

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/104412>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

© 2018 Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>.



Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

"The Consensus on Exercise Reporting Template (CERT) applied to exercise interventions in musculoskeletal trials demonstrated good rater agreement and incomplete reporting"

Authors

Slade SC, PhD*^{1,2}, Finnegan S, BSc³, Dionne CE, PhD⁴, Underwood M, PhD³, Buchbinder R, PhD^{5,6}

Institutional Affiliations and Addresses:

* Corresponding author

Email: s.slade2@latrobe.edu.au

Telephone: 61 0418 598 595

¹La Trobe Centre for Sport and Exercise Medicine Research, School of Allied Health/College of Science, Health and Engineering, La Trobe University

²Department of Physiotherapy, Monash University, Melbourne, Australia

³Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, Coventry, UK

⁴Centre de recherche du CHU de Québec-Université Laval, Québec, QC, Canada

⁵Monash Department of Clinical Epidemiology, Cabrini Institute, Melbourne, Australia

⁶Department of Epidemiology and Preventive Medicine, School of Public Health & Preventive Medicine, Monash University, Melbourne, Australia

Email: SCS s.slade2@latrobe.edu.au; susan.slade@monash.edu

SF Susanne.Finnegan@warwick.ac.uk

CED clermont.dionne@crchudequebec.ulaval.ca

MU M.Underwood@warwick.ac.uk

RB rachelle.buchbinder@monash.edu

Word count: 3370

Tables: 3

References: 64

Ethics approval: N/A

Conflicts of interest: none

Abstract

Objective: To determine inter-rater agreement and utility of the Consensus on Exercise Reporting Template (CERT) for evaluating reporting of musculoskeletal exercise trials.

Study Design and Setting: Two independent reviewers applied the CERT to a random sample of 20 exercise trials published 2010 to 2015 identified from searches of PEDro, CENTRAL and PubMed. Reviewers recorded whether each item criterion was met, detailed missing data and appraisal time. Percent agreement and the Prevalence and Bias Adjusted Kappa (PABAK) statistic measured inter-rater agreement.

Results: The trials included a range of musculoskeletal conditions (back/neck pain, hip/knee osteoarthritis, tendinopathies). For percent agreement, inter-rater agreement was high (13 items $\geq 80\%$) and for PABAK substantial (9 items: 0.61 – 0.80) and excellent (3 items: 0.81-1.0). Agreement was lower for starting level decision rule (percent agreement: 55%, PABAK 0.30); tailoring of exercise (%A: 65%, PABAK 0.40 (95% CI: 0.00 to 0.80)); exercise equipment (percent agreement: 70%, PABAK 0.30); and motivation strategies (percent agreement: 70%, PABAK 0.40). Sixty percent of descriptions were missing information for $\geq 50\%$ of CERT items. Mean appraisal time was 30 minutes and the majority of interventions required access to other published papers.

Discussion and Conclusions: The CERT has good inter-rater agreement and can comprehensively evaluate reporting of exercise interventions. Most trials do not adequately report intervention details and information can be difficult to obtain. Incomplete reporting of effective exercise programs may be remedied by using the CERT when constructing, submitting, reviewing and publishing manuscripts.

Key Words: Exercise; Musculoskeletal; Reporting guidelines; Publication quality

Introduction

The Consensus on Exercise Reporting Template (CERT) is a reporting guideline for describing exercise interventions in clinical trials and other evaluative study designs.^{1, 2} It was developed to address the problem of incomplete reporting of exercise interventions in clinical trials,³⁻⁸ and has been endorsed by an international group of experts.² It is an extension of

item 5 of the CONSORT Statement,⁹ and item 11 of the SPIRIT Statement.¹⁰ Its intended users include researchers, authors, journal editors, peer reviewers, clinicians, research ethics committees and funding bodies.²

The CERT includes 16 items considered essential to fully describe an exercise intervention to enable an informed interpretation of the trial results. A complete description of the intervention allows better interpretation of the results, and others to properly replicate the intervention in routine care or other studies. Items include description of exercise equipment, provider expertise, supervision, adherence measures, motivation strategies, individual or group exercise, decision rules for starting level and progression, replicable description of exercises, home program, non-exercise components, adverse events, setting, exercise dosage, generic or tailored programs, and program adherence.²

An Explanation and Elaboration Statement has been published that fully describes the CERT, the rationale for each item, and provides examples of good reporting.¹¹ To be useful, the CERT should be easily understood and interpreted in the same way, irrespective of who is applying it. Using a random sample of 20 musculoskeletal randomised controlled trials (RCTs) published between 2010 and 2015 we determined whether a novice to the CERT, trained in its use, understood and interpreted the items in the same way as one of the developers of the CERT or if some changes to improve clarity were needed. The assessment of the comprehensiveness of reporting of exercise interventions in the 20 trials is also presented.

