Sheffield Hallam University

Quality and acceptability of measures of exercise adherence in musculoskeletal settings: a systematic review

MCLEAN, Sionnadh <http://orcid.org/0000-0002-9307-8565>, HOLDEN, Melanie, POTIA, Tanzila, GEE, Melanie <http://orcid.org/0000-0001-9149-4314>, MALLETT, Ross, BHANBHRO, Sadiq <http://orcid.org/0000-0003-0771-8130>, PARSONS, Helen and HAYWOOD, Kirstie

Available from Sheffield Hallam University Research Archive (SHURA) at:

https://shura.shu.ac.uk/14544/

This document is the Published Version [VoR]

Citation:

MCLEAN, Sionnadh, HOLDEN, Melanie, POTIA, Tanzila, GEE, Melanie, MALLETT, Ross, BHANBHRO, Sadiq, PARSONS, Helen and HAYWOOD, Kirstie (2017). Quality and acceptability of measures of exercise adherence in musculoskeletal settings: a systematic review. Rheumatology (Oxford, England), 56 (3), 426-438. [Article]

Copyright and re-use policy

See http://shura.shu.ac.uk/information.html

Sheffield Hallam University Research Archive
http://shura.shu.ac.uk

RHEUMATOLOGY

Rheumatology 2017;56:426–438
doi:10.1093/rheumatology/kew422
Advance Access publication 23 December 2016

Original article

Quality and acceptability of measures of exercise adherence in musculoskeletal settings: a systematic review

Sionnadh McLean[1], Melanie A. Holden[2], Tanzila Potia[1], Melanie Gee[3], Ross Mallett[1], Sadiq Bhanbhro[3], Helen Parsons[4] and Kirstie Haywood[5]

Abstract

Objective. To recommend robust and relevant measures of exercise adherence for application in the musculoskeletal field.

Method. A systematic review of measures was conducted in two phases. Phase 1 sought to identify all reproducible measures used to assess exercise adherence in a musculoskeletal setting. Phase 2 identified published evidence of measurement and practical properties of identified measures. Eight databases were searched (from inception to February 2016). Study quality was assessed against the Consensus-based Standards for the Selection of Health Measurement Instruments guidelines. Measurement quality was assessed against accepted standards.

Results. Phase 1: from 8511 records, 326 full-text articles were reviewed; 45 reproducible measures were identified. Phase 2: from 2977 records, 110 full-text articles were assessed for eligibility; 10 articles provided evidence of measurement/practical properties for just seven measures. Six were exercise adherence-specific measures; one was specific to physical activity but applied as a measure of exercise adherence. Evidence of essential measurement and practical properties was mostly limited or not available. Assessment of relevance and comprehensiveness was largely absent and there was no evidence of patient involvement during the development or evaluation of any measure.

Conclusion. The significant methodological and quality issues encountered prevent the clear recommendation of any measure; future applications should be undertaken cautiously until greater clarity of the conceptual underpinning of each measure is provided and acceptable evidence of essential measurement properties is established. Future research should seek to engage collaboratively with relevant stakeholders to ensure that exercise adherence assessment is high quality, relevant and acceptable.

Key words: acceptability, adherence, exercise, measurement, musculoskeletal, physical activity, quality, systematic review

CLINICAL SCIENCE

Rheumatology key messages

- Current measures of exercise adherence for musculoskeletal populations are of poor quality.
- New measures of exercise adherence for musculoskeletal populations require a collaborative approach.

[1]Faculty of Health and Wellbeing, Collegiate Campus, Sheffield Hallam University, Sheffield, [2]Arthritis Research UK Primary Care Centre, Keele University, Keele, [3]Centre for Health and Social Care Research, Collegiate Campus, Sheffield Hallam University, Sheffield, [4]Clinical Trials Unit and [5]Royal College of Nursing Research Institute, Warwick Medical School, Warwick University, Coventry, UK

Submitted 31 July 2015; revised version accepted 18 October 2016

Correspondence to: Sionnadh McLean, Faculty of Health and Wellbeing, Collegiate Campus, Sheffield Hallam University, Sheffield S10 2BP, UK. E-mail: s.mclean@shu.ac.uk

Introduction

Musculoskeletal (MSK) disorders are burdensome [1]. For many, the associated progressive functional limitation in everyday activities, including paid employment, results in significant financial costs for individuals and society [2, 3]. Increasing age and lifestyle factors such as obesity and physical inactivity negatively impact MSK disorders [4, 5]; the ageing population and increasingly

© The Author 2016. Published by Oxford University Press on behalf of the British Society for Rheumatology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

sedentary lifestyles suggest that the disease burden will continue to increase [4].

Exercise and physical activity (EPA) can reduce pain, improve physical dysfunction and enhance quality of life for individuals with MSK disorders [6–10]; clinical guidelines advocate EPA within long-term management strategies [3, 11–13]. Physical activity is defined as any bodily movement produced by skeletal muscle that results in energy expenditure and includes occupational, sporting and household activities [14]. Exercise, a subset of physical activity, is specific, structured, planned and repetitive [14]. In this article exercise indicates therapeutic EPA aimed at reducing MSK symptoms.

An individual's ability to adhere to recommended exercise, defined as the extent to which a person's behaviour corresponds with agreed recommendations from a healthcare provider, is important for success [15, 16]. Patients who adhere to regular exercise are less likely to progress to recurrent, persistent or disabling problems [17, 18]. Increasing adherence may give greater patient benefit than improving aspects of the intervention itself [16]. Adherence to prescribed exercise is frequently reported as $<50\%$ [19–22]. Non-adherence may negatively impact treatment effectiveness and efficiency, therapeutic relationships, waiting times and cost of care [23–25]. Numerous strategies for increasing exercise adherence have been identified but their effectiveness is uncertain and guidance for best practice does not exist [26, 27]. Consequently, development and evaluation of exercise adherence interventions is essential [28]; however, guidance for the assessment of exercise adherence in MSK clinical trials or routine practice settings does not exist.

There is wide variation in the assessment of exercise adherence [29, personal communication, R. Frost, Glasgow Caledonian University]. Where large numbers of assessment approaches exist, structured reviews of the quality and acceptability of different approaches are essential to informing selection [30, 31]. This review seeks to identify all clearly reported and reproducible measures of exercise adherence applied in published studies of patients with MSK disorders, and to evaluate these measures against a transparent appraisal framework.

Methods

This two-phase systematic review was reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [32]. Phase 1 identified clearly reported and reproducible measures of exercise adherence in published MSK studies. Phase 2 reviewed published and unpublished evidence of measurement and practical properties for shortlisted measures. Study and measurement quality were assessed against the Consensus-based Standards for the Selection of health Measurement Instruments (COSMIN) checklist [31, 33, 34], and a transparent appraisal framework [35], respectively.

