If exercise is medicine, why don’t we know the dose? An overview of systematic reviews assessing reporting quality of exercise interventions in health and disease

Harrison J Hansford, Michael A Wewege, Aidan G Cashin, Amanda D Hagstrom, Briana K Clifford, James H McAuley, Matthew D Jones.

ABSTRACT
Objective To determine how well exercise interventions are reported in trials in health and disease.
Design Overview of systematic reviews.
Data sources PubMed, EMBASE, CINAHL, SPORTDiscus and PsycINFO from inception until June 2021.
Eligibility criteria Reviews of any health condition were included if they primarily assessed quality of exercise intervention reporting using the Consensus on Exercise Reporting Template (CERT) or the Template for Intervention Description and Replication (TIDieR). We assessed review quality using a modified version of A MeaSurement Tool to Assess systematic Reviews.
Results We identified 7804 studies and included 28 systematic reviews. The median (IQR) percentage of CERT and TIDieR items appropriately reported was 24% (19%) and 49% (33%), respectively. TIDieR items 1, Brief name (median (IQR) 100%, IQR 4) and 2, Why (median=98%, IQR 6), as well as CERT item 4, Supervision and delivery (median=68%, IQR 89), were the best reported. For replication of exercise interventions, TIDieR item 8, When and how much, was moderately well reported (median=62%, IQR 68) although CERT item 8, Description of each exercise to enable replication (median=23%, IQR 44) and item 13, Detailed description of the exercise intervention (median=24%, IQR 66) were poorly reported. Quality of systematic reviews ranged from moderate to critically low quality.
Conclusion Exercise interventions are poorly reported across a range of health conditions. If exercise is medicine, then how it is prescribed and delivered is unclear, potentially limiting its translation from research to practice.
PROSPERO registration number CRD42021261285; Open Science Framework: osf.io/my3ec.

INTRODUCTION
The benefits of exercise for preventing and managing chronic disease have been well described. Indeed, “with the possible exception of diet modification, we know of no single intervention with greater promise than physical exercise to reduce the risk of virtually all chronic diseases simultaneously.” Systematic reviews and meta-analyses show exercise to be similarly effective to medications for managing several chronic conditions, adding credibility to the notion that “exercise is medicine.” However, compared with trials of medicines, exercise trials tend to be of lower quality, at higher risk of bias and are less likely to report adverse events. Admittedly, some contribution to the reduced quality and higher risk of bias of exercise trials arises from difficulties in blinding participants and intervention providers. However, other important methodological features (eg, allocation concealment, analysis using intention-to-treat and blinding of assessors for objective measures) that should be used, often are not. Together, these methodological weaknesses limit confidence in the findings of exercise trials.

Notably, the reporting of interventions in exercise trials is also often poor, especially when compared with similar trials of medicines. This is significant because poor reporting of interventions in clinical trials impairs quality appraisal, evidence synthesis and replication, and limits the ability of stakeholders (eg, patients, clinicians, policymakers) to implement them into clinical practice. If the reporting of an intervention is poor, the intervention itself, or its ‘dose’ is unknown. To combat poor reporting of exercise interventions in clinical trials, the Consensus on Exercise Reporting Template (CERT)—a 16-item minimum data set considered necessary to report exercise interventions—was developed in 2016. The CERT adds to other intervention-specific reporting guidelines such as the Template for Intervention Description and Replication (TIDieR), developed in 2014, which aim to improve the reporting of interventions in clinical trials. Despite the advent of these checklists, and several others, reporting of exercise interventions remains poor and does not appear to have improved over time.

To illustrate the importance of reporting from research to clinical practice, take the example of a clinician who wants to prescribe an evidence-based exercise programme for their patient with patellofemoral pain. They find a methodologically rigorous systematic review showing that, based on moderate certainty evidence, exercise reduces patellofemoral pain compared with usual care. Hence, they deemed the exercise to be effective, and wish to replicate the intervention in practice. However, on reading the review, they discover the...
exercise interventions were poorly described, with little information on the type and dose, level of supervision and what co-interventions, if any, were delivered. The clinician is now uncertain what the ‘effective’ exercise programme was. The same could be said for clinicians working in exercise oncology, where key principles of training (eg, progression and reversibility) and prescriptive components of exercise (eg, frequency, intensity, time and type) are poorly reported.21

Several systematic reviews have been published that investigated the reporting quality of exercise interventions for various health conditions (eg, cardiovascular, musculoskeletal, neurological),19 22–24 but these are yet to be synthesised. These individual systematic reviews provide evidence for quality of reporting of specific health conditions; however, they do not inform on the quality of reporting across the entire field more broadly. Hence, the quality of reporting across exercise medicine literature remains unknown. The aim of this overview of systematic reviews was to determine how well exercise interventions have been reported in clinical trials of exercise for health and disease. For clinical research to be translated into practice, clinicians must be able to identify the intervention components in sufficient detail to replicate them. This is particularly important for complex interventions like exercise given the many modifiable variables that may impact its effectiveness.25–27

METHODS

This overview of systematic reviews was conducted in accordance with the Cochrane Handbook for Systematic Reviews of Interventions (Chapter V—Overviews) recommendations28 and reported in line with the Preferred Reporting Items for Systematic review and Meta-Analysis (PRISMA) 2020 statement.29 The protocol was registered prospectively on the Open Science Framework (OSF) (osf.io/my3ec/) and PROSPERO (CRD42021261285) prior to conducting searches. All data and code are publicly available on OSF (osf.io/my3ec/).

Eligibility criteria

We included systematic reviews of exercise interventions that specifically examined, as a primary aim, how well the exercise interventions were reported.

Searches

We searched electronic databases (PubMed, EMBASE, CINAHL, SPORTDiscus and PsycINFO) up to June 2021, using search terms relating to ‘exercise’ or ‘physical activity’ and ‘reporting’ (eg, CERT or TIDieR). We did not restrict the inclusion of reviews by year of publication, publication status or language. The search strategy for PubMed was as follows: (exercis*[Title] OR sport*[Title] OR physical activity*[Title] OR train*[Title] OR aerobic*[Title] OR resistance*[Title] OR physical training*[Title] OR active*[Title] OR move*[Title] OR rehab*[Title]) AND (cert| TiAb) OR tidier[TiAb] OR “template for intervention description and replication”[TiAb] OR report*[Title] OR complet*[TiAb] OR describ*[TiAb] OR replic*[TiAb] OR characteristics[TiAb] OR design[TiAb] OR program[TiAb]) OR (consensus on exercise reporting template) with a filter for systematic reviews in medicine. The search strategy for all other databases is shown in Table 1.

Results of electronic database searches were imported to Covidence19 where duplicate records were removed automatically. Two reviewers independently conducted two stages of eligibility screening: (1) title and abstract; (2) full text. Any disagreements on screening were resolved through discussion, with arbitration of a third author if required.

Record management and screening

Data from included reviews were extracted in duplicate by independent authors using Covidence extraction V.2.0.31 Disagreements were resolved through discussion or arbitration from a third author if required.

Data items: characteristics of included reviews

We extracted data relating to the review characteristics (included study design, health condition(s), sample size, number of interventions, exercise intervention type, comparator(s)) and reporting guideline used (CERT and/or TIDieR). If investigated by an included review, we also extracted data related to changes in reporting quality over time.

