Manual Therapy 22 (2016) 22–30

Contents lists available at ScienceDirect

Manual Therapy

journal homepage: www.elsevier.com/math

Review article

Patient-Reported Outcome (PRO) questionnaires for people with pain in any spine region. A systematic review

Edmund Leahy [a, b, *], Megan Davidson [a], Deenika Benjamin [a], Henry Wajswelner [a]

[a] Department of Rehabilitation, Nutrition and Sport, School of Allied Health, Level 5, HS3, La Trobe University, Bundoora, Vic 3086, Australia
[b] Physiotherapy Department, Northern Health, 185 Cooper St, Epping, Vic 3076, Australia

ARTICLE INFO

Article history:
Received 6 July 2015
Received in revised form
21 October 2015
Accepted 22 October 2015

Keywords:
Spine
Patient-reported
Outcome
Measurement properties

ABSTRACT

Background/objective: This systematic review investigates the measurement properties of Patient-Reported Outcome (PRO) questionnaires which evaluate disability associated with pain in any area of the spine.
Method: PRO questionnaires for people with pain in any spinal region were identified from existing systematic reviews and recent studies. Databases were searched for studies which evaluated the measurement properties of the included questionnaires to August 2015. Data synthesis used a levels of evidence approach which considered study methodological quality.
Results: The Extended Aberdeen Back Pain Scale (EA), Functional Rating Index (FRI) and Spine Functional Index (SFI) were identified as eligible for this review. The FRI was evaluated in 15 studies, with positive results for internal consistency, structural validity, hypothesis testing and responsiveness, negative results for measurement error and conflicting results for reliability. The SFI was evaluated in 3 studies with positive results for internal consistency, reliability, content validity, and structural validity. Conflicting results were found for hypothesis testing. The EA was evaluated in 3 studies which found negative results for internal consistency and structural validity.
Conclusions: The FRI is provisionally recommended for the assessment of disability in people with multi-area spinal pain. This conclusion is based on studies of mainly fair methodological quality.

1. Introduction

Patient-Reported Outcome (PRO) questionnaires provide an efficient and convenient method of assessing disability in people with spine pain. For the clinician, the benefits of using PRO questionnaires are not limited to monitoring treatment effectiveness. Through focussing on the patient's perspective, questionnaires can assist patient centred clinical reasoning as well as facilitating patient empowerment and self-management strategies (Kyte et al., 2015). There are many region-specific PRO questionnaires for low back and neck pain (Costa et al., 2007a; Schellingerhout et al., 2012), but none for upper back pain. Rather than developing a new questionnaire for upper back pain, other PRO questionnaires which are not restricted to one spine region could be used (Feise and Menke, 2010; Gabel et al., 2013). The advantage of these questionnaires is that only a single questionnaire needs to be used no matter where or how many areas of the spine are involved. The improved efficiency from using a single questionnaire may assist clinicians overcoming the time problem which is reported as a barrier to PRO questionnaire use (Duncan and Murray, 2012). As the prevalence of multi-region spine pain is reported to be 9.3% which is higher than the prevalence of neck pain alone at 4.4% (Strine and Hootman, 2007), these PRO questionnaires for any-region spine pain could be of great clinical value if they have sound measurement properties.

Well-designed systematic reviews enable researchers and clinicians to make informed decisions about which PRO questionnaires to use in specific populations. Systematic reviews of PRO questionnaires which adhere to the PRISMA guidelines and "COnsensus-based Standards for the Selection of health status Measurement INstruments" (COSMIN) checklist (Liberati et al., 2009; Mokkink et al., 2012) have been completed for questionnaires which evaluate the disability of patients with neck pain

* Corresponding author. Department of Rehabilitation, Nutrition and Sport, School of Allied Health, Level 5, HS3, La Trobe University, Bundoora, Vic 3086, Australia. Tel.: +61 3 3479 5857, +61 419001291 (mobile).
E-mail addresses: e.leahy@latrobe.edu.au (E. Leahy), m.davidson@latrobe.edu.au (M. Davidson), deenikabenjamin@optusnet.com.au (D. Benjamin), h.wajswelner@latrobe.edu.au (H. Wajswelner).

http://dx.doi.org/10.1016/j.math.2015.10.010
1356-689X/

(Schellingerhout et al., 2011, 2012). Although these reviews concluded that none of the questionnaires have been adequately assessed, they can still inform decisions regarding questionnaire selection. A similar review for PRO questionnaires for any-region spine pain has not been completed. Such a review would allow the comparison of available questionnaires and assist researchers and clinicians in choosing a suitable questionnaire.

The objective of this review was to evaluate the measurement properties of PRO questionnaires which evaluate disability associated with pain in any or multiple areas of the spine.