Phase 1: identifying measures of exercise adherence

A search strategy was developed to identify methods used to assess exercise adherence in MSK settings (see search strategy for phase 1 in Supplementary Data, available at *Rheumatology* Online, and study protocol [36]); all study types were included. Eight databases were searched (from inception to February 2016): Medline, SPORTDiscus, CINAHL Plus, PsycINFO, AMED, Cochrane Library, Embase and the Web of Science.

Titles, abstracts and full text articles were independently screened for inclusion by two reviewers from five (S.Mc., M.H., R.M., T.P., S.B.). Disagreement was discussed with a third independent reviewer from six (S.Mc., M.H., R.M., T.P., S.B., K.H.).

Articles were included if they focused on adults with an MSK disorder receiving therapeutic exercise in any setting, and for which assessments of adherence to exercise [patient- or clinician-reported or exercise diaries (if converted to an adherence scale)] were completed. Studies were excluded if they were not written in English or if participants were healthy volunteers, <18 years old, or with non-MSK conditions.

Reproducible measures of exercise adherence (i.e. the original measure could be located, had an appropriate citation or was reproducible based on information supplied by the author) [37, 38] were listed and categorized as clinician- or patient-reported. Performance measures (i.e. muscle strength, joint range of movement), performance of exercise technique and session attendance were excluded as proxy measures of adherence. Accelerometers and pedometers were excluded because they are primarily performance measures and measure motion rather than adherence.

Phase 2: evidence of quality and acceptability

Separate searches were conducted in the above databases for each shortlisted measure. Where the result set for a measure exceeded 50, a sensitive search filter for the identification of studies reporting evidence of measurement and/or practical properties was additionally applied [39] (Search strategy for phase 2 in Supplementary Data, available at *Rheumatology* Online). The developers of specific measures were also contacted to request additional evidence of measurement evaluation. Titles, abstracts and full text articles were independently assessed by two reviewers from four (M.H., T.P., R.M., S.Mc.); a third reviewer resolved any disagreements (K.H.). Reference lists of included articles were reviewed for additional published articles. English language articles were included if they provided evidence of assessment development and/or evaluation of the named measure(s) in an MSK population.

Data extraction and inter-rater reliability

A data extraction form informed by earlier reviews [35] and the COSMIN checklist [31, 34] was used to capture study-specific (population, intervention and setting) and measurement-specific information: reliability (internal consistency, test–retest, intra-/inter-tester, measurement

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

error); validity [content, structural validity (dimensionality), construct (evidence of explicit hypothesis testing); criterion]; responsiveness (criterion-/construct-based); interpretability (e.g. evidence of minimal important change); data precision (data quality, end effects); and evidence of where Item Response Theory models were applied. Extraction for practical properties included acceptability (relevance and respondent burden) and feasibility (clinician burden, including cost, time to complete and score) [30, 31, 34]. The extent of patient involvement in measurement development and/or application was also sought [35].

In accordance with the COSMIN checklist, study methodological quality was evaluated per measurement property and rated on a four-point scale (excellent, good, fair, poor); quality was determined by the lowest checklist rating per measurement property [31, 34]. Following a group training session, four primary reviewers (S.Mc., M.H., T.P., R.M.) independently undertook data extraction and applied the checklist. The reviewers were clinicians and/or researchers with little experience in assessing measurement properties and no previous exposure to the COSMIN checklist. The inter-rater agreement (percentage agreement) between two reviewers was evaluated for all included articles. Where disagreement existed, consensus was sought through a third, experienced reviewer (K.H.) who independently reviewed all articles.

Data synthesis

Data were qualitatively synthesized to determine the overall quality and acceptability of each measure [30, 33]. Synthesis considered the following: study methodological quality (COSMIN scores); number of studies reporting specific evidence per measure; results for each measurement and practical property per measure; and consistency between studies [33]. The overall quality of a measurement property was reported as: adequate (+), not adequate (−), conflicting (±), or unclear (?). Levels of evidence for the overall quality of each measurement property were further defined to indicate strong, moderate, limited, conflicting or unknown evidence [33].

Results

Identification of studies and measures

Phase 1

Following removal of duplicates, 8511 records were identified. Following title and abstract screening 326 full-text articles were retrieved and reviewed in full (Fig. 1).

A total of 234 approaches to measuring exercise adherence were identified. These included the following: exercise logs and diaries (n = 107); unnamed questionnaires or scales (n = 53); clearly described or named questionnaires or scales (n = 49); interviews (n = 17); and calendars or postcards (n = 8). Only the 49 clearly described and reproducible or named questionnaires or scales were included (Supplementary Table S1, available at *Rheumatology* Online).

Phase 2

Evidence for measurement and/or practical properties were sought for the 49 reproducible measures identified in phase 1. Following removal of duplicates, 2977 records were identified. Following title and abstract screening, 110 full-text articles were retrieved and reviewed in full and 10 retained for phase 2 (Fig. 2) [22, 40–48].

These 10 articles provide evidence for seven clearly defined measures of exercise adherence in an MSK population. Three are clinician-reported: Hopkins Rehabilitation Engagement Rating Scale (HRERS) [41], Pittsburgh Rehabilitation Participation Scale (PRPS) [42] and the Sport Injury Rehabilitation Adherence Scale (SIRAS) [45]. Four are patient-reported: Adherence to Exercise Scale for Older Patients (AESOP) [22], Community Healthy Activities Model Program for Seniors Activities Questionnaire for Older Adults (CHAMPS) [44]; the Modified Rehabilitation Adherence Questionnaire (RAQ-M) [42]; and the Rehabilitation Overadherence Questionnaire (ROAQ) [48]. Attempts to contact measurement developers for further information were unsuccessful.

Data extraction: inter-rater reliability

Evidence for 40/107 COSMIN items across 5/10 COSMIN domains (A, B, D, E, F) was extracted. Agreement exceeded 80% for only 20 items (50%) [31]. Disagreement was mainly due to poor reporting of evidence in the reviewed papers, associated interpretation difficulties, reading errors or difficulties applying the checklist.