Materials and Methods

Eligibility criteria

RCTs published from 1 January 2010 to 31 December 2015 that compared exercise to usual care or another intervention among adults with any musculoskeletal condition and reported outcomes such as pain, function or quality of life were eligible for inclusion. We selected this recent six-year period to reflect current practice. Trials could include interventions provided to participants in any setting (e.g. outpatient, at home, or in the community), and must have involved the prescription of a supervised or unsupervised exercise program, with or without the addition of other components (e.g. manipulation, lifestyle modification or counseling).

Search strategy

We searched the PEDro, Cochrane Central Register of Controlled Trials and PubMed databases from 1 January 2010 to 31 December 2015, using key search terms including exercise, physical therapy, training, rehabilitation, strengthening, resistance, aerobic, fitness and using methodological filters for RCTs. We used citation tracking and searched reference lists of included studies, and those of systematic reviews to identify further trials eligible for inclusion.

All search results were downloaded into a bibliographic management program and duplicates were deleted. One reviewer (SCS) screened the titles and abstracts. Any doubts about study eligibility were discussed among the authors and resolved via consensus. Included studies were transferred into Microsoft Excel (Microsoft Office 2007, Microsoft Corporation, Redmond, Washington, USA) and a computer-based random number generator (RAND function) was used to select a random sample of 20 papers. We chose 20 trials as we anticipated that this number would be sufficient to yield adequate and consistent results.

Development of data extraction guidelines and training of CERT novice

We developed data extraction guidelines to standardize the information to be extracted from each included paper. Descriptive data were systematically extracted into a spreadsheet, checked for consistency and merged into one document.

One of the developers of the CERT (SS) provided two CERT familiarization and training sessions via video-conferencing on two occasions and email follow-up on two occasions for the novice reviewer (SF). Both reviewers then independently pilot tested the data extraction form using two studies not included in the 20 to be reviewed.^{12, 13} A further video-conference meeting was held to discuss the process and findings, and make amendments to the CERT wording or data extraction form if needed for clarity.

Two changes were made to the CERT following this process (amended CERT is shown in Appendix 1):

1. Item 7 (decision rules for exercise progression) was subdivided into items: 7a (detailed description of the decision rule(s) for determining exercise progression) and 7b (detailed description of how the exercise program was progressed) as it included two questions and was difficult to rate.
2. The wording of item 16a (description of how adherence or *fidelity* to the exercise intervention is assessed/measured) was revised to 'Describe how adherence or *compliance* to the intervention is assessed/measured' to clarify that this item refers to a description of how adherence to the intervention is assessed (e.g. an exercise diary) rather than an assessment of how fidelity to the intervention/program is measured.

Application of the CERT

Two blinded reviewers (SS and SF) independently extracted the intervention data from each included study using the standardized data extraction form and data extraction guidelines of

the CERT. This included information about any equipment used for exercises, the exercise instructor, core procedural and contextual elements of the exercise intervention that are required for replication, and information about whether, and how well, the intervention complied with what was planned. Examples of poor or incomplete reporting were also extracted and recommendations made for ways to improve these descriptions.

The Explanation and Elaboration Statement was used to guide the scope and interpretation of each CERT item.¹¹ Each CERT item was rated as ‘Yes’ (criterion met, indicating item clearly reported), ‘No’ (indicating item not reported or not clearly described) or ‘Unsure’ and an overall rating of the exercise description was also made. For ‘No’ or ‘Unsure’ responses, detailed comments about what was missing/unclear were recorded. Examples of poor or incomplete reporting were also extracted and a list of questions about elements of the intervention that were missing or unclear was generated. If information was missing, the reviewers searched for and retrieved any published protocols, online appendices and supplementary data and extracted relevant information.

Each reviewer also recorded whether or not the study was published in an open access journal, the appraisal time for each study, ease of access to the intervention description (available in the published paper or required additional data from other sources and whether or not this was open access), and the time taken to retrieve additional materials when it was evident that these data were provided elsewhere.

Following completion of the comprehensiveness of reporting assessment by both reviewers, any differences between reviewers were discussed. If agreement could not be reached, an independent arbiter from the research team was consulted.

Data analysis

Data were entered into Microsoft Excel and analysed using descriptive statistics and narrative summaries.

Inter-rater reliability

Inter-rater reliability of the CERT was assessed for each of the 19 CERT items (including sub-items a) and b) for items 7, 14 and 16) using percent agreement¹⁴ and the Prevalence and Bias Adjusted Kappa (PABAK) coefficient¹⁵. While kappa statistics measure chance-adjusted agreement and are therefore more robust than simple percent agreement, when the prevalence of one of the categories is much higher than that of the other, chance agreement will be high and kappa can have unexpectedly low values.¹⁴⁻¹⁶ For percent agreement a score of 75% or greater is considered acceptable and > 80% is considered high).¹⁶ For PABAK coefficients the strength of agreement is expressed by the following descriptors: 0=poor, 0.01-0.20=slight, 0.21-0.40=fair, 0.41-0.60=moderate, 0.61-0.80=substantial and 0.81-1=excellent.¹⁵

Adequacy of reporting of exercise interventions

The number of items that were well reported for each study was tabulated, and the consistency of each item across all the studies was evaluated. A completeness of reporting score (number reported as a percentage of all 19 CERT criteria), was calculated for each included paper.

