Title:

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

Authors:
Dr Sionnadh McLean
Dr Melanie Holden
Tanzila Potia
Melanie Gee
Ross Mallett
Sadiq Bhanbhro
Dr Helen Parsons
Dr Kirstie Haywood

Affiliations:
1Faculty of Health and Wellbeing, Collegiate Campus, Sheffield Hallam University, Sheffield, S10 2BP
2Arthritis Research UK Primary Care Centre, Keele University, ST5 5BG
3Centre for Health and Social Care Research, Collegiate Campus, Sheffield Hallam University, Sheffield, S10 2BP.
4Warwick Medical School; The University of Warwick; Coventry CV4 7AL
5Royal College of Nursing Research Institute, Warwick Medical School, Warwick University. Coventry. CV4 7AL

Corresponding Author:
Dr Sionnadh McLean, Faculty of Health and Wellbeing, Collegiate Campus, Sheffield Hallam University, Sheffield, S10 2BP. Email: s.mclean@shu.ac.uk
ABSTRACT

We reviewed the quality and acceptability of exercise adherence measures applied in published musculoskeletal studies to inform recommendations for use in research and/or routine practice settings. A systematic review of measures was conducted in two phases. Study and measurement quality was assessed against recommended criteria. Phase one identified 313 articles, from which 41 reproducible measures were identified. Published evidence of measurement and practical properties for these measures (phase 2) was limited or unavailable, resulting in nine articles for just six measures: three clinician-reported and three patient-reported. Four measures were specific to the assessment of exercise adherence and two specific to physical activity. Significant methodological and quality issues were identified, making assessment recommendations difficult. The conceptual underpinning of exercise adherence is poorly defined. Future research should seek to engage collaboratively with relevant stakeholders to ensure that the way in which exercise adherence is assessed is high quality, relevant and acceptable.

Funding: This work was supported by a grant from the Chartered Society of Physiotherapy Charitable Trust (grant number PRF 12/13)

Keywords: Acceptability, Adherence, Exercise, Measurement, Musculoskeletal, Physical activity, Quality, Systematic review
Introduction

Musculoskeletal (MSK) injuries or disorders such as arthritis and osteoporosis are common and costly.[1] The associated disability burden is high and second only to mental and behavioural problems.[2] For many, the associated progressive functional limitation in everyday activities, including paid employment, results in significant financial costs for individuals and society (Walsh et al 2008). Increasing age and lifestyle factors such as obesity and physical inactivity negatively impact MSK disorders:[3,4] the ageing population and increasingly sedentary lifestyles suggests that the disease burden will continue to increase.[3]

Exercise can reduce pain, improve physical dysfunction and enhance the quality of life of individuals with a range of MSK disorders:[5-9] clinical guidelines advocate the use of exercise programmes as part of a long-term management strategy.[10-13] However, an individual’s ability to adhere to a recommended exercise programme, defined as “the extent to which a person’s behaviour corresponds with agreed recommendations from a healthcare provider”, is important for success.[14,15] Patients who adhere to regular physical activity are less likely to progress to recurrent, persistent or disabling problems,[16,17] and increasing adherence may derive greater patient benefit than improving aspects of the intervention itself.[15]

However, poor adherence to prescribed exercise is common, with estimates of less than 50% adherence reported.[18-21] Non-adherence may negatively impact treatment effectiveness and efficiency, the therapeutic relationship, waiting times and
cost of care.[22-24] Few strategies for effectively increasing exercise adherence have been identified and guidance for best practice does not exist.[25,26] Consequently, further investigation to develop exercise adherence interventions has been prioritised.[27] Evaluations of the relative benefit of these interventions are essential to informing such guidance; however, guidance for the assessment of exercise adherence in MSK clinical trials or routine practice settings does not exist.

Recent evidence suggests that wide variation in the assessment of exercise adherence exists.[28,29] Such heterogeneity in outcome reporting is problematic across many healthcare settings,[30,31] limiting the conduct of systematic reviews and meta-analyses of evidence.[32,33] Where a large number of assessment approaches exist, structured reviews of the quality and acceptability of different approaches are an essential pre-requisite to informing selection.[34,35] This review sought to identify all clearly reported and reproducible measures of exercise adherence applied in published studies of patients with MSK disorders, and to evaluate the quality and acceptability of these measures against a transparent appraisal framework, thus providing guidance for assessment in clinical practice and research settings.
Methods

The systematic review was conducted in two phases and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[36] Phase one identified all clearly reported and reproducible measures used to assess exercise adherence in published MSK studies. Phase two reviewed published and unpublished evidence of measurement and practical properties for shortlisted measures. Study and measurement quality was assessed against the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist,[35,37,38] and a transparent appraisal framework.[39] respectively.

**Phase 1: Identifying measures of exercise adherence**

Search strategy

A search strategy was developed to identify methods used to assess exercise adherence in musculoskeletal settings (Appendix 1); all study types were included. Eight databases were searched (inception - May 2013): Medline, SPORTDiscus, CINAHL Plus, PsycINFO, AMED, Cochrane Library, Embase, and the Web of Science.

Study selection
Titles, abstracts and full text articles were independently screened for inclusion by two reviewers (SMc, MH, RM, TP, SB). Disagreement was discussed with a third independent reviewer (SMc, MH, RM, TP, SB, KH).

Inclusion and exclusion criteria

Articles were included if they focused on adults with a MSK disorder receiving a therapeutic exercise or physical activity intervention delivered in any therapeutic setting (inpatient, outpatient, community) and for which assessments of adherence to exercise or activity (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 participants were healthy volunteers, less than 18 years old, or with non-musculoskeletal conditions.

Reproducible methods of assessment - supported by an appropriate citation or sufficient text to allow reproduction,[30,31] were listed and categorised 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.

Phase 2: Evidence of quality and acceptability

Search strategy
The names of short-listed measures were combined with a search filter specific to the identification of studies reporting evidence of measurement and/practical properties.[40] The search was modified for application in the databases identified above. The developers of specific measures were contacted to locate additional evaluative evidence. Titles, abstracts and full text articles were independently assessed by two reviewers (MH, TP, RM, SMc); a third reviewer resolved any disagreements (KH). Reference lists of included articles were reviewed for additional published articles.

Inclusion criteria
Published, English language articles were included if they provided evidence of assessment development and/or evaluation of the named measure(s) in a MSK population.

Data extraction and inter-rater reliability
A data extraction form informed by earlier reviews and the COSMIN checklist was used to capture study- (population, intervention, and setting) and measurement-specific information: reliability (internal consistency, test-retest, intra-/inter-tester, measurement error); validity (content, construct, convergent/divergent, and known group differences); explicit hypothesis testing; conceptual underpinning and aspects of exercise adherence assessed; responsiveness (criterion-/construct-based); interpretation (minimal important difference); and precision (data quality, end effects). Extraction for practical properties included acceptability (relevance and
respondent burden) and feasibility.[34,35,38] The extent of patient involvement in measurement development and/or application was also sought.[39]

In accordance with the COSMIN checklist, each measurement property reported by the study was rated on a 4-point scale (excellent, good, fair, poor).[38] Study methodological quality was evaluated per measurement property and determined by the lowest checklist rating.[35,38] Following a group training session, four primary reviewers (SMc, MH, TP, RM) 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. Following this initial evaluation, where disagreement existed, consensus was sought through discussion with a third, experienced reviewer (KH) who independently reviewed all articles.

Data synthesis
Data was qualitatively synthesised to determine the overall quality and acceptability of each measure.[34,37] The synthesis considered the following factors: i) study methodological quality (COSMIN scores); ii) number of studies reporting specific evidence per measure; iii) results for each measurement/practical property per measure; and iv) consistency between studies.[37] The synthesis score had two elements: 1) the overall quality of a measurement property was reported as: adequate (+), not adequate (-), conflicting (+/-), or unclear (?); 2) levels of evidence
for the overall quality of each measurement property was further defined to indicate ‘strong’, ‘moderate’, ‘limited’, ‘conflicting’, or ‘unknown’ evidence.[37]

Results

Identification of studies and measures

Phase 1
Following removal of duplicates, 11981 records were identified. Following title and abstract screening 313 full-text articles were retrieved and reviewed in full.

A total of 243 approaches to the assessment of exercise adherence were identified. In order of frequency of reporting these were: exercise logs and diaries (n=102); unnamed questionnaires/scales (n=51); reproducible or named questionnaires/scales (n=41); pedometers, accelerometers and other objective devices (n=24); interviews (n=17); and calendars/postcards (n=8). From this total, only the 41 named questionnaires/scales had an appropriate citation or sufficient detail to allow reproduction.

