

**Table 1 Summary of reviewers main comments**

<b>Key Issues Identified</b>	<b>Explanation</b>
<b>Starting point for the process</b>	Reviewers need to be explicit on whether evidence from DTAR is being rated from a PICO* perspective or from a test accuracy perspective as this can lead to different judgments on quality of evidence
<b>Clinical question</b>	A PICO styled key question formulation was not an explicit part of the exercise. A clear PICO styled key question was especially important in DTARs that compared multiple index tests or when different patient spectrums existed (1)
<b>Evidence base (outcomes, number of studies, number of patients)</b>	In one review, unit of analysis was not based on no. of patients (Virgili et al (2) analysed 'no. of eyes')
<b>Lack of clarity on how to assess following GRADE domains:</b>	
Risk of bias (RoB)	Difficult to make an assessment on RoB when QUADAS items were labeled as "unclear"
Indirectness	i. Absence of linkage of index test to patient care pathway in at least one review (1) ii. Given that test accuracy is inherently indirect evidence for PIO, would this warrant a downgrading to "serious" on the basis of indirectness?
Inconsistency	Different rationales used to rate this factor such as extent of confidence interval (CI) overlap or lack thereof, unexplained heterogeneity, inconsistent use of test threshold positivity and variable reference standard definitions
Imprecision	Different rationales used to rate this factor such as small study number and wide CI
Publication bias	Reviewers were unclear on how to assess this factor
<b>Across all GRADE domains</b>	Reviewers had to be conscious to not double downgrade on a single factor e.g. representativeness of study populations could be downgraded for RoB or Indirectness (Abba et al (3) had only 54% of a representative patient population, van der Windt (1) where only 1 study was in primary care setting)
<b>Additional points for comparative test reviews (3)</b>	i. For an indirect comparison of 2 index tests, each test needed to be assessed first against its reference standard and then relative to each other. This created the need for three separate tables ii. When making the relative comparison, the score for each GRADE domain (e.g. RoB, indirectness, etc.) was determined as the lower of the 2 scores for that domain for each of index tests compared to its ref standard iii. The overall quality of evidence (for an indirect comparison of 2 index tests) be further downgraded by one level for indirectness

1 Optical coherence tomography (OCT) for detection of macular edema in patients with diabetic retinopathy. Virgili G, Menchini F, Murro V, Peluso E, Rosa F, Casazza G. The Cochrane Library 2011, Issue 7

2 Physical examination for lumbar radiculopathy due to disc herniation in patients with low-back pain (Review). van der Windt DAWM, Simons E, Riphagen II, Ammendolia C, Verhagen AP, Laslett M, Devillé W, Deyo RA, Bouter LM, de Vet HCW, Aertgeerts B. The Cochrane Library 2011, Issue 2.

3 Rapid diagnostic tests for diagnosing uncomplicated P.falciparum malaria in endemic countries (Review). Abba K, Deeks JJ, Olliaro PL, Naing CM, Jackson SM, Takwoingi Y, Donegan S, Garner P. The Cochrane Library 2012, Issue 1.

\*PIO refers to patient important outcomes