Article:

Archives of Physical Medicine and Rehabilitation

Matching the outcomes to treatment targets of exercise for low back pain: does it make a difference? Results of secondary analyses from individual patient data of randomised controlled trials and pooling of results across trials in comparative meta-analyses

--Manuscript Draft--

Manuscript Number:	ARCHIVES-PMR-D-22-00146R2
Article Type:	Original Research
Keywords:	Low back pain; Exercise; treatment targets; secondary analysis; randomised controlled trials; composite outcomes
Corresponding Author:	Lianne Wood, Ph.D Nottingham University Hospitals NHS Trust UNITED KINGDOM
First Author:	Lianne Wood, Ph.D
Order of Authors:	Lianne Wood, Ph.D
	Nadine E Foster, DPhil
	Martyn Lewis, PhD
	Gert Bronfort, PhD
	Erik J Groessl, PhD
	Catherine E Hewitt, PhD
	Gisela C Miyamoto, PhD
	Silje E Reme, PhD
	Annette Bishop, PhD
Abstract:	Complex interventions, like exercise for non-specific low back pain (NSLBP), have many treatment targets. In randomised controlled trials (RCTs), matching the primary outcome to the exercise target(s) may provide greater standardised mean differences (SMDs) than using unmatched primary outcomes. Objective These secondary analyses of previous RCTs aimed to explore whether using a single matched or composite outcome might impact the results of previous RCTs testing exercise for NSLBP. The first objective was to explore whether a single matched outcome generated a greater SMD when compared to the original unmatched primary outcome SMD. The second objective was to explore whether a composite measure, comprised of matched outcomes, generated a greater SMD when compared to combining the original primary outcome SMD. Design, Setting and Participants We conducted exploratory secondary analyses of data from 1) five RCTs (n=1,033) that used an unmatched primary outcome but included (some) matched outcomes as secondary outcomes, and 2) four RCTs (n=864) that included multiple matched outcomes by developing composite outcomes. Intervention: Exercise compared to no exercise. Main Outcome Measures: The composite consisted of standardised averaged matched outcomes. All analyses replicated the RCTs' primary outcome analyses. Results Of five RCTs, three had greater SMDs with matched outcomes (pooled effect SMD 0.30 (95% CI 0.04, 0.56), p=0.02) compared to an unmatched primary outcome (pooled effect SMD 0.19 (95% CI -0.03, 0.40) p=0.09). Of four composite outcome

analyses, three RCTs had greater SMDs in the composite outcome (pooled effect SMD 0.28 (95%CI 0.05, 0.51) p=0.02) compared to the primary outcome (pooled effect SMD 0.24 (95%CI -0.04, 0.53) p=0.10).

Conclusion
These exploratory analyses suggest that using an outcome matched to exercise treatment targets in NSLBP RCTs may produce greater SMDs than an unmatched primary outcome. Composite outcomes could offer a meaningful way of investigating superiority of exercise than single domain outcomes.

Versus Arthritis Primary Care Centre

School of Medicine

Keele University

Keele

Staffordshire

ST5 5BG

l.wood2@keele.ac.uk

+44 7449732744

27-01-2022

Dr Leighton Chan and Dr Allen Heinemann

Editor-in-Chief

Archives of Physical Medicine and Rehabilitation

Dear Drs Chan and Heinemann,

Thank you for considering the included manuscript for publication in the Archives of Physical Medicine and Rehabilitation. The paper is entitled:

"Matching the outcomes to treatment targets of exercise for low back pain: does it make a difference? Results of secondary analyses from individual patient data of randomised controlled trials and pooling of results across trials in comparative meta-analyses."

Exercise is a core treatment for persistent non-specific low back pain, but the use of a single primary outcome may not be sufficient to capture the often multiple treatment targets identified within an exercise intervention. This paper describes the results of two secondary analyses of individual participant data from existing RCTs to explore whether firstly, matching the primary outcome to the identified treatment targets, and secondly, whether a composite matched outcome in comparison to the original primary outcome, may change the results and conclusions of existing RCTs in persistent non-specific low back pain. These results suggest that exercise prescribers and trial developers should consider the treatment targets of their exercise intervention when selecting the most appropriate outcome.

I hereby certify that this paper consists of original, unpublished work which is not under consideration for publication elsewhere. All authors have read and confirmed that specified requirements for co-authorship are fulfilled. All authors are listed, and have contributed significantly to this work.

Yours sincerely,

Lianne Wood (on behalf of the author team)

Dear Dr Rundell,

Ms. Ref. No.: ARCHIVES-PMR-D-22-00146

Title: Matching the outcomes to treatment targets of exercise for low back pain: does it make a difference? Results of secondary analyses from individual patient data of randomised controlled trials and pooling of results across trials in comparative meta-analyses

Archives of Physical Medicine and Rehabilitation

We are very grateful for the constructive feedback provided by the editors and each of the external reviewers of this manuscript. We feel our manuscript has been improved as a result, and hope you will agree. We have addressed the Reviewers' comments point-by-point below. We provide a clean and a highlighted version to demonstrate changes in the revised manuscript. Our responses below are shown in blue to distinguish from the Reviewers comments. Page numbers mentioned in responses refer to the manuscript version with highlighted changes.

We hope the revised manuscript is suitable for publication and look forward to hearing from you soon.

Yours sincerely

The author team

Reviewers' comments:

Reviewer #1: Thank you for providing more clarity on the study as requested. The authors did a nice job of addressing/adding to the details in the methodology and analysis.
In the Discussion - Implications for Clinicians and Researchers section, where it is stated: "We recommend that developers of exercise interventions consider logic models or programme development theory 36,37 in order to map and guide assessment of the mechanisms of action of their intervention, and the most likely outcomes to accurately measure the changes expected."...

1. Can the authors expand on this in more simple terms and provide an example of this approach? I would imagine that most of the clinical readership, and even many researchers would know very little about what the authors are referring to here.

Thank you for your comments. In response to this additional sentences have been added with an example figure to improve the understanding of the readership. "Previous intervention development has been exemplified by Hurley et al.37 and Kjaer et al.53 who provided detailed descriptions of their self-management and exercise programs (please see Figure 4 as an example program model), including the 'active' components of the intervention, the proposed determinants of change and the corresponding outcomes to capture the intended change. It should be noted that we do not suggest all RCTs need to consider this level of intervention development. However, considering the trial intervention through a visual model can help to alleviate research waste by ensuring capture of the most important outcomes, and may contribute to future knowledge of how these interventions may work." (In 379-388, page 17).

Figure 4: An example program model of the GLA:D Back intervention, the proposed patient achievements and the outcomes through the GLA:D Back program, and their theoretical links (reproduced from Kjaer et al., 2018, with permission)
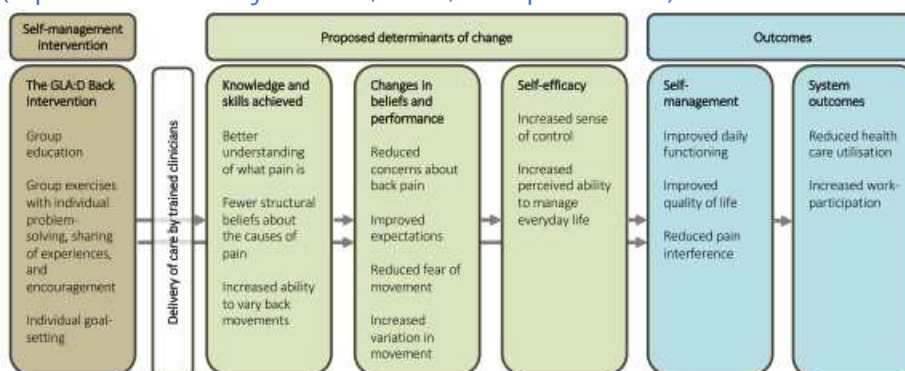
2. I would suggest being more specific in the tables with regard to the outcome of "Pain" - I appreciate that the authors are referring to Pain Intensity (VAS), rather than say Pain Interference, or Pain Behavior, however, I recommend being more complete/thorough. Likewise - for Physical Function, I am presuming this means self-reported/patient-reported physical function rather than observer-rated/physical capacity testing of physical function -

but again, would suggest being more explicit between what is 'self-reported' and what is 'physical capacity' based measures...as the tables are listed now, both of these forms of tests are intermingled in the list, and it would be more helpful to see these broken down into self-reported measures and physical capacity measures.

Thank you for this comment. To improve the ease of understanding the tables, we have separated self-reported outcome measures from objectively reported outcome measures, in Table 1. We have also clarified pain and physical function scores for all included trials in Tables 1, 2 and 3. Table 1 and 2 outcomes have been further clarified to distinguish between self-reported outcomes and objectively recorded outcomes to improve transparency.

3. Reviewer #2: Thank you for addressing my original comments. I have one further query. For Table 3 it appears for the Moffett et al 2006 trial the primary outcome and composites did not detect a significant change; however, it is reported in the final column that using the composite resulted in a change in results. Could you please confirm this is correct?

Many thanks for your comment and identifying this error! Table 3 has been amended to reflect that "no change" occurred in the results of the first two matched trials. This has also been reinforced in the accompanying text as follows: "Three of the four analyses showed results with the composite outcome variable that had greater SMDs in favour of the exercise intervention25,26,28, **of which two 25,28** were (more) statistically significant in comparison to the original RCTs' primary outcome results. All analyses showed a smaller standard error when using the composite outcome. The use of the co-primary composite generated greater SMDs than the composite outcome. However, the co-primary composite generated greater SMDs (**not statistically significant**) than the primary outcome in one RCT,26 but this was not reproduced in the other RCT analysis.27"(ln 232-240 page 11)

1. **Running Head**: Matching the outcomes to treatment targets of exercise

2. **Title:** Matching the outcomes to treatment targets of exercise for low back pain: does it make a difference? Results of secondary analyses from individual patient data of randomised controlled trials and pooling of results across trials in comparative meta-analyses

3. **Authors:** Lianne Wood[1,2], PhD**;** Nadine E Foster[1*], DPhil; Martyn Lewis[1], PhD; Gert Bronfort[3] PhD; Erik J Groessl[4], PhD; Catherine Hewitt[5], PhD; Gisela C Miyamoto[6] PhD; Silje E. Reme, PhD[7] Annette Bishop[1], PhD.

4. **Authors Institutions at time of study:** [1]Primary Care Centre Versus Arthritis**,** School of Medicine, Faculty of Medicine, Keele University, Newcastle-under-Lyme, UK; [2]Nottingham University Hospitals NHS Trust, Queens Medical Centre, Derby Road, Nottingham, UK; [3]Earl E Bakken Centre for Spirituality and Healing, University of Minnesota, USA; [4]University of California San Diego, Herbert Wertheim School of of Public Health and UCSD Health Services Research Centre; [5]York Trials Unit, Department of Health Sciences, University of York, UK; [6]Master's and Doctoral Program in Physical Therapy, Universidade Cidade de São Paulo, São Paulo, Brazil; [7]Department of Psychology, University of Oslo, Oslo, Norway *NEF has changed affiliation to the STARS Education and Research Alliance, Surgical Treatment and Rehabilitation Service (STARS), The University of Queensland and Metro North Health, Herston, Brisbane, Queensland, Australia

5. **Previous Presentation:** The contents of this paper have been published as part of a doctoral thesis (examined by Viva November 2020), awarded June 2021; presented as a poster at The Society for Back Pain Research conference, Groningen, The Netherlands in November 2019.

7. **Conflicts of Interest:** There are none to declare.

8. **Corresponding author**: L Wood, email: l.wood2@keele.ac.uk, Present address: Nottingham University Hospitals NHS Trust, Queens Medical Centre, Derby Road, Nottingham, NG 2UH, UK; +441159249924 ext 86217

9. **Clinical trial registration numbers:**

This is a secondary analysis of the following clinical trials:

- Miyamoto et al (2018): NCT02241538

- Bronfort et al. (2011) No registration number

- Moffett et al. (2006): ISRCTN48919562

- Harris et al. (2017) No registration number

- Tilbrook et al. (2011) Protocol published, no registration number

- Groessl et al. (2017) NCT02524158

- Shirato et al. (2010) Protocol published, no trial registration.

10. **Author Contributions:**

The conceptualisation of this study was developed by AB, NEF, ML and LW;

Highlights

- Exercise has multiple proposed treatment targets. Few RCTs match their outcomes to these targets.

