ELSEVIER

Journal of Clinical Epidemiology 67 (2014) 416–424

Journal of Clinical Epidemiology

The methodological quality of diagnostic test accuracy studies for musculoskeletal conditions can be improved

Nicholas Henschke[a,b,*], Julia Keuerleber[a], Manuela Ferreira[a], Christopher G. Maher[a], Arianne P. Verhagen[c]

[a]*Musculoskeletal Division, The George Institute for Global Health, Level 13, 321 Kent Street, Sydney NSW 2000, Australia*
[b]*Institute of Public Health, University of Heidelberg, Im Neuenheimer Feld 324, 69120 Heidelberg, Germany*
[c]*Department of General Practice, Erasmus MC, Dr. Molewaterplein 50, 3015 GJ Rotterdam, The Netherlands*

Accepted 18 November 2013

Abstract

Objectives: To provide an overview of reporting and methodological quality in diagnostic test accuracy (DTA) studies in the musculoskeletal field and evaluate the use of the QUality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist.

Study Design and Setting: A literature review identified all systematic reviews that evaluated the accuracy of clinical tests to diagnose musculoskeletal conditions and used the QUADAS checklist. Two authors screened all identified reviews and extracted data on the target condition, index tests, reference standard, included studies, and QUADAS items. A descriptive analysis of the QUADAS checklist was performed, along with Rasch analysis to examine the construct validity and internal reliability.

Results: A total of 19 systematic reviews were included, which provided data on individual items of the QUADAS checklist for 392 DTA studies. In the musculoskeletal field, uninterpretable or intermediate test results are commonly not reported, with 175 (45%) studies scoring ''no'' to this item. The proportion of studies fulfilling certain items varied from 22% (item 11) to 91% (item 3). The interrater reliability of the QUADAS checklist was good and Rasch analysis showed excellent construct validity and internal consistency.

Conclusion: This overview identified areas where the reporting and performance of diagnostic studies within the musculoskeletal field can be improved.

Keywords: Musculoskeletal; Diagnosis; QUADAS; Study quality; Rasch analysis; Systematic review

1. Introduction

Making an accurate diagnosis is essential to guide management and determine prognosis [1]. A common diagnostic research study is the test accuracy study, which generates a comparison of measurements made by an index test against those of an accepted ''gold standard'' or reference test [2]. Within the musculoskeletal field, arriving at a specific diagnosis is seen to be particularly challenging [3]. For various conditions, this difficulty has been attributed to a lack of adequate reference tests [4], poor reliability of index tests [5], or methodological weaknesses in test accuracy studies [6]. An increasing amount of published literature on the accuracy of index tests for musculoskeletal conditions is now available and systematic reviews of diagnostic test accuracy (DTA) have become more common. These reviews draw together the available evidence on test accuracy in light of potential methodological biases of the primary studies to provide clinicians with recommendations on diagnosis of specific conditions [7].

When incorporated into systematic reviews, assessment of methodological quality is necessary to allow potential biases and sources of variation that might lead to heterogeneity to be identified [8]. The QUality Assessment of Diagnostic Accuracy Studies (QUADAS) initiative provides a tool to assess the quality of test accuracy studies [9]. The tool combines features derived from empirical evidence and expert opinion into a checklist of 14 items assessing risk of bias, applicability, and reporting quality. The QUADAS checklist is designed to be tailored to specific review questions by authors of DTA reviews, allowing flexibility across numerous fields of research and study designs [9]. Since its development, the QUADAS checklist has been used in more than 200 DTA reviews [10]. The frequent use of this checklist provides an opportunity to evaluate the typical biases within the musculoskeletal field and provide recommendations on how to avoid these when designing test accuracy studies.

Conflict of interest: The authors declare no conflict of interest.

* Corresponding author. Tel.: +49-6221-56-5215; fax: +49-6221-56-5948.

E-mail address: henschke@uni-heidelberg.de (N. Henschke).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2013.11.008

What is new?