Data items: reporting characteristics

From the included systematic reviews, we extracted the completeness of reporting (primary outcome), expressed as the percentage of interventions that reported each item in sufficient detail for replication, according to the relevant checklist, TIDieR or CERT. All items of each of these checklists were considered in this overview (Table 1). If a combined or hybrid checklist was used, we separated the items from CERT and TIDieR and considered these separate checklists. We ourselves did not evaluate the completeness of reporting, rather, completeness of reporting was judged by the primary authors of the included systematic review. As a result, we relied on the level of reporting which the authors chose to obtain data from the primary trial manuscripts (ie, from the primary publication and supplemental material only). Similarly, we ourselves did not evaluate the change in reporting quality over time but instead used the judgements of the primary authors of the included systematic reviews. We contacted authors of the included systematic reviews when items on the relevant guideline were not reported. If a review included studies of multiple interventions, we extracted the completeness of reporting relative to the number of interventions, rather than the number of studies. We did not extract risk of bias ratings of individual studies as these were unlikely to affect the quality of reporting.

Data synthesis

We used R32 to conduct all analyses. From each review, we extracted the number or percentage of studies that appropriately described each item of the respective scale(s) (CERT or TIDieR). When data were presented in systematic reviews as the number of studies (ie, 6 of 24 studies reported the item sufficiently), it was converted to a percentage of studies for analysis to allow comparability between reviews. Data were synthesised using simple descriptive statistics (median, IQR and range) for each item of each relevant tool. Data were visually inspected for normality. Most data were not normally distributed; therefore, for consistency, the median was chosen as the summary statistic. We performed subgroup analyses on the completeness of reporting within different health and disease areas when >3 reviews of the same area were identified (eg, cardiovascular, musculoskeletal, neurological, etc). Studies were grouped into the above areas
Table 1  Items of included reporting guidelines, TiDIEr (A) items of TiDIEr checklist14 and CERT (B) items of CERT checklist13

<table>
<thead>
<tr>
<th>TiDIEr items</th>
<th>Item description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Brief name: provide the name or a phrase that describes the intervention</td>
</tr>
<tr>
<td>2</td>
<td>Why: describe any rationale, theory or goal of the elements essential to the intervention</td>
</tr>
<tr>
<td>3</td>
<td>What (materials): describe any physical or informational materials used in the intervention, including those provided to participants or used in intervention delivery or in training of intervention providers. Provide information on where the materials can be accessed (for example, online appendix, URL)</td>
</tr>
<tr>
<td>4</td>
<td>What (procedures): describe each of the procedures, activities, and/or processes used in the intervention, including any enabling or support activities</td>
</tr>
<tr>
<td>5</td>
<td>Who provided: for each category of intervention provider (for example, psychologist, nursing assistant), describe their expertise, background and any specific training given</td>
</tr>
<tr>
<td>6</td>
<td>How: describe the modes of delivery (such as face to face or by some other mechanism, such as internet or telephone) of the intervention and whether it was provided individually or in a group</td>
</tr>
<tr>
<td>7</td>
<td>Where: describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features</td>
</tr>
<tr>
<td>8</td>
<td>When and how much: describe the number of times the intervention was delivered and over what period of time including the number of sessions, their schedule, and their duration, intensity or dose</td>
</tr>
<tr>
<td>9</td>
<td>Tailoring: if the intervention was planned to be personalised, titrated or adapted then describe what, why, when and how</td>
</tr>
<tr>
<td>10</td>
<td>Modifications: if the intervention was modified during the course of the study, describe the changes (what, why, when and how)</td>
</tr>
<tr>
<td>11</td>
<td>How well (planned): if intervention adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them</td>
</tr>
<tr>
<td>12</td>
<td>How well (actual): if intervention adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CERT items</th>
<th>Item description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Detailed description of the type of exercise equipment (eg, weights, exercise equipment such as machines, treadmill, bicycle ergometer, etc)</td>
</tr>
<tr>
<td>2</td>
<td>Detailed description of the qualifications, teaching/supervising expertise, and/or training undertaken by the exercise instructor</td>
</tr>
<tr>
<td>3</td>
<td>Describe whether exercises are performed individually or in a group</td>
</tr>
<tr>
<td>4</td>
<td>Describe whether exercises are supervised or unsupervised and how they are delivered</td>
</tr>
<tr>
<td>5</td>
<td>Detailed description of how adherence to exercise is measured and reported</td>
</tr>
<tr>
<td>6</td>
<td>Detailed description of motivation strategies</td>
</tr>
<tr>
<td>7a</td>
<td>Detailed description of the decision rule(s) for determining exercise progression</td>
</tr>
<tr>
<td>7b</td>
<td>Detailed description of how the exercise programme was progressed</td>
</tr>
<tr>
<td>8</td>
<td>Detailed description of each exercise to enable replication (eg, photographs, illustrations, video, etc)</td>
</tr>
<tr>
<td>9</td>
<td>Detailed description of any home programme component (eg, other exercises, stretching, etc)</td>
</tr>
<tr>
<td>10</td>
<td>Describe whether there are any non-exercise components (eg, education, cognitive–behavioural therapy, massage, etc)</td>
</tr>
<tr>
<td>11</td>
<td>Describe the type and number of adverse events that occurred during exercise</td>
</tr>
<tr>
<td>12</td>
<td>Describe the setting in which the exercises are performed</td>
</tr>
<tr>
<td>13</td>
<td>Detailed description of the exercise intervention including, but not limited to, number of exercise repetitions/sets/sessions, session duration, intervention/programme duration, etc</td>
</tr>
<tr>
<td>14a</td>
<td>Describe whether the exercises are generic (one size fits all) or tailored to the individual</td>
</tr>
<tr>
<td>14b</td>
<td>Detailed description of how exercises are tailored to the individual</td>
</tr>
<tr>
<td>15</td>
<td>Describe the decision rule for determining the starting level at which people commence an exercise programme (such as beginner, intermediate, advanced, etc)</td>
</tr>
<tr>
<td>16a</td>
<td>Describe how adherence or fidelity to the exercise intervention is assessed/measured</td>
</tr>
<tr>
<td>16b</td>
<td>Describe the extent to which the intervention was delivered as planned</td>
</tr>
</tbody>
</table>

CERT, Consensus on Exercise Reporting Template; TiDIEr, Template for Intervention Description and Replication.

based on the domains described in Exercise and Sports Science Australia’s standards.33 As several methods were used by review authors to analyse changes in quality of reporting over time (eg, correlations, linear regression, completeness of reporting across different time periods or before and after the introduction of CERT and TiDIEr), we did not pool these results. Instead, our analysis of changes in reporting over time was described narratively. We did not assess the certainty of evidence as this was not relevant to the purpose of our overview of systematic reviews. CERT and TiDIEr do not define ‘good’ or ‘poor’ reporting13,14; however, post hoc, we categorised reporting quality as ‘good’ when ≥80% of interventions included in the reviews reported the item(s) sufficiently, ‘moderate’ when 79% to 50% reported the items sufficiently and ‘poor’ when <50% reported items sufficiently; in line with included reviews’ cut-offs.34

Quality of systematic reviews

Review quality was assessed independently and in duplicate using a modified version of A Measurement Tool to Assess systematic Reviews (AMSTAR 2)36 (online supplemental appendix 3). In our modified version, we excluded the items pertaining to meta-analysis or risk of bias within individual studies (items 9, 11, 12, 13, 14 and 15) as these were not relevant to our review question. The quality of each review was deemed ‘high’, ‘moderate’, ‘low’ or ‘critically low’ based on the number of critical flaws (a rating of ‘no’ in items 2, 4 and 7)36 or non-critical weaknesses (a ‘no’ or ‘partial yes’ in any other domain) with the review (online supplemental appendix 4).