- These analyses suggest that outcomes matched to exercise treatment targets may produce greater SMDs than outcomes that are not matched to exercise treatment targets

- Composite outcomes may generate greater SMDs and less uncertain estimates

1   **Title:** Matching the outcomes to treatment targets of exercise for low back pain: does

2   it make a difference? Results of secondary analyses from individual patient data of

3   randomised controlled trials and pooling of results across trials in comparative meta-

4   analyses

5

6   **<u>Abstract</u>**

7

8   **Objective**

9   To explore whether using a single matched or composite outcome might impact the

10   results of previous randomised controlled trials (RCTs) testing exercise for non-

11   specific low back pain (NSLBP). The first objective was to explore whether a single

12   matched outcome generated a greater standardised mean differences (SMD) when

13   compared to the original unmatched primary outcome SMD. The second objective was

14   to explore whether a composite measure, comprised of matched outcomes, generated

15   a greater SMD when compared to the original primary outcome SMD.

16

17   **Design**

18   We conducted exploratory secondary analyses of data.

19

20   **Setting**

21   Seven RCTs were included, of which two were based in the USA (University research

22   clinic, Veterans Affairs medical centre) and the UK (primary care clinics, nonmedical

23   centres). One each were based in Norway (clinics), Brazil (primary care), and Japan

24   (outpatient clinics).

25

26 **Participants**

27 The first analysis comprised 1) five RCTs (n=1,033) that used an unmatched primary

28 outcome but included (some) matched outcomes as secondary outcomes, and the

29 second analysis comprised 2) four RCTs (n=864) that included multiple matched

30 outcomes by developing composite outcomes.

31

32 **Intervention:**

33 Exercise compared to no exercise.

34

35 **Main Outcome Measures:**

36 The composite consisted of standardised averaged matched outcomes. All analyses

37 replicated the RCTs' primary outcome analyses.

38

39 **Results**

40 Of five RCTs, three had greater SMDs with matched outcomes (pooled effect SMD

41 0.30 (95% CI 0.04, 0.56), p=0.02) compared to an unmatched primary outcome

42 (pooled effect SMD 0.19 (95% CI -0.03, 0.40) p=0.09). Of four composite outcome

43 analyses, three RCTs had greater SMDs in the composite outcome (pooled effect

44 SMD 0.28 (95%CI 0.05, 0.51) p=0.02) compared to the primary outcome (pooled effect

45 SMD 0.24 (95%CI -0.04, 0.53) p=0.10).

46

47 **Conclusions**

48 These exploratory analyses suggest that using an outcome matched to exercise

49 treatment targets in NSLBP RCTs may produce greater SMDs than an unmatched

50 primary outcome. Composite outcomes could offer a meaningful way of investigating

51 superiority of exercise than single domain outcomes.

52

53

56

57

Abbreviations:

NSLBP   non-specific low back pain

RCT   randomised controlled trial

SMD   standardised mean difference

ANOVA   analysis of variance

ANCOVA   analysis of covariance

WOMAC   Western Ontario and McMaster Universities Osteoarthritis Index

Introduction

Persistent non-specific low back pain (NSLBP) is the leading cause of disability globally,[1,2] with an estimated 540 million people worldwide experiencing NSLBP.[3] Therapeutic exercise is the most widely recommended treatment for persistent NSLBP[4,5] with moderate certainty evidence that it has clinically important benefits for pain but small benefits for function.[6–9]

Exercise is a complex intervention with numerous components, such as biological,[10] psychological and social,[11] as well as treatment interaction components.[12] Therefore, there may be multiple potential treatment targets, where a treatment target is defined as the goal or intention the treatment aims to influence.[13] Most randomised controlled trials (RCTs) of exercise for persistent NSLBP do not specify their treatment targets.[14] Literature regarding RCT design stipulates that the primary outcome should match the rationale of the intervention,[15,16] yet outcome measures are often selected based on core outcome domains[17] and/or patient preference. A recent systematic review[18]

82 demonstrated that most (74%) of the included RCTs of exercise in persistent NSLBP

83 used primary outcomes not reflective of the RCT's specified exercise treatment targets.

84 Further, most RCTs demonstrate only small differences between exercise and control

85 arms,[7] and therefore clinically important interventions may be overlooked, if these

86 benefits are related to the selection of the primary outcome.

87

88 In complex interventions, such as exercise, which frequently have more than one

89 treatment target, the selection of a single primary outcome measure may be insufficient

90 to capture the benefits that can be achieved.[19] Watt et al.,[19] suggest that nominating a

91 single primary outcome in a RCT of a complex intervention may distort the overall

92 purpose. Composite outcomes, including two or more component outcome domains,[20]

93 may be more suitable than a single primary outcome in such RCTs, and may be better

94 able to demonstrate the effects of complex interventions. In addition, more meaningful

95 results of exercise RCTs for persistent NSLBP may be derived. However, due to the

96 limited evidence on composite measures available for NSLBP, future research in this

97 area has been recommended.[21]

98

99 It is unknown whether using a matched primary outcome or composite outcome

100 (comprised of the specified treatment targets) might alter the findings of previous

101 RCTs.[22] This secondary analysis aimed to explore whether using a single matched or

102 composite outcome might impact the results of previous RCTs testing exercise for

103 persistent NSLBP. The first objective was to explore whether a single outcome,

104 matched to the identified exercise treatment targets, generated a greater standardised

105 mean difference (SMD) when compared to the original unmatched primary outcome

106 SMD. The second objective was to explore whether a composite measure, comprised

107 of more than one outcome matched to the identified exercise treatment targets,

108 generated a greater SMD when compared to the original primary outcome SMD.

109

110
111 **Methods**

112

113 *Design*

114 Exploratory secondary analyses of seven previous RCTs. A random effects meta-

115 analysis (generated with RevMan 5.3) was used to compare: i) the overall effect of

116 using an unmatched primary outcome with the first reported matched outcome, and ii)

117 the overall effect of using a single primary outcome (matched or unmatched) with a

118 composite (matched) outcome.

119

120 *Data Source*

121 A recently completed systematic review of RCTs of exercise interventions compared

122 to no exercise in persistent NSLBP[18] informed the RCT sample for this study.

123 Treatment targets were extracted verbatim from the RCT published texts, where it was

124 clear the authors had described a rationale for how the exercise intervention was

125 proposed to work, or what they had designed the exercise intervention to target. In the

126 review, RCTs were categorised into: a matched group, where the primary outcome

127 reflected one of the identified treatment targets; or an unmatched group, where the

128 primary outcome did not reflect one of the identified treatment targets. The matching

129 process was subjective and performed by pairs of independent reviewers, as described

130 in Wood et al.[18] For each analysis, the authors of the identified RCTs were contacted

131 and the dataset requested. The first analysis identified RCTs within the unmatched

132 group that included secondary outcomes matched to the treatment targets. The second

133 analysis identified RCTs within both the matched and unmatched groups, where more

134 than one outcome reflected more than one stated exercise treatment target.

135
136
137 *Data Extraction*

138 Information pertinent to these analyses was extracted as part of the systematic review

139 process[18] by pairs of independent reviewers (see appendix 1). The stated treatment

140 target(s) of the exercise intervention, the primary and secondary outcomes for each

141 RCT, the outcomes that matched the stated exercise treatment targets, and the method

142 of analysis performed on primary and secondary outcomes were extracted for each

143 RCT (see Table 1).

144

145 *Data Analysis*

146 Both Analyses:

147 SMDs and 95% confidence intervals were calculated for each primary and matched

148 secondary outcome for between-arm differences at the primary outcome time-point

149 designated by the trial authors, or if no primary time-point was specified by the authors,

150 then the earliest time-point post-exercise-intervention. SMD statistics for all between-

151 arm differences were reported as intervention minus control: positive SMDs indicating

152 higher values for the exercise intervention (lower for the control), and by contrast,

153 negative SMDs indicating lower values for the intervention (higher for the control).

154 Where some variables had point estimates scoring in the opposite direction to other

155 included variables, these were transformed so that all variables scored in the same

156 direction.[23,24]

157

158 For linear mixed models[25–28] the data were transformed from wide to long format by

159 transforming the variables to cases and computing a new variable consisting of all

160 time-points relevant to that outcome. All outcomes of interest were converted to a

161 standardised variable (standardised z-score). Initial analyses aimed to replicate the

162 published data used for the primary outcome(s) and/or targeted secondary outcomes

163 where possible to do so. The replicated analysis was applied to the matched

164 secondary outcome(s). Linear mixed model analyses include all time-points available

165 for the relevant outcome. Therefore values for all available time-points for the matched

166 secondary outcomes were also used and reported[25–28].

167

168 <u>Second Analysis Only:</u>

169 The second analysis created a composite outcome, comprised of multiple outcomes

170 matched to the specified exercise treatment targets. For the creation of the composite

171 outcome, standardised composite outcomes were derived by computing a new

172 variable of the mean of the standardised outcome scores, matched to the treatment

173 targets, for each time-point.[29] A further analysis was performed where two primary

174 outcomes were specified, and both were matched to the treatment targets: a co-

175 primary composite was developed by creating a new variable of the mean of the

176 standardised primary outcomes at each time point. Exploratory analysis compared the

177 results of the first nominated primary outcome in comparison to a targeted composite

178 outcome and the co-primary outcome composite. The method of analysis of between-

179 arm standardised differences replicated the initial primary time-point analysis. All

180 analyses used Statistical Package for Social Science (SPSS) Statistics 24.

181

Results

A summary of dataset acquisition and analysis is displayed in Figure 1, and details of included trials are presented in Table 1.

Figure 1: Processes of identification of suitable trials for inclusion and analysis

Table 1: Included Trial Datasets

First Analysis: The Difference between Matched and Unmatched Outcome SMDs

In the first analysis, lead authors from five RCTs[25,28,30–32] were contacted, and three datasets acquired. Two RCTs provided sufficient information within their published papers, resulting in five RCTs analysed (1,033 participants). Two RCTs compared yoga to usual care,[30] and a waitlist control,[28] three RCTs tested supervised exercise programs in comparison to a brief intervention[32], a home exercise and manipulative arm[25], and prescribed NSAIDS[31].

Of the five RCTs included, three had greater SMDs and statistical significance in favour of exercise compared to a control-arm when a matched secondary outcome was used in comparison to an unmatched primary outcome[25,28,31] (see Table 2). Of the three full datasets analysed, two demonstrated larger, statistically significant effects in favour of exercise with at least one matched secondary outcome at the primary time-point(s), compared to an unmatched primary outcome[25,28]. The analysis

205 of Harris et al.[32] did not demonstrate any statistically significant differences using any

206 of the outcomes, but the use of the matched secondary outcome generated a greater

207 SMD in favour of the exercise group than when using the unmatched primary outcome.

208 The analysis of Tilbrook et al.[30] was the only trial analysed to demonstrate greater

209 between-arm differences when using an unmatched primary outcome.

210

211 ***Table 2: First analysis results demonstrating the difference between matched***

212 ***and unmatched outcome SMDs***

213

214 The original results and secondary analyses of the five RCTs are summarised in Figure

215 2: a pooled SMD of 0.19 (95% CI -0.03, 0.40; p=0.09) was seen for the unmatched

216 primary outcome, in comparison to the SMD of 0.30 (95% CI 0.04, 0.56; p=0.02) for

217 the first reported matched outcome. The subgroup differences (primary outcome

218 compared to the first matched outcome) were not statistically significant (SMD 0.11;

219 95% CI -0.34, 0.57; p=0.51).

220

221 ***Figure 2: Forest plot to demonstrate the pooled effect of the SMD for unmatched***

222 ***primary outcomes in comparison to matched secondary outcomes***

223

224 Second Analysis: Composite SMD calculations in comparison to Primary Outcome

225 SMDs

226 In the second analysis, lead authors from seven RCTs[25–28,33–35] were contacted, and

227 four authors shared their datasets.[25–28] Four RCTs were analysed (864 participants):

228   one compared differing Pilates dosages plus advice versus advice alone,[27] one

229   compared yoga to a waitlist,[28] one tested supervised exercise programs in a home

230   exercise versus a manipulative arm,[25] and one compared McKenzie exercises versus

231   a physiotherapy intervention.[26] The composite outcomes varied in composition with

232   three composite outcomes formed of six outcomes[25–27] and one composite comprised

233   of three outcomes[28]. For example, Groessl et al.[28] measured the outcomes of strength,

234   flexibility and pain relief in their RCT which were matched to the treatment targets of

235   increasing strength and flexibility and improving pain tolerance. Please see Table 3 for

236   more detail regarding composition of composite outcomes.