Phase 2
Evidence for measurement and/or practical properties were sought for the 36 measures. From a total of 3735 unique records, 105 full-text articles were reviewed in full and nine retained for phase two (see figure 1).[21,41-48]
These nine articles provide evidence for six clearly defined measures applied as measures of exercise adherence in an MSK population. Three are clinician-reported: Hopkins Rehabilitation Engagement Rating Scale (HRERS),[42] Pittsburgh Rehabilitation Participation Scale (PRPS),[43] and the Sport Injury Rehabilitation Adherence Scale (SIRAS).[46] Three are patient-reported: Adherence to Exercise Scale for Older Patients (AESOP),[21] Community Healthy Activities Model Program for Seniors Activities Questionnaire for Older Adults (CHAMPS);[45] and the modified Rehabilitation Adherence Questionnaire (RAQ-M).[44] Attempts to contact measurement developers for further information or examples of the original measure were largely 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 exceed 80% for only 20 items (50%).[35] Disagreement was most often due to poor reporting of evidence in the reviewed papers, associated interpretation difficulties, reading errors or difficulties applying the checklist.

Study characteristics
Although five studies were adequately sized for evaluative purposes (range 145-249),[38] four included fewer than 100 patients.[21,41,46,48] The ages of patients’ ranged from 18 to 96 years (see Table 1). Studies covered a wide range of MSK settings: athletes with acquired knee injuries;[44,46-48] general MSK disorders in outpatient settings;[41] older patients with generalized MSK conditions;[21,43,45] and acute inpatient populations.[42]
Adherence measures

Five of the six measures were originally developed as measures of exercise adherence: the SIRAS and RAQ-M purport to measure adherence to rehabilitation following sport-related injury; the HRERS and PRPS purport to measure adherence to or participation in rehabilitation in acute in-patients; the AESOP seeks to evaluate the ability of older adults to adhere to prescribed home exercise programmes (HEP).[21] Although a measure of physical activity in older adults, the CHAMPS has been used a proxy measure of exercise adherence and hence is included in this review.[45] With the exception of the RAQ-M which was evaluated in Korean athletes, all measures were developed and evaluated in the United States of America. The characteristics and measurement properties of all reviewed measures are summarised below and in tables 1 to 4. Study methodological quality and the qualitative synthesis is summarised in Table 5.

Clinician-reported

The HRERS is a 5-item questionnaire used to assess the therapist’s perception of an individual’s engagement in acute in-patient rehabilitation. There is limited evidence of reliability and validity following completion in an acute inpatient population of patients with spinal cord injury, stroke, amputation or hip/knee replacement.[42] The unidimensional structure 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 (Cronbachs alpha=0.91) and acceptable inter-rater agreement (ICC=0.73) was reported.[42] 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 (Table 3). 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 \((r=0.22)\) (Table 3);[42] although the authors suggest that hypothesized associations were supported, these were not clearly stated, hence limiting interpretation in support of measurement validity.

The PRPS is a single item rating of the extent of patient participation (effort and motivation) during each treatment session of acute inpatient rehabilitation.[43] Item development involved therapist interviews and therapy session observation of older patients with generalized MSK problems. There is limited evidence of reliability and validity following completion with older people with generalized MSK conditions.[43] High values of inter-rater reliability (range ICC=0.91 to 0.96) were reported.[43] 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 (LOS) were reported \((r=-0.13; p<0.05)\) (Table 5);[43] 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 greater than 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 SIRAS is a simple 3-item scale with which the therapist rates their perception of the degree to which a patient exerts them self, follows practitioner’s instructions and advice, and is receptive to changes in the rehabilitation program during a given rehabilitation session (Table 1). The single factor structure of the SIRAS (‘exercise adherence’) is supported by several studies following completion by athletes and the general MSK population.[41,47] Internal consistency evaluations further support reporting the SIRAS as a single index valu).[47] Acceptable levels of internal consistency supports application in groups of patients (Cronbach’s alpha range 0.82-0.8).[47,48] Poor to high levels of inter-rater (ICC range=0.57-0.77; RAI range=0.84-0.94) and acceptable one-week test-retest reliability has been reported (range=0.63-0.77).[41,47] Evidence in support of known-groups validity is provided following the assessment of standardized vignettes describing three levels of adherence in athletes(Table 3).[41,48]

**Patient-reported measures**

The AESOP is a 42-item interview-administered questionnaire, developed to assess exercise adherence in older patients.[21] The measure constitutes three domains, informed by social cognitive theory: self-efficacy expectations (15 items), outcome expectations (16 items) and 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).[21] All correlations between the three AESOP domains and the Short Form 12-item Health Survey (SF-12, version 2) physical (PCS) and mental component scales (MCS) were very small (Table 3); 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 which has been evaluated for use as a proxy measure of exercise adherence of daily life.[45] The CHAMPS asks about ‘activities that you may have done in the past 4-weeks’. The information is used to calculate a) frequency of activities - the number of minutes of physical activity per week and b) the calories expended per week in all physical activities. Each score can be calculated for 1) moderate and greater activity levels; and b) 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 two-week test-retest reliability.[45] 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 statistically significantly lower CHAMPS scores when compared to more active patients (p<0.001).[45] 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 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 (ES) = 0.54 and 0.64) when compared to the change in caloric expenditure (ES = 0.38 and 0.42), suggesting moderate levels of responsiveness.

A 25-item modified-version of the RAQ (RAQ-M) has recently been proposed (Shin et al, 2010). The original 40-item RAQ developed by fisher and colleagues was excluded from phase 1 of the review due to insufficient information to support reproduction.[49] Moreover, evidence of poor reliability and validity have underpinned recommendations for significant re-development.[46] The RAQ-M includes six domains of adherence: perceived exertion (3 items), pain tolerance during exercise (5 items), self-motivation (5 items), support from significant others (5 items), scheduling (4 items), and environmental conditions (3 items). The revised six-domain structure was informed by an exploratory and subsequent confirmatory factor analysis.[44] An initial analysis of the internal consistency reliability of the six-domains ranged from 0.66 (perceived exhaustion) to 0.87 (scheduling). Acceptable two-week test-retest reliability values were reported, ranged from 0.64 (pain tolerance) to 0.81 (support from significant others); however, the relative stability of these athletes was not reported. Small to moderate levels of association were reported between the RAQ-M domains and three adherence measures, including the SIRAS.[44] 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.
Discussion

Despite the large number of approaches to assessing exercise adherence reported in published MSK studies, clear recommendations for the assessment of exercise adherence in this population cannot be made because of poor reporting, inadequate quality and meagre conceptual underpinnings of reviewed measures. Routine practice and evaluative studies of interventions to enhance adherence to exercise require robust and relevant measures with acceptable evidence of essential measurement and practical properties with which to inform decision-making.[50] However, evidence for the six short-listed measures was mostly limited or not available.

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 or clinician 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.[38] Only two measures had limited evidence of structural validity; and just two had (poor) evidence describing measurement responsiveness. There was no evidence of the active involvement of patients as research partners during the development or evaluation of any measure. This is a finding reported in other reviews,[34,39] but increasingly viewed as an important consideration in enhancing the relevance and validity of patient-centred outcome
Only three 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. Although not evident within the development of the reviewed measures, it is likely that patients have different views to clinicians with regards to what is good adherence, the barriers encountered, and hence what should be included in an assessment of adherence. Further qualitative, collaborative exploration of the views of key stakeholders with regards to what should be assessed, by whom, when and in what context is essential to the further development of assessment in this field. A patient-centred, collaborative approach to developing a new measure appropriate to the assessment of exercise adherence in MSK settings is essential to enable a better understanding of the challenges and burden of adhering to exercise and the relative success of interventions designed to enhance adherence to be comprehensively evaluated.

The review is strengthened by use of the PRISMA guidelines. The methodological and quality concerns highlighted by the review were underpinned by a transparent evaluation of study (COSMIN) and measurement quality. This is the first study to evaluate the intra-rater reliability of COSMIN 4-point check-list: poor intra-reviewer agreement between trained, but relatively inexperienced, reviewers was found. Disagreement was often due to poor quality reporting, associated interpretation difficulties and challenges applying the checklist; discussion with an experienced reviewer was essential. These findings highlight the challenge for reviewers of 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 utilised 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 all included studies was adults with MSK conditions, and hence our results are not necessarily applicable to non-MSK populations.