RESULTS

We identified 7804 studies and included 28 systematic reviews18-20 22 24 34 35 37-57 (figure 1). These 28 reviews included 1467 studies comprised of 1724 interventions. We found only one article published in a language other than English—German—which was translated using Google Translate.58 A list of the studies excluded during full-text assessment, with reasons, is provided in online supplemental appendix 5. Ten reviews only used CERT,35 38 41 42 44 48 52-54 56 11 reviews only
used TIDieR, and used both CERT and TIDieR (table 2). One review used a hybrid tool comprised of items 1–5, 7, and 9–12 from TIDieR, complemented and expanded upon by items 6 and 8 of CERT. The median number of interventions included in the reviews was 24 (range 3–287, IQR 67). Twelve reviews assessed quality of reporting in musculoskeletal conditions, four in neurological conditions, six in cardiovascular conditions, four in cancer and five in ‘other’ conditions including urinary dysfunction, pelvic organ prolapse, organ transplant patients and older adults. See table 2 for all characteristics of included reviews. We contacted two review authors, who provided data not reported in the review manuscripts.

Quality of included reviews (AMSTAR 2)
Nine reviews were rated moderate quality, 11 low quality and 8 critically low quality. The most common methodological shortcomings were item 10, Reporting sources of funding of included studies, where 28 reviews (100%) did not report the item sufficiently, and item 3, Rationale for selection of study designs, where 22 reviews (79%) did not report the item sufficiently. The most adhered to item was item 6, Conducting data extraction in duplicate, with 23 reviews (82%) reporting this sufficiently (see online supplemental appendix 6 for the complete results of the AMSTAR 2 assessment).

Quality of reporting: CERT
Sixteen reviews used CERT to assess quality of reporting (n=643 studies, n=1157 interventions). The median percentage of all CERT items appropriately reported was 24% (range 5%–68%, IQR 19). The median score for each CERT item across the 16 reviews can be seen in figure 2. Item 4, Describe whether exercises are supervised or unsupervised and how they are delivered (median=68%, range 0%–100%, IQR 89) and Item 14, Describe whether the exercises are generic (one size fits all) or tailored to the individual (median=25%, range 0%–100%, IQR 70) were the best reported. In contrast, item 16a, Describe how adherence or fidelity to the exercise intervention is assessed/measured (median=5%, range 0%–93%, IQR 21) and item 16b, Describe the extent to which the intervention was delivered as planned (median=5%, range 0%–77%, IQR 38) were the most poorly reported. Some of the items most important for replication, item 8, Description of each exercise to enable replication (median=23%, range 0%–95%, IQR 44) and item 13, Detailed description of the exercise intervention, including sets, reps, duration, etc (median=24%), range 0%–100%, IQR 66) were also poorly reported. There were no obvious differences in CERT scores between health condition subgroups (online supplemental appendix 7).

Quality of reporting: TIDieR
Eighteen reviews used TIDieR to assess quality of reporting (n=1099 studies, n=1353 interventions). The median percentage of all TIDieR items appropriately reported was 49% (range 0%–100, IQR 33). The median score for each TIDieR item across the 18 reviews can be seen in figure 3. Item 1, Brief name (median=100%, range 0%–100%, IQR 0) and item 2, Why (median=98%, range 0%–100%, IQR 6) were the best reported. In contrast, item 10, Modifications (median=0%, range 0%–55%, IQR 12) and item 11, How well (planned) (median=23%, range 0%–70%, IQR 26) were the most poorly reported. The most relevant item to the ‘dose’ of exercise, item 8, When and how much, was moderately well reported (median=62%, range 0%–100%, IQR 68). Subgroup analyses (online supplemental appendix 8) showed the neurological area had the highest median score (65% (range 2%–100%, IQR 62)), followed by the cardiovascular area (48% (range 0%–100%, IQR 23)) and the ‘other’ area (43% (range 0%–100%, IQR 30)).

Changes in reporting over time
Five reviews investigated changes in reporting quality over time, but the findings were mixed. Three reviews found no changes over time. One review found slight decreases in reporting quality over time, whereas another found improvements in reporting quality over time (table 3).

