516 records, 12 were considered relevant. Nine of them had low risk of bias. All studies were randomized controlled trials conducted on small samples with none or one-month follow-up after

the end of intervention. Of them, six were included in quantitative analysis. Pooled effect sizes of meta-analyses on fast and comfortable walking speed, 6-min walking test, Timed Up and Go test and maximal oxygen consumption were below the level of minimal clinical significance."

Reviewer 3 (Dr. Melissa Raymond)

Comment 1

a) Please clarify how your systematic review is different to the other recently completed reviews that you have referenced? I am aware that all aren't systematic reviews, however this is not evident to the reader without further accessing the referenced papers. Please discuss in context as the reader may question why you chose to do another systematic review when other recent very reviews have been published.
b) Why did you feel that you found more results than previous recent studies ?
c) Out of interest, why were some studies not included in your review included in the other review studies ?

Response 1

a) Please, see our response #1 to the reviewer #2. We have now modified the text of Introduction distinguishing systematic and narrative reviews.
b) Brienesse et al. analysed 5 studies. Lima et al. analysed 4 studies. We analysed 12 studies.
c) When conducting a systematic review, maybe the most important point is a strict adherence to the pre-agreed PICO criteria and protocol of selection. If so, then different research teams end up with different lists of selected papers when PICO and protocol are even slightly different. If the reviewer would like, we would be happy to provide her with additional information about the exclusion reasons of some particular studies.

Comment 2

Exclusions of 6 papers – please explain 'outcome diversity' further – meta-analysis may be difficult with few data on outcomes, however may highlight areas for further research. What outcomes did these papers include that you excluded?

Response 2

A sentence has now been added to the end of the Discussion as follows:

"Further studies may also reveal the effects of resistance training on such important outcome measures as quality of life, activities of daily living, cost-effectiveness, and muscle strength left out of the scope of this review."

Comment 3

Page 5 – Please include references in evidence for lines 15-20

Response 3

The references have been included

[3. Taylor NF, Dodd KJ, Damiano DL. Progressive resistance exercise in physical therapy: a summary of systematic reviews. Physical therapy 2005;85(11):1208-23.

4. Valenzuela T. Efficacy of progressive resistance training interventions in older adults in nursing homes: a systematic review. Journal of the American Medical Directors Association 2012;13(5):418-28.]

Comment 4

Page 7 line 10 – 'small number of repetitions' - please specify. 'Small' number is subjective and non-specific, 20 repetitions may be considered not small but if progressed to maintain intensity may still be considered low intensity progressive resistance strength training.

Response 4

We agree with the reviewer. "Small number" is indeed non-specific. In our review, we have used this criterion as described by DeLorme and Watkins. Probably the more precise criterion suggested by the American College of Sports Medicine could also be used as well. Unfortunately, at this point, the PICO cannot be altered due to the demands of systematic approach. However, we have now added a sentence to the Discussion as follows:

"In this review 'small number of repetitions' was defined according to the classic work of DeLorme and Watkins in 1948. The use of more precise definition given by the American College of Sports Medicine, defining 'small number of repetitions' as <12 repetitions, might alter our results."

Comment 5

Page 10 – line 50 perhaps the word you may be meaning is frequency not intensity.

Response 5

The word 'intensity' has now been changed to 'frequency'.

Comment 6

Page 10 - line 57- appears to be word missing after treadmill e.g. walking

Response 6

We have now added 'training' at the end of sentence.

Comment 7

Page 12 – line 22 –the units for the previously suggested level of minimal detectable change are different to the pooled effect size units, thus not enabling the reader to draw any conclusions from

this. Looking directly at the Steffan study (ref 23 cited) perhaps this should be 11 seconds?

Response 7

We apologize for the typo. The text has now been corrected as follows:

"The effect of PRT on Timed Up and Go test was assessed by pooling samples of three studies. The size of pooled sample was 26 cases versus 29 controls. The pooled effect size was statistically insignificant -0.71 s (95%CI -1.47 to 0.06) in favor of intervention and below the minimal detectable change of 3.5 s previously suggested".

Comment 8

Discussion – page 13 – line 15 'gate' – perhaps 'gait' or 'measures of gait'?

Response 8

All typos with 'gate' have now been corrected to 'gait'.

Comment 9

Page 13 – line 24 – what were the limits? 'almost without limits'?

Response 9

We agree that this may sound unclear and decided to shorten the sentence as follows:

"However, we used very wide search clauses and performed the rest of the search and selection manually in order to avoid missing the potentially relevant reports."

Comment 10

Page 14 – line 3 – consider 'disease severity' versus 'disease's severity'

Response 10

The text has been changed as suggested by the reviewer.