Study characteristics

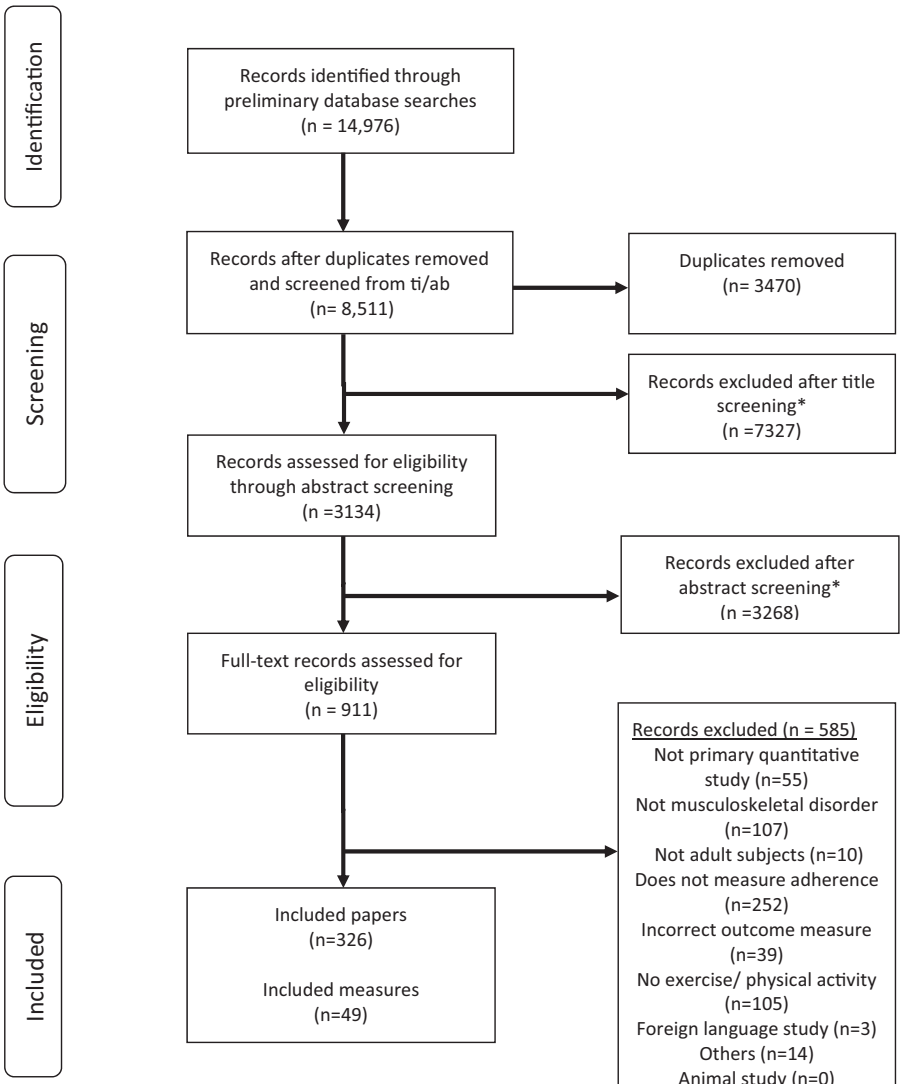
Although six studies were adequately sized for evaluative purposes (range 145–249) [34], four included fewer than 100 patients [22, 40, 45, 47]. The ages of patients ranged from 13 to 96 years (see Table 1). Studies covered a wide range of MSK settings: athletes with acquired knee injuries [43, 45–47]; general MSK disorders in outpatient settings [40]; older patients with generalized MSK conditions [22, 42, 44]; acute inpatient populations [41]; and athletic adolescents with MSK injuries [48].

Adherence measures

Six of the seven measures were originally developed as measures of exercise adherence, including the following: sports injury rehabilitation (SIRAS, RAQ-M, ROAQ); acute MSK inpatient rehabilitation (HRERS, PRPS); and MSK home exercise programmes (AESOP). Although originally developed as a measure of physical activity, the CHAMPS has subsequently been evaluated as a measure of exercise adherence and hence is included in this review [43]. With the exception of the RAQ-M which was evaluated in Korean athletes, all measures were developed and evaluated in the USA. The characteristics and measurement properties of all reviewed measures are summarized in Tables 1 and 2 and Supplementary Tables S2 and S3, available at *Rheumatology* Online. Study methodological quality and the qualitative synthesis is summarized in Table 3.