Whilst not reporting extensive search strategies or transparent appraisal processes, recent reviews of self-report measures of exercise adherence completed by patients with long-term health problems and undertaking unsupervised home-based exercise programmes and used in the assessment of adherence to home-based rehabilitation have similarly concluded that measures are largely un-reproducible with extremely limited evidence of essential psychometric properties, thus preventing any clear recommendations for assessment.[28,29] The lack of transparency in outcome reporting highlighted in these reviews must be addressed: from the large number of approaches purportedly used to assess exercise adherence, only 15% were taken forward to phase 2 of the review due to inadequate detail or lack of supporting reference. Appropriate reporting of assessment approaches is essential to ensuring that outcomes data is appropriately utilized. Moreover, good reporting contributes to the evidence-base with which to inform measurement selection. The CONSORT (Consolidated Standards of Reporting Trials) statement,[58,59] and recent PRO-
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. The transparency and detail of our review provides a critical insight into the many failings of ‘published’ measures of exercise adherence. In particular, the conceptual underpinnings of what should be assessed, by whom, when and in what context is poorly considered and is an essential requirement for future research. Moreover, the transparency in outcome reporting must be improved.

Key messages: (3 key messages, no more than 15 words each)

- The poor conceptualization and quality of available measures of exercise adherence limits current recommendations.

- The poor reporting of assessment approaches limits interpretation and must be addressed in future research.

- A collaborative understanding of what to assess, by whom, when and in what context is required.

Acknowledgements: none
Conflicts of interest: The authors have declared no conflict of interest

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

References:


29. Frost R, Williams B, McClurg D, Brady M, Levati S. A systematic review of adherence measurement methods currently used in randomised controlled trials of home-based rehabilitation interventions. Glasgow Caledonian University, Glasgow: Nursing, Midwifery and Allied Health Professions Research Unit; 2014.
33. Gargon E, Gurung B, Medley N, Altman DG, Blazey JM, Clarke M, Williamson PR. Choosing important health outcomes for comparative effectiveness research: A systematic review. 2014: e99111
39. Haywood KL, Staniszewska S, Chapman S. Quality and acceptability of patient-reported outcome measures used in chronic fatigue syndrome/myalgic


50. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use, Oxford University Press; 2014.


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

<table>
<thead>
<tr>
<th>Measure (Developer, Year)</th>
<th>Evaluations (n)</th>
<th>Construct</th>
<th>Domains (items)</th>
<th>Response options</th>
<th>Recall</th>
<th>Score range</th>
<th>Admin (time)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinician-completed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Hopkins Rehabilitation Engagement Rating Scale (HRERS)  
(Kortte et al. 2007)[42] | 1               | Behavioral observations of patients during acute inpatient rehabilitation | 5 items:  
1. Attendance at rehabilitation session (1)  
2. Frequency of required Verbal/Physical Prompts (1)  
3. Perceived Positive attitude to exercise (2)  
4. Perceived need for and benefit from rehabilitative exercise  
5. Active participation in rehabilitative exercise (1) | 6-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 (page 2) | Simple summation: range 5 to 30, where 5 is poor and 30 is best engagement in the therapy process | NR |
| Pittsburgh Rehabilitation Participation Scale (PRPS)  
(Lenze et al. 2004)[43] | 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)[46] | 8               | Adherence during rehabilitation sessions | 3 items:  
1. Perceived Intensity/Effort/Exertion (1)  
2. Frequency of following therapist instructions (1)  
3. Receptive to change in rehabilitation exercise (1) | 5-point numerical rating scale:  
Anchors:  
1. Minimum effort (1) to  
Maximum effort (5).  
2. Never (1) to Always (5)  
3. 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 |
<table>
<thead>
<tr>
<th>Scale/Questionnaire</th>
<th>Items</th>
<th>Construct(s) for predicting adherence</th>
<th>5-point agreement:</th>
<th>Time Frame</th>
<th>Scoring</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>AESOP (Hardage et al. 2007)</td>
<td>42</td>
<td>Social cognitive theory constructs for predicting HEP adherence in older adults: self-efficacy expectations, outcome expectations, and outcome expectancies</td>
<td>Strongly disagree (1); Disagree (2); No opinion (3); Agree (4); Strongly agree (5)</td>
<td>2 weeks</td>
<td>Simple summation: 1. Range 15-75; 2. Range 16-80; 3. Range 11-55</td>
<td>Lower scores suggest lower levels of adherence.</td>
</tr>
<tr>
<td>CHAMPS (Stewart et al. 2001)</td>
<td>41</td>
<td>Types and intensity levels of physical activity</td>
<td>Strongly disagree (1); Disagree (2); No opinion (3); Agree (4); Strongly agree (5)</td>
<td>4 weeks</td>
<td>Simple summation: 1. Frequency of activities per week: number of minutes of physical activity per week across all activities. 2. Calorie expenditure: per week multiply estimated duration of each activity by the MET value and summing across all activities. Both can be calculated for: A. Moderate and greater activity measures. B. All activity measures Therefore, four scores possible.</td>
<td>NR</td>
</tr>
<tr>
<td>RAQ-M (Shin et al. 2010)</td>
<td>25</td>
<td>Rehabilitation adherence in injured athletes</td>
<td>Strongly disagree (1); Disagree (2); Agree (3); Strongly agree (4)</td>
<td>1 week</td>
<td>Simple item summation for each domain: 1. range 3-12; 2. range 5-20; 3. range 5-20; 4. range 5-20; 5. range 4-16; 6. range 3-12 Higher scores reflect greater levels of adherence</td>
<td>NR</td>
</tr>
<tr>
<td>ROAQ (Podlog et al. 2013)</td>
<td>10</td>
<td>Assessment of over-adherence behaviours and beliefs in injured athletes</td>
<td>Never or strongly disagree (1) to Always or Strongly agree (5)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>
n= number of studies evaluating the measurement and practical properties of each measure, NR= Not reported, MET= Metabolic Energy Equivalent, 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, Rehabilitation Over-adherence Questionnaire (ROAQ)
Table 2. Methodological quality and investigated measurement and practical properties per measure per reviewed article

<table>
<thead>
<tr>
<th>Articles (n=9)</th>
<th>Population (n)</th>
<th>Age (years) Mean (SD); range</th>
<th>Measures</th>
<th>Reliability</th>
<th>Validity</th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brewer et al. 2002[48]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study 1</td>
<td>43 (practitioners)</td>
<td>range 20-43</td>
<td>SIRAS</td>
<td>-</td>
<td>Poor</td>
<td>-</td>
</tr>
<tr>
<td>Study 2</td>
<td>12 (rehab patients)</td>
<td>29.33 (11.44)</td>
<td>SIRAS</td>
<td>-</td>
<td>Poor</td>
<td>-</td>
</tr>
<tr>
<td>Brewer et al. 2000[47]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study 1</td>
<td>145 (orthopaedic outpatients)</td>
<td>43.95 (15.54)</td>
<td>SIRAS</td>
<td>Fair</td>
<td>Poor</td>
<td>Fair</td>
</tr>
<tr>
<td>Study 2</td>
<td>31 (sport related knee injury)</td>
<td>NR</td>
<td>SIRAS</td>
<td>-</td>
<td>Fair</td>
<td>-</td>
</tr>
<tr>
<td>Study 3</td>
<td>43 (rehab post ACL repair)</td>
<td>NR</td>
<td>SIRAS</td>
<td>-</td>
<td>Fair</td>
<td>-</td>
</tr>
<tr>
<td>Brewer et al. 1999[46]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hardage et al. 2007[22]</td>
<td>50</td>
<td>79.9 range 65-91</td>
<td>AESOP SF-12 GDS mMSE</td>
<td>-</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td>Kolt et al. 2007[41]</td>
<td>60 (physiotherapists = raters)</td>
<td>NR</td>
<td>SIRAS</td>
<td>-</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td>Study 2</td>
<td>45 patients (general MSK)</td>
<td>&gt;18 yrs</td>
<td>SIRAS</td>
<td>-</td>
<td>Poor</td>
<td>-</td>
</tr>
<tr>
<td>Kortte et al. 2007[42]</td>
<td>208</td>
<td>56.7 (17.52); range 18-91</td>
<td>HRRERS FIM BSI</td>
<td>Poor</td>
<td>Poor</td>
<td>Fair</td>
</tr>
<tr>
<td>Study</td>
<td>n</td>
<td>Age</td>
<td>Measure</td>
<td>PRPS</td>
<td>FIM-motor</td>
<td>Study 1</td>
</tr>
<tr>
<td>------------------</td>
<td>-------</td>
<td>------</td>
<td>------------------------------</td>
<td>------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>Lenze et al. 2004[43]</td>
<td>242</td>
<td>70.8 (14.8); range 20-96</td>
<td>PRPS</td>
<td>FIM-motor</td>
<td>Fair</td>
<td>Poor</td>
</tr>
<tr>
<td>Podlog et al. 2013[49]</td>
<td>118</td>
<td>16.0 (1.4); Range 13-18</td>
<td>RAOQ</td>
<td>SPSQ AIMS I-PRRS</td>
<td>Fair</td>
<td>-</td>
</tr>
<tr>
<td>Shin et al., 2010[44]</td>
<td>240</td>
<td>NR</td>
<td>RAOQ</td>
<td>SPSQ AIMS I-PRRS</td>
<td>Fair</td>
<td>Poor</td>
</tr>
<tr>
<td>Stewart et al. 2001[45]</td>
<td>249</td>
<td>74.1 range 65-90</td>
<td>CHAMPS BMI domains SPPB</td>
<td>SF-36</td>
<td>SPPB</td>
<td>Good</td>
</tr>
</tbody>
</table>