Comment 11 Table 3- please check raw mean difference data – e.g Combs 2013 walking speed. Some values may be missing the negative signs or be calculated incorrectly.
Response 11 We have checked the original article by Combs et al. It seems that our Table 3 contains the same figures on TUG as the Table 2 in the article by Combs et al.

Comment 12 Table 3 – figure inserted into table – is this accidental? Otherwise unreadable at A4 page size.
Response 12 Yes, it was an accident and it has now been corrected.

Comment 13 PRST and vitamins versus vitamins – in your PICO statement, you stated that it your Intervention would be PRST versus nothing/placebo/intervention, but not PRST PLUS another intervention versus another intervention. Can you please make this clearer or state why you made exception to your initial statement.
Response 13 We think that including that study is not controversial regarding our PICO. We have now added a text to the Results as follows: *“It is self-evident that most of the patients with PD have more than one treatment. Thus, when comparing PRT and vitamins against only vitamins, omitting vitamins was accepted by us as approximation and the study by DiFrancisco-Donoghue et al. could be included into this review”.*

Comment 14 2010 Allen paper (as per summary table)– please comment on the potential for delay in progression for exercises may not have been not progressed for 4 weeks?
Response 14

We apologize for not understanding the comment. Would it be possible to get a clarification on that?

Comment 15

Discussion – you may consider commenting on importance of intensity of strength training as this appears to have varied between studies.

Response 15

We have now added a sentence to the Discussion as follows:

"The case and control treatments, as well as intensity, duration, and frequency of PRT, employed in the selected studies were diverse and sometimes hardly comparable."

Comment 16

Page 32- comfortable walking speed effect size tables – consider showing all of confidence intervals in graph as you have for the large intervals of 6 minute walk. Consider altering graph for fast walking speed to same x axis units if you are wanting to keep the same units on the x-axis.

Response 16

We have considered the suggestions of the reviewer. We would like, however, to preserve the current form of our presentation.

Comment 17

Please note if you have/have not pooled the two results from the same study (Ref 17) in your effect size graph – this may need to be reviewed by statistician as to whether this is statistically appropriate (i.e. if you have pooled 2 samples from the same study in your meta-analysis).

Response 17

In that study, 36 subjects were assigned to 4 groups: 1) vitamins, 2) vitamins + exercise, 3) exercise, and 4) none. We have calculated 2 effect sizes for that study: 1) group 1 vs. group 2 and 2) group 3 vs. group 4.

Reviewer 4 (Dr. Rod Taylor)

Comment 1
The author's definition of resistance training is poor and need better precision. As a consequence it is not clear that the studies included in this systematic review appropriately capture the full body of RCT evidence in this area
Response 1
Unfortunately, there is no precise and standardized definitions of resistance training existed. We have mentioned this fact in our “Strengths and weaknesses” section: *“Due to uncertain definition of PRT, it is possible that some relevant studies remained undetected.”*

Comment 2
Scope of outcomes considered - the review focuses on measures of exercise capacity and misses what are key outcomes for patients - quality of life, activities of daily living and policy maker - healthcare utilisation
Response 2
We have now added a sentence to the end of the Discussion section as follows: *“Further studies may also reveal the effects of resistance training on such important outcome measures as quality of life, activities of daily living, cost-effectiveness, and muscle strength left out of the scope of this review.”*

Comment 3
- abstract: no mention of RCTs, dates of search cut off, effect sizes not defined
Response 3
The Abstract has now been modified as suggested by the Reviewer: *“Design: Systematic review and meta-analysis. Data sources: Central, Medline, Embase, Cinahl, Web of Science, Pedro* ***until May 2014****.* ***Randomized controlled or controlled clinical trials.*** *The methodological quality of studies was assessed according to the Cochrane Collaboration’s domain-based evaluation framework. Data synthesis: random effects meta-analysis with test for heterogeneity*

*using the I² and pooled estimate as **raw mean difference**."*

Comment 4

- introduction: pico should be methods not hear

Response 4

Pico criteria have now been placed in the Methods section.

Comment 5

- methods: clarify study design inclusion, why random effects meta-analysis?, need to systematic review of within person correlation coeff,

Response 5

Please, see out modified text in "Data synthesis and analysis" section:

"We used a random effects meta-analysis to quantify the pooled effect size of included studies as a more natural choice than fixed effects in the context of multiple clinical trials conducted in diverse settings. In addition, test for heterogeneity supported this choice."

We apologize for not understanding the comment on "…need to systematic review of within person correlation coeff…" Would it be possible to get some clarification on that?