

# Understanding the Difference between Risk of Bias and Concerns Regarding Applicability Using QUADAS-2: A Methodological Review

Rutjes Anne W.S.<sup>(1)</sup>, Lachi Alessio<sup>(1)</sup>, Holmes Jude<sup>(2)</sup>, Tomlinson Eve<sup>(3)</sup>, Leeflang Mariska M. G.<sup>(4)</sup>, Whiting Penny<sup>(3)</sup>, Mallett Sue<sup>(2)</sup>, Davenport Clare<sup>(5,6)</sup>

(1) UniCamillus International Medical University, Rome, Italy

(2) Centre for Medical Imaging, University College London, London, UK

(3) Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.

(4) Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

(5) Test and Prediction Group, Department of Applied Health Science, University of Birmingham, Birmingham, B15 2TT, UK

(6) NIHR Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham B15 2TT, UK

Joint first authors

**CORRESPONDING AUTHOR:** Rutjes Anne W.S., [anne.rutjes@unicamillus.org](mailto:anne.rutjes@unicamillus.org)

**KEYWORDS:** Risk of bias assessment, applicability, diagnostic test accuracy, systematic review

## INTRODUCTION

Diagnostic test accuracy (DTA) systematic reviews bring together findings from DTA studies to summarise the accuracy of a diagnostic test. Studies included in a DTA review should be assessed for risk of bias and applicability concerns, because a review of biased studies, or studies that do not apply directly to the review question, could result in misleading conclusions. Studies are most commonly assessed with the QUADAS-2 tool. Anecdotal evidence has suggested that researchers sometimes struggle to differentiate between risk of bias and applicability. Here, we investigate this distinction for the patient selection and index test domains.

## OBJECTIVES

To develop a framework for assessing the applicability of the study target population and the study index test to the review defined target populations and index test(s). We aimed to explore review authors' applicability assessments for the QUADAS-2 patient selection and index test domains, to inform the framework.

## METHODS

DTA reviews were eligible for inclusion if they were published in the Cochrane Library, had used QUADAS-2, and had at least one study rated as "high concerns" for appli-

cability of the patient selection or index test domain. Review selection was checked by a second reviewer. From each review, we extracted article identifiers such as title, authors and publication date, extracted the primary objectives and elements of review questions: population, index test(s), target condition, reference standard. For each review, we extracted the rationale provided by the authors for "high concerns" applicability judgements for the respective domains. We also extracted author's rules how to assess applicability concern with QUADAS-2, whenever it was tailored to the review topic. One reviewer assessed these rationales and rules, which were verified by another. We also recorded any other issues that arose as part of the applicability assessment, such as sub-optimal reporting or erroneous applicability assessment. Two reviewers categorized the rationales inductively into themes, which will be discussed in the QUADAS steering group. The final framework will be informed by the identified themes and thorough group discussions.

## RESULTS

This review is in progress. Of the 186 available Cochrane DTA reviews, 123 met our inclusion criteria: 110 for the patient selection domain, 75 the index test domain. The data extraction process is ongoing and final results will be presented at the conference. The majority of the more recent Cochrane DTA reviews include guidance tailored to the review's topic. Review authors typically have a broad review question, whereas signaling questions tailored to the review's topic are

more restrictive. Several themes emerged in the patient selection domain. These include: study setting not matching the review question, study unit not matching the review study unit (lesion level versus patient level); study's target condition not meeting the review's target condition (in-vitro versus in-vivo); study's disease spectrum not covering the review's targeted disease spectrum (i.e. due to sampling methods, choice of selection criteria; inappropriate exclusions) and study's indication for testing not sufficiently matching the reviews indication. In the index test domain, themes identified so far include a mismatch in the study and the review with respect to: index test technology; test protocol; thresholds to define the target condition; clinical background of the examiner; experience of the examiner; use of consensus test interpretation rather than use of the interpretation of a single examiner, use of clinical information. In both domains, a number of reviews misunderstood QUADAS-2 applicability or provided insufficient information on the rationale for high concerns. Once all themes are identified, the framework will be developed and presented as SISMEC.

## CONCLUSIONS

Clear sources of applicability concerns are identifiable, but several Cochrane review authors struggle to adequately identify and report them. At SISMEC, we will present the applicability framework to guide review authors in their assessment of applicability concerns for the QUADAS patient selection and index test domains