n= population size in included study, SD=standard deviation, NR=Not reported, 6-min walk=Six-minute walking test, BMI=Body Mass Index, 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 Over-adherence Questionnaire, I-PRRS=Modified Injury Psychological Readiness to Return to Sport Scale
Table 3: Overall quality of measurement properties per reviewed measure of exercise adherence for MSK populations.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Evaluations (n)</th>
<th>Reliability</th>
<th>Validity</th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Test-retest (intra/inter)</td>
<td>Internal consistency</td>
<td>Measurement error</td>
</tr>
<tr>
<td>Therapist-completed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRERS</td>
<td>1</td>
<td>+ limited</td>
<td>+ limited</td>
<td>Nil</td>
</tr>
<tr>
<td>PRPS</td>
<td>1</td>
<td>+ limited</td>
<td>Nil</td>
<td>Nil</td>
</tr>
<tr>
<td>SIRAS</td>
<td>8</td>
<td>+ limited</td>
<td>+ limited</td>
<td>Nil</td>
</tr>
<tr>
<td>Patient-completed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AESOP</td>
<td>1</td>
<td>- limited</td>
<td>Nil</td>
<td>Nil</td>
</tr>
<tr>
<td>CHAMPS</td>
<td>1</td>
<td>- limited</td>
<td>Nil</td>
<td>Nil</td>
</tr>
<tr>
<td>RAQ-M</td>
<td>1</td>
<td>+ limited</td>
<td>+ limited</td>
<td>Nil</td>
</tr>
<tr>
<td>ROAQ</td>
<td>2</td>
<td>Nil</td>
<td>+ limited</td>
<td>Nil</td>
</tr>
</tbody>
</table>

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.
Appendix 1

Search strategy for phase 1

The search strategies used title/abstract words and relevant indexing to capture the concept of exercise adherence in the context of musculoskeletal rehabilitation, for adult patients. The strategies also contained the following exclusions: "cardiac rehabilitation", "pulmonary rehabilitation", "neuro* rehabilitation", "stroke".

To capture exercise adherence: search terms/synonyms for adherence [see below] were searched in proximity (within 3 words) to terms for exercise, in the title/abstract fields; secondly, search terms/synonyms for adherence were searched in combination (AND) with database subject headings for exercise/therapeutic exercise etc; thirdly, search terms/synonyms for exercise were searched in combination (AND) with database subject headings for patient compliance; finally, the database subject headings for exercise/therapeutic exercise etc. were searched in combination (AND) with database subject headings for patient compliance.

To capture musculoskeletal rehabilitation, the above searches were combined (AND) with the search terms/synonyms in the title/abstract fields and the database headings listed below.

**Adherence terms:** adher*, nonadher*, compliant*, noncompliant*, concordan*, cooperat*, co-operat*, uncooperat*, unco-operat*, engag*, disengag*, behaviour#, MeSH: "Patient Compliance"

**Exercise terms:** activ*, exercis*, physical n3 train*, weight n3 train*, sport#, rehab*, MeSH: "Therapeutic Exercise+", "Exercise Therapy+", "Exercise+", "Physical Activity", "Motor Activity"

# Indicative search strategy (Medline/CINAHL Plus with Fulltext, via EBSCOHost)

S1

TI ((adher* or nonadher*) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train* or sport# or rehab*)) OR AB ((adher* or nonadher*) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*))

S2

TI ((complian* or noncomplian*) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*)) OR AB ((complian* or noncomplian*) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*))

S3

TI (concordan* n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*)) OR AB (concordan* n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*))

S4

TI ((cooperat* or co-operat* or uncooperat* or unco-operat*) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*)) OR AB ((cooperat* or co-operat* or uncooperat* or unco-operat*) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*))

S5

TI ((engag* or disengag*) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*)) OR AB ((engag* or disengag*) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*))

S6

TI ((behaviour# or behavior#) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*)) OR AB ((behaviour# or behavior#) n3 (activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*))

S7

S1 OR S2 OR S3 OR S4 OR S5 OR S6

S8

TI (adher* or nonadher*) OR AB (adher* or nonadher*)

S9

TI (complian* or noncomplian*) OR AB (complian* or noncomplian*)

S10

TI (concordan*) OR AB (concordan*)

S11

TI ((cooperat* or co-operat* or uncooperat* or unco-operat*) OR AB ((cooperat* or co-operat* or uncooperat* or unco-operat*)

S12

TI (engag* or disengag*) OR AB (engag* or disengag*)

S13

TI (behaviour# or behavior#) OR AB (behaviour# or behavior#)

S14

S8 OR S9 OR S10 OR S11 OR S12 OR S13

S15

(MH "Therapeutic Exercise+") OR (MH "Exercise Therapy+")

S16

(MH "Exercise+")

S17

(MH "Physical Activity") or (MH "Motor Activity")

S18

S15 OR S16 OR S17

S19

S14 AND S18

S20

TI ((activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*) OR AB ((activ* or exercis* or (physical n3 train*) or (weight n3 train*) or sport# or rehab*)))
(MH "Patient Compliance")
S20 AND S21
S18 AND S21
S7 OR S19 OR S22 OR S23
(MH "Physical Therapy") or (MH "Exercise Movement Techniques") OR (MH "Exercise Therapy")
MH ("Osteopathy") OR (MH "Osteopathic Medicine")
(MH "Chiropractic") OR (MH "Manipulation, Chiropractic")
TI (osteopath* or chiropract* or musculoskeletal or msk ) OR AB (osteopath* or chiropract* or musculoskeletal or msk)
TI (physiotherap* or rehabilitat* ) OR AB (physiotherap* or rehabilitat* )
(MH "Musculoskeletal Diseases")
TI (osteoarthrit* or spondyl* or osteitis or osteochondritis ) OR AB (osteoarthrit* or spondyl* or osteitis or osteochondritis)
TI (arthropathy or bursitis or "shoulder impingement" or myalgia ) OR AB (arthropathy or bursitis or "shoulder impingement" or myalgia)
TI (lordosis or sacroiliac or sciatica or cervicogenic ) OR AB (lordosis or sacroiliac or sciatica or cervicogenic)
(MH "Sciatica")
(MH "Tendinopathy")
(MH "Allodynia")
TI (dyskinesis or tendinitis or tendinopathy or alldynia) OR AB (dyskinesis or tendinitis or tendinopathy or alldynia)
(MH "Hyperalgesia")
(MH "Subluxation")
TI (hyperalgesia or subluxation or disc or misalignment ) OR AB (hyperalgesia or subluxation or disc or misalignment)
TI ("osteopathic lesion" or "frozen shoulder" or "degenerative joint disease") OR AB ("osteopathic lesion" or "frozen shoulder" or "degenerative joint disease")
TI muscular n3 pain OR AB muscular n3 pain
TI (back or lumbar or lumbo* or spine or spinal) n3 pain ) OR AB (back or lumbar or lumbo* or spine or spinal) n3 pain)
TI (neck or cervical) n3 pain ) OR AB (neck or cervical) n3 pain)
TI (knee* or hip or hips or shoulder*) n3 pain ) OR AB (knee* or hip or hips or shoulder*) n3 pain)
TI (ankle# or foot or feet or elbow# or hand#) n3 pain ) OR AB (ankle# or foot or feet or elbow# or hand#) n3 pain)
S47 TI ("flank pain" or "buttock pain" or "joint pain" or "radicular pain") OR AB ("flank pain" or "buttock pain" or "joint pain" or "radicular pain")

S48 (MH "Back Pain+")

S49 (MH "Neck Pain")

S50 (MH "Neuralgia+")

S51 (MH "Elbow Pain")

S52 (MH "Arthralgia+")

S53 TI (neuralgia or lumbago or arthralgia) OR AB (neuralgia or lumbago or arthralgia)

S54 TI ("adverse neural tension" or "muscle tear") OR AB ("adverse neural tension" or "muscle tear")

S55 TI ((sprain* or strain*) n5 musc*) OR AB ((sprain* or strain*) n5 musc*)

S56 (MH "Musculoskeletal System+")

S57 (MS "Pain+")

S58 S56 AND S57

S59 S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S58

TI ("cardiac rehabilitation" or "pulmonary rehabilitation" or "neuro* rehabilitation" or stroke or cancer or carcinoma) OR AB ("cardiac rehabilitation" or "pulmonary rehabilitation" or "neuro* rehabilitation" or stroke or cancer)

S60 TI (child* NOT adult*) OR AB (child* NOT adult*)

S61 TI (infan* NOT adult*) OR AB (infan* NOT adult*)

S62 (MH "Child+") NOT (MH "Adult+")

S63 S60 OR S61 OR S62 OR S63

S64 S59 not S64

S65 S24 AND S65
Appendix 2

Search strategy for phase 2

For each shortlisted named measure, the name was searched as a word/phrase in the title/abstract fields of each database. Where the results set exceeded 50 records, the 'Sensitive search filter for measurement properties' found in Appendix 2 of Terwee et al. (2009) [1] was additionally applied.