Results

The search identified 1581 papers and 1471 were excluded because they were duplicates or did not meet the inclusion criteria. Reasons for exclusion included trial not for a musculoskeletal condition, participants not adults (children and adolescents), publication

prior to 2010, and physical therapy without an exercise component or no exercise intervention. The random sample of 20 trials selected from the remaining 110 studies is described in Appendix 2.¹⁷⁻³⁶

Details of what supplementary materials were identified when needed, whether the trial report and supplementary papers were published in an open access journal, the time taken to appraise each study and time to retrieve additional data by reviewer is shown in Table 1. In general, the CERT novice took longer to appraise studies. Typically, appraisal took longer (ranging from 40 to 55 minutes) when the description of the exercise intervention was in the text of the primary paper rather than summarised in tables/figures (2/20 trials),^{30, 31} or when additional information had to be accessed from a protocol or other published papers and/or online appendices (11/20 trials).^{17, 19, 22, 24-26, 32, 33, 36} When relevant data could be easily accessed by a direct web link or were clearly presented in a figure or table (7/20 trials), the appraisal took less time (ranging from 20 to 30 minutes).^{18, 23, 27, 30, 31, 34, 35}

There were limitations to obtaining the exercise program details from the primary trial paper and other sources such as supplementary data and published literature due to lack of open access or restricted or inactive web links (Table 1). Nine published papers were behind a pay wall and required, for example, a subscription or payment (one of these was available via the Research Gate website). For other sources, such as referenced papers containing the exercise intervention or published protocol, 12 were behind a paywall. For three of the exercise interventions, it was necessary to source and read in detail more than two publications and/or a book. Three interventions had dedicated websites. However these were not apparent from reading the report, and were identified by general internet searches (Table 1)³⁷⁻⁵³.

For 10 (50%) trials, a relevant cited website, such as an open access journal, contained required information about the exercise interventions.^{19-26, 32, 33} A databank such as www.researchgate.net or www.academia.edu also provided access for 20% of these^{20, 22, 23, 31, 35} but one was not accessible.³¹ Ten (50%) trials lacked sufficient detail in the primary report and required reference to a total of 16 published papers,³⁷⁻⁵² of which only 12 were open access.^{38-44, 47-51} Five of the open access reports also provided information about, or links to, a published protocol,^{17-19, 122, 23} but one of these was not open access.¹⁹

Inter-rater agreement of CERT assessment

Table 2 presents percent agreement and PABAK coefficients for each item. Overall inter-rater agreement was high according to percent agreement (13/19 items $\geq 80\%$); with PABAK it was substantial (9 items: 0.61 – 0.80) or excellent (3 items: 0.81-1.0). Lowest or fair agreement was seen for the description of a starting level decision rule (55% agreement, PABAK 0.30); tailoring of exercise (65% agreement, PABAK 0.40); description of equipment (65% agreement, PABAK 0.30); and motivation strategies (70% agreement, PABAK 0.40).

Twelve of the possible 19 CERT items had a substantial or excellent PABAK coefficient (0.61 – 1.00); three items had a moderate coefficient (0.41- 0.60) and four items had a fair coefficient (0.21 – 0.40). Where there was disagreement, the experienced (SS) more frequently rated an item as not reported (40 occasions) compared with the novice (SF) (9 occasions) and this may reflect greater familiarity with the CERT contents and decision rules. Consensus was reached on all the dissonant item and trial scores without the need for an independent arbiter.

The novice reviewer was unable to make a decision about whether or not a criterion had been met on 12 occasions and consensus was reached following discussion without the need for

independent arbitration. Where there were differences in opinion, the final consensus was twice as likely to be the expert reviewer's assessment (n=37) compared with the novice reviewer's assessment (n=18).

The results of the CERT for each study by reviewer and by final consensus are presented in the supplementary online appendix. Not including the items for which the novice reviewer could not make a decision (11 occasions), there was perfect agreement across all items between reviewers for one trial and disagreement on one, two, four or five out of the 19 items in two trials each, and disagreement for three items for nine trials.

Final consensus CERT results

Based upon the final consensus, none of the trials fulfilled criteria for all CERT items. The mean number of CERT items that were fulfilled was 9.9 with a range of 4 to 13. Well reported items included supervision (19 trials, 95%), intervention duration (18 trials, 90%) and setting (17 trials, 85%). The most poorly reported items were description of motivation strategies, which was only reported in one trial,³² adherence to the intervention (four trials, 20%),^{21, 24, 30, 36} decision rule for starting level (five trials, 25%),^{25-28, 33} progression decision rule (seven trials, 35%),^{17, 20, 25, 26, 33, 35, 36} tailoring exercises to the individual (seven trials, 35%),^{17, 21, 22, 25, 27, 29, 33} and description of the home exercise program (8 trials, 40%).^{17, 18, 20, 23-25, 29, 35} Less than half (9 trials, 45%) described the exercise equipment used, provided a detailed description of each exercise, described exercise adherence or whether it was individual or group exercise.