DISCUSSION
Our overview of systematic reviews revealed that exercise interventions are poorly reported across all health and disease
<table>
<thead>
<tr>
<th>Study (author, year)</th>
<th>Trial designs included</th>
<th>Health condition</th>
<th>Number of participants</th>
<th>Number of interventions</th>
<th>Type of exercise</th>
<th>Template used</th>
<th>AMSTAR 2 rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abell et al(^\text{18}) 2015</td>
<td>RCTs</td>
<td>Coronary artery disease</td>
<td>NR</td>
<td>74</td>
<td>Supervised or unsupervised exercise programme, with or without the addition of lifestyle modification and counselling</td>
<td>TIDieR</td>
<td>Critically low quality</td>
</tr>
<tr>
<td>Barros et al(^\text{37}) 2020</td>
<td>RCTs</td>
<td>Low back pain</td>
<td>NR</td>
<td>10</td>
<td>Pilates</td>
<td>Both</td>
<td>Low quality</td>
</tr>
<tr>
<td>Bartholdy et al(^\text{38}) 2019</td>
<td>Clinical guidelines</td>
<td>Knee osteoarthritis</td>
<td>9819</td>
<td>130</td>
<td>Exercise*</td>
<td>Hybrid</td>
<td>Moderate quality</td>
</tr>
<tr>
<td>Berti et al(^\text{39}) 2020</td>
<td>Randomised, controlled and uncontrolled clinical trials, and case reports</td>
<td>Parkinson’s disease</td>
<td>386</td>
<td>21</td>
<td>Argentine tango</td>
<td>TIDieR</td>
<td>Moderate quality</td>
</tr>
<tr>
<td>Burgess et al(^\text{40}) 2021</td>
<td>RCTs</td>
<td>Hip osteoarthritis</td>
<td>707</td>
<td>14</td>
<td>Exercise*</td>
<td>Both</td>
<td>Low quality</td>
</tr>
<tr>
<td>Charlton et al(^\text{41}) 2017</td>
<td>RCT, case study, case series, risk stratified prevention</td>
<td>Prevention or treatment of groin pain in athletes</td>
<td>2042</td>
<td>14</td>
<td>Exercise*</td>
<td>CERT</td>
<td>Moderate quality</td>
</tr>
<tr>
<td>Davidson et al(^\text{42}) 2021</td>
<td>RCTs</td>
<td>Low back pain</td>
<td>NR</td>
<td>180</td>
<td>Exercise*</td>
<td>Both</td>
<td>Moderate quality</td>
</tr>
<tr>
<td>Dichiavi et al(^\text{43}) 2021</td>
<td>RCTs</td>
<td>Patellofemoral pain</td>
<td>NR</td>
<td>19</td>
<td>Exercise*</td>
<td>CERT</td>
<td>Low quality</td>
</tr>
<tr>
<td>Giagio et al(^\text{44}) 2021</td>
<td>RCTs</td>
<td>Pelvic organ prolapse</td>
<td>4076</td>
<td>26</td>
<td>Pelvic floor muscle training</td>
<td>Both</td>
<td>Moderate quality</td>
</tr>
<tr>
<td>Hall et al(^\text{45}) 2018</td>
<td>Any study type</td>
<td>Men with urinary dysfunction</td>
<td>NR</td>
<td>108</td>
<td>Pelvic floor muscle training</td>
<td>CERT</td>
<td>Moderate quality</td>
</tr>
<tr>
<td>Holden et al(^\text{46}) 2018</td>
<td>RCTs</td>
<td>Patellofemoral pain</td>
<td>NR</td>
<td>38</td>
<td>Exercise*</td>
<td>TIDieR</td>
<td>Moderate quality</td>
</tr>
<tr>
<td>Kattakad et al(^\text{47}) 2020</td>
<td>RCTs</td>
<td>Juvenile idiopathic arthritis</td>
<td>NR</td>
<td>10</td>
<td>Physical activity interventions</td>
<td>Both</td>
<td>Low quality</td>
</tr>
<tr>
<td>Keene et al(^\text{48}) 2020</td>
<td>RCT or quasi-RCT</td>
<td>Fragility fracture of lower limb or pelvis</td>
<td>3564</td>
<td>65</td>
<td>Exercise*</td>
<td>TIDieR</td>
<td>Moderate quality</td>
</tr>
<tr>
<td>Knols et al(^\text{49}) 2018</td>
<td>RCTs</td>
<td>Lung transplant patients</td>
<td>369</td>
<td>7</td>
<td>Exercise*</td>
<td>TIDieR</td>
<td>Moderate quality</td>
</tr>
<tr>
<td>Knols et al(^\text{50}) 2019</td>
<td>RCTs</td>
<td>Traumatic brain injury</td>
<td>81</td>
<td>3</td>
<td>Early mobilisation</td>
<td>TIDieR</td>
<td>Critically low quality</td>
</tr>
<tr>
<td>Lohse et al(^\text{51}) 2018</td>
<td>RCTs</td>
<td>Stroke</td>
<td>12,847</td>
<td>287</td>
<td>Any physical or occupational therapy intervention that required active movement on the part of the participant</td>
<td>TIDieR</td>
<td>Low quality</td>
</tr>
<tr>
<td>Mack et al(^\text{52}) 2018</td>
<td>RCTs</td>
<td>Osteoporosis</td>
<td>NR</td>
<td>23</td>
<td>Exercise*</td>
<td>CERT</td>
<td>Critically low quality</td>
</tr>
<tr>
<td>Major et al(^\text{53}) 2019</td>
<td>RCTs</td>
<td>Rotator cuff disease</td>
<td>NR</td>
<td>34</td>
<td>Strengthening, scapula stabilising, stretching and ‘corrective’ exercises</td>
<td>CERT</td>
<td>Critically low quality</td>
</tr>
<tr>
<td>McEwen et al(^\text{54}) 2019</td>
<td>RCTs</td>
<td>Stroke</td>
<td>1267</td>
<td>16</td>
<td>Group circuit class therapy</td>
<td>Both</td>
<td>Low quality</td>
</tr>
<tr>
<td>McGregor et al(^\text{55}) 2018</td>
<td>Randomised and non-randomised controlled trials, non-controlled prospective observational studies</td>
<td>Pulmonary hypertension</td>
<td>NR</td>
<td>19</td>
<td>Aerobic or strength exercise</td>
<td>CERT</td>
<td>Low quality</td>
</tr>
<tr>
<td>Meneses-Echavez et al(^\text{56}) 2019</td>
<td>RCTs</td>
<td>Cancer</td>
<td>NR</td>
<td>138</td>
<td>Aerobic, resistance and flexibility training, as well as yoga, Qi-gong and Tai-Chi</td>
<td>TIDieR</td>
<td>Low quality</td>
</tr>
<tr>
<td>Neele et al(^\text{57}) 2016</td>
<td>RCTs</td>
<td>Traumatic brain injury</td>
<td>NR</td>
<td>58</td>
<td>Exercise*</td>
<td>TIDieR</td>
<td>Critically low quality</td>
</tr>
<tr>
<td>Raje et al(^\text{58}) 2021</td>
<td>RCTs</td>
<td>Solid organ transplant recipients</td>
<td>NR</td>
<td>21</td>
<td>Aerobic, resistance or combination</td>
<td>CERT</td>
<td>Low quality</td>
</tr>
<tr>
<td>Tew et al(^\text{59}) 2016</td>
<td>RCTs</td>
<td>Peripheral arterial disease</td>
<td>NR</td>
<td>76</td>
<td>Supervised exercise</td>
<td>TIDieR</td>
<td>Critically low quality</td>
</tr>
<tr>
<td>Torres-Pareja et al(^\text{60}) 2019</td>
<td>RCTs and uncontrolled studies</td>
<td>Multiple sclerosis</td>
<td>163</td>
<td>10</td>
<td>Flexibility/range of motion</td>
<td>CERT</td>
<td>Critically low quality</td>
</tr>
</tbody>
</table>
areas. This was true regardless of the reporting template used, though completeness of reporting was slightly higher according to TIDieR than CERT. Completeness of reporting does not appear to have improved over time, and most reviews were of low quality. Based on these findings, if exercise is medicine, then how it is prescribed and delivered is unclear, potentially limiting its translation from research to practice.

Maintaining a high quality of intervention reporting is important in all fields of medicine, including exercise. Poor reporting of interventions may limit the ability of clinicians and policymakers to implement interventions in clinical practice, as it may be unclear how interventions should be delivered.\textsuperscript{11,12} For example, if any intervention was shown to improve an important aspect of health (eg, blood pressure) or fitness (eg, aerobic capacity or muscle strength), it is important to know the characteristics of the intervention that led to this improvement. Further, in an increasingly global field, it can be confusing with many different naming conventions of exercise(s) within and between disciplines both nationally and internationally. Descriptions of exercises, including pictures, could help combat this issue and enhance the quality of reporting.\textsuperscript{19,20} Evidence synthesis is also impaired by poor reporting as comparators and interventions may not be pooled for meta-analysis if the content of the treatments is unclear.\textsuperscript{21,22} High-quality reporting is needed in the field of exercise in order to promote clinical translation, evidence synthesis and clear appraisal of studies.

Poor reporting of interventions is not unique to exercise. Indeed, similar issues have been observed across a range of medical interventions,\textsuperscript{62} but exercise studies appear to more poorly report interventions.\textsuperscript{11} Our results show that the names of, and rationale for, exercise (TIDieR items 1 and 2) were very well reported, but this is of little use for researchers or clinicians trying to replicate the intervention. In contrast, key intervention components needed to optimise translation to practice, for example, detailed description of exercises to enable replication and, perhaps most crucially, detailed description of the exercise prescription, were poorly reported (figure 2). Moreover, items crucial to assessing intervention fidelity, adherence and adverse events were also poorly reported (figure 2). Intervention fidelity has important implications for the internal validity of a study,\textsuperscript{63} whereas reporting of adherence and adverse events is crucial to enable assessment of how tolerable and feasible the intervention was. To improve quality appraisal, evidence synthesis, replication and translation of exercise interventions to practice, reporting of exercise interventions must improve.