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

Fig. 1 PRISMA flowchart for phase 1 of the systematic review



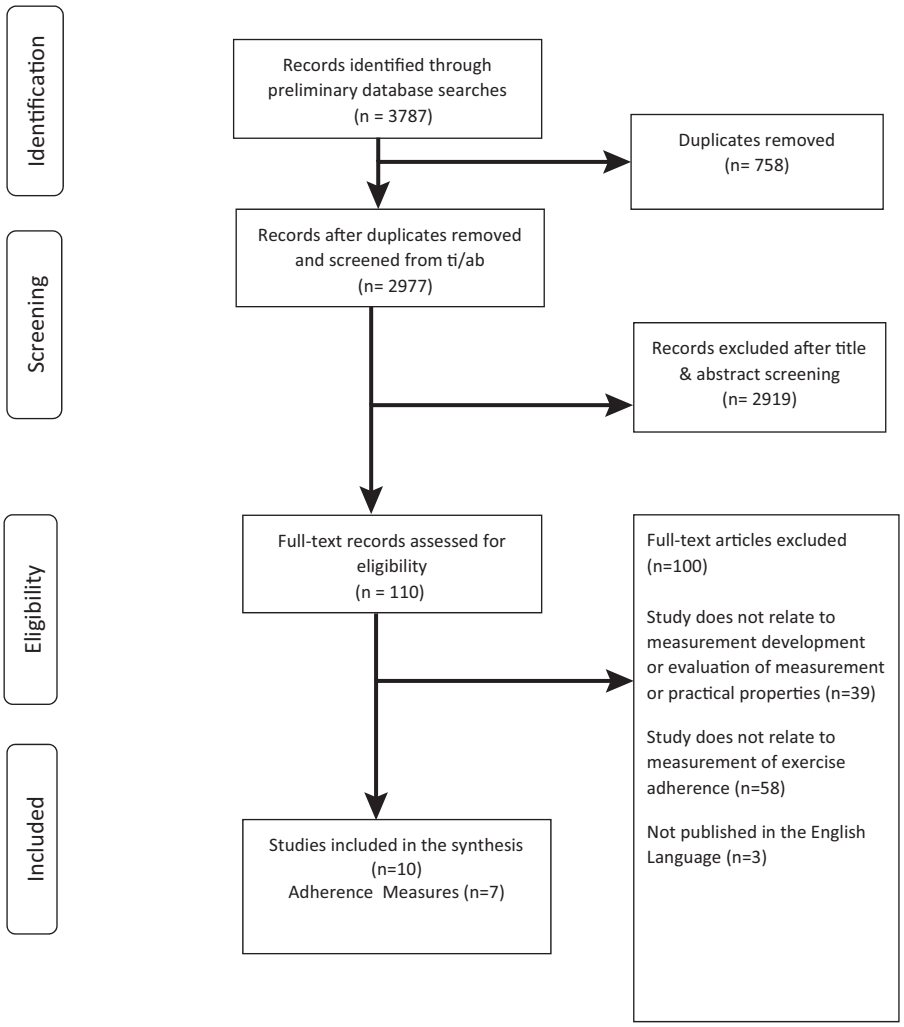
Clinician-reported

The five-item HRERS assesses the therapist's perception of an individual's engagement in acute inpatient rehabilitation. There is limited evidence of reliability and validity following completion in a population of patients with spinal cord injury, stroke, amputation or hip/knee replacement [41]. The unidimensional structure (structural validity) of the HRERS as a measure of engagement was supported by principal component factor analysis across the different diagnostic groups. A high level of internal consistency for this single dimension (Cronbach's $\alpha = 0.91$) and acceptable inter-rater agreement (intraclass correlation coefficient (ICC) $= 0.73$) was reported [41]. Evidence of known-groups validity was provided against groups defined by a range of external criteria hypothesized to be associated with engagement including scores on the Functional Impact Measures (FIM) and rates of therapy absenteeism. Small correlations were reported between the HRERS and a range of clinical variables including depression ($r = 0.24$), denial of illness ($r = 0.30$), self-rated negative affect ($r = -0.23$) and level of functioning ($r = 0.22$) [41]; although the authors suggest that hypothesized associations were supported, these were not clearly stated, hence limiting interpretation in support of measurement validity.

The single-item PRPS is used to rate patient participation during each treatment session of acute inpatient rehabilitation [42]. Item development involved therapist interviews and therapy session observation of older patients with generalized MSK problems. There is limited

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

Fig. 2 PRISMA flowchart for phase 2 of the systematic review



evidence of reliability and validity following completion with older people with generalized MSK conditions [42]. High values of inter-rater reliability (range of ICC = 0.91–0.96) were reported [42]. Small correlations between the PRPS and the FIM-motor (range r = 0.38), with change in FIM-motor (r = 0.32) and length of stay were reported (r = −0.13; $P < 0.05$) (Supplementary Table S3) [41]; however, the absence of *a priori* hypothesized associations between variables limits interpretation. Similarly, although a statistically significant score improvement was reported in those inpatients with a length of stay >9 days [score increase from 4.29 ± 0.93 to 4.67 ± 1.04; $p < 0.001$], external anchors against which change in participation may be judged or suggestions for interpretation of score change are not provided.

The three-item SIRAS is used by therapists to rate the degree to which patients exert themselves, follow the practitioner's instructions and advice, and are receptive to changes in the rehabilitation programme during a given rehabilitation session. The single factor structure of the SIRAS (exercise adherence) is supported by several studies following completion by athletes and the general MSK population [40, 46]. Internal consistency evaluations further support reporting the SIRAS as a single index value [46]. Acceptable levels of internal consistency supports application in groups of patients (Cronbach's α range 0.82–0.8) [46, 47]. Poor to high levels of inter-rater (ICC range = 0.57–0.77; Rater Agreement Index range = 0.84–0.94) and acceptable 1-week test–retest reliability has been reported (range = 0.63–0.77) [39, 45]. Evidence in support of known-groups validity is provided following the assessment of standardized vignettes describing three levels of adherence in athletes [40, 47].

Patient-reported measures

The AESOP is a 42-item interview-administered questionnaire, developed to assess exercise adherence in older patients [22]. The measure constitutes three domains, informed by social cognitive theory: self-efficacy expectations (15 items), outcome expectations (16 items) and

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

Table 1 Characteristics of reviewed measures used to assess exercise adherence in patients with MSK problems

Measure (developer, year)	Evaluations (n)	Construct	Domains (items)	Response options	Recall	Score range	Admin (time)
Clinician-completed							
Hopkins Rehabilitation Engagement Rating Scale (HRERS) (Kortte *et al.*, 2007) [41]	1	Behavioural observations of patients during acute inpatient rehabilitation	Five items: • Attendance at rehabilitation session (1) • Frequency of required verbal/physical prompts (1) • Perceived positive attitude to exercise (2) • Perceived need for and benefit from rehabilitative exercise • Active participation in rehabilitative exercise (1)	Six-point descriptive: • Never (1) • Seldom (2) • Some of the time (3) • Most of the time (4) • Nearly always (5) • Always (6)	At the time of the rehabilitation session/ at time of discharge to represent a summary of observations during patients' inpatient stay	Simple summation: range 5–30, where 5 is poor and 30 is best engagement in the therapy process	NR
Pittsburgh Rehabilitation Participation Scale (PRPS) (Lenze *et al.*, 2004) [42]	1	Observed patient 'participation' in a therapy session	Single item to assess patient participation in a therapy session	Detailed 6-point Likert scale, ranging from: • None (1): patient refused entire session or did not participate in exercises to • Excellent (6): patient participated in all exercises with max effort, finished all exercises and actively took interest in exercises and/or future therapy sessions	At the time of the rehabilitation session	One response is selected: range 1 (poor) to 6 excellent participation	NR
Sport Injury Rehabilitation Adherence Scale (SIRAS) (Brewer *et al.* 1999) [45]	8	Adherence during rehabilitation sessions	Three items: • Perceived intensity/ effort/exertion (1) • Frequency of following therapist instructions (1) • Receptive to change in rehabilitation exercise (1)	Five-point numerical rating scale:Anchors: • Minimum effort (1) to maximum effort (5) • Never (1) to always (5) • Very unreceptive (1) to very receptive (5)	1 week	Index (composite) score—summation of score for the three items: range 0–15, where 1 is lower adherence, and 15 is maximal adherence	NR

(continued)