- Less than half of all primary diagnostic studies in the musculoskeletal field fulfilled item 11 (blinding of reference standard), item 4 (time between tests), item 13 (intermediate tests results), or item 14 (study withdrawals) of the QUality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist.
- Rasch analysis showed that the QUADAS checklist has excellent construct validity and internal consistency.
- The reporting and performance of diagnostic test accuracy (DTA) studies in the musculoskeletal field can be improved, and authors should be encouraged to follow recommended guidelines to do so.
- These findings provide empirical evidence supporting the use of a summary score for the QUADAS checklist.
- The development of a summary score for the reporting and methodological quality of DTA studies should be considered and may be a useful feature for comparability across studies within the same field and to facilitate interpretation.

The evaluation of quality checklists such as QUADAS is a continuous process as new methodological developments arise. Recently, in light of feedback from review authors and methodological research, the QUADAS checklist has been updated to QUADAS-2, which offers additional features [10]. Nevertheless, the available information in the original QUADAS tool provides a unique opportunity to evaluate the checklist and describe methodological quality in specific fields. Previous studies have evaluated both the inter- and intrarater reliability [11] of QUADAS and found a strong relationship between fulfillment of the items and test accuracy results [12,13]. Although the checklist appears to be reliable and useful, the use of an overall quality score derived from the QUADAS checklist has been discouraged [14]. However, several DTA reviews have calculated and used an overall quality score as this presumably eases the process of drawing conclusions. One method to evaluate the overall utility of QUADAS that has not yet been applied is Rasch analysis. An advantage of Rasch analysis is that it offers a sophisticated method for assessing whether an instrument measures a single construct, in this case, the methodological quality of DTA studies [15]. Accordingly, Rasch analysis will provide an empirical basis for judging whether or not it is meaningful to sum item scores to create a summary score.

The present study aims to: (1) describe the application of QUADAS in systematic DTA reviews for common musculoskeletal conditions, (2) identify the proportion of test accuracy studies that fulfill each item of the QUADAS, and (3) evaluate the construct validity and internal reliability of the QUADAS checklist using Rasch analysis.

2. Methods

2.1. Data sources

A comprehensive search of the literature was performed to identify all published systematic reviews that evaluated diagnostic accuracy of clinical tests to diagnose common musculoskeletal conditions. As this study aimed to describe the use of the QUADAS checklist in DTA reviews, an electronic forward citation search was performed using the Web of Science database to identify all studies citing the original publication of the QUADAS checklist by Whiting et al. [9]. This was supplemented with a search of MEDLINE and the Cochrane Database of Systematic Reviews to identify more recent reviews that used the QUADAS checklist. Finally, a search was performed on the MEDION database of diagnostic studies and diagnostic systematic reviews (www.mediondatabase.nl).

2.2. Study selection

Two authors (N.H. and J.K.) independently screened the titles and abstracts of all reviews identified by the search to exclude those that were clearly outside the scope of the study. To determine eligibility for the analysis, reviews were included if they satisfied the following criteria: (1) were systematic DTA reviews for musculoskeletal conditions; (2) reported on the accuracy of diagnostic imaging or index tests from the clinical history and physical examination, (3) used the QUADAS checklist to evaluate quality of the original test accuracy studies, and (4) provided data on each item of the QUADAS fulfilled by the original studies. Where data were not reported in the systematic reviews, attempts were made to contact the review authors for further information.

2.3. Data extraction and analysis

Two authors (N.H. and J.K.) independently extracted the following data from each eligible review: target condition, index tests evaluated, reference standard, number of included studies, and individual ratings for each item (yes/no/unclear) of the QUADAS checklist for each included study. All disagreements were resolved via discussion and consensus.

The QUADAS checklist from each review was tabulated; and for each item, the proportion of included studies in each category (yes, no, or unclear) was determined. From each included review, descriptions of how each item was scored and modifications made to the original checklist were recorded. Where primary test accuracy studies were assessed by more than one review, the scores for each item were extracted to a separate checklist for assessment of interrater reliability by calculating the kappa statistic. For