Several templates have been developed to assist in improving the reporting of exercise interventions. These include condition-specific tools (eg, CERT-PFMT\textsuperscript{64}) for pelvic floor muscle training and more general templates.\textsuperscript{13,16} We chose to use CERT and TIDieR for this overview of systematic reviews as they are valid and reliable\textsuperscript{13,14} and focus on key intervention variables such as how, how much and how well, among others.\textsuperscript{13,14} The CERT was designed to build upon TIDieR to provide additional detail of important exercise intervention components.\textsuperscript{13} Interestingly, while the included reviews scored poorly well on TIDieR (median=49%, IQR 33) they scored much worse on CERT (median=24%, IQR 19). This disparity may be explained by the broad nature of TIDieR whereby, in trying to cover all healthcare interventions,\textsuperscript{62} it is too general for exercise. Based on the specificity of CERT to exercise, we recommend that authors use CERT to guide reporting of their exercise interventions. Our overview of systematic reviews showed that when important intervention components are defined and examined with more
scrutiny, as exercise is with CERT, items crucial to the replicability of exercise interventions are poorly reported.

Despite the advent of TIDieR and CERT, there has been little change in the quality of reporting of exercise interventions over time. The reason for this is not clear. It may be that authors are unaware of these templates. Indeed, it can be difficult to navigate the hundreds of reporting guidelines available on the Enhancing the QUAlity and Transparency Of health Research Network (equator-network.org). Alternatively, authors may be aware of these templates but simply choose not to use them. We acknowledge that full adherence to reporting guidelines can be difficult, particularly with the stringent word limits of many journals. In this instance, we suggest authors provide as much detail as possible within the manuscript and provide all other additional information required by CERT or TIDieR (or other relevant reporting guideline) as supplemental material. There may also be other methods to improve reporting of exercise trials. Journals have previously mandated the use of reporting guidelines such as Consolidated Standards of Reporting Trials66 and PRISMA,67 which significantly improved the reporting of trials and systematic reviews, respectively.68–70 Therefore, exercise medicine journals may be well positioned to improve the reporting quality of the research they publish by encouraging, or preferably requiring, submission of a completed CERT checklist when exercise trials are submitted. Without this, the quality of reporting of exercise interventions may remain poor, limiting the possibility of potentially impactful exercise interventions being implemented in clinical practice.

A noted limitation of the evidence included in our overview of systematic reviews was that, using a modified AMSTAR 2 tool,
Exercise is an intervention with widespread positive effects on many health conditions. But, across all fields involving exercise medicine, the quality of exercise intervention reporting is poor. High-quality reporting is needed to improve quality appraisal, enable evidence synthesis and replication, and improve translation in clinical settings. There has been little change in quality of reporting over time despite the presence of reporting checklists. Researchers, and the journals they submit to, have the opportunity to improve intervention reporting in exercise medicine by following TIDieR or CERT and encouraging or requiring inclusion of a completed checklist as part of standard practice when do not sufficiently report comparators,2 which is important for assessing the internal validity of the trial.

**CONCLUSION**

Exercise is effective for improving a range of health conditions, although exercise interventions are often poorly reported.

⇒ Poor reporting of interventions can reduce the ability for readers and researchers to assess quality, synthesise evidence, replicate and implement potentially effective interventions into practice.

⇒ The quality of reporting across studies of exercise medicine is unknown.
What are the new findings

⇒ Exercise interventions are poorly reported across all health areas of exercise medicine.
⇒ The quality of intervention reporting has not improved over time.
⇒ If exercise is medicine, then how is it prescribed and delivered is unclear, potentially limiting its translation from research to practice.

submitting exercise studies. This would likely lead to improved reporting quality over time, and a better understanding of the ‘dose’ of exercise medicine needed to optimise health outcomes.

Correction notice This article has been corrected since it published Online First. Author names have been updated.

Twitter Harrison J Hansford @HJHansford, Aidan G Cashin @AidanCashin, Amanda D Hagstrom @ADHagstrom, James H McAuley @pain_neura and Matthew D Jones @Mattjones0203

Contributors MDI is the guarantor and conceived the study. HJH and MDI wrote the first draft of the protocol and manuscript. HJH, MAW, AGC and MDI conducted article screening. HJH, MAW, AGC, AH, BKC and MDI extracted data. MDI assisted in interpreting the data. HJH conducted data analysis. All authors contributed importantly to the content and style of the protocol and manuscript. All authors approved the final version.

Funding MAW was supported by a Postgraduate Scholarship from the National Health and Medical Research Council of Australia, a School of Medical Sciences Top-Up Scholarship from the University of New South Wales, and a PhD Supplementary Scholarship from Neuroscience Research Australia.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study does not involve human participants.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Harrison J Hansford http://orcid.org/0000-0002-5942-8509
Michael A Weege http://orcid.org/0000-0002-3283-2149
Aidan G Cashin http://orcid.org/0000-0003-4190-7912
Amanda D Hagstrom http://orcid.org/0000-0002-8036-9216
Matthew D Jones http://orcid.org/0000-0002-5534-755X

REFERENCES
25 Clark JE. The impact of duration on effectiveness of exercise, the implication for periodization of training and goal setting for individuals who are overweight, a meta-analysis. Biomed Sport 2016;3:309–33.
31 Covidence systematic review software [program]. Melbourne Australia: Veritas Health Innovation 2019.
33 Exercise and sports science Australia. accredited exercise physiologist professional standards. Albion, Queensland 2015.


9 of 10
Review

Supplementary Material

**Appendix 1: Search strategy**

**PubMed – 28 Jun. 21**


- Filter systematic review, humans

**EMBASE – 22 Jun. 21**

((exercis* or sport* or physical activity or train* or aerobic or resistance or physical training or active or move* or rehab*) and (cert or "consensus on exercise reporting template" or tidier or "template for intervention description and replication" or report* or complet* or describ* or replic* or characteristics or design or program) and systematic review).m_titl.

**CINAHL & SPORTDiscus – 22 Jun. 21**

TI (exercise or physical activity or fitness) AND TI (report* or cert or tidier or "consensus on exercise reporting template" or "template for intervention description and replication") AND TI systematic review

**PSYCHInfo – 28 Jun. 21**

ab(systematic review) AND ab(exercise OR physical activity OR fitness) AND ab(report* OR cert OR tidier OR "consensus on exercise reporting template" OR "template for intervention description and replication")

Filter -systematic review

**Appendix 2.** Modified AMSTAR 2 tool used in this overview of reviews

1. Did the research questions and inclusion criteria for the review include the components of PICO?

<table>
<thead>
<tr>
<th>For Yes:</th>
<th>Optional (recommended)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Timeframe for follow up</td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
</tr>
<tr>
<td>Comparator group</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td></td>
</tr>
</tbody>
</table>

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

For Partial Yes:
The authors state that they had a written protocol or guide that included ALL the following:

- review question(s)
- a search strategy
- inclusion/exclusion criteria
- a risk of bias assessment

For Yes:
As for partial yes, plus the protocol should be registered and should also have specified:

- a meta-analysis/synthesis plan, if appropriate, and
- a plan for investigating causes of heterogeneity
- a plan for investigating causes of heterogeneity

3. Did the review authors explain their selection of the study designs for inclusion in the review?

For Yes, the review should satisfy ONE of the following:

- Explanation for including only RCTs
  
- OR Explanation for including only NRSI

OR Explanation for including both RCTs and NRSI

4. Did the review authors use a comprehensive literature search strategy?

For Partial Yes (all the following):
searched at least 2 databases (relevant to research question)
provided key word and/or search strategy
justified publication restrictions (e.g. language)

For Yes, should also have (all the following):
searched the reference lists / bibliographies of included studies
searched trial/study registries
included/consulted content experts in the field
where relevant, searched for grey literature
conducted search within 24 months of completion of the review
5. Did the review authors perform study selection in duplicate?
For Yes, either ONE of the following:

- at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include

OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.