237

238   The composite analysis impacted the results of three of four RCTs,[25,26,28] as seen in

239   Table 3. Three of the four analyses showed results with the composite outcome

240   variable that had greater SMDs in favour of the exercise intervention[25,26,28], of which

241   two[25,28] were (more) statistically significant in comparison to the original RCTs' primary

242   outcome results. All analyses showed a smaller standard error when using the

243   composite outcome. The use of the co-primary composite generated greater SMDs

244   than the composite outcome. However, the co-primary composite generated greater

245   SMDs (not statistically significant) than the primary outcome in one RCT,[26] but this was

246   not reproduced in the other RCT analysis.[27]

247

248   ***Table 3: Second analysis results of composite SMD calculations compared to***

249   ***primary outcome SMDs***

250

251 This is summarised in Figure 3 whereby a pooled SMD of 0.24 (95% CI -0.04, 0.53;

252 p=0.10) was seen for the primary outcome in comparison to the SMD of 0.28 (95% CI

253 0.05, 0.51; p=0.02) for the matched composite outcome. The subgroup differences

254 (primary outcome compared to matched composite) were not statistically significant

255 (SMD 0.03 (95% CI -0.13, 0.20) p=0.86).

256 **_Figure 3: Summary plot to demonstrate pooled SMD of primary outcome in_**

257 **_comparison to composite outcome_**

258

259 **Discussion**

260

261 The results of these exploratory secondary analyses of previous RCTs of exercise for

262 NSLBP suggest that it is possible that using a primary outcome matched to the

263 treatment targets of exercise may generate greater SMDs than a single unmatched

264 primary outcome. Further, using a composite outcome, matched to multiple exercise

265 treatment targets, may give greater power to detect superiority of exercise over a non-

266 exercise control. In three of five RCTs, a single matched outcome measure generated

267 a greater SMD than the original unmatched primary outcome SMD, and would impact

268 the results of four RCTs. In two of four RCTs, a composite matched outcome would

269 impact the results in favour of exercise versus control. Our analyses provide some

270 support for matching the primary outcome to the treatment targets of the exercise

271 intervention, and for considering the use of a composite outcome in comparison to a

272 single outcome when multiple exercise targets are identified. Using a matched outcome

273 may provide more clinically meaningful results, and will allow for identification of

274 treatment interventions that may be more effective than previously supposed.

13

275

Treatment targets may be described as intermediate variables or surrogate outcomes, as they may sit on the pathway to a patient relevant outcome such as pain or function. However, this may not always be the case, and the treatment targets reported by the authors of these RCTs may not have been based on clear programme development theory or logic modelling.[36,37] Many of the treatment targets identified by the RCT authors were captured by some of their outcomes, but there were no published intervention development or programme evaluation[38] papers for any of the included RCTs within which to test the degree that these treatment targets were indeed the focus of their intervention. Thus, it is difficult to identify which of the treatment targets may have been prioritised, or which may have been changed by the exercise interventions. In exercise, where multiple treatment targets are common, it is challenging without clear intervention theory, to understand how the exercise intervention may have exerted its effect. Heneghan et al.[39] caution against the use of surrogate outcomes as primary outcomes, without a clear understanding of the impact and effect of these upon patient-relevant outcomes. In the field of exercise and NSLBP, the effect surrogate outcomes have on important patient outcomes like pain, function and quality of life is poorly understood. Furthermore, there is a lack of understanding as to what mechanisms of effect underpin exercise interventions for NSLBP.[40,41]

The results of these exploratory secondary data analyses provide some support for considering the use of a composite matched outcome rather than a single unmatched outcome in trials of exercise for NSLBP. The results contrast with those from Parkes et al.[42] who compared a composite outcome (the Western Ontario and McMaster

Universities Osteoarthritis Index [WOMAC] score, pain and rescue medication) to a single outcome (WOMAC pain) in knee osteoarthritis. Their composite outcome demonstrated modest improvements in responsiveness when compared to WOMAC pain alone, but these were not statistically significant. While composite outcomes are uncommon as primary outcome measures in RCTs in the field of NSLBP, they are frequently used in cardiovascular medicine, and have both advantages and disadvantages. The use of a composite outcome can reduce the sample size,[43,44] which is beneficial both for the recruitment period and associated costs of RCTs.[45,46] However, in cardiovascular disease when a composite outcome included the outcome measures of most importance to patients, composite outcomes were less likely to demonstrate a moderate treatment effect.[46] Moreover, there is a risk of overestimation of treatment impact and effect when using composite outcomes if the component outcomes are not reported completely, leading to incorrect interpretation of the results.[39] If the use of composite outcomes is to be considered in NSLBP, composite outcomes would need to be chosen based on sound rationale. Furthermore, all outcomes selected to be included in the composite should individually be expected to demonstrate an important effect, as any outcome that does not will dilute the overall effect. Hence, composites make sense if the targeted outcomes all contribute to an important treatment effect and are responsive to change. This proposal is supported by our results that show the co-primary (matched) analysis produced the overall highest SMDs (greater than the composite).

Most RCTs of exercise for LBP appear to use a recommended core outcome domain[47] as a primary outcome.[18] Core outcome domains are necessary to allow for comparison of results across multiple datasets, and are useful for combined evidence approaches

325 such as meta-analysis. However, the authors of the LBP core outcome set highlight

326 that the agreed domains do not restrict measurement or the choice of primary outcome,

327 but "mandate collection and reporting of the core outcome set alongside the outcomes

328 of interest".[17] It could be argued that prioritising pain or back-related disability as the

329 primary outcome domain in RCTs testing exercise for persistent NSLBP may not

330 accurately reflect the benefits of exercise, if these outcome domains do not match the

331 range of treatment targets of the intervention. The challenge of outcome measure

332 selection is encapsulated by Coster et al.,[48] *"The ultimate value of a RCT …will be*

333 *directly tied to how well the selected outcome measure matches the researcher's*

334 *understanding of what he or she expects to change, to what degree it is expected to*

335 *change, over what period of time this change will happen and how that change can*

336 *best be identified".* As exercise is a complex intervention with multiple potential

337 treatment targets, there are multiple possible outcomes that could be used, but multiple

338 outcomes should be interpreted with caution.[49] The proposed treatment targets of the

339 intervention should influence the selection of the primary outcome, from which the

340 minimally important difference is used to calculate the sample size.[49] Literature

341 regarding RCT design stipulates that the primary outcome should match the rationale

342 of the intervention.[16,50] The results of this analysis suggest that matching the primary

343 outcome to the treatment targets of the intervention may generate greater SMDs in

344 favour of exercise, and that a composite outcome comprised of the most important

345 treatment targets could generate greater SMDs with smaller standard errors in favour

346 of exercise. A matched 'targeted' composite or single outcome may provide the RCT

347 team with the best chance of detecting the benefits of exercise compared to a control

348 or comparator, as well as providing a clear framework for future testing of how exercise

349 may potentially achieve its effects. This may have clinical implications given we have

350 limited understanding of what components or targets of exercise are most influential in

351 creating change in outcomes of importance.

352

353 **Strengths and Limitations**

354

355 This is the first study to explore the relationship between matched outcomes or

356 composite outcomes and the treatment targets of the exercise intervention in RCT

357 datasets of exercise for NSLBP. A strength of this study is the individual patient data

358 acquisition of seven previously published RCTs which allowed secondary analysis of

359 the data and generation of new composite variables. The analysis methods replicated

360 the primary analysis method used by the trial teams of the individual RCTs, and this

361 ensured the data were comparable, strengthening the results of this analysis. These

362 RCTs were selected from a sample of RCTs included in a systematic review,[18] which

363 may have been subject to publication bias. The main limitation is that this was an

364 exploratory secondary analysis of a small number of RCT datasets. SMDs were chosen

365 as a means to compare outcome estimates of different outcomes, but this may limit the

366 interpretability of the results as the SMD can be highly influenced by the SD of the

367 outcome data.[51]

368

369 **Implications for Clinicians and Researchers**

370

371 Greater SMDs in favour of exercise interventions in RCTs for persistent NSLBP may

372 be derived from a combination of outcome measures rather than one alone in

373 determining treatment success, similar to the approach in the field of osteoarthritis. [52,21]

17

374   Greater SMD results may help to identify clinically meaningful treatments that may have

375   previously been overlooked due to selection of an unmatched primary outcome.

376   Validation of these results is required in a larger sample of exercise trials in NSLBP,

377   and it would be interesting to explore the same issues for other complex interventions

378   for NSLBP, and for other conditions. Clinicians and developers of exercise

379   interventions may wish to consider what their exercise intervention targets, in order to

380   select the most appropriate outcomes for that intervention. Further, it may be more

381   beneficial for developers of RCT interventions to use a composite outcome comprised

382   of the most important outcomes targeted to the intervention being tested. We

383   recommend that developers of exercise interventions consider logic models or

384   programme development theory[36,37] in order to map and guide assessment of the

385   mechanisms of action of their intervention, and the most likely outcomes to accurately

386   measure the changes expected. Previous intervention development has been

387   exemplified by Hurley et al.[37] and Kjaer et al.[53] who provided detailed descriptions of

388   their self-management and exercise programs (please see Figure 4 as an example

389   program model), including the 'active' components of the intervention, the proposed

390   determinants of change and the corresponding outcomes to capture the intended

391   change. It should be noted that we do not suggest all RCTs need to consider this level

392   of intervention development. However, considering the trial intervention through a

393   visual model can help to alleviate research waste by ensuring capture of the most

394   important outcomes, and may contribute to future knowledge of how these

395   interventions may work.

396   ***Figure 4: An example program model of the GLA:D Back intervention, the***

397   ***proposed patient achievements and the outcomes through the GLA:D Back***

398 *program, and their theoretical links (reproduced from Kjaer et al.,*[53] *with*

399 *permission)*

400

Conclusion

402

403 This study provides initial support that using i) a primary outcome matched to the

404 treatment targets of the intervention may generate greater SMDs, and using ii) a

405 composite outcome comprised of several outcomes matched to the exercise treatment

406 targets, may generate greater SMDs and tighter estimates in favour of exercise

407 interventions in comparison to a non-exercise arm in persistent NSLBP. Exercise

408 prescribers and developers should consider the treatment targets of their intervention

409 when selecting the most appropriate outcome(s).

410

References

1. Buchbinder R, van Tulder M, Öberg B, et al. Low back pain: a call for action. *Lancet*. 2018;391(10137):2384-2388.

2. Hoy D, March L, Brooks P, et al. Measuring the global burden of low back pain. *Best Pract Res Clin Rheumatol*. 2010;24:155-165.

3. Hoy D, March L, Brooks P, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis*. 2014;0:1-7.

4. Stochkendahl MJ, Kjaer P, Hartvigsen J, et al. National Clinical Guidelines for non-surgical treatment of patients with recent onset low back pain or lumbar radiculopathy. *Eur Spine J*. 2018;27:60-75.

5. National Institute for Health and Care Excellence. *Low Back Pain and Sciatica in over 16s: Assessment and Management Assessment and Non-Invasive Treatments Low Back Pain and Sciatica in over 16s*.; 2016.

6. Babatunde OO, Jordan JL, Van der Windt DA, Hill JC, Foster NE, Protheroe J. Effective treatment options for musculoskeletal pain in primary care: A systematic overview of current evidence. Fleckenstein J, ed. *PLoS One*. 2017;12(6):e0178621.

7. Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MMW. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev*. 2021;CD009790:in press.

8. Hayden JA, Wilson MN, Stewart S, et al. Exercise treatment effect modifiers in persistent low back pain: an individual participant data meta-analysis of 3514 participants from 27 randomised controlled trials On behalf of Chronic Low Back Pain IPD Meta-Analysis Group. *Br J Sport Med*. 2019;0:1-16.

9. Searle A, Spink M, Ho A, Chuter V. Exercise interventions for the treatment of

436   chronic low back pain: a systematic review and meta-analysis of randomised

437   controlled trials. *Clin Rehabil.* 2015;29(12):1155-1167.

438   10.   Naugle KM, Naugle KE, Riley JL, III. Reduced Modulation of Pain in Older

439   Adults After Isometric and Aerobic Exercise. *J Pain.* 2016;17(6):719-728.

440   11.   Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical

441   activity and exercise for chronic pain in adults: an overview of Cochrane

442   Reviews. *Cochrane Database Syst Rev.* 2017;4(4):CD011279.

443   12.   Steiger F, Wirth B, de Bruin ED, Mannion AF. Is a positive clinical outcome

444   after exercise therapy for chronic non-specific low back pain contingent upon a

445   corresponding improvement in the targeted aspect(s) of performance? A

446   systematic review. *Eur Spine J.* 2012;21(4):575-598.

447   13.   Justice L, Sofka A, McGinty A. Targets, Techniques, and Treatment Contexts

448   in Emergent Literacy Intervention. *Semin Speech Lang.* 2007;28(1):014-024.

449   14.   Wood L, Ogilvie R, Hayden JA. Specifying the treatment targets of exercise

450   interventions: do we? *Br J Sports Med.* 2020;54(20):1235-1236.

451   15.   Chiarotto A, Terwee CB, Ostelo RW. Choosing the right outcome

452   measurement instruments for patients with low back pain. *Best Pract Res Clin*

453   *Rheumatol.* 2016;30(6):1003-1020.