Search strategy for measurement properties filter (Medline/CINAHL Plus with Fulltext, via EBSCOHost)

S1  (MH "Methods")
S2  Validation Studies
S3  Comparative Study
S4  (MH "Psychometrics")
S5  TI psychometr* OR AB psychometr*
S6  TI ( clinimetr* OR clinometr* ) OR AB ( clinimetr* OR clinometr* )
S7  (MH "Outcome Assessment (Health Care)")
S8  TI "outcome assessment" OR AB "outcome assessment"
S9  TX "outcome measure**
S10 (MH "Observer Variation")
S11 TI "observer variation" OR AB "observer variation"
S12 (MH "Health Status Indicators")
S13 (MH "Reproducibility of Results")
S14 TI reproducib* OR AB reproducib*
S15 (MH "Discriminant Analysis")
S16 TI ( reliab* OR unreliab* ) OR AB ( reliab* OR unreliab* )
S17 TI valid* OR AB valid*
S18 TI coefficient OR AB coefficient
S19 TI ( homogeneity OR homogeneous ) OR AB ( homogeneity OR homogeneous )
S20 TI "internal consistency" OR AB "internal consistency"
S21 TI ( cronbach* AND (alpha OR alphas) ) OR AB ( cronbach* AND (alpha OR alphas) )
S22  TI ( item AND (correlation* OR selection* OR reduction*) ) OR AB ( item AND (correlation* OR selection* OR reduction*) )
S23  TI agreement OR AB agreement
S24  TI ( precision OR imprecision ) OR AB ( precision OR imprecision )
S25  TI "precise values" OR AB "precise values"
S26  TI test-retest OR AB test-retest
S27  TI ( test AND retest ) OR AB ( test AND retest )
S28  TI ( reliab* AND (test OR retest) ) OR AB ( reliab* AND (test OR retest) )
S29  TI stability OR AB stability
S30  TI ( interrater OR inter-rater ) OR AB ( interrater OR inter-rater )
S31  TI ( intrarater OR intra-rater ) OR AB ( intrarater OR intra-rater )
S32  TI ( intertester OR inter-tester ) OR AB ( intertester OR inter-tester )
S33  TI ( intratester OR intra-tester ) OR AB ( intratester OR intra-tester )
S34  TI ( interobserver OR inter-observer ) OR AB ( interobserver OR inter-observer )
S35  TI ( intraobserver OR intra-observer ) OR AB ( intraobserver OR intra-observer )
S36  TI ( intertechnician OR inter-technician ) OR AB ( intertechnician OR inter-technician )
S37  TI ( intratechnician OR intra-technician ) OR AB ( intratechnician OR intra-technician )
S38  TI ( interexaminer OR inter-examiner ) OR AB ( interexaminer OR inter-examiner )
S39  TI ( intraexaminer OR intra-examiner ) OR AB ( intraexaminer OR intra-examiner )
S40  TI ( interassay OR inter-assay ) OR AB ( interassay OR inter-assay )
S41  TI ( intraassay OR intra-assay ) OR AB ( intraassay OR intra-assay )
S42  S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41
S43  TI ( interindividual OR inter-individual ) OR AB ( interindividual OR inter-individual )
S44  TI ( intraindividual OR intra-individual ) OR AB ( intraindividual OR intra-individual )
S45  TI ( interparticipant OR inter-participant ) OR AB ( interparticipant OR inter-participant )
TI ( intraparticipant OR intra-participant ) OR AB ( intraparticipant OR intra-participant )

TI ( kappa OR kappa's OR kappas ) OR AB ( kappa OR kappa's OR kappas )

TI repeatab* OR AB repeatab*

TI ( replicab* AND (measure OR measures OR findings OR result OR results OR test OR tests) ) OR AB ( replicab* AND (measure OR measures OR findings OR result OR results OR test OR tests) )

TI ( repeated AND (measure OR measures OR findings OR result OR results OR test OR tests) ) OR AB ( repeated AND (measure OR measures OR findings OR result OR results OR test OR tests) )

TI ( generaliza* OR generalisa* ) OR AB ( generaliza* OR generalisa* )

TI concordance OR AB concordance

TI ( intraclass AND correlation* ) OR AB ( intraclass AND correlation* )

TI discriminative OR AB discriminative

TI "known group" OR AB "known group"

TI ( "factor analysis" OR "factor analyses" ) OR AB ( "factor analysis" OR "factor analyses" )

TI dimension* OR AB dimension*

TI subscale* OR AB subscale*

TI ( multitrait AND scaling AND (analysis or analyses) ) OR AB ( multitrait AND scaling AND (analysis or analyses) )

TI "item discriminant" OR AB "item discriminant"

TI "interscale correlation" OR AB "interscale correlation"

TI ( error OR errors ) OR AB ( error OR errors )

TI "individual variability" OR AB "individual variability"

TI ( variability AND (analysis OR values) ) OR AB ( variability AND (analysis OR values) )

TI ( uncertainty AND (measurement or measuring) ) OR AB ( uncertainty AND (measurement or measuring) )

TI "standard error of measurement" OR AB "standard error of measurement"

TI sensitiv* OR AB sensitiv*

TI responsive* OR AB responsive*

TI ( (minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference) ) OR AB ( (minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference) )
OR clinically) AND (important OR significant OR detectable) AND (change OR difference) )

S70  TI ( small* AND (real OR detectable) AND (change OR difference) ) OR AB ( small* AND (real OR detectable) AND (change OR difference) )

S71  TI "meaningful change" OR AB "meaningful change"

S72  TI "ceiling effect" OR AB "ceiling effect"

S73  TI "floor effect" OR AB "floor effect"

S74  TI "item response model" OR AB "item response model"

S75  TI IRT OR AB IRT

S76  TI ( "differential item functioning" OR DIF ) OR AB ( "differential item functioning" OR DIF )

S77  TI Rasch OR AB Rasch

S78  TI "computer adaptive testing" OR AB "computer adaptive testing"

S79  TI "item bank" OR AB "item bank"

S80  TI "cross-cultural equivalence" OR AB "cross-cultural equivalence"

S81  S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR S62 OR S63 OR S64 OR S65 OR S66 OR S67 OR S68 OR S69 OR S70 OR S71 OR S72 OR S73 OR S74 OR S75 OR S76 OR S77 OR S78 OR S79 OR S80

S82  S42 OR S81

References:

### Appendix 3: Summary of all Exercise Adherence Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Developer</th>
<th>Primary purpose (e.g. adherence, physical activity/ population etc)</th>
<th>Brief description of domains measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Rapid Assessment of Physical Activity (RAPA)</td>
<td>Topolski et al. 2006 [1]</td>
<td>Amount and intensity of physical activity of older adult patients</td>
<td>2 sections: 1 that tests aerobic activities and another for strength and flexibility. 9 yes/ no questions.</td>
</tr>
<tr>
<td>Community Health Activities Model Program for Seniors (CHAMPS)</td>
<td>Stewart et al. 1998 [3]</td>
<td>Types and intensity levels of Physical activity</td>
<td>41 items measuring activities of daily living, work-related, social activities and leisure activities</td>
</tr>
<tr>
<td>Sport Injury Rehabilitation Adherence Scale (SIRAS)</td>
<td>Brewer et al. 2000 [5]</td>
<td>Adherence during clinic-based rehabilitation programmes</td>
<td>3 item measuring 1) Intensity of effort on rehabilitation exercise; 2) Frequency of following practitioner’s instructions and advice and 3) receptivity to changes in the physical therapy programme.</td>
</tr>
<tr>
<td>Short Questionnaire to Assess health enhancing physical activity (SQUASH)</td>
<td>Wendel-Vos et al. 2003 [6]</td>
<td>Walking and bicycling habits</td>
<td>Asks how many days per week walking and cycling activities were performed and how much time on average was engaged in this, and (if applicable) how strenuous this activity was.</td>
</tr>
<tr>
<td>Tegner activity scale</td>
<td>Tegner and</td>
<td>Activity level post Knee</td>
<td>Indicate the HIGHEST level of activity that</td>
</tr>
<tr>
<td>Tool Name</td>
<td>Reference</td>
<td>Type of Measurement</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>The Physical Activity Scale for the Elderly questionnaire (PASE)</td>
<td>Washburn et al. 1993 [8]</td>
<td>Physical activity questionnaire for elderly</td>
<td>Frequency and duration of leisure time, household and work related activity</td>
</tr>
<tr>
<td>Minnesota Leisure Time Physical Activity questionnaire</td>
<td>Periera et al. 1997 [9]</td>
<td>Leisure time activities</td>
<td>2 sections: walking and miscellaneous and conditioning exercise. Yes or No for each activity</td>
</tr>
<tr>
<td>Yale Physical Activity Survey (YPAS)</td>
<td>Depietro et al. 1993 [10]</td>
<td>Physical activity of older adults</td>
<td>Frequency and duration of activities</td>
</tr>
<tr>
<td>Stanford Brief Physical Activity Survey</td>
<td>Taylor-Piliae et al. 2006</td>
<td>Physical activity questionnaire</td>
<td>Physical activity on-the-job and during leisure-time during the past year,</td>
</tr>
<tr>
<td>Canadian Occupational Performance Measure (COPM)</td>
<td>Law et al. 1990 [12]</td>
<td>Individualized outcome measure designed to detect change in a client's self-perception of occupational performance over time.</td>
<td>NA</td>
</tr>
<tr>
<td>London Health and Fitness Questionnaire</td>
<td>Rowland et al. 1994 [14]</td>
<td>Physical activity in older adults</td>
<td>Measured peoples exercise knowledge using a series of positive and negative statements on a 5 point likert scale</td>
</tr>
<tr>
<td>MAARS model -</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Motivation to adopt and maintain regular physical activity</td>
<td>Adherence to Exercise Scale for Older Patients (AESOP)</td>
<td>Self-efficacy expectations, outcome expectations, and outcome expectancies for predicting adherence</td>
<td>42 items measuring 1. Self-efficacy expectations (15 items); 2. Outcome expectations (16 items); 3. Outcome expectancies (11 items)</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>---------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Freiburg Questionnaire of Physical Activity</td>
<td>Frey et al. 1998 [16]</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>PRISCUS Physical Activity Questionnaire</td>
<td>Trampisch et al. 2010 [17]</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Exercise Self-Efficacy</td>
<td>McAuley 1993 [18]</td>
<td>Beliefs in the ability to continue exercise</td>
<td>8 questions that assess beliefs in one’s ability to continue exercising on a three time per week basis at moderate intensities (upper end of one’s perceived exertion range), for 40+ minutes per session in the future.</td>
</tr>
<tr>
<td>Confidence in ability to adhere (adaptation of Lorig’s self-efficacy scale)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>UCLA Activity Score</td>
<td>Zahiri et al.1998 [19]</td>
<td>Current activity level</td>
<td>Check one box out of 10 that best describes activity level</td>
</tr>
<tr>
<td>Home Exercise Compliance assessment (HECA)</td>
<td>NA</td>
<td>Adherence to home exercise</td>
<td>Patients record the number of exercise sessions completed during the previous week.</td>
</tr>
<tr>
<td>International Physical Activity Questionnaires (IPAQ) short form</td>
<td>Craig et al. 2003 [20]</td>
<td>Physical activity in young and middle aged adults</td>
<td>4 generic items tested</td>
</tr>
<tr>
<td>Physical Activity Recall Items</td>
<td>Sallis et al. 1985 [21]</td>
<td>Physical activity</td>
<td>Sleep (2 items) and physical activities (7 items) assessed for the past 7 days</td>
</tr>
<tr>
<td>Attitudes towards ACL rehabilitation questionnaire</td>
<td>Niven et al 2012 [22]</td>
<td>Attitudes and adherence behaviours to a recommended ACL</td>
<td>Assesses intention, attitude, subjective norm, perceived behavioral control, self-efficacy and adherence</td>
</tr>
<tr>
<td>Health Professional Compliance Evaluation</td>
<td>rehabilitation programme</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>--------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habitual Physical Activity Questionnaire</td>
<td>Physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baecke et al. 1982 [23]</td>
<td>29 items concerning the following five components: occupation, movement, sport, leisure time activities excluding sport, and sleeping habits.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correctness of Exercise Performance Scale</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning for Exercise Scale</td>
<td>Commitment to a plan of physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pender 1996 [24]</td>
<td>11 items measuring commitment and strategies to carry out exercise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stages of Exercise Change questionnaire</td>
<td>Intention to change or maintain exercise behaviour</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reed et al. 1997 [25]</td>
<td>Consists of five items representing one of five primary stages of the Trans-theoretical Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hopkins Rehabilitation Engagement Rating Scale</td>
<td>Used in rating behavioural observation during acute in-patient rehabilitation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kortte et al. 2007 [26]</td>
<td>5 items measuring 1.Attendance at rehabilitation session (1); 2.Frequency of required Verbal/Physical Prompts (1); 3.Perceived Positive attitude to exercise (2) ;4.Active participation in rehabilitative exercise (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Physical Activity Scale for Individuals with Physical Disabilities</td>
<td>Physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washburn et al. 2002 [27]</td>
<td>13 questions record the number of days per week and hours per day for participation in leisure time, household, and occupational physical activities over the past 7 days</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Activity Level Index</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>the summed energy expenditure of all reported activities divided by 168, the number of hours per week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthritis self-management behaviour scale</td>
<td>Evidence suggesting that a chronic disease self-management program</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lorig et al 1985 [28]</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerobics Centre Longitudinal Study Physical Activity Questionnaire (ACLS)</td>
<td>Stofan et al 1998 [29]</td>
<td>Leisure and physical activities</td>
<td>10 questions which assess participation in 10 specific exercise related activities within the last 3 months</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Longitudinal Ageing Study Amsterdam Physical Activity Questionnaire (LAPAQ)</td>
<td>Voorips et al. 1991 [31]</td>
<td>Physical activity</td>
<td>Examines the frequency and duration of specific types of activity in the past two weeks</td>
</tr>
<tr>
<td>Stanford Exercise Behavior Scale</td>
<td>Lorig et al. 1996 [33]</td>
<td>Exercise behaviour</td>
<td>6 items which assess amount of exercise activities undertaken during the past week</td>
</tr>
<tr>
<td>Health promoting lifestyle profile</td>
<td>Walket et al. 1987 [34]</td>
<td>Health related questionnaire</td>
<td>48 item measures health promoting behaviours in 6 domains: nutrition, exercise, health responsibility, stress management, interpersonal support, and self-actualization</td>
</tr>
<tr>
<td>Health promoting lifestyle II</td>
<td>Walker &amp; Hill-Polerecky 1996 [35]</td>
<td>Health related questionnaire</td>
<td>52 items in a total scale and six subscales to measure behaviors in the theorized dimensions of health-promoting lifestyle: spiritual growth, interpersonal relations, nutrition, physical activity, health responsibility, and stress management</td>
</tr>
<tr>
<td>Compliance Behaviour Index</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Modified - Rehabilitation Adherence</td>
<td>Shin et al 1988 [36]</td>
<td>Rehabilitation adherence in injured athletes</td>
<td>40 items measuring Self-report inventory with subscales designed to assess 1) perceived</td>
</tr>
</tbody>
</table>
Questionnaire (RAQ-M) | exertion, 2) pain tolerance, 3) self-motivation, 4) support from significant others, 5) scheduling and 6) environmental conditions.
---|---
Rehabilitation Compliance Scale (RCS) | Rheiner 1994 [37] Rehabilitation compliance | Measures patients’ compliance to a rehabilitation program.