6. Did the review authors perform data extraction in duplicate?
For Yes, either ONE of the following:

- at least two reviewers achieved consensus on which data to extract from included studies

OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.

7. Did the review authors provide a list of excluded studies and justify the exclusions?
For Partial Yes:
- provided a list of all potentially relevant studies that were read in full-text form but excluded from the review

For Yes, must also have:
- Justified the exclusion from the review of each potentially relevant study

8. Did the review authors describe the included studies in adequate detail?
For Partial Yes (ALL the following):
- described populations
- described interventions
- described comparators
- described outcomes
- described research designs

For Yes, should also have ALL the following:
- described population in detail
- described intervention in detail (including doses where relevant)
- described comparator in detail (including doses where relevant)
- described study’s setting
- timeframe for follow-up
10. Did the review authors report on the sources of funding for the studies included in the review?
   
   For Yes
   
   Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies
   
   Yes
   
   No

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

   For Yes:
   
   The authors reported no competing interests OR
   
   The authors described their funding sources and how they managed potential conflicts of interest

   Yes
   
   No

**Appendix 3.** Rating overall confidence in the results of the systematic review (AMSTAR 2)

**High**

- No or one non-critical weakness: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest

**Moderate**

- More than one non-critical weakness*: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review

**Low**
- One critical flaw with or without non-critical weaknesses: the review has a critical flaw and may not provide an accurate and comprehensive summary of the available studies that address the question of interest

**Critically low**

- More than one critical flaw with or without non-critical weaknesses: the review has more than one critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies

* Multiple non-critical weaknesses may diminish confidence in the review and it may be appropriate to move the overall appraisal down from moderate to low confidence.

Note: A critical flaw was a rating was deemed a ‘no’ in any of the critical domains, as specified by the AMSTAR 2 guidance document. These domains were Items 2, 4, 7. A non-critical weakness was a no, or partial yes in any other domain.
Appendix 4: Excluded studies, with reasons

**Wrong outcomes:**


Allen NE, Suriyarachchi GD, Paul SS, et al. Exercise and motor training in people with Parkinson's disease: A systematic review of participant characteristics, intervention delivery, retention rates, adherence, and adverse events in clinical trials. *Parkinson's Disease* 2012((Allen, Suriyarachchi, Paul, Song, Canning) Clinical and Rehabilitation Research Group, Faculty of Health Sciences, University of Sydney, P.O. Box 170, Lidcombe, Sydney, NSW 1825, Australia(Sherrington) Musculoskeletal Division, George Institute for Global) 854328. doi: http://dx.doi.org/10.1155/2012/854328


Bonafiglia JT, Islam H, Preobrazenski N, et al. Risk of bias and reporting practices in studies comparing VO2max responses to sprint interval vs. continuous training: A systematic review and meta-analysis. *Journal of Sport and Health Science* 2021((Bonafiglia, Islam, Preobrazenski, Gurd) School of Kinesiology and Health Studies, Queen's University, Kingston, ON K7L 3N6, Canada) doi: http://dx.doi.org/10.1016/j.jshs.2021.03.005


Goff AJ, Page WS, Clark NC. Reporting of acute programme variables and exercise descriptors in rehabilitation strength training for tibiofemoral joint soft tissue
injury: A systematic review. *Physical Therapy in Sport* 2018;34((Goff) Health and Social Sciences, Singapore Institute of Technology, 10 Dover Drive 138683, Singapore(Page, Clark) Faculty of Sport, Health and Applied Sciences, St Mary’s University, Waldegrave Road, Twickenham TW1 4SX, United Kingdom).227-37. doi: http://dx.doi.org/10.1016/j.ptsp.2018.10.012


Sauzet O, Kleine M, Exner A-K, et al. Longitudinal randomised controlled trials in rehabilitation post-stroke: A systematic review on the quality of reporting and use of baseline outcome values. *BMC Neurology* 2015(Sauzet, Kleine, Exner) AG Epidemiology and International Public Health, School of Public Health, Bielefeld University, PO. Box 10 01 31, Bielefeld, 33501 Germany(Menzel-Begemann) Faculty of Nursing and Health, University of Applied Science, Munster, Leon) doi: http://dx.doi.org/10.1186/s12883-015-0344-y


Reporting not primary aim


Not systematic review

Wrong intervention


### Appendix 5: AMSTAR Assessment

<table>
<thead>
<tr>
<th>Study</th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
<th>Item 7</th>
<th>Item 8</th>
<th>Item 10</th>
<th>Item 16</th>
<th>Overall Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abell 2015</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Critically Low</td>
</tr>
<tr>
<td>Barros 2020</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Low</td>
</tr>
<tr>
<td>Bartholdy 2019</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Berti 2020</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>/</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Burgess 2021</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>+</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Low</td>
</tr>
<tr>
<td>Charlton 2017</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Davidson 2021</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Moderate</td>
</tr>
<tr>
<td>Dischiavi 2021</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>+</td>
<td>Low</td>
</tr>
<tr>
<td>Giagio 2021</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Hall 2018</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Holden 2018</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td>Kattackal 2020</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Keene 2020</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td>Knols 2018</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td>Knols 2019</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td>Critically Low</td>
</tr>
<tr>
<td>Lohse 2018</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Mack 2018</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Critically Low</td>
</tr>
<tr>
<td>Major 2019</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Critically Low</td>
</tr>
<tr>
<td>McEwen 2019</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>McGregor 2018</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Meneses-Echavez 2019</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Study</td>
<td>Yes</td>
<td>/ Partial</td>
<td>No</td>
<td>Yes</td>
<td>/ Partial</td>
<td>No</td>
<td>Yes</td>
<td>/ Partial</td>
<td>No</td>
<td>Yes</td>
<td>/ Partial</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----</td>
<td>------------</td>
<td>----</td>
<td>-----</td>
<td>------------</td>
<td>----</td>
<td>-----</td>
<td>------------</td>
<td>----</td>
<td>-----</td>
<td>------------</td>
</tr>
<tr>
<td>Neele 2016</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Raje 2021</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tew 2016</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Torres-Pareja 2019</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Vasquez-Araneda 2021</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Yamato 2016</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zhang 2021</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

+ Yes, / Partial yes, - No
Appendix 6: CERT Subgroups

6a. Consensus on Exercise Reporting Template (CERT) summary of reviews which included participants with ‘other’ health conditions (n=4). Items names are fully described in Appendix 2. Equip = description of exercise equipment, Quals = description of qualifications of instructor, Group = description of whether exercise performed in group or individually, Superv. = description of whether exercises are supervised and how they are delivered, Adher = description of how adherence is measured and reported, Motiv = description of motivation strategies, Rules = description of rules for determining exercise progression, Prog = description of how exercise was progressed, Replic = description of exercise to enable replication, Home = description of any home program, Non-ex = description of non-exercise components, Adv. Ev = description of type and number of adverse events, Setting = description of setting of exercise, Interv = description of exercise intervention (i.e. sets, reps, duration etc.), Generic = description of whether exercises are generic or tailored, Tailor = description of how exercises are tailored, Starting = description of rule for the starting level of participants, Fidelity = how adherence or fidelity to intervention is measured, Planned = description of the extent to which the intervention was delivered as planned.
6b. Consensus on Exercise Reporting Template (CERT) summary of reviews which included participants with musculoskeletal health conditions (n=10). Items names are fully described in Appendix 2. Equip = description of exercise equipment, Quals = description of qualifications of instructor, Group = description of whether exercise performed in group or individually, Superv. = description of whether exercises are supervised and how they are delivered, Adher = description of how adherence is measured and reported, Motiv = description of motivation strategies, Rules = description of rules for determining exercise progression, Prog = description of how exercise was progressed, Replic = description of exercise to enable replication, Home = description of any home program, Non-ex = description of non-exercise components, Adv. Ev = description of type and number of adverse events, Setting = description of setting of exercise, Interv = description of exercise intervention (i.e. sets, reps, duration etc.), Generic = description of whether exercises are generic or tailored, Tailor = description of how exercises are tailored, Starting = description of rule for the starting level of participants, Fidelity = how adherence or fidelity to intervention is measured, Planned = description of the extent to which the intervention was delivered as planned.
Appendix 7: TIDieR Subgroups