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

Table 1 Continued

Measure (developer, year)	Evaluations (n)	Construct	Domains (items)	Response options	Recall	Score range	Admin (time)
Patient-completed							
Adherence to Exercise Scale for Older Patients (AESOP) (Hardage *et al.* 2007) [22]	1	Social cognitive theory constructs for predicting home exercise programme (HEP) adherence in older adults: self-efficacy expectations, outcome expectations and outcome expectancies	42 items: • Self-efficacy expectations (15) • Outcome expectations (16) • Outcome expectancies (11)	Five-point agreement: • Strongly disagree (1) • Disagree (2) • No opinion (3) • Agree (4) • Strongly agree (5)	2 weeks	Simple summation: • Range 15–75 • Range 16–80 • Range 11–55Lower scores suggest lower levels of adherence	NR
Community Healthy Activities Model Program for Seniors (CHAMPS) (Stewart *et al.* 2001) [44]	1	Types and intensity levels of physical activity	41 items: • Ranging over activities of daily living, work related activities, social activities and leisure activities	Five-point agreement: • Strongly disagree (1) • Disagree (2) • No opinion (3) • Agree (4) • Strongly agree (5)	4 weeks	• Frequency of activities per week: number of minutes of physical activity per week across all activities • Calorie expenditure: per week multiply esti mated duration of each activity by the MET value and summing across all activities • Both can be calculated for: • (i) Moderate and greater activity measures • (ii) All activity measuresTherefore, four scores possible	NR
Modified Rehabilitation Adherence Questionnaire (RAQ-M) (Shin *et al.* 2010) [43]	1	Rehabilitation adherence in injured athletes	25 items: • Perceived exertion (3) • Pain tolerance (5) • Self-motivation (5) • Support from significant others (5) • Scheduling (4) • Environmental conditions (3)	Four-point agreement: • Strongly disagree (1) • Disagree (2) • Agree (3) • Strongly agree (4)	1 week	Simple item summation for each domain: • Range 3–12 • Range 5–20 • Range 5–20 • Range 5–20 • Range 4–16 • Range 3–12Higher scores reflect greater levels of adherence	NR
Rehabilitation Overadherence Questionnaire (ROAQ) (Podlog *et al.* 2013) [48]	2	Assessment of overadherence behaviours and beliefs in injured athletes	Two domains (10 items): • Ignoring practitioner recommendations (6) • Attempting an expe dited rehabilitation (4)	Five-point agreement: • Never or strongly disagree (1) to • Always or strongly agree (5)	NR	NR	NR

n: number of studies evaluating the measurement and practical properties of each measure; NR: not reported; MET: metabolic energy equivalent.

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

Table 2 Methodological quality and investigated measurement and practical properties per measure per reviewed article

Article (n = 9)	Population (n)	Age mean (SD), years; range	Measures	Reliability		Validity			Responsiveness
				Internal reliability	Test–retest	Convergent/ divergent	Known groups	Structural	Responsiveness
Brewer *et al.* 2002 [47] Study 1	43 (practitioners)	Range 20–43	SIRAS	—	Poor	—	Poor	—	—
Study 2	12 (rehab patients)	29.33 (11.44)	SIRAS	—	Poor	—	—	—	—
Brewer *et al.* 2000 [46] Study 1	145 (orthopaedic outpatients)	43.95 (15.54)	SIRAS	Fair	—	Poor	—	Fair	—
Study 2	31 (sport related knee injury)	NR	SIRAS	—	Fair	—	—	—	—
Study 3	43 (rehab post-ACL repair)	NR	SIRAS	—	Fair	—	—	—	—
Brewer *et al.* 1999 [45]	31	NR	RAQ SIRAS	Poor	Poor	Poor	—	—	—
Hardage *et al.* 2007 [22]	50	79.9; range 65–91	AESOP SF-12 mMSE GDS	—	Poor	Poor	Poor	—	—
Kolt *et al.* 2007 [40] Study1	60 (physiotherapists = raters)	NR	SIRAS	—	Poor	—	Poor	Poor	—
Study 2	45 patients (general MSK)	>18 years	SIRAS	—	Poor	—	—	Poor	—
Kortte *et al.* 2007 [41]	208	56.7 (17.52); range 18–91	HRERS FIM BSI L-DIQ PANAS CHART	Poor	Poor	Fair	Fair	Fair	—
Lenze *et al.* 2004 [42]	242	70.8 (14.8); range 20–96	PRPS FIM-motor	—	Fair	Poor	—	—	Poor
Podlog *et al.* 2013 [48] Study 1	118 injured adolescent athletes	16.0 (1.4); range 13–18	RAOQ • SPSQ • AIMS • I-PRRS	Fair	—	Fair	—	Fair	—
Study 2	105 injured collegiate athletes	NR	RAOQ • SPSQ • AIMS • I-PRRS	Fair	—	Fair	—	Fair	—
Shin *et al.*, 2010 [43]	240 injured athletes	NR	RAQ-M SIRAS	Fair	Poor	Poor	Poor	Poor	—
Stewart *et al.* 2001 [44]	249	74.1; range 65–90	CHAMPS BMI SF-36 domains SPPB 6-min walk	—	Good	Good	Good	—	Fair

n: population size in included study; NR: not reported; 6-min walk: six-minute walking test; BSI: Brief Symptom Inventory; CHART: Craig Handicap Assessment and Reporting Technique; FIM: Functional Impact Measure; GDS: Geriatric Depression Scale; L-DIQ: Levine's Denial of Illness Questionnaire; mMSE: mini-Mental State Examination; PANAS: Positive and Affective Negative State; SF-12: Short-Form 12-item Health Survey; SF-36: Short-Form 36-item Health Survey; SPPB: Short Physical Performance Battery; HRERS: Hopkins Rehabilitation Engagement Rating Scale; PRPS: Pittsburgh Rehabilitation Participation Scale; SIRAS: Sport Injury Rehabilitation Adherence Scale; AESOP: Adherence to Exercise Scale for Older Patients; CHAMPS: Community Healthy Activities Model Program for Seniors; RAQ-M: Modified Rehabilitation Adherence Questionnaire; SPSQ: Self-Presentation in Sport Questionnaire; AIMS: Athletic Identity Measurement Scale; ROAQ: Rehabilitation Overadherence Questionnaire; I-PRRS: Modified Injury Psychological Readiness to Return to Sport Scale.