2. Methods

2.1. Questionnaires selection

A list of spinal PRO questionnaires was compiled from the content and reference lists of recent neck and back reviews of PRO questionnaires (Grotle et al., 2005; Costa et al., 2007a; Ferreira et al., 2010; Terwee et al., 2011; Schellingerhout et al., 2012) and from a recent report of the development of a PRO questionnaire for any-region spinal pain (Gabel et al., 2013).

Selection of PRO questionnaires was completed using pre-determined inclusion criteria. For inclusion, disability related to the neck, upper back and low back region pain needed to be evaluated in one questionnaire. Questionnaires needed to be available in English and independently completed by patients. Those requiring an interview for completion were excluded.

The list of potential PRO questionnaires was screened for eligibility by two authors (EL, MD). Questionnaires that had 'low back' or 'neck' in the name were excluded. For the remaining questionnaires additional information was sought from studies describing their development to determine eligibility. The final included questionnaires were decided through consensus agreement.

2.2. Design

A systematic review was completed for the selected questionnaires using a pre-determined protocol based on the PRISMA statement (Liberati et al., 2009) and COSMIN checklist guidelines (Mokkink et al., 2012; Terwee et al., 2012).

2.3. Search strategy

A title and abstract search was completed for included questionnaires from inception to 5th August 2015 using Pubmed and Cinahl (EBSCO platform) databases. Search keywords were the questionnaire names. No language limits or other filters were used. Articles were imported into an EndNote x6 file and duplicates removed. Citation searches were completed for the earliest publication for each questionnaire using Web of Science and Pubmed databases. Reference lists of included articles were hand searched for additional articles.

Title and abstracts were independently evaluated for inclusion by two authors (EL, HW) with disagreements resolved through discussion. If required a third reviewer made the final decision. Full text of remaining articles was then independently assessed for eligibility.

2.4. Identification of eligible studies

Studies were included if they were full text original journal articles published in English and evaluated the measurement properties of the selected questionnaires in any population. Measurement properties evaluated were any of those defined by the COSMIN checklist (Mokkink et al., 2010c): internal consistency, reliability, measurement error, construct validity, hypothesis testing, cross-cultural validity, criterion validity, content validity or responsiveness. Review articles were excluded.

2.5. Data extraction

Measurement property and descriptive data was independently extracted by two authors (EL, DB). Disagreements were resolved by discussion. Descriptive data extracted was based on the interpretability and generalisability sections of the COSMIN checklist as recommended by the developers (Terwee et al., 2012). This data included setting, sample number, patient characteristics, instrument language, gender and age.

2.6. Quality assessment

Two reviewers (EL, DB) independently assessed and rated each study's methodological quality using the COSMIN checklist (Mokkink et al., 2012; Terwee et al., 2012) which has been recommended for use in systematic reviews of PRO questionnaires (Mokkink et al., 2010b). Disagreements were settled through discussion, with a third reviewer (MD) making the final decision. The COSMIN checklist is made up of 12 sections each of which has between 5 and 18 items. Nine of these sections evaluate study quality with respect to specific measurement properties. Studies are rated as either poor, fair, good or excellent for each measurement property. The 9 measurement properties fit into one of 3 domains which are reliability, validity and responsiveness. Measurement property definitions are described elsewhere (Mokkink et al., 2012). Cohen's Kappa was calculated to determine the level of inter-rater agreement for quality assessment.

2.7. Results synthesis

Evaluation of each questionnaire's measurement properties was completed using levels of evidence approach previously used in systematic reviews of PRO questionnaires (Schellingerhout et al., 2011, 2012). This approach was adopted from the Cochrane Back Review Group (van Tulder et al., 2003) and enables synthesis of measurement property data with the studies' methodological quality. Measurement properties are rated as having strong, moderate, limited, conflicting or unknown level of evidence. Strength of evidence can be in a positive or negative direction. Criteria used to evaluate the strength and direction of evidence is described in Tables 1 and 2. Levels of evidence evaluation were completed separately for each of the questionnaire language versions.

Where required by the levels of evidence approach (Schellingerhout et al., 2012), values were calculated from other measures reported. Smallest Detectable Change (SDC) was calculated from the Standardised Error of Measurement (SEM) using the formula $SDC = 1.96 \times \sqrt{2} \times SEM$ (de Vet et al., 2006; Terwee et al., 2007, 2009). For evaluation of measurement error, the levels of evidence approach requires the Minimal Important Change (MIC) to be compared to the SDC to determine whether there is a positive or negative result (Schellingerhout et al., 2011, 2012). The MIC (Terwee et al., 2009) has the same definition as the Minimally Clinically Important Difference (MCID) (Jaeschke et al., 1989) therefore the terms were considered equivalent for this review.

3. Results

3.1. Selection of questionnaires

Fifty-nine PRO questionnaires were identified, of which 3 fulfilled the inclusion criteria. These were the Extended Aberdeen