7a. Template for intervention description and replication (TIDieR) summary of reviews which included participants with cardiovascular health conditions (n=5). Items names are fully described in Appendix 2. Mater = what (materials), Proced = what (procedures), When = when and how much, Tailor = tailoring, Mods = modifications, How w. (p) = How well (planned), How w. (a) = How well (actual)
7b. Template for intervention description and replication (TIDieR) summary of reviews which included participants with musculoskeletal health conditions (n=7). Items names are fully described in Appendix 2. Mater = what (materials), Proced = what (procedures), When = when and how much, Tailor = tailoring, Mods = modifications, How w. (p) = How well (planned), How w. (a) = How well (actual)
7c. Template for intervention description and replication (TIDieR) summary of reviews which included participants with neurological health conditions (n=3). Items names are fully described in Appendix 2. Mater = what (materials), Proced = what (procedures), When = when and how much, Tailor = tailoring, Mods = modifications, How w. (p) = How well (planned), How w. (a) = How well (actual)
Appendix 1: Search strategy

PubMed – 28 Jun. 21
(exerci* [Title] OR sport* [Title] OR physical activity [Title] OR train* [Title] OR aerobic [Title] OR resistance [Title] OR physical training [Title] OR active [Title] OR move* [Title] OR rehab* [Title]) AND (cert [TiAb] OR tidier [TiAb] OR "template for intervention description and replication" [TiAb] OR report* [Title] OR complet* [TiAb] OR describ* [TiAb] OR replic* [TiAb] OR characteristics [TiAb] OR design [TiAb] OR program [TiAb]) OR (consensus on exercise reporting template)

- Filter systematic review, humans

EMBASE – 22 Jun. 21
((exerci* or sport* or physical activity or train* or aerobic or resistance or physical training or active or move* or rehab*) and (cert or "consensus on exercise reporting template" or tidier or "template for intervention description and replication" or report* or complet* or describ* or replic* or characteristics or design or program) and systematic review).m_titl.

CINAHL & SPORTDiscus – 22 Jun. 21
TI (exercise or physical activity or fitness) AND TI (report* or cert or tidier or "consensus on exercise reporting template" or "template for intervention description and replication") AND TI systematic review

PSYCHInfo – 28 Jun. 21
ab(systematic review) AND ab(exercise OR physical activity OR fitness) AND ab(report* OR cert OR tidier OR "consensus on exercise reporting template" OR "template for intervention description and replication")

Filter -systematic review

Appendix 2. Modified AMSTAR 2 tool used in this overview of reviews

1. Did the research questions and inclusion criteria for the review include the components of PICO?

<table>
<thead>
<tr>
<th>For Yes:</th>
<th>Optional (recommended)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Timeframe for follow up</td>
</tr>
<tr>
<td>Intervention</td>
<td></td>
</tr>
<tr>
<td>Comparator group</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td></td>
</tr>
</tbody>
</table>
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

<table>
<thead>
<tr>
<th>For Partial Yes:</th>
<th>For Yes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>The authors state that they had a written protocol or guide that included ALL the following:</td>
<td>As for partial yes, plus the protocol should be registered and should also have specified:</td>
</tr>
<tr>
<td>review question(s)</td>
<td>a meta-analysis/synthesis plan, if appropriate, and</td>
</tr>
<tr>
<td>a search strategy</td>
<td>a plan for investigating causes of heterogeneity</td>
</tr>
<tr>
<td>inclusion/exclusion criteria</td>
<td>a plan for investigating causes of heterogeneity</td>
</tr>
<tr>
<td>a risk of bias assessment</td>
<td></td>
</tr>
</tbody>
</table>

3. Did the review authors explain their selection of the study designs for inclusion in the review?

For Yes, the review should satisfy ONE of the following:

- Explanation for including only RCTs
- Explanation for including only NRSI
- Explanation for including both RCTs and NRSI

4. Did the review authors use a comprehensive literature search strategy?

For Partial Yes (all the following): For Yes, should also have (all the following):

- searched at least 2 databases (relevant to research question)
- provided key word and/or search strategy
- justified publication restrictions (e.g. language)
- searched the reference lists / bibliographies of included studies
- searched trial/study registries
- included/consulted content experts in the field
- where relevant, searched for grey literature
- conducted search within 24 months of completion of the review
5. Did the review authors perform study selection in duplicate?
For Yes, either ONE of the following:

- at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include  
  Yes  No

- OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.

6. Did the review authors perform data extraction in duplicate?
For Yes, either ONE of the following:

- at least two reviewers achieved consensus on which data to extract from included studies  
  Yes  No

- OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.

7. Did the review authors provide a list of excluded studies and justify the exclusions?
For Partial Yes:

- provided a list of all potentially relevant studies that were read in full-text form but excluded from the review

For Yes, must also have:

- Justified the exclusion from the review of each potentially relevant study  
  Yes  Partial Yes  No

8. Did the review authors describe the included studies in adequate detail?
For Partial Yes (ALL the following):

- described populations
- described interventions
- described comparators
- described outcomes
- described research designs

For Yes, should also have ALL the following:

- described population in detail
- described intervention in detail (including doses where relevant)
- described comparator in detail (including doses where relevant)
- described study’s setting
- timeframe for follow-up  
  Yes  Partial Yes  No
10. Did the review authors report on the sources of funding for the studies included in the review?

For Yes

Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies

Yes
No

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

For Yes:

The authors reported no competing interests OR

The authors described their funding sources and how they managed potential conflicts of interest

Yes
No

Appendix 3. Rating overall confidence in the results of the systematic review (AMSTAR 2)

High

- No or one non-critical weakness: the systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest

Moderate

- More than one non-critical weakness*: the systematic review has more than one weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review

Low
- One critical flaw with or without non-critical weaknesses: the review has a critical flaw and may not provide an accurate and comprehensive summary of the available studies that address the question of interest

**Critically low**

- More than one critical flaw with or without non-critical weaknesses: the review has more than one critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies

* Multiple non-critical weaknesses may diminish confidence in the review and it may be appropriate to move the overall appraisal down from moderate to low confidence.