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

Table 3 Overall quality of measurement properties per reviewed measure of exercise adherence for MSK populations

Measure	Evaluations (n)	Reliability			Validity			Responsiveness
		Test–retest (intra/inter)	Internal consistency	Measurement error	Content	Convergent/ divergent	Known groups	Structural	Responsiveness
Therapist-completed									
HRERS	1	+limited	+limited	Nil	Nil	+limited	+limited	+limited	Nil
PRPS	1	+limited	Nil	Nil	Nil	+limited	Nil	Nil	−limited
SIRAS	8	+limited	+limited	Nil	Nil	+limited	+limited	Nil	Nil
Patient-completed									
AESOP	1	−limited	Nil	Nil	Nil	+limited	?limited	Nil	Nil
CHAMPS	1	−limited	Nil	Nil	Nil	+limited	?limited	Nil	−limited
RAQ-M	1	+limited	+limited	Nil	Nil	?limited	?limited	+limited	Nil
ROAQ	2	Nil	+limited	Nil	−limited	+limited	Nil	+limited	Nil

n: number of studies evaluating the measurement and practical properties of each measure; the overall quality of a measurement property is reported as adequate (+), not adequate (−), conflicting (±), or unclear (?); levels of evidence for the overall quality of each measurement property is strong, moderate, limited, conflicting, or unknown evidence. HRERS: Hopkins Rehabilitation Engagement Rating Scale; PRPS: Pittsburgh Rehabilitation Participation Scale; SIRAS: Sport Injury Rehabilitation Adherence Scale; AESOP: Adherence to Exercise Scale for Older Patients; CHAMPS: Community Healthy Activities Model Program for Seniors; RAQ-M: Modified Rehabilitation Adherence Questionnaire; ROAQ: Rehabilitation Overadherence Questionnaire.

outcome expectancies (11 items). Although acceptable test–retest reliability was reported for two domains—self-efficacy expectations (ICC = 0.80) and outcome expectations (ICC = 0.77)—low levels were reported for outcome expectancies (ICC = 0.33) [22]. All correlations between the three AESOP domains and the Short Form 12-item Health Survey (SF-12, version 2) physical and mental component scales were very small; the absence of *a priori* hypothesized associations between variables limits interpretation in support of measurement validity.

The CHAMPS activities questionnaire is a 41-item patient-reported or interview-administered questionnaire. The CHAMPS is a measure of physical activity that has been evaluated for use as a measure of exercise adherence in daily life [44]. The CHAMPS asks about activities that you may have done in the past 4 weeks. The information is used to calculate frequency of activities—the number of minutes of physical activity per week and the calories expended per week in all physical activities. Each score can be calculated for moderate and greater activity levels, and all activity levels. Hence, four scores are possible. Data from an intervention trial to increase activity levels among community-dwelling older people (CHAMPS trial) was assessed for score stability at 6 months (for participants in the non-active treatment or control group and hence not expected to change) and 2-week test–retest reliability [43]. Moderate levels of test–retest reliability were reported across the different CHAMP scores (range = 0.58–0.67); the authors suggest that the low levels could be influenced by the difficulty in recalling activities. As hypothesized, patients who were classified as being inactive had significantly lower CHAMPS scores when compared with more active patients ($P < 0.001$) [44]. Correlations between the CHAMPS scores and a range of health measures supported *a priori* stated hypotheses, providing acceptable evidence in support of the CHAMPS as a measure of physical activity in older people. Evidence suggests that the CHAMPS can detect improvement in physical activity levels in a large group of participants receiving an active intervention to facilitate increased activity. These changes were greater for the frequency measures [effect size = 0.54 and 0.64) when compared with the change in caloric expenditure (effect size = 0.38 and 0.42), suggesting moderate levels of responsiveness.

The 25-item RAQ-M was developed to evaluate exercise adherence in injured athletes [43]. The original 40-item RAQ developed by Fisher [48] was excluded from phase 1 of the review due to insufficient information to support reproduction. Moreover, evidence of poor reliability and validity have underpinned recommendations for significant redevelopment [45]. The RAQ-M includes six domains of adherence: perceived exertion (three items), pain tolerance during exercise (five items), self-motivation (five items), support from significant others (five items), scheduling (four items) and environmental conditions (three items). The revised six-domain structure was informed by an exploratory and subsequent confirmatory

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

factor analysis [43]. An initial analysis of the internal consistency reliability of the six domains ranged from 0.66 (perceived exhaustion) to 0.87 (scheduling). Acceptable 2-week test–retest reliability values were reported, and ranged from 0.64 (pain tolerance) to 0.81 (support from significant others); however, the relative stability of these athletes' injuries was not reported. Small to moderate levels of association were reported between the RAQ-M domains and three adherence measures, including the SIRAS [43]; however, the absence of *a priori* hypothesized associations between variables limits interpretation. A process of forward and backward translation facilitated translation of the measure from English into Korean.

The 10-item ROAQ purports to measure the tendency for an athlete to be overly adherent to a rehabilitation regime, ignore practitioner recommendations and attempt an expedited rehabilitation and return to sport [48]. Items were generated following a review of the literature for indicators of over-adherence and discussion with experts in sports psychology and clinical rehabilitation of athletes. Young athletes were not consulted. The two-domain factor structure was supported following ROAQ completion by two independent groups of athletes, the first aged 13–18 years (study 1) and the second older adolescents (study 2). Acceptable levels of internal consistency reliability ($\alpha > 0.70$) were reported for both domains in both groups. There is limited evidence in support of the construct validity of the measure; the absence of *a priori* hypothesized associations between variables limits interpretation. The ROAQ has only been evaluated by the developers.

Discussion

Despite the large number of reported approaches to assessing exercise adherence, clear recommendations for the assessment of exercise adherence in MSK populations cannot be made because of poor reporting, inadequate quality and meagre conceptual underpinnings of reviewed measures. Evidence for the seven short-listed measures was mostly limited or not available. Although originally developed as a measure of physical activity in older adults, the CHAMPS has been applied and evaluated as a measure of exercise adherence [44]. Application of a measure for a purpose other than that for which it was developed undermines the validity of the results and limits meaningful interpretation with which to inform decision-making.

Evidence of measurement error, content or face validity, data quality, precision and score interpretation was not identified for any of the reviewed measures. None of the studies explored the relevance, acceptability or appropriateness of measures to the target population, or considered respondent burden. Although all measures had limited evidence of construct validity (convergent; known groups), the absence of *a priori* hypothesized associations between variables limits interpretation and undermines the quality of evidence [34]. Only three measures had limited evidence of structural validity; and just two had (poor) evidence describing measurement responsiveness. There was no evidence of involvement of patients as research partners during the development of any measure. This is a finding reported in other reviews [32, 37], but increasingly viewed as an important consideration in enhancing the relevance and validity of patient-centred outcome assessment [50–52]. Only four of the reviewed measures were patient-reported; the additional measures were clinician-reported. Discrepancies between patients and health professionals with regards to understanding or defining a good outcome have been widely reported [53–56]. It is likely that patients and clinicians have different views about what constitutes adherence. A collaborative exploration of the views of stakeholders, including patients, health professionals and rehabilitation experts, with regards to what should be assessed, by whom, when and in what context is essential to the development of assessment in this field. A new patient-derived measure with a clear conceptual underpinning that reflects the needs of key stakeholders is essential to ensure meaningful investigation of the challenges and burden of adhering to exercise [52].