Notes: NA= details not available

Reference:


Appendix 4: Evidence of reliability for measures of exercise adherence following completion by patients with MSK problems

<table>
<thead>
<tr>
<th>Measure</th>
<th>Evaluations (n)</th>
<th>Internal consistency reliability</th>
<th>COSMIN</th>
<th>Test-retest reliability (inter-rater; intra-rater; test-retest)</th>
<th>COSMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinician-completed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRERS</td>
<td>1</td>
<td>Poor</td>
<td></td>
<td>Inter-rater agreement</td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 raters: 1 PT and 1 OT. n=206 patients. Assessment taken at a similar time prior to discharge (specifics not reported) ICC=0.73 (Kortte et al. 2007)[1]</td>
<td></td>
</tr>
<tr>
<td>PRPS</td>
<td>1</td>
<td>Poor</td>
<td></td>
<td>Inter-rater agreement</td>
<td>Fair</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total of 5 therapists (3 PT and 2 OT): 2 therapists independently assessed each session: 20 OT sessions and 25 PT sessions. Therapist pairs were masked to each-others scores. ICC for OT and PT ratings: OT: 0.91 PT: 0.96 (Lenze et al. 2004)[2]</td>
<td></td>
</tr>
<tr>
<td>SIRAS</td>
<td>8</td>
<td>Fair</td>
<td></td>
<td>Intra-rater agreement/ Test-retest: 2 raters: 1 treating PT and 1 observing PT. n=28 patients. Re-test period 1 week (stability not reported). Test-retest reliability (weighted kappa): 1) Treating physiotherapist: 0.76 (95% CI 0.61 to 0.90) 2) Observing physiotherapist: 0.63 (95% CI 0.39 to 0.88) (Kolt et al 2007 - study 2)[4]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Poor</td>
<td></td>
<td>Number of raters is unclear; n= 31 patients. Re-test period 1-week (stability unclear). ICC[2,1]=0.77 (Brewer et al. 2000- study 2)[3]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intra-rater agreement: 2 qualified rehabilitation practitioners; n=12 patients. SIRAS completed after four consecutive appointments (Unclear: re-test period, stability). Rater Agreement Indices (RAI): 0.94 (Brewer et al. 2002- study 2)[5]</td>
<td></td>
</tr>
</tbody>
</table>

Cronbach’s alpha 0.91 (Kortte et al. 2007)[1]

Cronbach’s alpha 0.82 (n=145) (Brewer et al. 2000 – Study 1)[3]

Cronbach’s alpha 0.86 for multiple administrations (n= 43) (Brewer et al. 2000 – Study 1)[3]
43 student rehabilitation practitioners completed the SIRAS to rate three vignettes of exercise adherence (highly / moderately / minimally adherent) (stability = set vignettes):

**Rater Agreement Indices (RAI):**
1. High Adherence 0.90
2. Moderate adherence 0.86
3. Low adherence 0.84
4. Aggregate 0.84

(Brewer et al, 2002 – Study 1)[5]

19 raters; n=43 patients. Inter-rater reliability between primary (n=43 assessments) and secondary provider (n= 39 assessments):

ICC=0.57

(Brewer et al 2000 - study 3)[3]

Three video illustrations of ‘exercise adherence’ categorized as i) High adherence; ii) Moderate adherence; iii) Low adherence:

**Inter-rater agreement (Rater Agreement Indices)** n= 60 raters:

i) 0.93; ii) 0.87; iii) 0.92

(Kolt et al. 2007 - study 1)[4]

2 raters: 1 treating PT and 1 observing PT. n= 28 patients. Re-test period 1 week.

**Inter-rater agreement (weighted kappa) assessed at:**

i) first clinical session 0.76 (95% CI 0.61 to 0.90)
ii) second clinical session 0.89 (95% CI 0.82 to 0.97)

(Kolt et al. 2007- study 2)[4]

---

<table>
<thead>
<tr>
<th>AESOP</th>
<th>Test-retest (interview completion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n=28 patients. Re-test period 2-weeks (stability not reported). AESOP domains ICC (3,1):</td>
</tr>
<tr>
<td></td>
<td>1. Self-efficacy expectations 0.796</td>
</tr>
<tr>
<td></td>
<td>Outcome expectations 0.771</td>
</tr>
<tr>
<td></td>
<td>3. Outcome expectancies 0.328</td>
</tr>
<tr>
<td></td>
<td>(Hardage et al. 2007)[6]</td>
</tr>
</tbody>
</table>

---

Patient-completed

Fair

Fair

Poor

Poor
| CHAMPS     | 1 | Test-retest (self-completion)  
| n=173 patients. Re-test period 6 months; (stability expected - non-intervention or control group)  
| ICC (2,1):  
| A. Moderate and greater intensity  
| Caloric expenditure 0.67  
| Frequency per week 0.58  
| B. All activities  
| Caloric expenditure 0.66  
| Frequency per week 0.62  
| (Stewart et al. 2001)[7] | Poor |
| RAQ-M 1 | Cronbach’s alpha (n=120):  
| range –  
| Perceived exertion 0.66  
| Pain tolerance 0.79  
| Environmental conditions 0.79  
| Support from significant others 0.82  
| Self-motivation 0.83  
| Scheduling 0.87  
| (Shin et al, 2010)[8] | Poor |
| ROAQ 2 | Cronbach’s alpha (n=118):  
| Ignoring practitioner recommendations 0.83  
| Attempting an expedited rehabilitation 0.70  
| (Podlog et al 2013 – study 1)[8]  
| Cronbach’s alpha (n=105):  
| Ignoring practitioner recommendations 0.86  
| Attempting an expedited rehabilitation 0.75  
| (Podlog et al 2013 – study 2)[8] | Fair |
(n)= number of studies evaluating the measurement and practical properties of each measure, PT=Physical Therapist, OT= Occupational Therapist, ICC=Intraclass Correlation Coefficient, CI= Confidence Interval, BMI= Body Mass Index, LBF= Lower Body Functioning, 6MW= 6 Minute Walk, SRPF= Self-Reported Physical Functioning, SREF= Self-Reported Energy/Fatigue, SRP= Self-Reported Pain, SRPWB= Self-Reported Psychological Well-Being, 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, PAF=Principal Axis Factoring, CFA=Confirmatory Factor Analysis, RAOQ – Rehabilitation Over-adherence Questionaire

Reference:

Appendix 5: Evidence of validity for measures of exercise adherence following completion by patients with MSK problems