Table 3 provides verbatim examples of poor or incomplete reporting of exercise interventions and the reasons why the items were considered to be poorly or incompletely reported.

Discussion

Overall, the inter-rater reliability of the CERT scoring was high when measured by both percent agreement and PABAK coefficient. Items of lower inter-rater agreement concerned reporting of the descriptions of exercise equipment, what constituted non-exercise components of an intervention, how the exercises were tailored to the individual, how a starting level of exercise was determined and reporting of adherence to the intervention.

The initial construction of item 7 about exercise progression was problematic for raters in reaching agreement in the pilot and was amended to 7a (progression decision rule) and b (how progressed) for clarification. Following full data extraction the independent raters made suggestions to clarify/amend the wording of items 1 (exercise equipment), 10 (non-exercise components), 14b (how tailored to the individual), 15 (starting level decision rule), and 16b (extent of intervention adherence) (Appendix 1). These changes informed the Explanation and Elaboration Statement that is designed to guide implementation of the CERT.¹¹

The mean time to appraise the reporting of interventions was 30 minutes when no additional sources of information were required. However, this time doubled when data extraction extended to additional sources of information and does not include time taken to source the various links and websites for relevant data describing the exercise and comparison interventions. Our review was hampered by finding that some of the websites and online appendices containing these data were no longer active. In addition, the more complex items, such as explicit descriptions of the exercises and exercise dosage and progression, which require greater explanation, were also more likely to be poorly reported. Clinicians wanting to fully replicate an exercise intervention are unlikely to have the time, resources and/or access to obtain the necessary details and materials required.

We suggest that authors should be obliged to make a detailed description of their intervention available at the time of publication of trial results. If not detailed in the paper, it should be easily and freely accessible, such as in a separate intervention design paper in an open access journal, a stable ‘intervention bank’ or as an on-line supplement to the main results paper. We concur with Hoffman et al⁷ that the provision of a stable ‘intervention bank’ (e.g. videos, manuals, and fidelity tools linked to trial registration number) would overcome the problem of word restrictions in journals. Many universities, and some funders, now maintain open access repositories that would be able to securely curate such material.

We found a high rate of incomplete reporting of exercise interventions in the musculoskeletal trials we reviewed. The omission of essential information about exercise interventions is a barrier to valid interpretation of the trial results, implementation of these programs in clinical practice and replication in research. Incomplete reporting of key features of the exercise intervention in primary papers also has important consequences for the conduct, interpretation and use of systematic reviews, in particular determining whether or not the interventions are sufficiently homogenous to be able to pool data for meta-analysis. A significant omission in many reports was explicit description of the actual exercises and the exercise dosages and progressions. This is a core component of exercise and we propose that this is an easily remediable factor.

Although rater 2 was a novice to the checklist, its utility was reflected in competence following two video-conferencing training sessions and consensus without the need for independent arbitration from the research team. While the novice did receive training and benefited from a pilot period, others who have used the CERT have reported no difficulties in using the Explanation and Elaboration Statement and Checklist.^{54, 55}

Completeness of reporting has been investigated for the CONSORT Statement⁵⁶⁻⁶⁰ and the TIDieR Checklist.^{7, 56} Inter-rater reliability has not been routinely evaluated for reporting guidelines. We located one example for the TIDieR Checklist⁶¹ in which the authors reported that the PABAK coefficient was ‘substantial’ or ‘almost perfect’ (i.e. 0.61 or higher) for just 13 of the 34 TIDieR items. This improved following discussion and consensus. Other methodological studies have investigated instrument characteristics such as internal validity.⁶²⁻⁶⁴

The CERT provides a user-friendly framework for comprehensive reporting of exercise interventions and can be completed by a novice reviewer. Completion of the CERT reporting guideline would aid in ensuring that all important elements of the exercise intervention have been described in trial protocols and results papers. We recommend that authors read and use the Explanation and Elaboration Statement to supplement the CERT checklist when they are designing and reporting exercise interventions. Several journals have already included consideration of the CERT as a mandatory component of the manuscript submission process. In addition to journal endorsement, routine publication of protocols and access to intervention manuals, trial websites and graphic or video tools via links in the primary paper would also be worthwhile.

Strengths and limitations

We used an internationally endorsed reporting guideline, two independent raters assessed intervention reporting and there was very good to almost perfect agreement for the majority of CERT items. We used a random sample of RCTs that varied across journals, conditions and contexts. Missing intervention details were obtained by sourcing reference materials e.g. previous publications, online and supplementary data.