The review is strengthened by use of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [32]. The methodological and quality concerns highlighted by the review were underpinned by a transparent evaluation of study (COSMIN) and measurement quality [33–35]. This is the first study to evaluate the intra-rater reliability of the COSMIN four-point check-list: poor intra-reviewer agreement between trained, but relatively inexperienced, reviewers was found. These findings highlight the challenge for reviewers of patient reported outcome measure (PROM) quality: poor quality reporting often fails to match the rigors of the COSMIN gold standard checklist and inexperienced reviewers may struggle to unpack complicated or poor quality papers. We recommend that all reviews include an experienced reviewer to guide extraction and/or act as arbiter. Moreover, clear guidance for transparent reporting of PROM quality in published papers is required.

Our extensive search strategy utilized multiple major databases and although limited to English-language publications, English-language abstracts for non-English publications were reviewed and, with the exception of three articles excluded due to language, were excluded due to irrelevance. It is unlikely that any selection bias resulted. The focus of our review was adults with MSK conditions, and our results are not necessarily applicable to non-MSK populations.

A recent review of self-report measures of exercise adherence completed by patients with long-term health problems and undertaking unsupervised home-based exercise programmes similarly concluded that measures are largely unreproducible with extremely limited evidence of essential psychometric properties, thus preventing any clear recommendations for assessment [29]. Another review related to home exercise adherence concluded that there were no valid measures of home exercise adherence for chronic low back pain [57]. The lack of good quality measures and transparency in adherence

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

reporting highlighted in these review must be addressed [29, 57]. In our review only 15% (7 from 45) of the measures purportedly used to assess exercise adherence were taken forward from phases 1 to 2 of the review due to inadequate detail or lack of supporting reference. Appropriate reporting of assessment approaches is essential to ensuring that adherence data are appropriately utilized. Moreover, good reporting contributes to the evidence base with which to inform measurement selection. The Consolidated Standards of Reporting Trials statement [58, 59], and recent patient-reported outcome extension seek to encourage more complete and transparent reporting of assessment approaches and outcome data [60].

In conclusion, we cannot recommend any measure of exercise adherence for MSK settings due to the inadequacy of essential measurement and practical properties for clearly defined measures. Our review provides a critical insight into the many failings of published measures of exercise adherence, though this may reflect the difficulty of measuring adherence. In particular, the conceptual underpinnings of what should be assessed, by whom, when and in what context are poorly considered and essential for future research. Moreover, the transparency in adherence measure reporting must be improved.

Funding: This work was supported by the Chartered Society of Physiotherapy Charitable Trust [grant number PRF/12/13].

Disclosure statement: The authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 United States Bone and Joint Initiative. The Burden of Musculoskeletal Diseases in the United States. 2014. http://www.boneandjointburden.org/2014-report (16 May 2016, date last accessed).

2 Hoy DG, Smith E, Cross M *et al*. Reflecting on the global burden of musculoskeletal conditions: Lessons learnt from the global burden of disease 2010 study and the next steps forward. Ann Rheum Dis 2015;74:4–7.

3 Walsh NE, Brooks P, Hazes JM *et al*. Standards of care for acute and chronic musculoskeletal pain: The bone and joint decade (2000–2010). Arch Phys Med Rehabil 2008;89:1830–45.

4 Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bull World Health Organ 2003;81:646–56.

5 Reginster JY. The prevalence and burden of arthritis. Rheumatology 2002;41(Suppl 1):3–6.

6 Fransen M, McConnell S. Exercise for osteoarthritis of the knee. Cochrane Database Syst Rev 2008;(4):CD004376.

7 Rodrigues EV, Gomes ARS, Tanhoffer AIP, Leite N. Effects of exercise on pain of musculoskeletal disorders: A systematic review. Acta Ortop Bras 2014;22:334–8.

8 Bertozzi L, Gardenghi I, Turoni F *et al*. Effect of therapeutic exercise on pain and disability in the management of chronic nonspecific neck pain: Systematic review and meta-analysis of randomized trials. Phys Ther 2013;93:1026–36.

9 Oesch P, Kool J, Hagen KB, Bachmann S. Effectiveness of exercise on work disability in patients with non-acute non-specific low back pain: systematic review and meta-analysis of randomized controlled trials. J Rehabil Med 2010;42:193–205.

10 Fuentes CJP, Armijo-Olivo S, Magee DJ, Gross DP. Effects of exercise therapy on endogenous pain-relieving peptides in musculoskeletal pain: a systematic review. Clin J Pain 2011;27:365–74.

11 Delitto A, George SZ, Van Dillen L *et al*. Low back pain clinical practice guidelines linked to the international classification of functioning, disability, and health from the orthopaedic section of the American Physical Therapy Association. J Orthop Sports Phys Ther 2012;42:A1–A57.

12 Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: The Chronic Osteoarthritis Management Initiative of the US Bone and Joint Initiative. Semin Arthritis Rheum 2014;43:701–12.

13 Childs JD, Cleland JA, Elliott JM *et al*. Neck pain: Clinical practice guidelines linked to the international classification of functioning, disability, and health from the orthopaedic section of the American Physical Therapy Association. J Orthop Sports Phys Ther 2008;38:A1–A34.

14 Casperson CJ, Powell KE, Christenson GM. Physical activity, exercise and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100:126.

15 Roddy E, Zhang W, Doherty M *et al*. Evidence-based recommendations for the role of exercise in the management of osteoarthritis of the hip or knee—the MOVE consensus. Rheumatology 2005;44:67–73.

16 World Health Organization. Adherence to Long Term Therapies: Evidence for Action. Geneva: World Health Organization, 2003.

17 McLean SM, May S, Moffett JK, Sharp DM, Gardiner E. Prognostic factors for progressive non-specific neck pain: A systematic review. Phys Ther Rev 2007;12:207–20.

18 Hayden JA, Van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. Ann Intern Med 2005;142:776–85.

19 McLean SM, Moffett JAK, Sharp DM, Gardiner E. A randomised controlled trial comparing graded exercise treatment and usual physiotherapy for patients with non-specific neck pain (the GET UP neck pain trial). Man Ther 2013;18:199–205.

20 Kolt G, McEvoy J. Adherence to rehabilitation in patients with low back pain. Man Ther 2003;8:110–6.

21 Forkan R, Pumper B, Smyth N *et al*. Exercise adherence following physical therapy intervention in older adults with impaired balance. Phys Ther 2006;86:401–10.

22 Hardage J, Peel C, Morris D *et al*. Adherence to exercise scale for older patients (AESOP): A measure for predicting

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

exercise adherence in older adults after discharge from home health physical therapy. J Geriatr Phys Ther 2007;30:69–78.