454   16.   Craig P, Matthews L, Moore L, Simpson S, Skivington K. Updated guidance:

455   developing and evaluating complex interventions [draft of updated guidance for

456   consultation]. 2019:99.

457   17.   Chiarotto A, Deyo RA, Terwee CB, et al. Core outcome domains for clinical

458   trials in non-specific low back pain. *Eur Spine J.* 2015;24(6):1127-1142.

459   18.   Wood L, Foster NE, Lewis M, Bishop A. Exercise interventions for persistent

460   non-specific low back pain – does matching outcomes to treatment targets

461   make a difference? A systematic review and meta-analysis. *J Pain*.

462   2021;22(2):107-126.

463 19. Watt H, Harris M, Noyes J, et al. Development of a composite outcome score

464   for a complex intervention - measuring the impact of Community Health

465   Workers. *Trials*. 2015;16(1):107.

466 20. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition,

467   reporting, and interpretation of composite outcomes in clinical trials:

468   Systematic review. *BMJ*. 2010;341(7769):381.

469 21. Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH Task Force on

470   research standards for chronic low back pain. *J Pain*. 2014;15(6):569-585.

471 22. Campbell N, Murray E. Designing and evaluating complex interventions to

472   improve health care. *BMJ*. 2007;334(7591):455-459.

473 23. Pogue J, Devereaux PJ, Thabane L, Yusuf S. Designing and analyzing clinical

474   trials with composite outcomes: Consideration of possible treatment

475   differences between the individual outcomes. *PLoS One*. 2012;7(4).

476 24. Sankoh AJ, D'Agostino RB, Huque MF. Efficacy endpoint selection and

477   multiplicity adjustment methods in clinical trials with inherent multiple endpoint

478   issues. *Stat Med*. 2003;22(20):3133-3150.

479 25. Bronfort G, Maiers MJ, Evans RL, et al. Supervised exercise, spinal

480   manipulation, and home exercise for chronic low back pain: A randomized

481   clinical trial. *Spine J*. 2011;11(7):585-598.

482 26. Moffett JK, Jackson DA, Gardiner ED, et al. Randomized trial of two

483   physiotherapy interventions for primary care neck and back pain patients:

484   "McKenzie" vs brief physiotherapy pain management. *Rheumatology*.

485   2006;45(12):1514-1521.

486   27.   Miyamoto GC, Franco KFM, van Dongen JM, et al. Different doses of Pilates-

487   based exercise therapy for chronic low back pain: a randomised controlled trial

488   with economic evaluation. *Br J Sports Med*. 2018;52:859-868.

489   28.   Groessl EJ, Liu L, Chang DG, et al. Yoga for Military Veterans with Chronic

490   Low Back Pain: A Randomized Clinical Trial. *Am J Prev Med*. 2017;53(5):599-

491   608.

492   29.   Song M-K, Lin F-C, Ward S, Fine J, Hill C. Composite Variables: When and

493   How. *Nurs Res*. 2013;62(1):45-49.

494   30.   Tilbrook HE, Cox H, Hewitt CE, et al. Yoga for Chronic Low Back Pain. *Ann

495   Intern Med*. 2011;155(9):569-578.

496   31.   Shirado O, Doi T, Akai M, et al. Multicenter randomized controlled trial to

497   evaluate the effect of home-based exercise on patients with chronic low back

498   pain: the Japan low back pain exercise therapy study. *Spine (Phila Pa 1976)*.

499   2010;35(17):E811-9.

500   32.   Harris A, Moe TF, Eriksen HR, et al. Brief intervention, physical exercise and

501   cognitive behavioural group therapy for patients with chronic low back pain

502   (The CINS trial). *Eur J Pain (United Kingdom)*. 2017;21(8):1397-1407.

503   33.   Maul I, Läubli T, Oliveri M, Krueger H. Long-term effects of supervised physical

504   training in secondary prevention of low back pain. *Eur Spine J*.

505   2005;14(6):599-611.

506   34.   Hildebrandt VH, Roper KI, Van den B, Douwes M, Van den Heuvel SG, Van

507   Buuren S. Cesar therapy is temporarily more effective than a standard

508   treatment from the general practitioner in patients with chronic aspecific lower

509   back pain; randomized, controlled and blinded study with a I year follow-up.

510   *Ned Tijdschr Geneeskd*. 2000;144(47 PG-2258-2264):2258-2264.

511   35.   Chen HM, Wang HH, Chen CH, Hu HM. Effectiveness of a stretching exercise

512   program on low back pain and exercise self-efficacy among nurses in Taiwan:

513   A randomized clinical trial. *Pain Manag Nurs*. 2014;15(1):283-291.

514   36.   Rohwer A, Pfadenhauer L, Burns J, et al. Logic models help make sense of

515   complexity in systematic reviews and health technology assessments. *J Clin*

516   *Epidemiol*. 2017;83:37-47.

517   37.   Hurley DA, Murphy LC, Hayes D, et al. Using intervention mapping to develop

518   a theory-driven, group-based complex intervention to support self-

519   management of osteoarthritis and low back pain (SOLAS). *Implement Sci*.

520   2016;11(1):56.

521   38.   Moore GF, Audrey S, Barker M, et al. Process evaluation of complex

522   interventions: Medical Research Council guidance. *BMJ*. 2015;350(19

523   6):h1258-h1258.

524   39.   Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to

525   translate into benefits for patients. *Trials*. 2017;18(1):1-7.

526   40.   Helmhout PH, Staal JB, Maher CG, Petersen T, Rainville J, Shaw WS.

527   Exercise therapy and low back pain: insights and proposals to improve the

528   design, conduct, and reporting of clinical trials. *Spine (Phila Pa 1976)*.

529   2008;33(16):1782-1788.

530   41.   Rainville J, Hartigan C, Martinez E, Limke J, Jouve C, Finno M. Exercise as a

531   treatment for chronic low back pain. *Spine J*. 2004;4(1):106-115.

532   42.   Parkes MJ, Callaghan MJ, Tive L, Lunt M, Felson DT. Responsiveness of

533   Single versus Composite Measures of Pain in Knee Osteoarthritis. *J*

534   *Rheumatol*. 2018;45(9):1308-1315.

535   43.   Ross S. Composite outcomes in randomized clinical trials: arguments for and

536   against. *Am J Obstet Gynecol*. 2007;196(2):119.e1-119.e6.

537   44.   Vaanholt MCW, Kok MM, von Birgelen C, Weernink MGM, van Til JA. Are

538   component endpoints equal? A preference study into the practice of composite

539   endpoints in clinical trials. *Heal Expect*. 2018;21(6):1046-1055.

540   45.   Ferreira-González I, Permanyer-Miralda G, Busse JW, et al. Methodologic

541   discussions for using and interpreting composite endpoints are limited, but still

542   identify major concerns. *J Clin Epidemiol*. 2007;60:651-657.

543   46.   Ferreira-González I, Busse JW, Heels-Ansdell D, et al. Problems with use of

544   composite end points in cardiovascular trials: Systematic review of randomised

545   controlled trials. *BMJ*. 2007;334(7597):786-788.

546   47.   Chiarotto A, Terwee CB, Deyo RA, et al. A core outcome set for clinical trials

547   on non-specific low back pain: study protocol for the development of a core

548   domain set. *Trials*. 2014;15(1):511.

549   48.   Coster WJ. Making the Best Match: Selecting Outcome Measures for Clinical

550   Trials and Outcome Studies MeSH TERMS clinical trials as topic decision

551   making guidelines as topic outcome assessment (health care) treatment

552   outcome. *Am J Occup Ther*. 2013;67:162-170.

553   49.   van Tulder M, Malmivaara A, Hayden J, Koes B. Statistical significance versus

554   clinical importance: trials on exercise therapy for chronic low back pain as

555   example. *Spine (Phila Pa 1976)*. 2007;32(16):1785-1790.

556   50.   Chiarotto A, Ostelo RW, Turk DC, Buchbinder R, Boers M. Core outcome sets

557   for research and clinical practice. *Brazilian J Phys Ther*. 2017;21(2):77-84.

558   51.   Faraone S V. Interpreting estimates of treatment effects: Implications for

559   managed care. *P T*. 2008;33(12).

560   52.   Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an

561   international initiative to improve outcome measurement in rheumatology.

562   *Trials*. 2007;8:38.

563 53. Kjaer P, Kongsted A, Ris I, et al. GLA:D ® Back group-based patient education

564   integrated with exercises to support self-management of back pain -

565   Development, theories and scientific evidence - Development, t. *BMC*

566   *Musculoskelet Disord*. 2018;19(1):1-21.

567

568 **Figure Legends**

569

570 Figure 1: Processes of identification of suitable trials for inclusion and analysis

571 Figure 2: Forest plot to demonstrate the pooled effect of the SMD for unmatched

572 primary outcomes in comparison to matched secondary outcomes

573 Figure 3: Summary plot to demonstrate pooled SMD of primary outcome in

574 comparison to composite outcome

575 Figure 4: An example program model of the GLA:D Back intervention, the proposed

576 patient achievements and the outcomes through the GLA:D Back program, and their

577 theoretical links (reproduced from Kjaer et al., 2018, with permission)

578

579

1   **Title:** Matching the outcomes to treatment targets of exercise for low back pain: does

2   it make a difference? Results of secondary analyses from individual patient data of

3   randomised controlled trials and pooling of results across trials in comparative meta-

4   analyses

5

6   **<u>Abstract</u>**

7

8   **Objective**

9   To explore whether using a single matched or composite outcome might impact the

10   results of previous randomised controlled trials (RCTs) testing exercise for non-

11   specific low back pain (NSLBP). The first objective was to explore whether a single

12   matched outcome generated a greater standardised mean differences (SMD) when

13   compared to the original unmatched primary outcome SMD. The second objective was

14   to explore whether a composite measure, comprised of matched outcomes, generated

15   a greater SMD when compared to the original primary outcome SMD.

16

17   **Design**

18   We conducted exploratory secondary analyses of data.

19

20   **Setting**

21   Seven RCTs were included, of which two were based in the USA (University research

22   clinic, Veterans Affairs medical centre) and the UK (primary care clinics, nonmedical

23   centres). One each were based in Norway (clinics), Brazil (primary care), and Japan

24   (outpatient clinics).

25

Participants

The first analysis comprised 1) five RCTs (n=1,033) that used an unmatched primary outcome but included (some) matched outcomes as secondary outcomes, and the second analysis comprised 2) four RCTs (n=864) that included multiple matched outcomes by developing composite outcomes.

Intervention:

Exercise compared to no exercise.

Main Outcome Measures:

The composite consisted of standardised averaged matched outcomes. All analyses replicated the RCTs' primary outcome analyses.

Results

Of five RCTs, three had greater SMDs with matched outcomes (pooled effect SMD 0.30 (95% CI 0.04, 0.56), p=0.02) compared to an unmatched primary outcome (pooled effect SMD 0.19 (95% CI -0.03, 0.40) p=0.09). Of four composite outcome analyses, three RCTs had greater SMDs in the composite outcome (pooled effect SMD 0.28 (95%CI 0.05, 0.51) p=0.02) compared to the primary outcome (pooled effect SMD 0.24 (95%CI -0.04, 0.53) p=0.10).

Conclusions

These exploratory analyses suggest that using an outcome matched to exercise treatment targets in NSLBP RCTs may produce greater SMDs than an unmatched

50   primary outcome. Composite outcomes could offer a meaningful way of investigating

51   superiority of exercise than single domain outcomes.

52

53   **Key words**:   Low back pain, exercise, treatment targets, secondary analysis,

54   randomised controlled trials, composite outcomes.

55

56

64

Introduction

65

66

Persistent non-specific low back pain (NSLBP) is the leading cause of disability

67

globally,[1,2] with an estimated 540 million people worldwide experiencing NSLBP.[3]

68

Therapeutic exercise is the most widely recommended treatment for persistent

69

NSLBP[4,5] with moderate certainty evidence that it has clinically important benefits for

70

pain but small benefits for function.[6–9]

71

72

Exercise is a complex intervention with numerous components, such as biological,[10]

73

psychological and social,[11] as well as treatment interaction components.[12] Therefore,

74

there may be multiple potential treatment targets, where a treatment target is defined

75

as the goal or intention the treatment aims to influence.[13] Most randomised controlled

76

trials (RCTs) of exercise for persistent NSLBP do not specify their treatment targets.[14]

77

Literature regarding RCT design stipulates that the primary outcome should match the

78

rationale of the intervention,[15,16] yet outcome measures are often selected based on

79

core outcome domains[17] and/or patient preference. A recent systematic review[18]

80

81   demonstrated that most (74%) of the included RCTs of exercise in persistent NSLBP

82   used primary outcomes not reflective of the RCT's specified exercise treatment targets.

83   Further, most RCTs demonstrate only small differences between exercise and control

84   arms,[7] and therefore clinically important interventions may be overlooked, if these

85   benefits are related to the selection of the primary outcome.

86

87   In complex interventions, such as exercise, which frequently have more than one

88   treatment target, the selection of a single primary outcome measure may be insufficient

89   to capture the benefits that can be achieved.[19] Watt et al.,[19] suggest that nominating a

90   single primary outcome in a RCT of a complex intervention may distort the overall

91   purpose. Composite outcomes, including two or more component outcome domains,[20]

92   may be more suitable than a single primary outcome in such RCTs, and may be better

93   able to demonstrate the effects of complex interventions. In addition, more meaningful

94   results of exercise RCTs for persistent NSLBP may be derived. However, due to the

95   limited evidence on composite measures available for NSLBP, future research in this

96   area has been recommended.[21]

97

98   It is unknown whether using a matched primary outcome or composite outcome

99   (comprised of the specified treatment targets) might alter the findings of previous

100   RCTs.[22] This secondary analysis aimed to explore whether using a single matched or

101   composite outcome might impact the results of previous RCTs testing exercise for

102   persistent NSLBP. The first objective was to explore whether a single outcome,

103   matched to the identified exercise treatment targets, generated a greater standardised

104   mean difference (SMD) when compared to the original unmatched primary outcome

105 SMD. The second objective was to explore whether a composite measure, comprised

106 of more than one outcome matched to the identified exercise treatment targets,

107 generated a greater SMD when compared to the original primary outcome SMD.

108

109
110 **Methods**

111

112 *Design*

113 Exploratory secondary analyses of seven previous RCTs. A random effects meta-

114 analysis (generated with RevMan 5.3) was used to compare: i) the overall effect of

115 using an unmatched primary outcome with the first reported matched outcome, and ii)

116 the overall effect of using a single primary outcome (matched or unmatched) with a

117 composite (matched) outcome.

118

119 *Data Source*

120 A recently completed systematic review of RCTs of exercise interventions compared

121 to no exercise in persistent NSLBP[18] informed the RCT sample for this study.

122 Treatment targets were extracted verbatim from the RCT published texts, where it was

123 clear the authors had described a rationale for how the exercise intervention was

124 proposed to work, or what they had designed the exercise intervention to target. In the

125 review, RCTs were categorised into: a matched group, where the primary outcome

126 reflected one of the identified treatment targets; or an unmatched group, where the

127 primary outcome did not reflect one of the identified treatment targets. The matching

128 process was subjective and performed by pairs of independent reviewers, as described

129 in Wood et al.[18] For each analysis, the authors of the identified RCTs were contacted

130 and the dataset requested. The first analysis identified RCTs within the unmatched

131 group that included secondary outcomes matched to the treatment targets. The second

132 analysis identified RCTs within both the matched and unmatched groups, where more

133 than one outcome reflected more than one stated exercise treatment target.

134
135
136 *Data Extraction*

137 Information pertinent to these analyses was extracted as part of the systematic review

138 process[18] by pairs of independent reviewers (see appendix 1). The stated treatment

139 target(s) of the exercise intervention, the primary and secondary outcomes for each

140 RCT, the outcomes that matched the stated exercise treatment targets, and the method

141 of analysis performed on primary and secondary outcomes were extracted for each

142 RCT (see Table 1).

143

144 *Data Analysis*

145 <u>Both Analyses:</u>

146 SMDs and 95% confidence intervals were calculated for each primary and matched

147 secondary outcome for between-arm differences at the primary outcome time-point

148 designated by the trial authors, or if no primary time-point was specified by the authors,

149 then the earliest time-point post-exercise-intervention. SMD statistics for all between-

150 arm differences were reported as intervention minus control: positive SMDs indicating

151 higher values for the exercise intervention (lower for the control), and by contrast,

152 negative SMDs indicating lower values for the intervention (higher for the control).

153 Where some variables had point estimates scoring in the opposite direction to other

7

154   included variables, these were transformed so that all variables scored in the same

155   direction.[23,24]

156

157   For linear mixed models[25–28] the data were transformed from wide to long format by

158   transforming the variables to cases and computing a new variable consisting of all

159   time-points relevant to that outcome. All outcomes of interest were converted to a

160   standardised variable (standardised z-score). Initial analyses aimed to replicate the

161   published data used for the primary outcome(s) and/or targeted secondary outcomes

162   where possible to do so. The replicated analysis was applied to the matched

163   secondary outcome(s). Linear mixed model analyses include all time-points available

164   for the relevant outcome. Therefore values for all available time-points for the matched

165   secondary outcomes were also used and reported[25–28].

166

167   <u>Second Analysis Only:</u>

168   The second analysis created a composite outcome, comprised of multiple outcomes

169   matched to the specified exercise treatment targets. For the creation of the composite

170   outcome, standardised composite outcomes were derived by computing a new

171   variable of the mean of the standardised outcome scores, matched to the treatment

172   targets, for each time-point.[29] A further analysis was performed where two primary

173   outcomes were specified, and both were matched to the treatment targets: a co-

174   primary composite was developed by creating a new variable of the mean of the

175   standardised primary outcomes at each time point. Exploratory analysis compared the

176   results of the first nominated primary outcome in comparison to a targeted composite

177   outcome and the co-primary outcome composite. The method of analysis of between-

8

178   arm standardised differences replicated the initial primary time-point analysis. All

179   analyses used Statistical Package for Social Science (SPSS) Statistics 24.

180

Results

A summary of dataset acquisition and analysis is displayed in Figure 1, and details of included trials are presented in Table 1.

Figure 1: Processes of identification of suitable trials for inclusion and analysis

Table 1: Included Trial Datasets

First Analysis: The Difference between Matched and Unmatched Outcome SMDs

In the first analysis, lead authors from five RCTs[25,28,30–32] were contacted, and three datasets acquired. Two RCTs provided sufficient information within their published papers, resulting in five RCTs analysed (1,033 participants). Two RCTs compared yoga to usual care,[30] and a waitlist control,[28] three RCTs tested supervised exercise programs in comparison to a brief intervention[32], a home exercise and manipulative arm[25], and prescribed NSAIDS[31].

Of the five RCTs included, three had greater SMDs and statistical significance in favour of exercise compared to a control-arm when a matched secondary outcome was used in comparison to an unmatched primary outcome[25,28,31] (see Table 2). Of the three full datasets analysed, two demonstrated larger, statistically significant effects in favour of exercise with at least one matched secondary outcome at the primary time-point(s), compared to an unmatched primary outcome[25,28]. The analysis

204   of Harris et al.[32] did not demonstrate any statistically significant differences using any

205   of the outcomes, but the use of the matched secondary outcome generated a greater

206   SMD in favour of the exercise group than when using the unmatched primary outcome.

207   The analysis of Tilbrook et al.[30] was the only trial analysed to demonstrate greater

208   between-arm differences when using an unmatched primary outcome.

209

210   ***Table 2: First analysis results demonstrating the difference between matched***

211   ***and unmatched outcome SMDs***

212

213   The original results and secondary analyses of the five RCTs are summarised in Figure

214   2: a pooled SMD of 0.19 (95% CI -0.03, 0.40; p=0.09) was seen for the unmatched

215   primary outcome, in comparison to the SMD of 0.30 (95% CI 0.04, 0.56; p=0.02) for

216   the first reported matched outcome. The subgroup differences (primary outcome

217   compared to the first matched outcome) were not statistically significant (SMD 0.11;

218   95% CI -0.34, 0.57; p=0.51).

219

220   ***Figure 2: Forest plot to demonstrate the pooled effect of the SMD for unmatched***

221   ***primary outcomes in comparison to matched secondary outcomes***

222

223   Second Analysis: Composite SMD calculations in comparison to Primary Outcome

224   SMDs

225   In the second analysis, lead authors from seven RCTs[25–28,33–35] were contacted, and

226   four authors shared their datasets.[25–28] Four RCTs were analysed (864 participants):

227   one   compared differing Pilates dosages plus advice versus advice alone,[27] one

228   compared yoga to a waitlist,[28] one tested supervised exercise programs in a home

229   exercise versus a manipulative arm,[25] and one compared McKenzie exercises versus

230   a physiotherapy intervention.[26] The composite outcomes varied in composition with

231   three composite outcomes formed of six outcomes[25–27] and one composite comprised

232   of three outcomes[28]. For example, Groessl et al.[28] measured the outcomes of strength,

233   flexibility and pain relief in their RCT which were matched to the treatment targets of

234   increasing strength and flexibility and improving pain tolerance. Please see Table 3 for

235   more detail regarding composition of composite outcomes.

236

237   The composite analysis impacted the results of three of four RCTs,[25,26,28] as seen in

238   Table 3. Three of the four analyses showed results with the composite outcome

239   variable that had greater SMDs in favour of the exercise intervention[25,26,28], of which

240   two[25,28] were (more) statistically significant in comparison to the original RCTs' primary

241   outcome results. All analyses showed a smaller standard error when using the

242   composite outcome. The use of the co-primary composite generated greater SMDs

243   than the composite outcome. However, the co-primary composite generated greater

244   SMDs (not statistically significant) than the primary outcome in one RCT,[26] but this was

245   not reproduced in the other RCT analysis.[27]

246

247   ***Table 3: Second analysis results of composite SMD calculations compared to***

248   ***primary outcome SMDs***

249

250   This is summarised in Figure 3 whereby a pooled SMD of 0.24 (95% CI -0.04, 0.53;

251   p=0.10) was seen for the primary outcome in comparison to the SMD of 0.28 (95% CI

252   0.05, 0.51; p=0.02) for the matched composite outcome. The subgroup differences

253   (primary outcome compared to matched composite) were not statistically significant

254   (SMD 0.03 (95% CI -0.13, 0.20) p=0.86).

255   ***Figure 3: Summary plot to demonstrate pooled SMD of primary outcome in***

256   ***comparison to composite outcome***

257

258   **<u>Discussion</u>**

259

260   The results of these exploratory secondary analyses of previous RCTs of exercise for

261   NSLBP suggest that it is possible that using a primary outcome matched to the

262   treatment targets of exercise may generate greater SMDs than a single unmatched

263   primary outcome. Further, using a composite outcome, matched to multiple exercise

264   treatment targets, may give greater power to detect superiority of exercise over a non-

265   exercise control. In three of five RCTs, a single matched outcome measure generated

266   a greater SMD than the original unmatched primary outcome SMD, and would impact

267   the results of four RCTs. In two of four RCTs, a composite matched outcome would

268   impact the results in favour of exercise versus control. Our analyses provide some

269   support for matching the primary outcome to the treatment targets of the exercise

270   intervention, and for considering the use of a composite outcome in comparison to a

271   single outcome when multiple exercise targets are identified. Using a matched outcome

272   may provide more clinically meaningful results, and will allow for identification of

273   treatment interventions that may be more effective than previously supposed.

13

274

275   Treatment targets may be described as intermediate variables or surrogate outcomes,

276   as they may sit on the pathway to a patient relevant outcome such as pain or function.

277   However, this may not always be the case, and the treatment targets reported by the

278   authors of these RCTs may not have been based on clear programme development

279   theory or logic modelling.[36,37] Many of the treatment targets identified by the RCT

280   authors were captured by some of their outcomes, but there were no published

281   intervention development or programme evaluation[38] papers for any of the included

282   RCTs within which to test the degree that these treatment targets were indeed the

283   focus of their intervention. Thus, it is difficult to identify which of the treatment targets

284   may have been prioritised, or which may have been changed by the exercise

285   interventions. In exercise, where multiple treatment targets are common, it is

286   challenging without clear intervention theory, to understand how the exercise

287   intervention may have exerted its effect. Heneghan et al.[39] caution against the use of

288   surrogate outcomes as primary outcomes, without a clear understanding of the impact

289   and effect of these upon patient-relevant outcomes. In the field of exercise and

290   NSLBP, the effect surrogate outcomes have on important patient outcomes like pain,

291   function and quality of life is poorly understood. Furthermore, there is a lack of

292   understanding as to what mechanisms of effect underpin exercise interventions for

293   NSLBP.[40,41]

294

295   The results of these exploratory secondary data analyses provide some support for

296   considering the use of a composite matched outcome rather than a single unmatched

297   outcome in trials of exercise for NSLBP. The results contrast with those from Parkes et

298   al.[42] who compared a composite outcome (the Western Ontario and McMaster

299 Universities Osteoarthritis Index [WOMAC] score, pain and rescue medication) to a

300 single outcome (WOMAC pain) in knee osteoarthritis. Their composite outcome

301 demonstrated modest improvements in responsiveness when compared to WOMAC

302 pain alone, but these were not statistically significant. While composite outcomes are

303 uncommon as primary outcome measures in RCTs in the field of NSLBP, they are

304 frequently used in cardiovascular medicine, and have both advantages and

305 disadvantages. The use of a composite outcome can reduce the sample size,[43,44] which

306 is beneficial both for the recruitment period and associated costs of RCTs.[45,46]

307 However, in cardiovascular disease when a composite outcome included the outcome

308 measures of most importance to patients, composite outcomes were less likely to

309 demonstrate a moderate treatment effect.[46] Moreover, there is a risk of overestimation

310 of treatment impact and effect when using composite outcomes if the component

311 outcomes are not reported completely, leading to incorrect interpretation of the

312 results.[39] If the use of composite outcomes is to be considered in NSLBP, composite

313 outcomes would need to be chosen based on sound rationale. Furthermore, all

314 outcomes selected to be included in the composite should individually be expected to

315 demonstrate an important effect, as any outcome that does not will dilute the overall

316 effect. Hence, composites make sense if the targeted outcomes all contribute to an

317 important treatment effect and are responsive to change. This proposal is supported

318 by our results that show the co-primary (matched) analysis produced the overall highest

319 SMDs (greater than the composite).

320

321 Most RCTs of exercise for LBP appear to use a recommended core outcome domain[47]

322 as a primary outcome.[18] Core outcome domains are necessary to allow for comparison

323 of results across multiple datasets, and are useful for combined evidence approaches

324 such as meta-analysis. However, the authors of the LBP core outcome set highlight

325 that the agreed domains do not restrict measurement or the choice of primary outcome,

326 but "mandate collection and reporting of the core outcome set alongside the outcomes

327 of interest".[17] It could be argued that prioritising pain or back-related disability as the

328 primary outcome domain in RCTs testing exercise for persistent NSLBP may not

329 accurately reflect the benefits of exercise, if these outcome domains do not match the

330 range of treatment targets of the intervention. The challenge of outcome measure

331 selection is encapsulated by Coster et al.,[48] *"The ultimate value of a RCT …will be*

332 *directly tied to how well the selected outcome measure matches the researcher's*

333 *understanding of what he or she expects to change, to what degree it is expected to*

334 *change, over what period of time this change will happen and how that change can*

335 *best be identified".* As exercise is a complex intervention with multiple potential

336 treatment targets, there are multiple possible outcomes that could be used, but multiple

337 outcomes should be interpreted with caution.[49] The proposed treatment targets of the

338 intervention should influence the selection of the primary outcome, from which the

339 minimally important difference is used to calculate the sample size.[49] Literature

340 regarding RCT design stipulates that the primary outcome should match the rationale

341 of the intervention.[16,50] The results of this analysis suggest that matching the primary

342 outcome to the treatment targets of the intervention may generate greater SMDs in

343 favour of exercise, and that a composite outcome comprised of the most important

344 treatment targets could generate greater SMDs with smaller standard errors in favour

345 of exercise. A matched 'targeted' composite or single outcome may provide the RCT

346 team with the best chance of detecting the benefits of exercise compared to a control

347 or comparator, as well as providing a clear framework for future testing of how exercise

348 may potentially achieve its effects. This may have clinical implications given we have

349   limited understanding of what components or targets of exercise are most influential in

350   creating change in outcomes of importance.

351

352   **Strengths and Limitations**

353

354   This is the first study to explore the relationship between matched outcomes or

355   composite outcomes and the treatment targets of the exercise intervention in RCT

356   datasets of exercise for NSLBP. A strength of this study is the individual patient data

357   acquisition of seven previously published RCTs which allowed secondary analysis of

358   the data and generation of new composite variables. The analysis methods replicated

359   the primary analysis method used by the trial teams of the individual RCTs, and this

360   ensured the data were comparable, strengthening the results of this analysis. These

361   RCTs were selected from a sample of RCTs included in a systematic review,[18] which

362   may have been subject to publication bias. The main limitation is that this was an

363   exploratory secondary analysis of a small number of RCT datasets. SMDs were chosen

364   as a means to compare outcome estimates of different outcomes, but this may limit the

365   interpretability of the results as the SMD can be highly influenced by the SD of the

366   outcome data.[51]

367

368   **Implications for Clinicians and Researchers**

369

370   Greater SMDs in favour of exercise interventions in RCTs for persistent NSLBP may

371   be derived from a combination of outcome measures rather than one alone in

372   determining treatment success, similar to the approach in the field of osteoarthritis.

373   [52,21] Greater SMD results may help to identify clinically meaningful treatments that

374   may have previously been overlooked due to selection of an unmatched primary

375   outcome. Validation of these results is required in a larger sample of exercise trials in

376   NSLBP, and it would be interesting to explore the same issues for other complex

377   interventions for NSLBP, and for other conditions. Clinicians and developers of

378   exercise interventions may wish to consider what their exercise intervention targets,

379   in order to select the most appropriate outcomes for that intervention. Further, it may

380   be more beneficial for developers of RCT interventions to use a composite outcome

381   comprised of the most important outcomes targeted to the intervention being tested.

382   We recommend that developers of exercise interventions consider logic models or

383   programme development theory[36,37] in order to map and guide assessment of the

384   mechanisms of action of their intervention, and the most likely outcomes to

385   accurately measure the changes expected. Previous intervention development has

386   been exemplified by Hurley et al.[37] and Kjaer et al.[53] who provided detailed

387   descriptions of their self-management and exercise programs (please see Figure 4

388   as an example program model), including the 'active' components of the intervention,

389   the proposed determinants of change and the corresponding outcomes to capture

390   the intended change. It should be noted that we do not suggest all RCTs need to

391   consider this level of intervention development. However, considering the trial

392   intervention through a visual model can help to alleviate research waste by ensuring

393   capture of the most important outcomes, and may contribute to future knowledge of

394   how these interventions may work.

395   *Figure 4: An example program model of the GLA:D Back intervention, the*

396   *proposed patient achievements and the outcomes through the GLA:D Back*

397 *program, and their theoretical links (reproduced from Kjaer et al.[53] with*

398 *permission)*

399

Conclusion

401

402 This study provides initial support that using i) a primary outcome matched to the

403 treatment targets of the intervention may generate greater SMDs, and using ii) a

404 composite outcome comprised of several outcomes matched to the exercise treatment

405 targets, may generate greater SMDs and tighter estimates in favour of exercise

406 interventions in comparison to a non-exercise arm in persistent NSLBP. Exercise

407 prescribers and developers should consider the treatment targets of their intervention

408 when selecting the most appropriate outcome(s).

409

References

410

411 1. Buchbinder R, van Tulder M, Öberg B, et al. Low back pain: a call for action.

412 *Lancet.* 2018;391(10137):2384-2388.

413 2. Hoy D, March L, Brooks P, et al. Measuring the global burden of low back

414 pain. *Best Pract Res Clin Rheumatol.* 2010;24:155-165.

415 3. Hoy D, March L, Brooks P, et al. The global burden of low back pain: estimates

416 from the Global Burden of Disease 2010 study. *Ann Rheum Dis.* 2014;0:1-7.

417 4. Stochkendahl MJ, Kjaer P, Hartvigsen J, et al. National Clinical Guidelines for

418 non-surgical treatment of patients with recent onset low back pain or lumbar

419 radiculopathy. *Eur Spine J.* 2018;27:60-75.

420 5. National Institute for Health and Care Excellence. *Low Back Pain and Sciatica*

421 *in over 16s: Assessment and Management Assessment and Non-Invasive*

422 *Treatments Low Back Pain and Sciatica in over 16s.*; 2016.

423 6. Babatunde OO, Jordan JL, Van der Windt DA, Hill JC, Foster NE, Protheroe J.

424 Effective treatment options for musculoskeletal pain in primary care: A

425 systematic overview of current evidence. Fleckenstein J, ed. *PLoS One.*

426 2017;12(6):e0178621.

427 7. Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MMW. Exercise

428 therapy for chronic low back pain. *Cochrane Database Syst Rev.*

429 2021;CD009790:in press.

430 8. Hayden JA, Wilson MN, Stewart S, et al. Exercise treatment effect modifiers in

431 persistent low back pain: an individual participant data meta-analysis of 3514

432 participants from 27 randomised controlled trials On behalf of Chronic Low

433 Back Pain IPD Meta-Analysis Group. *Br J Sport Med.* 2019;0:1-16.

434 9. Searle A, Spink M, Ho A, Chuter V. Exercise interventions for the treatment of

435   chronic low back pain: a systematic review and meta-analysis of randomised

436   controlled trials. *Clin Rehabil.* 2015;29(12):1155-1167.

437 10. Naugle KM, Naugle KE, Riley JL, III. Reduced Modulation of Pain in Older

438   Adults After Isometric and Aerobic Exercise. *J Pain.* 2016;17(6):719-728.

439 11. Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical

440   activity and exercise for chronic pain in adults: an overview of Cochrane

441   Reviews. *Cochrane Database Syst Rev.* 2017;4(4):CD011279.

442 12. Steiger F, Wirth B, de Bruin ED, Mannion AF. Is a positive clinical outcome

443   after exercise therapy for chronic non-specific low back pain contingent upon a

444   corresponding improvement in the targeted aspect(s) of performance? A

445   systematic review. *Eur Spine J.* 2012;21(4):575-598.

446 13. Justice L, Sofka A, McGinty A. Targets, Techniques, and Treatment Contexts

447   in Emergent Literacy Intervention. *Semin Speech Lang.* 2007;28(1):014-024.

448 14. Wood L, Ogilvie R, Hayden JA. Specifying the treatment targets of exercise

449   interventions: do we? *Br J Sports Med.* 2020;54(20):1235-1236.

450 15. Chiarotto A, Terwee CB, Ostelo RW. Choosing the right outcome

451   measurement instruments for patients with low back pain. *Best Pract Res Clin*

452   *Rheumatol.* 2016;30(6):1003-1020.

453 16. Craig P, Matthews L, Moore L, Simpson S, Skivington K. Updated guidance:

454   developing and evaluating complex interventions [draft of updated guidance for

455   consultation]. 2019:99.

456 17. Chiarotto A, Deyo RA, Terwee CB, et al. Core outcome domains for clinical

457   trials in non-specific low back pain. *Eur Spine J.* 2015;24(6):1127-1142.

458 18. Wood L, Foster NE, Lewis M, Bishop A. Exercise interventions for persistent

459   non-specific low back pain – does matching outcomes to treatment targets

460   make a difference? A systematic review and meta-analysis. *J Pain*.

461   2021;22(2):107-126.

462 19. Watt H, Harris M, Noyes J, et al. Development of a composite outcome score

463   for a complex intervention - measuring the impact of Community Health

464   Workers. *Trials*. 2015;16(1):107.

465 20. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition,

466   reporting, and interpretation of composite outcomes in clinical trials:

467   Systematic review. *BMJ*. 2010;341(7769):381.

468 21. Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH Task Force on

469   research standards for chronic low back pain. *J Pain*. 2014;15(6):569-585.

470 22. Campbell N, Murray E. Designing and evaluating complex interventions to

471   improve health care. *BMJ*. 2007;334(7591):455-459.

472 23. Pogue J, Devereaux PJ, Thabane L, Yusuf S. Designing and analyzing clinical

473   trials with composite outcomes: Consideration of possible treatment

474   differences between the individual outcomes. *PLoS One*. 2012;7(4).

475 24. Sankoh AJ, D'Agostino RB, Huque MF. Efficacy endpoint selection and

476   multiplicity adjustment methods in clinical trials with inherent multiple endpoint

477   issues. *Stat Med*. 2003;22(20):3133-3150.

478 25. Bronfort G, Maiers MJ, Evans RL, et al. Supervised exercise, spinal

479   manipulation, and home exercise for chronic low back pain: A randomized

480   clinical trial. *Spine J*. 2011;11(7):585-598.

481 26. Moffett JK, Jackson DA, Gardiner ED, et al. Randomized trial of two

482   physiotherapy interventions for primary care neck and back pain patients:

483   "McKenzie" vs brief physiotherapy pain management. *Rheumatology*.

484   2006;45(12):1514-1521.

485   27.   Miyamoto GC, Franco KFM, van Dongen JM, et al. Different doses of Pilates-

486   based exercise therapy for chronic low back pain: a randomised controlled trial

487   with economic evaluation. *Br J Sports Med*. 2018;52:859-868.

488   28.   Groessl EJ, Liu L, Chang DG, et al. Yoga for Military Veterans with Chronic

489   Low Back Pain: A Randomized Clinical Trial. *Am J Prev Med*. 2017;53(5):599-

490   608.

491   29.   Song M-K, Lin F-C, Ward S, Fine J, Hill C. Composite Variables: When and

492   How. *Nurs Res*. 2013;62(1):45-49.

493   30.   Tilbrook HE, Cox H, Hewitt CE, et al. Yoga for Chronic Low Back Pain. *Ann*

494   *Intern Med*. 2011;155(9):569-578.

495   31.   Shirado O, Doi T, Akai M, et al. Multicenter randomized controlled trial to

496   evaluate the effect of home-based exercise on patients with chronic low back

497   pain: the Japan low back pain exercise therapy study. *Spine (Phila Pa 1976)*.

498   2010;35(17):E811-9.

499   32.   Harris A, Moe TF, Eriksen HR, et al. Brief intervention, physical exercise and

500   cognitive behavioural group therapy for patients with chronic low back pain

501   (The CINS trial). *Eur J Pain (United Kingdom)*. 2017;21(8):1397-1407.

502   33.   Maul I, Läubli T, Oliveri M, Krueger H. Long-term effects of supervised physical

503   training in secondary prevention of low back pain. *Eur Spine J*.

504   2005;14(6):599-611.

505   34.   Hildebrandt VH, Roper KI, Van den B, Douwes M, Van den Heuvel SG, Van

506   Buuren S. Cesar therapy is temporarily more effective than a standard

507   treatment from the general practitioner in patients with chronic aspecific lower

508   back pain; randomized, controlled and blinded study with a I year follow-up.

509   *Ned Tijdschr Geneeskd*. 2000;144(47 PG-2258-2264):2258-2264.

510 35. Chen HM, Wang HH, Chen CH, Hu HM. Effectiveness of a stretching exercise

511 program on low back pain and exercise self-efficacy among nurses in Taiwan:

512 A randomized clinical trial. *Pain Manag Nurs*. 2014;15(1):283-291.

513 36. Rohwer A, Pfadenhauer L, Burns J, et al. Logic models help make sense of

514 complexity in systematic reviews and health technology assessments. *J Clin*

515 *Epidemiol*. 2017;83:37-47.

516 37. Hurley DA, Murphy LC, Hayes D, et al. Using intervention mapping to develop

517 a theory-driven, group-based complex intervention to support self-

518 management of osteoarthritis and low back pain (SOLAS). *Implement Sci*.

519 2016;11(1):56.

520 38. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex

521 interventions: Medical Research Council guidance. *BMJ*. 2015;350(19

522 6):h1258-h1258.

523 39. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to

524 translate into benefits for patients. *Trials*. 2017;18(1):1-7.

525 40. Helmhout PH, Staal JB, Maher CG, Petersen T, Rainville J, Shaw WS.

526 Exercise therapy and low back pain: insights and proposals to improve the

527 design, conduct, and reporting of clinical trials. *Spine (Phila Pa 1976)*.

528 2008;33(16):1782-1788.

529 41. Rainville J, Hartigan C, Martinez E, Limke J, Jouve C, Finno M. Exercise as a

530 treatment for chronic low back pain. *Spine J*. 2004;4(1):106-115.

531 42. Parkes MJ, Callaghan MJ, Tive L, Lunt M, Felson DT. Responsiveness of

532 Single versus Composite Measures of Pain in Knee Osteoarthritis. *J*

533 *Rheumatol*. 2018;45(9):1308-1315.

534 43. Ross S. Composite outcomes in randomized clinical trials: arguments for and

535   against. *Am J Obstet Gynecol.* 2007;196(2):119.e1-119.e6.

536   44.   Vaanholt MCW, Kok MM, von Birgelen C, Weernink MGM, van Til JA. Are

537   component endpoints equal? A preference study into the practice of composite

538   endpoints in clinical trials. *Heal Expect.* 2018;21(6):1046-1055.

539   45.   Ferreira-González I, Permanyer-Miralda G, Busse JW, et al. Methodologic

540   discussions for using and interpreting composite endpoints are limited, but still

541   identify major concerns. *J Clin Epidemiol.* 2007;60:651-657.

542   46.   Ferreira-González I, Busse JW, Heels-Ansdell D, et al. Problems with use of

543   composite end points in cardiovascular trials: Systematic review of randomised

544   controlled trials. *BMJ.* 2007;334(7597):786-788.

545   47.   Chiarotto A, Terwee CB, Deyo RA, et al. A core outcome set for clinical trials

546   on non-specific low back pain: study protocol for the development of a core

547   domain set. *Trials.* 2014;15(1):511.

548   48.   Coster WJ. Making the Best Match: Selecting Outcome Measures for Clinical

549   Trials and Outcome Studies MeSH TERMS clinical trials as topic decision

550   making guidelines as topic outcome assessment (health care) treatment

551   outcome. *Am J Occup Ther.* 2013;67:162-170.

552   49.   van Tulder M, Malmivaara A, Hayden J, Koes B. Statistical significance versus

553   clinical importance: trials on exercise therapy for chronic low back pain as

554   example. *Spine (Phila Pa 1976).* 2007;32(16):1785-1790.

555   50.   Chiarotto A, Ostelo RW, Turk DC, Buchbinder R, Boers M. Core outcome sets

556   for research and clinical practice. *Brazilian J Phys Ther.* 2017;21(2):77-84.

557   51.   Faraone S V. Interpreting estimates of treatment effects: Implications for

558   managed care. *P T.* 2008;33(12).

559   52.   Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an

560   international initiative to improve outcome measurement in rheumatology.

561   *Trials*. 2007;8:38.

562   53.   Kjaer P, Kongsted A, Ris I, et al. GLA:D ® Back group-based patient education

563   integrated with exercises to support self-management of back pain -

564   Development, theories and scientific evidence - Development, t. *BMC*

565   *Musculoskelet Disord*. 2018;19(1):1-21.

566

567   **Figure Legends**

568

569   Figure 1: Processes of identification of suitable trials for inclusion and analysis

570   Figure 2: Forest plot to demonstrate the pooled effect of the SMD for unmatched

571   primary outcomes in comparison to matched secondary outcomes

572   Figure 3: Summary plot to demonstrate pooled SMD of primary outcome in

573   comparison to composite outcome

574   Figure 4: An example program model of the GLA:D Back intervention, the proposed

575   patient achievements and the outcomes through the GLA:D Back program, and their

576   theoretical links (reproduced from Kjaer et al., 2018, with permission)

577

578

Click here to access/download
Appendix
Appendix 1.docx

Figure 1: Process of identification of suitable trials for inclusion and process of analysis

Figure 2: Forest plot to demonstrate the pooled effect of the SMD for unmatched primary outcomes in comparison to matched secondary outcomes

				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Std. Mean Difference	SE	Weight	IV, Random, 95% CI	IV, Random, 95% CI
17.1.1 Unmatched Primary Outcomes					
Bronfort et al. (2011)	0.21	0.1429	10.2%	0.21 [-0.07, 0.49]	
Groessl et al. (2017)	0.14	0.1633	9.3%	0.14 [-0.18, 0.46]	
Harris et al. (2017)	-0.16	0.1633	9.3%	-0.16 [-0.48, 0.16]	
Shirado et al. (2010)	0.18	0.1531	9.8%	0.18 [-0.12, 0.48]	
Tilbrook et al. (2011)	0.5	0.1225	11.1%	0.50 [0.26, 0.74]	
Subtotal (95% CI)			49.8%	**0.19 [-0.03, 0.40]**	
Heterogeneity: Tau² = 0.04; Chi² = 10.95, df = 4 (P = 0.03); I² = 63%					
Test for overall effect: Z = 1.72 (P = 0.09)					
17.1.2 Matched Secondary Outcomes					
Bronfort et al. (2011)	0.57	0.108	11.8%	0.57 [0.36, 0.78]	
Groessl et al. (2017)	0.23	0.1378	10.5%	0.23 [-0.04, 0.50]	
Harris et al. (2017)	0.1	0.2551	6.1%	0.10 [-0.40, 0.60]	
Shirado et al. (2010)	0.54	0.1429	10.2%	0.54 [0.26, 0.82]	
Tilbrook et al. (2011)	-0.01	0.1122	11.6%	-0.01 [-0.23, 0.21]	
Subtotal (95% CI)			50.2%	**0.30 [0.04, 0.56]**	
Heterogeneity: Tau² = 0.06; Chi² = 17.53, df = 4 (P = 0.002); I² = 77%					
Test for overall effect: Z = 2.29 (P = 0.02)					
Total (95% CI)			100.0%	**0.24 [0.08, 0.41]**	
Heterogeneity: Tau² = 0.05; Chi² = 29.66, df = 9 (P = 0.0005); I² = 70%					
Test for overall effect: Z = 2.98 (P = 0.003)					
Test for subgroup differences: Chi² = 0.43, df = 1 (P = 0.51), I² = 0%					

Favours non-exercise   Favours exercise

Std. is standard as part of SMD, SE is the standard error, IV is inverse variance, CI is confidence interval.

Figure 3: Summary plot to demonstrate pooled SMD of primary outcome in comparison to composite outcome

Study or Subgroup	Std. Mean Difference	SE	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
22.1.1 Primary					
Bronfort et al. (2011)	0.21	0.1429	10.9%	0.21 [-0.07, 0.49]	
Groessl et al. (2017)	0.14	0.1633	10.0%	0.14 [-0.18, 0.46]	
Miyamoto et al. (2018)	0.69	0.1633	10.0%	0.69 [0.37, 1.01]	
Moffett et al. (2006)	-0.01	0.1071	12.6%	-0.01 [-0.22, 0.20]	
Subtotal (95% CI)			43.5%	**0.24 [-0.04, 0.53]**	
Heterogeneity: Tau² = 0.07; Chi² = 12.97, df = 3 (P = 0.005); I² = 77%					
Test for overall effect: Z = 1.66 (P = 0.10)					
22.1.2 Composite					
Bronfort et al. (2011)	0.26	0.051	15.0%	0.26 [0.16, 0.36]	
Groessl et al. (2017)	0.3	0.0918	13.4%	0.30 [0.12, 0.48]	
Miyamoto et al. (2018)	0.6	0.102	12.9%	0.60 [0.40, 0.80]	
Moffett et al. (2006)	0	0.0408	15.3%	0.00 [-0.08, 0.08]	
Subtotal (95% CI)			56.5%	**0.28 [0.05, 0.51]**	
Heterogeneity: Tau² = 0.05; Chi² = 40.16, df = 3 (P < 0.00001); I² = 93%					
Test for overall effect: Z = 2.38 (P = 0.02)					
Total (95% CI)			100.0%	**0.26 [0.09, 0.43]**	
Heterogeneity: Tau² = 0.05; Chi² = 53.30, df = 7 (P < 0.00001); I² = 87%					
Test for overall effect: Z = 3.08 (P = 0.002)					
Test for subgroup differences: Chi² = 0.03, df = 1 (P = 0.86), I² = 0%					

Favours control   Favours exercise

Std. represents standard as part of SMD, SE is standard error, IV is inverse variance, CI is confidence interval.