<table>
<thead>
<tr>
<th>Measure</th>
<th>Number of evaluations (n)</th>
<th>Known Groups Validity (hypothesis – stated / deduced / not reported?)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>COSMIN</th>
<th>Construct Validity (structural; construct - divergent / convergent; hypothesis – stated / deduced / not reported?)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>COSMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinician-completed HRERS</td>
<td>1</td>
<td>Groups defined by: Racial differences (White or Other): N/S Gender (male or female): N/S Diagnostic groups (spinal cord injury (SCI), Stroke, Amputation, Orthopaedic): N/S (Kortte et al. 2007)&lt;sup&gt;1&lt;/sup&gt; Relationship between HRERS (three categories) and clinical variables ‘hypothesized to be associated with ‘engagement’ (but direction not stated) Functional Impact Measure (FIM) efficiency Number of total absences Number of refusals Number of non-refusal absences</td>
<td>Fair</td>
<td>Structural validity Factor structure: hypothesized uni-dimensional structure (‘engagement’) supported by principal component factor analysis (explored for each diagnostic group) (n=206). HRERS with clinical variables (hypothesized association between variables not stated but ‘supported’) Functional Impact Measure (FIM): r= 0.20 Brief Symptom Inventory (depression): r= 0.24 Levine’s Denial of Illness questionnaire: r= 0.30 Positive and Affective Negative State (PANAS): PANAS self-rated negative effect r= 0.23 PANAS self-rated positive effect r= 0.36 Craig Handicap Assessment and Reporting Technique (CHART) - level of functioning at 3 months post discharge: r= 0.22</td>
<td>Fair</td>
</tr>
<tr>
<td>HRERS (mean)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRERS (mean)</td>
<td></td>
<td>FIM efficiency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20 / 20-25 / &gt;25</td>
<td>1.25 / 1.87 / 2.03 (p= 0.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRERS (mean)</td>
<td></td>
<td>Total absence rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20 / 20-25 / &gt;25</td>
<td>28 / 15 / 9 (p &lt;0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRERS (mean)</td>
<td></td>
<td>Therapy refusal rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20 / 20-25 / &gt;25</td>
<td>14 / 7 / 2 (p &lt;0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRERS (mean)</td>
<td></td>
<td>Therapy non-refusal absence rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20 / 20-25 / &gt;25</td>
<td>14 / 9 / 7 (p&lt;0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Kortte et al. 2007) [1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| PRPS | 1 | PRPS with (hypothesized association not stated):  
korFunctional Independence Measure-Motor (FIM-motor)  
At treatment admission \( r = 0.38 \)  
At treatment discharge: PRPS with change in FIM-motor \( r = 0.32 \) (authors suggest the result supports hypothesized association – but this is not explicit)  
Gender \( r = -0.05 \)  
Length of stay (LOS) \( r = -0.13 \)  
Age \( r = -0.21 \)  
Race \( r = -0.01 \)  
Medical co-morbidity (count) \( r = -0.03 \) (Lenze et al. 2004)[2] | Poor |
| SIRAS | 7 | Three video evaluations describing three levels of adherence: 1) High adherence; 2) Moderate adherence; 3) Low adherence. n=60 assessments. (hypothesized associations not reported, but can be assumed): Mean (SD) scores higher for 1) High adherence (mean 13.53 (1.51)) versus 2) Moderate adherence (mean 8.02 (1.95) versus 3) Low adherence (mean 4.59 (1.57)). Statistical significance of group differences not reported. (Kolt et al, 2007)[3] | Poor | Structural validity Principal component analysis (PCA) supported single factor structure following completion by physiotherapy students (n=60): each student completed the SIRAS for a hypothetical patient across high, medium and low adherence conditions. (hypothesized structure not proposed) (Kolt et al, 2007 – study 1)[3]PCA carried out for two assessors at two sessions supported the single factor structure. (Kolt et al, 2007 - study 2)[3]PCA supported the hypothesized single factor structure (single factor 74% of variation) (eigenvalue 2.21) (n= 145) (Brewer et al, 2000 – study 1)[4] | Poor | Structural validity Principal component analysis (PCA) supported single factor structure following completion by physiotherapy students (n=60): each student completed the SIRAS for a hypothetical patient across high, medium and low adherence conditions. (hypothesized structure not proposed) (Kolt et al, 2007 – study 1)[3]PCA carried out for two assessors at two sessions supported the single factor structure. (Kolt et al, 2007 - study 2)[3]PCA supported the hypothesized single factor structure (single factor 74% of variation) (eigenvalue 2.21) (n= 145) (Brewer et al, 2000 – study 1)[4] |
| AESOP | 1 | Patients with 'low scores' on self-efficacy (n=24) and 'outcome expectations' (n=16) domains 'adhered' to the exercise regime; those with high scores (greater than the mean value)(n=8) did not adhere. However, the external marker for 'adherence' is not clarified. (Hardage et al, 2007)[6] | Poor | AESOP with SF-12 (version 2) (spearman correlations) (hypothesized association not stated): SF-12 Physical Component Score (PCS): Self-efficacy r=0.13 Outcome expectations r=-0.01 Outcome expectancies r=-0.04 SF-12 Mental Component Score (MCS): | Poor | AESOP with SF-12 (version 2) (spearman correlations) (hypothesized association not stated): SF-12 Physical Component Score (PCS): Self-efficacy r=0.13 Outcome expectations r=-0.01 Outcome expectancies r=-0.04 SF-12 Mental Component Score (MCS): |
| CHAMPS | Poor | Self-efficacy $r=0.01$  
Outcome expectations $r=-0.06$  
Outcome expectancies $r=-0.09$  
(Hardage et al. 2007)[6] | Poor |
|---|---|---|---|
| Three groups defined by known activity levels (via a detailed exploration of self-reported activities) (n=249):  
1. Not participating in any exercise or recreational sports. (inactive/initially sedentary)  
2. Participating in some exercise or recreational sports (according to ACSM criteria) (underactive).  
3. Participating in activities at levels that met ACSM guidelines (active)  
Hypothesis: levels of physical activity on the CHAMPS would be lowest for the least active.  
As hypothesised, the inactive group had statistically significant lower CHAMPS scores (all four values) when compared to the underactive and active groups ($p<0.001$).  
(Stewart et al. 2001)[7] | CHAMPS with several health measures (stated hypotheses and supported) (n=249). CHAMPS scores (code):  
A. Moderate and greater intensity  
A1. Caloric expenditure  
A2. Frequency per week  
B. All activities  
B1. Caloric expenditure  
B2. Frequency per week  
BMI: range -0.04 (B1) to -0.06 (A1)  
Short Physical Performance Battery (lower body functioning): range 0.15 (B2) to 0.28 (A1)  
6-minute walk: range 0.10 (B2) to 0.27 (A1)  
Short-Form 36-item Health Survey (SF-36) - four domains:  
Physical function: range 0.23 (B2) to 0.30 (A1,A2)  
Vitality: range 0.10 (B2) to 0.23 (A2)  
Body pain: range 0.08 (B2) to 0.17 (A2)  
Emotional well-being (range 0.02 (B2) to 0.14 (A2)  
(Stewart et al. 2001)[7] | Poor |
| RAQ-M | Poor | Structural validity  
Factor structure: six-domain structure informed by exploratory (n=102) and then confirmatory (n=120) factor analysis. Hypothesised structure not proposed/defined.  
RAQ-M with measures of adherence (hypothesized association not stated): ‘range 0.27 to -0.63’.  
Patient attendance at rehab sessions – specific | Poor |
| ROAQ | 2 | Structural validity  
Hypothesized a priori 2 factor structure:  
1) Ignoring practitioner recommendations and 2) Attempting an expedited rehabilitation  
Study 1: Following completion of the long-form 19-item measure by injured adolescent athletes (n=118): Principal Axis Factoring (PAF) supported a 2-factor, 10-item solution (2 factors with eigenvalues >1.0; explaining 53.17% of variance) (Podlog et al, 2013 – Study 1)[9]  
Study 2: Following completion of the 10-item measure by injured collegiate athletes (n= 105): Confirmatory Factor Analysis (CFA) supported the 2-factor structure (Podlog et al, 2013)[9]  
Correlation between the two domains r= 0.49 (Study 1 n= 118) and r= 0.58 (Study 2 n= 105).  
Construct Validity  
Study 1 (n= 118)(Podlog et al, 2013)[9]  
ROAQ domains with clinical variables |
<table>
<thead>
<tr>
<th>Domain 1: Ignoring Practitioner Recommendation with:</th>
<th>Subscale 2: Attempting an Expedited Return with:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Athletic Identity (AIMS): r = 0.23</td>
<td>Athletic Identity (AIMS): r = 0.46</td>
</tr>
<tr>
<td>Appearing Athletically Untalented (SPSQ subscale): r=0.29</td>
<td>Appearing Athletically Untalented (SPSQ subscale): r=0.18</td>
</tr>
<tr>
<td>Concerns about physical appearance (SPSQ subscale): r= 0.20</td>
<td>Mental composure inadequacies (SPSQ subscale): r=0.31</td>
</tr>
<tr>
<td>Appearing fatigued (SPSQ subscale): r=0.28</td>
<td>Concerns about Physical Appearance (SPSQ subscale): r=0.36</td>
</tr>
</tbody>
</table>

*Subscale 1: Ignoring Practitioner*

<table>
<thead>
<tr>
<th>Subscale 2: Attempting an Expedited Return with:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concerns about Physical Appearance (SPSQ subscale): r=0.26</td>
</tr>
</tbody>
</table>

*Study 2 (n= 105) (Podlog et al, 2013)[9]*

**Fair**

---

*a Hypothesis testing: 'hyp deduced' - hypothesis of association between variables can be deduced; 'No hyp' - hypothesis of association between variables not stated and cannot be deduced from the article text; b= At the end of patients’ first session, the treating Physiotherapist (PT) completed.*
the Sport Injury Rehabilitation Adherence Scale (SIRAS: Brewer et al., 2002) prediction form (n = 169) while SIRAS assessment form was completed at the 6th (n = 100), 12th and last rehabilitation sessions where applicable (n = 23).

n= number of studies evaluating the measurement and practical properties of each measure, N/S= Not Significant, SD= Standard deviation, r= correlation coefficient, BMI= Body Mass Index, LBF= Lower Body Functioning, 6MW= 6 Minute Walk, SF-12=Short-Form 12-item Health Survey, SRPF= Self-Reported Physical Functioning, SREF= Self-Reported Energy/Fatigue, SRP= Self-Reported Pain, SRPWB= Self-Reported Psychological Well-Being, 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, ACSM= American College of Sports Medicine, PAF=Principal Axis Factoring, CFA=Confirmatory Factor Analysis, SPSQ=Self-Presentation in Sport Questionnaire, AIMS=Athletic Identity Measurement Scale, ROAQ=Rehabilitation Over-adherence Questionnaire, I-PRRR=Modified Injury Psychological Readiness to Return to Sport Scale

Reference:
