Outcome measures evaluating physical functioning and their measurement properties in adolescent idiopathic scoliosis: a protocol for a systematic review

Samia Alamrani,1,2 Alison Rushton,1 Adrian Gardner,3 Deborah Falla1,1 Nicola R Heneghan1

ABSTRACT

Introduction Physical functioning (PF) is the ability to carry out the physical activity of daily living. It is an important outcome that provides a meaningful evaluation of individuals’ life. PF can be assessed using patient-reported outcome measures, performance-based outcome measures or body structure and function measure. Measures need to be valid, reliable and responsive to change to evaluate the effects of an intervention. Adolescent idiopathic scoliosis (AIS) is the most common deformity among the paediatric population and impacts on individuals’ lives. This systematic review will appraise evidence on the measurement properties of PF tools in individuals with AIS.

Methods/analysis A protocol for systematic review and meta-analysis informed by Cochrane guidelines is reported in line with Preferred Reporting Items for Systematic Reviews and Meta-Analysis-P. MEDLINE, PsycINFO, EMBASE, CINAHL, SPORTDiscus, Web of Science and PubMed will be searched in two stages, from inception until December 2019. Search 1 will inventory all studies that assessed PF in participants with AIS, without any limitations. The search terms will be scoliosis, adolescent and PF-related terms. Search 2 will include studies which investigated instrument measurement properties in the same population for measures identified in search one. Two reviewers will independently perform study selection, data extraction, risk of bias and overall quality assessment. The COnsensus-Based Standards for the selection of health Measurement INstruments (COSMIN) risk of bias tool and modified Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines will be used. A meta-analysis will be conducted if possible, or the evidence will be synthesised and summarised per measurement property per outcome measure per measurement type.

Ethics and dissemination This review will provide recommendations for practice and future research, considering psychometric properties of outcome measures of PF in AIS. The results of this study will be disseminated through a peer-reviewed publication and conference presentation.

PROSPERO registration number CRD42019142335.

Strengths and limitations of this study

- This review will synthesise evidence of patient-reported, performance-based or body structure and function outcome measures of physical functioning (PF) for use in practice or research involving individuals with adolescent idiopathic scoliosis (AIS).
- The search strategy of this review comprises two stages. The first stage will retrieve all studies that assessed PF in individuals with AIS, while the second stage will retrieve studies that investigated measurement properties of the instrument identified in the first search.
- This study will employ rigorous methods and uses COnsensus-Based Standards for the selection of health Measurement INstruments (COSMIN) risk of bias tool and modified Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach.
- This review will be limited to studies of the English language that assess measurement properties among adolescents with idiopathic scoliosis.

INTRODUCTION

Scoliosis is a complex three-dimensional deformity of the spine, including lateral curvature and rotation of the vertebrae, and characterised by a curve angle ≥10°. There are two main types of scoliosis, idiopathic and non-idiopathic with the latter arising from congenital, neuromuscular or mesenchymal causes. While the aetiology of idiopathic scoliosis remains unknown, genetic, hormonal and mechanical factors are involved. Adolescent idiopathic scoliosis (AIS) often develops between 10 and 16 years of age and represents ~85% of cases. AIS is the most common spinal deformity among the paediatric population, with a prevalence ranging from 2% to 3%. Nearly 80% of those affected presents with a curvature of the...
thoracic or thoracolumbar/lumbar region. While men and women are equally affected, women are reported to be at 10 times greater risk of curve progression.

A number of health-related problems are reported among individuals with AIS including lower quality of life, back pain, pulmonary dysfunction, stress and mental health disorders. A major component of health status and health-related quality of life is physical functioning (PF), which can be used to identify individuals at risk of disability and to predict health and social care use. Accordingly, PF is included in the core outcome set (COS) for use within clinical trials for many musculoskeletal conditions, including adolescents with spinal deformity. Where the COS study includes all types of spinal deformity, there is now need for a more specific systematic review of PF outcome measures for this unique population subset. Limitations in PF are reported by individuals with AIS, for example walking, moving around and maintaining body position. Additionally, pain is often reported in individuals with AIS, which may cause functional limitations.

PF can be assessed with patient-reported outcome measures (PROMs), performance-based outcome measures (PBOMs) or a measure of body structure and function. The most widely used PROM for assessment of the quality of life as well as PF of individuals with spinal deformity is the Scoliosis Research Society (SRS) questionnaire and its variants. The SRS is mostly used among surgically treated individuals with AIS, but may not be applicable to those treated conservatively. Although relevant, PROMs should be used cautiously, as it influenced by patients’ perception of their abilities to perform activities and lack sensitivity to change. Measures such as PBOMs have the potential to provide unbiased and reproducible assessments of PF during the performance of activities of daily living, such as walking speed and trunk endurance testing. Within individuals with AIS, little is known about the available PBOMs for evaluating PF. The body structure and function measures such as radiographs can give an indication about dysfunctions in structure but fail to fully capture functional limitations.

The SRS-22r questionnaire is the reference standard outcome measure of quality of life, which include PF items as recommended by the recent COS study for adolescents and young adults with spinal deformity. However, the SRS-22r fails to fully capture important aspects of PF for individuals with AIS, for example, self-care and mobility. The COS study included all forms of spinal deformities, the heterogeneity limits applicability to individuals with AIS as a discrete population.

Measurement properties of outcome measures are important to avoid the risk of bias and ensure accuracy in the evaluation of test results. The COmensus-based Standards for the selection of health Measurement INstruments (COSMIN) group developed a taxonomy of measurement properties to improve the selection of outcome measures. Three main domains identified reliability, validity and responsiveness with further subgrouping. A systematic review is needed to evaluate the measurement properties of PF outcome measure for individuals with AIS. Review findings will inform clinicians and researchers on the best available tools for the assessment of PF in AIS. Furthermore, findings will inform future research drawing on a range of measures of PF to investigate health status in AIS.

**Objective**

To identify outcome measures used to assess PF in individuals with AIS. A secondary aim is to evaluate the measurement properties of PF outcome measures in AIS.

**METHODS**

This protocol has been informed by experts in musculoskeletal orthopaedics including a consultant spine surgeon, musculoskeletal rehabilitation experts including physiotherapists and individuals with review, measurement properties and research experience. It has been designed in line with the COSMIN methodology for systematic reviews of PROMs and is reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis-P (PRISMA-P). The search for this systematic review will be conducted in two parts. Stage 1 to identify studies used an outcome measure to evaluate PF in individuals with AIS. This search will allow a list of outcome measures to be generated. Stage 2 will identify studies, which evaluated measurement properties of PF outcome measure identified in the first search.

**Search 1: inventory of outcome measures**

**Eligibility criteria**

**Study design**

All study designs including randomised clinical trials, cohort, observational studies and case studies will be included to identify all outcome measures of PF being used with individuals with AIS. No limitation on language or location will be applied at this stage.

**Participants**

Participants aged between 10 years and 18 years of age, with a diagnosis of idiopathic scoliosis and ≥10° Cobb angle, will be considered. No restrictions will be applied to the curve severity, evaluation settings and the type of treatment.

**Outcome**

Any study that includes assessments of the PF of AIS using a specific outcome measure will be included. PF is defined according to the Core Outcome Measures in Effectiveness Trials (COMET) taxonomy, as any physical activities of daily living such as the ability to walk, independence, self-care, performance status and disability index. The outcome measures are defined as any one of the following:

1. PROMs in the form of questionnaires or scales designed for AIS to evaluate PF or if it is included as a subscale within a questionnaire.

2. PBOMs a measure of PF by clinician while the individual is performing a functional task, for example, walking and/or
3. Body structure and function measures which means any dysfunction in a specific body part or system which may limit function such as range of motion.²⁶

Search strategy
A comprehensive, systematic and reproducible search strategy will be completed by one reviewer (SA). Databases will be searched to identify studies that assessed PF among individuals with AIS. To ensure that all relevant studies are included, the type of the outcome measure will not be specified at this stage (figure 1). Initial search terms will be developed for MEDLINE and then adapted with relevant syntax and subject headings for the other databases. An example of the search 1 strategy is available as an online supplementary file 1. As a result of this search, a list of outcome measures for PF used in AIS will be generated. Then, the outcome measures will be classified as, that is, PROM, PBOM or measure of body structure and function. The list will then be used to perform the search 2.

Search 2: measurement properties of outcome measures
Eligibility criteria
Study design
Any study that has evaluated one or more measurement properties of the identified outcome measures in search 1 will be eligible. Only full-text studies available in English will be included. Conference abstracts will be excluded due to the inability to effectively evaluate the quality of the study.

Participants
Participants aged between 10 and 18 years of age, with a diagnosis of idiopathic scoliosis and ≥10° Cobb angle, will be eligible. In studies with mixed cohorts, >50% of participants should be individuals with AIS for the study to be included. Authors of studies will be contacted in case of missing information about number of participants with AIS. Studies without original participant data (eg, systematic review) will be excluded.

Outcome
The outcomes of interest are the measurement properties: reliability including (internal consistency, test–retest, inter-rater and intrarater), measurement error, validity
(content validity, structural validity or criterion validity), hypothesis testing and responsiveness of the outcome measures identified in the search 1 will be eligible. Studies that provide indirect evidence on the measurement properties (by testing an alternative test against an outcome measure of interest and studies in which the outcome measure is used to measure an outcome) will be excluded. Also, studies that only provide normative data will be excluded.

Search strategy
Using the list of outcome measures determined from search 1, one reviewer (SA) will conduct the search. Each category of outcome measure will be searched separately. The search terms will be consisting of the name of the outcome measure/s, the AIS and the measurement properties (figure 1). The recommended search filters specifically designed for retrieving articles on measurement properties will be adapted and used at this stage.33 An example of the search 2 strategy is available as an online supplementary file 1.

Information sources
The electronic search of databases will be conducted including MEDLINE (1946–November 2019), PsycINFO (1967–December 2019) and EMBASE (1974–December 2019) through OVID interface, CINAHL (1937–December 2019) and SPORTdiscus (1800–December 2019) through EBSCO interface, Web of Science (1900–December 2019) and PubMed (1997–December 2019). No language limitations will be applied in search 1; however, the search 2 will be limited to the full-text article in English. The Web of Science database will be searched for conference proceedings for the last 5 years for search 1 only. A hand search in the key journals including Spine, the Spine Journal, Spine Deformity, Scoliosis and Spinal Disorders and European Spine Journal as well as contacting relevant leading researchers in the field. Further, searching for the grey literature, including British National Bibliography for report literature, open grey, dissertation abstracts and Electronic Thesis Online Service (EThOS) will be conducted.

Data management
Search records will be imported into Endnote V.X9 (Clarivate Analytics). Using Endnote, the abstracts and full texts will be stored. The duplicates will be identified through the Endnote software and exact duplicates will be removed.

Selection process
A standardised eligibility assessment will be performed by two independent reviewers (SA, Elena Bini (EB)). All studies identified by the search strategy will be assessed based on title/abstract for eligibility. If there is insufficient information to include/exclude study, full text will be retrieved and then screened for eligibility. The study selection (included and excluded studies) with the reasons for exclusion will be summarised in a PRISMA flow diagram.31 Articles will be included if both reviewers agreed that the eligibility criteria were met. Any disagreement will be first discussed and the third reviewer (NRH) will mediate situations of disagreement. At each assessment stage, agreement between reviewers will be estimated with percentage of agreement and the κ statistic using SPSS for Windows statistical software package (IBM SPSS Statistics V.25).

Data collection process
Two reviewers (SA, EB) will independently extract the data of eligible studies. A bespoke data extraction form will be used and piloted on three studies. Any disagreement between reviewers will be mediated through discussion with a third reviewer (NRH) if needed. If information is not clear or unavailable in the studies, corresponding authors will be contacted to request further details. A second and final reminder will then each be sent 2 weeks apart.

Data items
The data that will be extracted from each study at each stage is summarised in table 1. In the case of missing data, the authors of the study will be contacted.

Outcomes and prioritisation
The gold standard and the primary outcome measure for evaluation of body structure and function (eg, spinal curvature) are the radiographs using the Cobb method.2 However, no primary PROM or PBOM of PF for individuals with AIS, can be identified for this review.

Risk of bias in individual studies
The COSMIN checklist for assessment of risk of bias and methodological quality in individual studies will be used.28 It was revised and specifically designed for use in systematic reviews of PROMs to evaluate studies on measurement properties.34 The methodological quality of each study for each measurement property will be assessed separately.30 The items for each measurement property in the relevant standards box will be rated as either very good, adequate, doubtful or inadequate quality.30 Then, the overall methodological quality of the measurement property will be rated based on ‘the worst score counts principle’, that is, the overall quality of the study for a specific measurement property is based on the lowest rating of any items in the standards’ box.30 The result of each item and overall rating will be reported in the final results. The COSMIN group recommend researchers to adapt the checklist to other measures (ie, PBOMs, body structure and function measure) since it was originally developed for PROMs.30 Two independent reviewers (SA, EB) will assess the risk of bias for all included studies. Any disagreement will be resolved through discussion, and if no agreement is reached a third reviewer (NRH) will be consulted. The agreement between reviewers will be estimated with percentage agreement and the κ statistic using SPSS for Windows statistical software package (IBM SPSS Statistics V.25) and will be reported in the final results.
The COSMIN guidelines for systematic reviews will be followed for the synthesis of the results. Characteristics of the outcome measures, sample, measurement properties results, information about interpretability and feasibility of the scores of the included outcome measures will be presented in overview tables for each outcome measure. Each measurement property for each study per tool will be rated against the updated criteria for good measurement property per outcome measure in each category. Using the modified (GRADE) approach, each measure of measurement properties will be obtained by calculating weighted means and 95% CI. If deemed not possible to pool the results, a qualitative synthesis will be conducted, for example the percentage of confirmed hypotheses for construct validity will be provided. The pooled or summarised evidence will be rated as sufficient when at least 75% of the results met the criteria. For example, for structural validity, ‘at least 75% of the confirmatory factor analysis studies should found the same factor structure’. The recommendation of an outcome measure will be depending on the measurement properties, as well as interpretability and feasibility results. The tool should have sufficient content validity and at least low-quality evidence for sufficient internal consistency to be recommended for use and the results of this tool is trustworthy.

### Data synthesis

The COSMIN guidelines for systematic reviews will be followed for the synthesis of the results. Characteristics of the outcome measures, sample, measurement properties results, information about interpretability and feasibility of the scores of the included outcome measures will be presented in overview tables for each outcome measure. Each measurement property for each study per tool will be rated against the updated criteria for good measurement property as either sufficient (+), insufficient (–) or indeterminate (?). The result of rating of measurement property and its methodological quality rating will be added to the overview table. Then, the evidence will be pooled or summarised per measurement property per tool, with the overall result will be rated against the criteria for good measurement properties, and the quality of the evidence will be graded using a modified Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach.

The results on measurement properties from different studies will be pooled in a meta-analysis if there is enough clinical and methodological homogeneity. The data will be statistically pooled when: (1) individuals with AIS displayed similar characteristics in terms of curve severity and intervention; (2) similar base-line score; (3) same time interval; and (4) same statistical parameters. If inconsistent results of measurement properties were present due to different subgroups (ie, mild and severe curve), the consistent results will be separately summarised per subgroup. Pooled estimate of measurement properties will be obtained by calculating weighted means and 95% CI. If deemed not possible to pool the results, a qualitative synthesis will be conducted, for example the percentage of confirmed hypotheses for construct validity will be provided. The pooled or summarised evidence will be rated as sufficient when at least 75% of the results met the criteria. For example, for structural validity, ‘at least 75% of the confirmatory factor analysis studies should found the same factor structure’.

### Confidence in cumulative evidence

Two independent reviewers will assess the quality and strength of evidence for the pooled or summarised result. Using the modified (GRADE) approach, each measurement property per outcome measure in each category will be evaluated. The GRADE approach uses five factors to determine the quality of the evidence: risk of bias (quality of the studies), inconsistency (of the results of the studies), indirectness (evidence comes from different populations, interventions or outcomes than the ones included in the review), imprecision, and publication bias. The strength of evidence is graded as very low, low, moderate, high, or very high. The reviewers will also assess the quality of the evidence for each outcome measure using the modified (GRADE) approach.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Summary of items to be extracted from included studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study and participants characteristics</strong></td>
<td>Reference, year, country, design of study, age, gender, sample size (used in the analysis), curve size, curve type, type of intervention (bracing, physiotherapy, exercise or surgery)</td>
</tr>
<tr>
<td><strong>Outcome measure</strong></td>
<td>PROM: Name of outcome measure, means of scores (SD), mode of administration, recall period, subscale, number of items, response option, response rate, missing items, setting, target population, scoring, original language, available translation PBOM: Name of outcome measure, equipment needed, number of assessments, outcome (eg, time needed, ability/disability), setting, scoring Body structure and function measure: Name of outcome measure, equipment needed, mode of administration, setting, scoring, outcome (eg, time needed, ability/disability)</td>
</tr>
<tr>
<td><strong>Measurement properties</strong></td>
<td>Validity: Name of outcome measure, type of validity, descriptive statistics, missing value, comparator outcome or predictor outcome, hypothesis, statistics method, CI, validation results Reliability: Name of outcome measure, type of reliability, descriptive statistic, time interval, reliability coefficient, measurement error Responsiveness: Name of outcome measure, Method of testing: Hypothesis testing, Distribution based method (ES, SRM and MDC), hypothesis, time to follow-up. Anchor-based methods (MIC or MCIC or MID), anchor/s. Interpretability: Name of outcome measure, distribution of score in the study population, percentage of missing items, floor and ceiling effects, scores and change scores available for relevant (sub)groups, MIC Or MID, information on response shift Feasibility: Patient’s comprehensibility, Clinician’s comprehensibility, Type of administration, Length of instrument, Completion time, Patient’s required mental and physical ability level, Ease of standardisation, Ease of score calculation, Copyright, Cost of an instrument, Required equipment, Availability in different settings, Regulatory agency’s requirement for approval</td>
</tr>
</tbody>
</table>

ES, effects size; MCIC, minimal clinically important change; MDC, minimal detectable change; MIC, minimal important change; MID, minimal important difference; PBOM, performance-based outcome measure; PROM, patient-reported outcome measures; SRM, standardised response mean.
of interest in the review), imprecision (wide confidence intervals) and publication bias (negative results are less often published). For evaluating measurement properties in systematic reviews of PROMs, only four factors will be assessed as recommended by COSMIN group, while the fifth factor (publication bias) will be removed as there is no registry exists for measurement properties.

**DISCUSSION**

PF is considered as an important outcome domain in health-related quality of life. It can be used to predict future disability as well as health and social care use. Individuals with AIS reported a limitation in their PF. Thus, measurement of its impact is important in research and clinical practice. Numerous of tools are available for the assessment of PF, ranging from patient-reported to PBOMs. However, it is essential to confirm the psychometric properties of these tools before recommending for clinical use. The COS study for ‘all spine deformities’ identified the SRS-22r as the recommended PROM for assessment for PF among young adults with spinal deformities. However, there is still a need for a more specific review that evaluate the quality of all outcome measures used in the assessment of PF in AIS including patient-reported and performance-based as well as measures of body structure and function. This systematic review will retrieve all tools that have been used to assess PF among individuals with AIS. Then, it will evaluate and synthesise the quality of studies that report psychometric properties of PF outcome measures in AIS. This review will provide a comprehensive assessment of current evidence which may benefit: (1) health practitioners in selection of the most suitable tools to assess PF in AIS; (2) patients who need a good outcome measures that reflect their actual health status; and (3) researchers and policy maker who can use the recommend measures in designing research trials and defining the COS for individuals with AIS, which in turn will improve health assessment and patient care. Limitations of this review are a focus on individuals with AIS specifically, so recommendations cannot be generalised to other forms of scoliosis.

**Patient and public involvement**

A study question and systematic review protocol were informed following discussion at a patient and public involvement meeting at the Centre of Precision Rehabilitation for Spinal Pain at the University of Birmingham. The group consisted of individuals with different musculoskeletal and spinal complaints. They actively contributed to research question and to establish the need for systematic review. Since no patient data is needed, patients will not be involved in data collection or analysis. However, the results of the study will be shared at public engagement events.

**Implications of this study**

AIS is a complex deformity of the spine and causes a significant impact on physical activities of individuals’ daily living such as walking and maintaining body position. Therefore, PF gives an indication about the current health status and identifies people at risk of disability. Therefore, PF is considered as one of the outcomes that should be assessed and reported in clinical trials of musculoskeletal conditions. A systematic review is needed to evaluate current practice in the assessment of PF among individuals with AIS. The results of this review will inform clinicians and researchers on the best available tools for assessment of PF in AIS. This review could provide a research agenda that may highlight the gap in the literature around PF measure and their measurement properties among individuals with AIS.

**Ethics and dissemination**

No ethics approval is required for this systematic review. The results of this systematic review will be disseminated through peer-reviewed journals as well as international and national conferences presentation. The publications will be split into different publications according to the volume of data. Each category of outcome measures will be published in a separate article.

**Twitter** Alison Rushton @abrushton, Deborah Falla @Deb_Falla and Nicola R Heneghan @HeneghanNicola

**Contributors** All authors conceptualised and designed the protocol. SA is a PhD student and NRH (lead supervisor), AR and DF are supervisors and AG is a spinal surgeon. SA drafted the initial manuscript with NRH, AR, DF and AG providing guidance on design, topic, methodology and analyses. All authors reviewed and commented on each draft of the protocol. All authors have approved and contributed to the final manuscript.

**Funding** SA is a PhD student, supported by a scholarship form University of Tabuk, Tabuk, Saudi Arabia.

**Competing interests** None declared.

**Patient consent for publication** Patients and/or the public were involved in the design, conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**

Alison Rushton http://orcid.org/0000-0001-8114-7669
Deborah Falla http://orcid.org/0000-0003-1689-6190
Nicola R Heneghan http://orcid.org/0000-0001-7599-3674

**REFERENCES**