The random sample of included trials may not be representative of all RCTs of exercise for musculoskeletal conditions published between 2010 and 2015. Also the level of inter-rater agreement may be over-estimated due to the small sample size of 20. The precision of the PABAK estimates would have been higher with more studies. Overall, however, the high degree of agreement and consistency of findings, using two different statistics, provide strong support to our conclusions.

Implications for research

We recommend research to (1) explore and understand why authors do not provide complete descriptions of exercise interventions e.g. identify the barriers and enablers; (2) explore what clinicians consider that they need to implement the published interventions into their clinical practice; (3) evaluate the effect of journal implementation of the CERT reporting guideline on the quality of published reports and determinants of usefulness for peer review; and (4) establish a consensus about how to determine the starting level of an exercise program and include consideration of such factors as strength testing, aerobic capacity, physical ability, preference, exercise experience.

Implications for practice

We suggest a requirement for manuscripts to contain explicit descriptions of: (1) the actual exercises, for example in online appendices, as photographs or embedded videos; (2) exercise dosages and progressions in, for example, exertion or resistance; (3) how exercises are individualised. The absence of these components is a barrier to the translation and implementation of exercise interventions and has implications for patient outcomes.

We recommend that authors and reviewers use the CERT to evaluate the completeness of reporting of exercise interventions in trial protocols and reports. The CERT could also be

used as a template for extracting details about exercise interventions in systematic reviews of trials of exercise and be included in the Characteristics of Included Studies Tables in Cochrane reviews.

Competing interests

The authors declare there are no competing or conflicting interests.

Acknowledgements

Statistical analysis was provided by Dr Xia Li, Consultant Statistician, La Trobe University, Melbourne, Australia.

Authors' contribution

SCS and RB conceived the study and all authors contributed to its design and content. SCS and SF conducted data extraction and analysis with input from all authors. All authors have read and approved the final manuscript.

Funding

This research project was partially funded by a 2014 Arthritis Australia Philip Benjamin Grant, Number: 2014GIA03 – awarded to SS, RB and MU. RB is funded by an Australian National Health and Medical Research Council (NHMRC) Senior Principal Research Fellowship.

References

1. Slade SC, Dionne CE, Underwood M, Buchbinder R. Standardised method for reporting exercise programs: Protocol for a modified Delphi study. *BMJ Open*. 2014; **4**: e006682.
2. Slade SC, Dionne CE, Underwood M, Buchbinder R and the CERT Panel. Consensus on Exercise Reporting Template (CERT): a modified Delphi study. *Phys Ther*. 2016; **96**: 1514-24.
3. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med*. 2010; **8**:24.

4. Conn VS, Cooper PS, Ruppap TM, Russell CL. Searching for the intervention in intervention research reports. *J Nurs Scholarsh.* 2008; **40**:52–9.
5. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet.* 2014; **383**:166–75.
6. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet.* 2014; **383**:267–76.
7. Hoffmann TC, Eructi C, Glasziou PP. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. *BMJ.* 2013; **347**: f3755.
8. Abell B, Glasziou P, Hoffman T. Reporting and Replicating Trials of Exercise-Based Cardiac Rehabilitation Do We Know What the Researchers Actually Did? *Circulation: Cardiovasc Qual Outcomes.* 2015; **8**:187-94.
9. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Elbourne D, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010; **340**: c869.
10. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gotzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Ann Int Med.* 2013; **158**: 200-07.
11. Slade SC, Dionne CE, Underwood M, Buchbinder R. The Consensus on Exercise Reporting Template (CERT): Explanation and Elaboration Statement. *BJSM.* 2016; **50**: 1428-37.
12. Hebert JJ, Fritz jm, Thackeray A, Koppenhaver SL, Teyhen D. Early multimodal rehabilitation following lumbar disc surgery: a randomised clinical trial comparing the effects of two exercise programmes on clinical outcome and lumbar multifidus muscle function. *BJSM.* 2015; **49**: 100-06.
13. Peterson M, Butler S, Eriksson M, Svardsudd K. A randomized controlled trial of eccentric vs. concentric graded exercise in chronic tennis elbow (lateral elbow tendinopathy). *Clin Rehabil.* 2014; **28**: 862-72.
14. de Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epi.* 2006; **59**(10):1033–39.
15. Chen G, Faris P, Hemmelgarn B, et al. Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC Med Res Methodol.* 2009; **9**:5.