Note: A critical flaw was a rating deemed a ‘no’ in any of the critical domains, as specified by the AMSTAR 2 guidance document. These domains were Items 2, 4, 7. A non-critical weakness was a no, or partial yes in any other domain.
Appendix 4: Excluded studies, with reasons

Wrong outcomes:


Allen NE, Suriyarachchi GD, Paul SS, et al. Exercise and motor training in people with Parkinson’s disease: A systematic review of participant characteristics, intervention delivery, retention rates, adherence, and adverse events in clinical trials. *Parkinson’s Disease* 2012;((Allen, Suriyarachchi, Paul, Song, Canning) Clinical and Rehabilitation Research Group, Faculty of Health Sciences, University of Sydney, P.O. Box 170, Lidcombe, Sydney, NSW 1825, Australia(Sherrington) Musculoskeletal Division, George Institute for Global):854328. doi: http://dx.doi.org/10.1155/2012/854328


Bonafiglia JT, Islam H, Preobrazenski N, et al. Risk of bias and reporting practices in studies comparing VO2max responses to sprint interval vs. continuous training: A systematic review and meta-analysis. *Journal of Sport and Health Science* 2021;((Bonafiglia, Islam, Preobrazenski, Gurd) School of Kinesiology and Health Studies, Queen’s University, Kingston, ON K7L 3N6, Canada) doi: http://dx.doi.org/10.1016/j.jshs.2021.03.005


Goff AJ, Page WS, Clark NC. Reporting of acute programme variables and exercise descriptors in rehabilitation strength training for tibiofemoral joint soft tissue
injury: A systematic review. *Physical Therapy in Sport* 2018;34((Goff) Health and Social Sciences, Singapore Institute of Technology, 10 Dover Drive 138683, Singapore((Page, Clark) Faculty of Sport, Health and Applied Sciences, St Mary’s University, Waldegrave Road, Twickenham TW1 4SX, United Kingdom).227-37. doi: http://dx.doi.org/10.1016/j.ptsp.2018.10.012


Sauzet O, Kleine M, Exner A-K, et al. Longitudinal randomised controlled trials in rehabilitation post-stroke: A systematic review on the quality of reporting and use of baseline outcome values. *BMC Neurology* 2015(Sauzet, Kleine, Exner) AG Epidemiology and International Public Health, School of Public Health, Bielefeld University, PO. Box 10 01 31, Bielefeld, 33501 Germany(Menzel-Begemann) Faculty of Nursing and Health, University of Applied Science, Munster, Leon) doi: http://dx.doi.org/10.1186/s12883-015-0344-y


Reporting not primary aim


Not systematic review

Wrong intervention


## Appendix 5: AMSTAR Assessment

<table>
<thead>
<tr>
<th>Study</th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
<th>Item 7</th>
<th>Item 8</th>
<th>Item 10</th>
<th>Item 16</th>
<th>Overall Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abell 2015</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Critically Low</td>
</tr>
<tr>
<td>Barros 2020</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Low</td>
</tr>
<tr>
<td>Bartholdy 2019</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Berti 2020</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Burgess 2021</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>+</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Low</td>
</tr>
<tr>
<td>Charlton 2017</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Davidson 2021</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>Moderate</td>
</tr>
<tr>
<td>Dischiavi 2021</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>+</td>
<td>Low</td>
</tr>
<tr>
<td>Giagio 2021</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Hall 2018</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Holden 2018</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Kattachal 2020</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Low</td>
</tr>
<tr>
<td>Keene 2020</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Moderate</td>
</tr>
<tr>
<td>Knols 2018</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Moderate</td>
</tr>
<tr>
<td>Knols 2019</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Critically Low</td>
</tr>
<tr>
<td>Lohse 2018</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>/</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Low</td>
</tr>
<tr>
<td>Mack 2018</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Critically Low</td>
</tr>
<tr>
<td>Major 2019</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Critically Low</td>
</tr>
<tr>
<td>McEwen 2019</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Low</td>
</tr>
<tr>
<td>McGregor 2018</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Low</td>
</tr>
<tr>
<td>Meneses-Echavez 2019</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Low</td>
</tr>
</tbody>
</table>

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance on the information supplied by the author(s) placed on this supplemental material which has been supplied by the author(s).
<table>
<thead>
<tr>
<th>Study</th>
<th>+</th>
<th>-</th>
<th>+</th>
<th>-</th>
<th>-</th>
<th>+</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neele 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raje 2021</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tew 2016</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torres-Pareja 2019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vasquez-Araneda 2021</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yamato 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang 2021</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>/</td>
<td>+</td>
<td>+</td>
<td>/</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+ Yes, / Partial yes, - No

Critically Low
Low

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).

Br J Sports Med

Appendix 6: CERT Subgroups

6a. Consensus on Exercise Reporting Template (CERT) summary of reviews which included participants with ‘other’ health conditions (n=4). Items names are fully described in Appendix 2. Equip = description of exercise equipment, Quals = description of qualifications of instructor, Group = description of whether exercise performed in group or individually, Superv. = description of whether exercises are supervised and how they are delivered, Adher = description of how adherence is measured and reported, Motiv = description of motivation strategies, Rules = description of rules for determining exercise progression, Prog = description of how exercise was progressed, Replic = description of exercise to enable replication, Home = description of any home program, Non-ex = description of non-exercise components, Adv. Ev = description of type and number of adverse events, Setting = description of setting of exercise, Interv = description of exercise intervention (i.e. sets, reps, duration etc.), Generic = description of whether exercises are generic or tailored, Tailor = description of how exercises are tailored, Starting = description of rule for the starting level of participants, Fidelity = how adherence or fidelity to intervention is measured, Planned = description of the extent to which the intervention was delivered as planned.
6b. Consensus on Exercise Reporting Template (CERT) summary of reviews which included participants with musculoskeletal health conditions (n=10). Items names are fully described in Appendix 2. Equip = description of exercise equipment, Quals = description of qualifications of instructor, Group = description of whether exercise performed in group or individually, Superv. = description of whether exercises are supervised and how they are delivered, Adher = description of how adherence is measured and reported, Motiv = description of motivation strategies, Rules = description of rules for determining exercise progression, Prog = description of how exercise was progressed, Replic = description of exercise to enable replication, Home = description of any home program, Non-ex = description of non-exercise components, Adv. Ev = description of type and number of adverse events, Setting = description of setting of exercise, Interv = description of exercise intervention (i.e. sets, reps, duration etc.), Generic = description of whether exercises are generic or tailored, Tailor = description of how exercises are tailored, Starting = description of rule for the starting level of participants, Fidelity = how adherence or fidelity to intervention is measured, Planned = description of the extent to which the intervention was delivered as planned.
Appendix 7: TIDieR Subgroups

7a. Template for intervention description and replication (TIDieR) summary of reviews which included participants with cardiovascular health conditions (n=5). Items names are fully described in Appendix 2. Mater = what (materials), Proced = what (procedures), When = when and how much, Tailor = tailoring, Mods = modifications, How w. (p) = How well (planned), How w. (a) = How well (actual)
7b. Template for intervention description and replication (TIDieR) summary of reviews which included participants with musculoskeletal health conditions (n=7). Items names are fully described in Appendix 2. Mater = what (materials), Proced = what (procedures), When = when and how much, Tailor = tailoring, Mods = modifications, How w. (p) = How well (planned), How w. (a) = How well (actual)
7c. Template for intervention description and replication (TIDieR) summary of reviews which included participants with neurological health conditions (n=3). Items names are fully described in Appendix 2. Mater = what (materials), Proced = what (procedures), When = when and how much, Tailor = tailoring, Mods = modifications, How w. (p) = How well (planned), How w. (a) = How well (actual)