Scoring the quality of diagnostic accuracy studies: an example using QUADAS

THE UNIVERSITY *of* York

Penny Whiting[1], Roger Harbord[2] and Jos Kleijnen[1]

[1]Centre for Reviews and Dissemination, University of York, [2]Department of Social Medicine, University of Bristol

Background

■ Quality scores are a convenient method to incorporate quality into a systematic review.

■ The main problem with quality scores is determining how to weight each item. There is no objective way of doing this and different methods are likely to produce different scores.

■ We are not aware of any work which has looked at different weightings for the same quality assessment tool in the area of diagnostic accuracy studies.

Objective

■ To use different methods of weighting QUADAS[1] items to produce five quality scores, and to examine the effects of incorporating these into a systematic review.

Methods

■ *Scoring schemes:* We developed five schemes for weighting QUADAS to produce different quality scores. These are summarised in Table 1.

Table 1: QUADAS[1] and weighting for each item using each of the different scoring schemes

■ *Data set:* 28 studies on ultrasound for the diagnosis of vesico-ureteral reflux in children. The studies can be split into standard (12 studies) and contrast-enhanced (16 studies) ultrasound.

■ *Methods for investigating the effects of the quality scores on test performance:*

• We used four different methods to investigate the effects of quality scores on test performance.

• Each method was performed separately for the standard and contrast-enhanced ultrasound studies:

a. *Ranking of studies:* we ranked `studies according to quality score

b. *Difference in estimated diagnostic accuracy between high and low quality studies:* We stratified studies into 'high' and 'low' quality studies. Studies with scores higher than the median score were classified as 'high' quality, studies with the median quality score or lower were classified as 'low' quality. Relative diagnostic odds ratios (RDORs) comparing results from high and low quality studies were estimated.

c. *Quality score as a possible source of heterogeneity:* the SROC model was extended to include 'quality score' as a covariate.

Results

a. Ranking of studies

Table 2: ranking of studies according to each different quality score

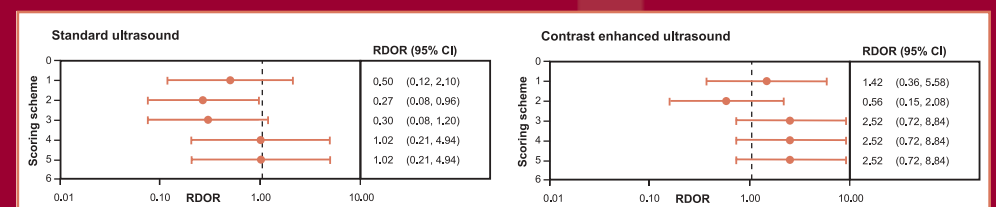### b. Difference in estimated diagnostic accuracy between high and low quality studies

Figure 1: Forest plots showing the RDOR in 'high' quality studies compared to 'low' quality studies for each of the five quality scoring schemes

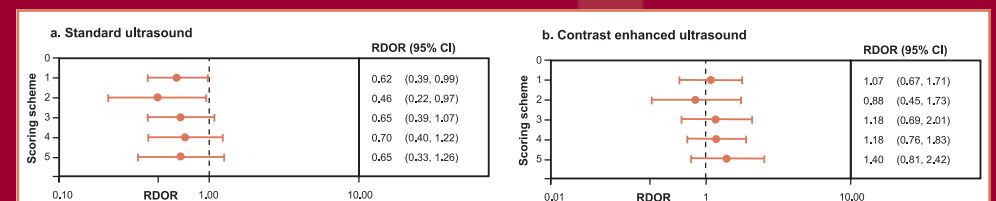### c. Quality score as a possible source of heterogeneity:

Figure 2: Forest plots showing the RDOR for a 10 point increase in quality for each of the 5 quality scoring schemes

Conclusions

■ Using different methods of weighting quality items produces different quality scores

■ Incorporating these different quality scores into the results of a review can produce conflicting results

■ Quality scores should not be used to summarise the results of the quality assessment

■ This work has limitations due to the small number of included studies

References

1. Whiting P, Rutjes A, Reitsma J, Bossuyt P, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology* 2003;3:25. Available from: http://www.biomedcentral.com/1471-2288/3/25