16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159–74.
17. Bennell KL, Hunt MA, Wrigley TV, Hunter DJ, McManus FJ, Hodges PW, et al. Hip strengthening reduces symptoms but not knee load in people with medial knee osteoarthritis and varus malalignment: a randomised controlled trial. *Osteoarthritis Cart.* 2010; **18**: 621-28.
18. Bennell KL, Matthews B, Greig A, Briggs A, Kelly A, Sherburn M, et al. Effects of an exercise and manual therapy program on physical impairments, function and quality-of-life in people with osteoporotic vertebral fracture: a randomised, single-blind controlled pilot trial. *BMC Musc Disord.* 2010; **11**: 36..
19. Childs JD, Teyhen DS, Casey PR, McCoy-Singh KA, Feldtman AW, Wright AC, et al. Effects of traditional sit-up training versus core stabilization exercises on short-term musculoskeletal injuries in US Army soldiers: a cluster randomized trial. *Phys Ther.* 2010; **90**: 1404-12.
20. Cleland JA, Mintken P, McDevitt A, Bieniek ML, Carpenter KJ, Kulp K, et al. Manual physical therapy and exercise versus supervised home exercise in the management of patients with inversion ankle sprain: a multicenter randomized clinical trial. *JOSPT.* 2013; **43**: 443-55.
21. Crawshaw DP, Helliwell PS, Hensor EMA, Hay EM, Aldous SJ, Conaghan PG. Exercise therapy after corticosteroid injection for moderate to severe shoulder pain: large pragmatic randomised trial. *BMJ (Clinical research ed.)*. 2010; **340**: c3037.
22. Da Luz MA, Costa LOP, Fuhro FF, Manzoni AC, Oliveira NT, Cabral CM. Effectiveness of mat Pilates or equipment-based Pilates exercises in patients with chronic nonspecific low back pain: a randomized controlled trial. *Phys Ther.* 2014; **94**: 623-31.
23. Dolak KL, Silkman C, Medina-McKeon J, Hosey RG, Lattermann C, Uhl TL. Hip strengthening prior to functional exercises reduces pain sooner than quadriceps strengthening in females with patella-femoral pain syndrome: a randomized clinical trial. *JOSPT.* 2011; **41**: 560-70.
24. French HP, Cusak T, Brennan A, Caffrey A, Conroy R, Cuddy V, et al. Exercise and Manual Therapy Arthritis Research Trial (EMPART) for osteoarthritis of the hip: a multicentre randomised controlled trial. *Arch Phys Med Rehabil.* 2013; **94**: 302-14.
25. Hallgren HCB, Holmgren T, Oberg B, Johansson K, Adolfsson LE. A specific exercise strategy reduced the need for surgery in subacromial pain patients. *BJSM.* 2014; **48**: 1431-36.

26. Ismail MM, Gammalaldeen MH, Hassa KA. Closed Kinetic Chain exercises with or without additional hip strengthening exercises in management of patellofemoral pain syndrome: A randomized controlled trial. *Eur J Phys Rehabil Med.* 2013; **49**: 687-98.
27. Jorge RTB, de Souza MC, Chiari A, Jones A, Fernandes Ada R, Lombardi Junior I, et al. Progressive resistance exercise in women with osteoarthritis of the knee: a randomized controlled trial. *Clin Rehabil.* 2015; **29**: 234-43.
28. Kell RT, Risi AD, Barden JM The response of persons with chronic nonspecific low back pain to three different volumes of periodized musculoskeletal rehabilitation. *J Strength Cond Res.* 2011; **25**: 1052-1064.
29. Kromer TO, de Bie RA, Bastiaenen CHG. Effectiveness of individualized physiotherapy on pain and functioning compared to a standard exercise protocol in patients presenting with clinical signs of subacromial impingement syndrome. A randomized controlled trial. *BMC Musc Disord.* 2010; **11**: 114.
30. Macedo LG, Latimer J, Maher CG, Hodgs PW, McAuley JH, Nicholas MK, et al. Effect of motor control exercises versus graded activity in patients with chronic nonspecific low back pain: a randomized controlled trial. *Phys Ther.* 2012; **92**: 363-77.
31. Marshall PWM, Kennedy S, Brooks C, Lonsdale C. Pilates exercise or stationary cycling for chronic non-specific low back pain: does it matter? A randomized controlled trial with 6-month follow-up. *Spine.* 2013; **38**: E952-59.
32. Pisters MF, Veenhof C, Schellevis FG, De Bakker DH, Dekker J. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized controlled trial comparing two different physical therapy interventions. *Osteoarthritis Cart.* 2010; **18**: 1019-26.
33. Rasotto C, Bergamin M, Sieverdes JC, Gobbo S, Alberton CL, Neunhaeuserer D, et al. A tailored workplace exercise program for women at risk for neck and upper limb musculoskeletal disorders: a randomized controlled trial. *J Occup Environ Med.* 2015; **57**: 178-83.
34. Shnayderman I, Katz-Leurer M. An aerobic walking program versus muscle strengthening program for chronic low back pain: a randomized controlled trial. *Clin Rehabil.* 2013; **27**: 207-14.
35. Soderberg J, Grooten WJ, Ang BO. Effects of eccentric training on hand strength in subjects with lateral epicondylalgia: a randomized-controlled trial. *Scand J Med Sci Sports.* 2012; **22**: 797-03.

36. Stasinopoulos D, Manias P. Comparing two eccentric exercise programmes for the management of Achilles tendinopathy. A pilot trial. *J Bodywork Movement Ther.* **2013****17**: 309-15.
37. Bennell KL, Hunt MA, Wrigley TV, Hunter DJ, Hinman RS. The effects of hip muscle strengthening on knee load, pain, and function in people with knee osteoarthritis: a protocol for a randomised, single-blind controlled trial. *BMC Musculoskel Dis.* 2007; **8**: 121.
38. Childs JD, Teyhen DS, Benedict TM, Morris JB, Fortenberry AD, McQueen RM, et al. Effects of sit-up training versus core stabilization exercises on sit-up performance. *Med Sci Sports Exerc.* 2009; **41**:2072–83.
39. Bassett SF, Prapavessis H. Home-based physical therapy intervention with adherence-enhancing strategies versus clinic-based management for patients with ankle sprains. *Phys Ther.* 2007; **87**:1132-43.
40. de Luz MA, Costa, LOP, Fuhro FF, Manzoni ACT, de Oliveira NTB, Cabral CMNI. Effectiveness of mat Pilates or equipment-based Pilates in patients with chronic low back pain: a protocol of a RCT. *BMC Musc Disord.*2013; **14**: 16.
41. French HP, Cusack T, Brennan A, Caffrey A, Conroy R, Cuddy V, et al. Exercise and Manual Therapy Arthritis Research Trial (EMPART) for osteoarthritis of the hip: a multicentre randomised controlled trial. *BMC Musc Disord.* 2009; **10**: 9.
42. Holmgren T, Hallgren H, Oberg B, Aldofsson L, Johansson K. Effect of specific exercise strategy on need for surgery in patients with subacromial impingement syndrome: randomised controlled study. *BMJ.* 2012; **344**:e787.
43. Miller MG, Berry DC, Bullard S, Gilders R. Comparisons of land-based and aquatic-based plyometric programs during an 8-week training period. *J Sports Rehab.* 2002.**11**:269–83.
44. Miller MG, Herniman JJ, Ricard MD, Cheatham CC, Michael TJ. The Effects of 6-week plyometric training program on agility. *J Sports Sci Med.* 2006. **5**:459-65.
45. Piper T, Erdmann LD. A 4 step plyometric program. *J Strength Cond.* 1998. **20**:72-3.
46. Hodges PW, Ferreira PH, Ferreira ML. Lumbar Spine: treatment of instability and disorders of movement control: In: Magee D, Zachazewski JE, Quillen WS, eds. *Pathology and Intervention in Musculoskeletal Rehabilitation.* Amsterdam, The Netherlands: Elsevier Science BV; 2007: 389-25.
47. Ferreira ML, Ferreira PH, Latimer J, Herbert RD, Hodges PW, Jennings MD, et al. Comparison of general exercise, motor control exercise and spinal manipulative therapy in chronic low back pain: a randomised controlled trial. *Pain.* 2007; **131**:31-7.

48. Costa LOP, Maher CG, Latimer J, Hodges PW, Herbert RD, Refshauge KM, et al. Motor control exercises for chronic low back pain: a randomised placebo controlled trial. *Phys Ther.* 2009; **89**: 1275-86.
49. Veenhof C, Koke A, Dekker J, Oostendorp RA, Bijlsma JW, van Tulder MW, et al. Effectiveness of behavioral graded activity in patients with osteoarthritis of hip and/or knee: a randomized controlled trial. *Arthritis Rheum.* 2006; **55**:925e34.
50. Pisters MF, Veenhof C, de Bakker DH, Schellevis FG, Dekker J. Behavioural graded activity results in better exercise adherence and more physical activity than usual care in people with osteoarthritis: a cluster randomised trial. *J Physiother.* 2010; **56**:41-7.
51. Alfredson H, Pietila T, Jonsson P, Lorentzon R. 1998. Heavy load eccentric calf muscle training for the treatment of chronic Achilles tendinosis. *Am J Sports Med.* **26**: e366.
52. Day ML, McGuigan MR, Brice G, Foster C. Monitoring exercise intensity during resistance training using the session RPE scale. *J Strength Cond Res.* 2004; **18**: 353-58.
53. Stanish W, Rubinovich R, Curwin S. Eccentric training in chronic tendinitis. *Clin Orthopaedics Relat Res.* 1986; **208**: 65e68.
54. Wang C, Redgrave J, Shafizadeh M, Majid A, Kilner K, Ali AN. Aerobic exercise interventions reduce blood pressure in patients after stroke or transient ischaemic attack: a systematic review and meta-analysis. *BJSM.* 2018: doi: 10.1136/bjsports-2017-098903.
55. Kelly GA, Kelley KS, Callahan LF. Community-deliverable exercise and anxiety in adults with arthritis and other rheumatic diseases: a systematic review with meta-analysis of randomised controlled trials. *BMJ Open.* 2018; 8(2): 1-17.
56. MacCarthy A, Kirtley S, de Beyer JA, Altman DG, Simera I. Reporting guidelines for oncology research: helping to maximise the impact of your research. *Brit J Cancer.* 2018; **118**: 619-28.
57. Ghimire S, Kyung E, Kang W, Kim E. Assessment of adherence to the CONSORT statement for quality of reports on randomized controlled trial abstracts from four high-impact general medical journals. *Trials.* 2012; **13**: 77.
58. Sut N, Senocak M, Koksalan H. Assessing the quality of randomized controlled trials from two leading cancer journals using the CONSORT statement. *Hematol Oncol Stem Cell Ther.* 2008; **1**(1): 38-43.