Figure 4: An example program model of the GLA:D Back intervention, the proposed patient achievements and the outcomes through the GLA:D Back program, and their theoretical links (reproduced from Kjaer et al., 2018, under the Creative Commons licence with permission)
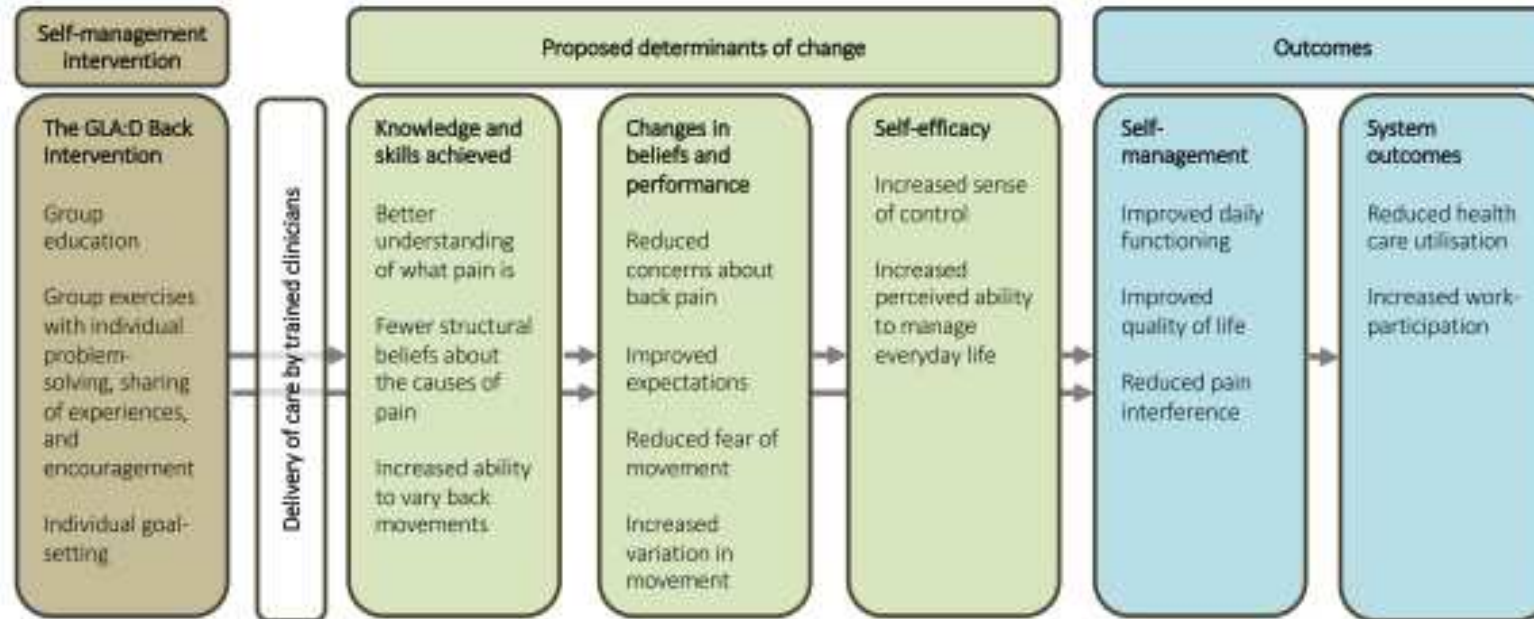
Table 1: Included Trial Datasets

Analysis		Trial	Intervention	Control	Exercise Treatment Targets	Outcome Domains		Primary Time-Point	Analysis Performed	
						All Primary	**Matched Secondary**		**Primary Outcome**	**Secondary Outcome**
FIRST ANALYSIS		Shirado et al., 2011[1]	Exercise	NSAIDs	Increasing overall physical activity; spinal mobility	*Self-reported:* Pain intensity (VAS), Physical function (RMDQ) and Health-related	*Objectively recorded:* Flexibility (finger floor distance)	8 weeks	Only SMD analysis performed	

					quality of life (JLEQ)				
	Tilbrook et al., 2011[2]	Yoga	Usual care	Improving mobility; strength; posture; reducing pain	*Self-reported:* Physical function (RMDQ)	*Self-reported:* Pain intensity (Aberdeen Back Pain Scale)	12 weeks		
	Harris et al., 2017[3]	Brief intervention with physical activity	Brief intervention	Fear avoidance and movement phobia; re-establish normal movement patterns	*Objectively recorded:* Increased work participation – change form full-time sick	*Self-reported:* Fear-avoidance behaviours (Fear-Avoidance Beliefs Questionnaire)	12 months	Differences between groups were measured with chi-square tests for each of	ANOVA

					Self-reported:	Objectively recorded:			
					leave to partial sick leave or full return to work			the 12 months	
SECOND ANALYSIS	Bronfort et al., 2011[4]	Supervised exercise	Spinal manipulation (Home exercise and advice)	Increase trunk muscle endurance; increase trunk stability	Self-reported: Pain intensity (11-point box scale)	Objectively recorded: Static endurance (flexion, extension), dynamic endurance (flexion, extension), isometric	12 weeks*	Analysis of covariance (ANCOVA) for differences between the three groups and linear	Change scores for trunk performance measures were used and then analysed for group differences

							strength (flexion, extension).		mixed-model	with analysis of variance (ANOVA)
SECOND ANALYSIS		Groessl et al., 2017[5]	Yoga	Waitlist control	Increase strength and flexibility; reduce stress; increased pain tolerance	*Self-reported:* Physical function (RMDQ)	*Self-reported:* Pain intensity (BPI) (reported); *Objectively recorded:* Range of motion (Saunders digital inclinometer) and core	12 weeks	Linear mixed-model	

							strength (prone and supine bridge) (not reported in RCT paper)		
		Miyamoto et al., 2018[6]	Pilates once a week, twice a week and three times a week plus advice	Advice alone	Improving disability; reducing absence from work; physical and functional recovery; reduce pain; improve	*Self-reported*: Pain intensity (NRS), Physical function (RMDQ)	*Self-reported*: Physical Function (PSFS), Global Perceived Effect, Catastrophizing (PCS), Kinesiophobi	6 weeks	Liner mixed-model

					catastrophising and kinesiophobia		a (TSK), Health-related Quality of Life (HRQoL) (SF6D)		
		Moffett et al., 2006[7]	McKenzie exercise	Solution finding approach	Fear of physical activity; relieve pain; reduce anxiety and depression; help them take control of their	*Self-reported*: Fear avoidance (TSK), Physical function (RMDQ)	*Self-reported*: Health control (Multidimensional health locus of control), Self-efficacy (PSEQ), Anxiety and	6 weeks*	Linear mixed-model

					situation; enable the individual to cope better; return to their normal activities sooner; prevent long-term disability		Depression (HADS)		

Legend: Only matched secondary outcomes are listed here. *Bronfort et al. [4] and Moffett et al. [7] did not specify their primary time-point, thus the first time-point post-treatment was used, as per the method used in the systematic review[8]. Abbreviations used: NSAIDs non-steroidal anti-inflammatories; VAS Visual Analogue Scale; RMDQ Roland and Morris Disability Questionnaire; JLEQ Japan Low Back Pain Evaluation Questionnaire; SMD Standardised Mean Difference; ANOVA Analysis of Variance; ANCOVA Analysis of Covariance; BPI Brief Pain Inventory; NRS Numeric Rating Scale; PSFS Patient Specific Functional Scale; PCS Pain Catastrophising Scale; TSK Tampa Scale of Kinesiophobia; SF6D Short-Form 6-Dimension questionnaire; PSEQ Pain Self-Efficacy Questionnaire; HADS Hospital Anxiety and Depression Scale.

Table 2: First analysis results demonstrating the difference between matched and unmatched outcome SMDs

Trial	Comparator	Outcome Domain (Primary Outcome Shaded)	Standardised Mean Difference (95% Confidence Interval)	Analysis Method
Shirado et al., 2010[30]	Exercise vs NSAIDS	Pain intensity	0.17 (-0.12, 0.47)	Published Data
		Physical function	0.27 (-0.02, 0.55)	
		Health-related quality of life	0.29 (-0.00, 0.57)	
		Forward finger distance*	0.54 (0.26, 0.83)	
Tilbrook et al., 2011[31]	Yoga vs Usual care	Physical function	0.50 (0.26, 0.74)	
		Pain intensity	-0.01 (-0.23, 0.22)	
Bronfort et al., 2011[25]	Exercise vs Manipulation	Pain intensity	0.21 (-0.07, 0.5)	Linear Mixed Model
		Static endurance flexion*	0.55 (0.32, 0.79)	
		Static endurance extension*	0.31 (0.09, 0.52)	

		Dynamic endurance flexion*	0.56 (0.34, 0.78)	
		Dynamic endurance extension*	0.84 (0.62, 1.05)	
		Isometric strength flexion*	0.15 (-0.00, 0.31)	
		Isometric strength extension*	0.17 (0.02, 0.32)	
Bronfort et al., 2011[25]	Exercise vs Manipulation	Pain intensity	0.21 (-0.07, 0.5)	
		Static endurance flexion*	0.57 (0.31, 0.83)	
		Static endurance extension*	0.32 (0.08, 0.57)	
		Dynamic endurance flexion*	0.59 (0.34, 0.83)	ANCOVA
		Dynamic endurance extension*	0.84 (0.61, 1.07)	
		Isometric strength flexion*	0.20 (0.01, 0.38)	
		Isometric strength extension*	0.19 (0.00, 0.37)	
Groessl et al., 2017[28]	Yoga vs Waiting list	Physical function	0.14 (-0.27, 0.55)	Linear Mixed
		Pain intensity	0.30 (0.08, 0.52)	

		Plank*	0.23 (-0.04, 0.51)	
		Flexion ROM*	0.27 (-0.08, 0.61)	
		Extension ROM*	0.08 (-0.28, 0.44)	
Harris et al., 2017[32]	Physical exercise vs Brief intervention only	Return to work*	-0.16 (-0.32, -0.00)	Chi²
		Fear avoidance (work)	-0.29 (-0.64, 0.06)	
		Fear avoidance (physical activity)	0.01 (-0.31, 0.33)	ANOVA

NSAIDS is non-steroidal anti-inflammatory drugs; ANOVA is analysis of variance; ANCOVA is analysis of covariance; ROM is range of motion; Outcomes shaded in grey are unmatched primary outcomes identified by trial authors. All outcomes were self-reported measures, apart from *, which were objectively measured.

Table 3: Second analysis results of composite SMD calculations compared to primary outcome SMDs

Primary Outcome Classification	Trial	Primary Time-Point	Outcome	SMD (Brackets denote 95% confidence intervals)	Sig. (at p<0.05)	Conclusion
Matched	Miyamoto et al. 2018[27]	6 weeks	Primary (Pain <mark>intensity</mark>)	0.69 (0.4, 1.0)	<0.0001	No change
			Composite*	0.60 (0.4, 0.8)	<0.0001	
			Co-primary composite	0.62 (0.37, 0.86)	<0.0001	
	Moffett et al. 2006[26]	6 weeks	Primary (Fear Avoidance Beliefs)	-0.01 (-0.22,0.20)	NS	<mark>No change</mark>
			Composite°	0.00 (-0.08,0.08)	NS	
			Co-primary composite	0.08 (-0.13,0.29)	NS	

Unmatched	Bronfort et al., 2011[25]	12 weeks	Primary (Pain Intensity)	0.21 (-0.07, 0.5)	Not reported	Changed results in favour of exercise
			Composite¥ (ANCOVA)	0.26 (0.16,0.36)	<0.0001	
			Composite¥ (LMM)	0.43 (0.31, 054)	<0.0001	
	Groessl et al., 2017[28]	12 weeks	Primary (Physical Function)	0.14 (-0.46,0.18)	NS	Changed results in favour of exercise
			Composite§	0.30 (0.08, 0.52)	0.007	

Where NS is non-significant, SMD is standardised mean difference, LMM is linear mixed model, ANCOVA is analysis of variance with co-variates. The composite outcomes were comprised of: *Miyamoto et al. pain, physical function, pain catastrophising, fear-avoidance beliefs, global perceived effect and a patient-specific functional scale); °Moffett et al. fear-avoidance beliefs, physical function, health control, self-efficacy, anxiety and depression; ¥Bronfort et al. dynamic endurance flexion and extension strength, static endurance flexion and extension strength, isometric flexion and extension strength; §Groessl et al. strength, flexibility and pain relief.

Click here to access/download
ICMJE Form
LW ICMJ form.docx

Click here to access/download
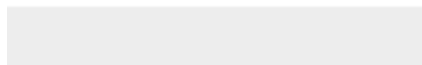**ICMJE Form**
GM ICMJ form.docx

Click here to access/download
ICMJE Form
CH ICMJ form.docx

Click here to access/download
ICMJE Form
EG ICMJ form.docx

Click here to access/download
ICMJE Form
GB ICMJ form.docx

Click here to access/download

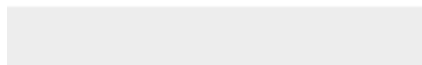**ICMJE Form**

ML ICMJ form.docx

Click here to access/download
ICMJE Form
AB ICMJ form.docx

Click here to access/download
ICMJE Form
NF ICMJ form.docx

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	1,2
Introduction			
Background and	2a	Scientific background and explanation of rationale	3,4
objectives	2b	Specific objectives or hypotheses	4,5
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	5,8, 9,10
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	NA
Participants	4a	Eligibility criteria for participants	5
	4b	Settings and locations where the data were collected	NA
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	NA
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	5,6
	6b	Any changes to trial outcomes after the trial commenced, with reasons	NA
Sample size	7a	How sample size was determined	NA
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomisation:			
Sequence	8a	Method used to generate the random allocation sequence	NA
generation	8b	Type of randomisation; details of any restriction (such as blocking and block size)	NA
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	NA
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	NA

		assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	8,9,10
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	6,7
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	6,7

Results

Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	8, figure 1,
	13b	For each group, losses and exclusions after randomisation, together with reasons	NA
Recruitment	14a	Dates defining the periods of recruitment and follow-up	NA
	14b	Why the trial ended or was stopped	NA
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	NA
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Figure 2
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Figure 2, 3
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	NA
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	NA
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	NA

Discussion

Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	14
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	15
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	15

Other information

Registration	23	Registration number and name of trial registry	Title page
Protocol	24	Where the full trial protocol can be accessed, if available	NA
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	Title page

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

Click here to access/download
Archives Submission Checklist
APMR_Checklist Final July 2017 (1).docx