23 Gucciardi E. A systematic review of attrition from diabetes education services: strategies to improve attrition and retention research. Can J Diabetes 2008;32:53–65.

24 Martin C, Perfect T, Mantle G. Non-attendance in primary care: the views of patients and practices on its causes, impact and solutions. Fam Pract 2005;22:638–43.

25 Weinger K, McMurrich SJ, Yi JP, Lin S, Rodriguez M. Psychological characteristics of frequent short-notice cancellers of diabetes medical and education appointments. Diabetes Care 2005;28:1791–3.

26 Jordan JL, Holden MA, Mason E, Foster NE. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. Cochrane Database Syst Rev 2010;(1):CD005956.

27 McLean SM, Burton M, Bradley L, Littlewood C. Interventions for enhancing adherence with physiotherapy: A systematic review. Man Ther 2010;15:514–21.

28 Rankin G, Rushton A, Olver P, Moore A. Chartered society of physiotherapy's identification of national research priorities for physiotherapy using a modified Delphi technique. Physiotherapy 2012;98:260–72.

29 Bollen JC, Dean SG, Siegert RJ, Howe TE, Goodwin VA. A systematic review of measures of self-reported adherence to unsupervised home-based rehabilitation exercise programmes, and their psychometric properties. BMJ Open 2014;4:e005044.

30 Haywood K, Collin S, Crawley E. Assessing severity of illness and outcomes of treatment in children with chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME): A systematic review of patient−reported outcome measures (PROMs). Child Care Health Dev 2014;40:806–24.

31 Mokkink LB, Terwee CB, Patrick DL *et al*. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737–45.

32 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Int J Surg 2010;8:336–41.

33 Elbers RG, Rietberg MB, van Wegen EE *et al*. Self-report fatigue questionnaires in multiple sclerosis, Parkinson's disease and stroke: a systematic review of measurement properties. Qual Life Res 2012;21:925–44.

34 Terwee CB, Mokkink LB, Knol DL *et al*. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21:651–7.

35 Haywood KL, Staniszewska S, Chapman S. Quality and acceptability of patient-reported outcome measures used in chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME): A systematic review. Qual Life Res 2012;21:35–52.

36 Holden MA, Haywood KL, Potia TA, Gee M, McLean S. Recommendations for exercise adherence measures in musculoskeletal settings: a systematic review and consensus meeting (protocol). Syst Rev 2014;3:10.

37 Whitehead L, Perkins G, Clarey A, Haywood K. A systematic review of the outcomes reported in cardiac arrest clinical trials: the need for a core outcome set. Resuscitation 2015;88:150–7.

38 Whistance R, Forsythe R, McNair A *et al*. A systematic review of outcome reporting in colorectal cancer surgery. Colorectal Dis 2013;15:e548–60.

39 Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 2009;18:1115–23.

40 Kolt GS, Brewer BW, Pizzari T, Schoo AM, Garrett N. The sport injury rehabilitation adherence scale: a reliable scale for use in clinical physiotherapy. Physiotherapy 2007;93:17–22.

41 Kortte KB, Falk LD, Castillo RC, Johnson-Greene D, Wegener ST. The Hopkins Rehabilitation Engagement Rating Scale: development and psychometric properties. Arch Phys Med Rehabil 2007;88:877–84.

42 Lenze EJ, Munin MC, Quear T *et al*. The Pittsburgh Rehabilitation Participation Scale: reliability and validity of a clinician-rated measure of participation in acute rehabilitation. Arch Phys Med Rehabil 2004; 85:380–4.

43 Shin JT, Park R, Song WI, Kim SH, Kwon SM. The redevelopment and validation of the rehabilitation adherence questionnaire for injured athletes. Int J Rehabil Res 2010;33:64–71.

44 Stewart AL, Mills KM, King AC *et al*. CHAMPS physical activity questionnaire for older adults: outcomes for interventions. Med Sci Sports Exerc 2001;33:1126–41.

45 Brewer BW, Daly JM, Van Raalte JL, Petitpas A, Sklar J. A psychometric evaluation of the rehabilitation adherence questionnaire. J Sport Exercise Psychol 1999;21: 167–73.

46 Brewer BW, Van Raalte JL, Petitpas AJ *et al*. Preliminary psychometric evaluation of a measure of adherence to clinic-based sport injury rehabilitation. Physical Therapy Sport 2000;1:68–74.

47 Brewer BW, Avondoglio JB, Cornelius AE *et al*. Construct validity and interrater agreement of the sport injury rehabilitation adherence scale. J Sport Rehab 2002;11:170–8.

48 Podlog L, Gao Z, Kenow L *et al*. Injury rehabilitation overadherence: preliminary scale validation and relationships with athletic identity and self-presentation concerns. J Athletic Training 2013;48:37250.

49 Fisher AC. Adherence to sports-injury rehabilitation programs. Phys Sports Med 1988;16:47–47.

50 Mayer MM. Seeking what matters. Patient 2012;5:71–4.

51 Haywood K, Brett J, Salek S *et al*. Patient and public engagement in health-related quality of life and patient-reported outcomes research: What is important and why should we care? Findings from the first ISOQOL patient engagement symposium. Qual Life Res 2015;24:1069–76.

52 Staniszewska S, Haywood KL, Brett J, Tutton L. Patient and public involvement in patient-reported outcome measures. Patient 2012;5:79–87.

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020

53 Hewlett SA. Patients and clinicians have different perspectives on outcomes in arthritis. J Rheumatol 2003;30:877–9.

54 Kessler L, Ramsey SD. The outcomes of the cancer outcomes research symposium: A commentary. Med Care 2002;40:III104–8.

55 Kvien TK, Heiberg T. Patient perspective in outcome assessments—perceptions or something more? J Rheumatol 2003;30:873–6.

56 Liang MH. Pushing the limits of patient-oriented outcome measurements in the search for disease modifying treatments for osteoarthritis. J Rheumatol Suppl 2004;70:61–5.

57 Beinart NA, Goodchild CE, Weinman JA, Ayis S, Godfrey EL. Individual and intervention-related factors associated with adherence to home exercise in chronic low back pain: a systematic review. Spine J 2013;13:1940–50.

58 Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. BMC Med 2010;8:18.

59 Moher D, Hopewell S, Schulz KF *et al*. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg 2012;10:28–55.

60 Calvert M, Brundage M, Jacobsen PB, Schunemann HJ, Efficace F. The CONSORT patient-reported outcome (PRO) extension: Implications for clinical trials and practice. Health Qual Life Outcomes 2013;11:184-7525-11-184.

Downloaded from https://academic.oup.com/rheumatology/article/56/3/426/2736411 by guest on 05 November 2020