59. Peters JPM, Hooft L, Grolman W, Stegeman I. Assessment of the Quality of Reporting of Randomised Controlled Trials in Otorhinolaryngologic Literature – Adherence to the CONSORT Statement. *PLoS ONE*. 2015;**10**(3): e0122328.
60. Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev*. 2012; **1**:60.
61. Abell B, Glasziou P, Hoffman T. Reporting and Replicating Trials of Exercise-Based Cardiac Rehabilitation Do We Know What the Researchers Actually Did? *Circulation: Cardiovasc Qual Outcomes*. 2015; **8**:187-94.
62. Prochaska, J. J., et al. Examination of the factor structure of physical activity behaviors. *J Clin Epidemiol*. 2000; **53**(8): 866-74.
63. Washburn, R. A., et al. The physical activity scale for the elderly (PASE): evidence for validity. *J Clin Epidemiol*. 1999; **52**(7): 643-51.
64. Weiss, T. W., et al. The validity of single-item, self-assessment questions as measures of adult physical activity. *J Clin Epidemiol*. 1990; **43**(11): 1123-29.

Appendices

Appendix 1: Consensus on Exercise Reporting Template

CERT ✓ Consensus on *E*xercise *R*eporting *T*emplate

A Checklist for what to include when reporting exercise programs

Section/Topic	Item #	Checklist item	Location **	
			Primary paper (page, table, appendix)	† Other (paper or protocol, website (URL))
WHAT: materials	1	Detailed description of the type of exercise equipment (e.g. weights, exercise equipment such as machines, treadmill, bicycle ergometer etc)	-	-
WHO: provider	2	Detailed description of the qualifications, teaching/supervising expertise, and/or training undertaken by the exercise instructor	-	-
HOW: delivery	3	Describe whether exercises are performed individually or in a group	-	-
	4	Describe whether exercises are supervised or unsupervised and how they are delivered	-	-
	5	Detailed description of how adherence to exercise is measured and reported	-	-
	6	Detailed description of motivation strategies	-	-
	7a	Detailed description of the decision rule(s) for determining exercise progression	-	-
	7b	Detailed description of how the exercise program was progressed	-	-
	8	Detailed description of each exercise to enable replication (e.g. photographs, illustrations, video etc)	-	-
9	Detailed description of any home program component (e.g. other exercises, stretching etc)	-	-	

	10	Describe whether there are any non-exercise components (e.g. education, cognitive behavioural therapy, massage etc)	-	-
	11	Describe the type and number of adverse events that occur during exercise	-	-
WHERE: location	12	Describe the setting in which the exercises are performed	-	-
WHEN, HOW MUCH: dosage	13	Detailed description of the exercise intervention including, but not limited to, number of exercise repetitions/sets/sessions, session duration, intervention/program duration etc	-	-
TAILORING: what, how	14a	Describe whether the exercises are generic (one size fits all) or tailored to the individual	-	-
	14b	Detailed description of how exercises are tailored to the individual	-	-
	15	Describe the decision rule for determining the starting level at which people commence an exercise program (such as beginner, intermediate, advanced etc)	-	-
HOW WELL: planned, actual	16a	Describe how adherence or compliance to the exercise intervention is assessed/measured	-	-
	16b	Describe the extent to which the intervention was delivered as planned	-	-

***It is recommended that this checklist is used in conjunction with the Explanation and Elaboration Statement which is a guide each item in the CERT Checklist**

The CERT Checklist is designed for reporting details of an exercise intervention. The CERT Checklist should be used in conjunction with a reporting checklist appropriate for the study type e.g. the CONSORT Statement (www.consort-statement.org) for randomised controlled trials, the SPIRIT Statement (www.spirit-statement.org) for a clinical trial protocol. For further guidance regarding reporting guidelines please consult the EQUATOR network (www.equator-network.org)

** Authors – please use N/A if an item is not applicable

Reviewers – please use “?” if information is not provided or not/insufficiently reported

† If the information is not provided in the primary paper that is under consideration, please provide details of where this information is available e.g. in a published protocol, published papers (provide citation details) or on a website (provide the URL).