

# Exploring the magnitude of verification bias in diagnostic accuracy studies

Alexis Llewellyn, Mark Simmonds. CRD, University of York. [alexis.llewellyn@york.ac.uk](mailto:alexis.llewellyn@york.ac.uk)

## Background

In diagnostic test accuracy studies patients with a negative test result are less likely to receive a 'gold standard' test to confirm the diagnosis if that test is invasive or potentially harmful. This may result in misclassifying undetected disease and overestimating diagnostic accuracy. This is called verification bias.

## Objective

To investigate the magnitude of verification bias and its implications for patients undergoing colposcopy tests.

## Methods

Data were taken from a recent review of the diagnostic accuracy of adjunctive colposcopy compared with standard colposcopy for detecting precancerous and cancerous lesions in women at risk of cervical cancer.<sup>1</sup> The 'gold standard' test was to take one or more biopsies from the cervix, but not all studies did this when the adjunctive colposcopy test was negative.

QUADAS-2 was used to assess the risk of bias and modified to account for the direction and magnitude of bias. We compared the sensitivity and specificity in studies where biopsies were performed (low risk of verification bias) to those where they were not performed (high risk of verification bias).

We then created simulated data to explore the likely effect of verification bias in studies where no biopsies were performed by varying the likely rate of misclassified undetected lesions based on epidemiological evidence.

## Patient involvement

The main relevant UK patient group informed the scope of the review and commented on the review report. Diagnostic accuracy was a key outcome for patients.

## Conclusions

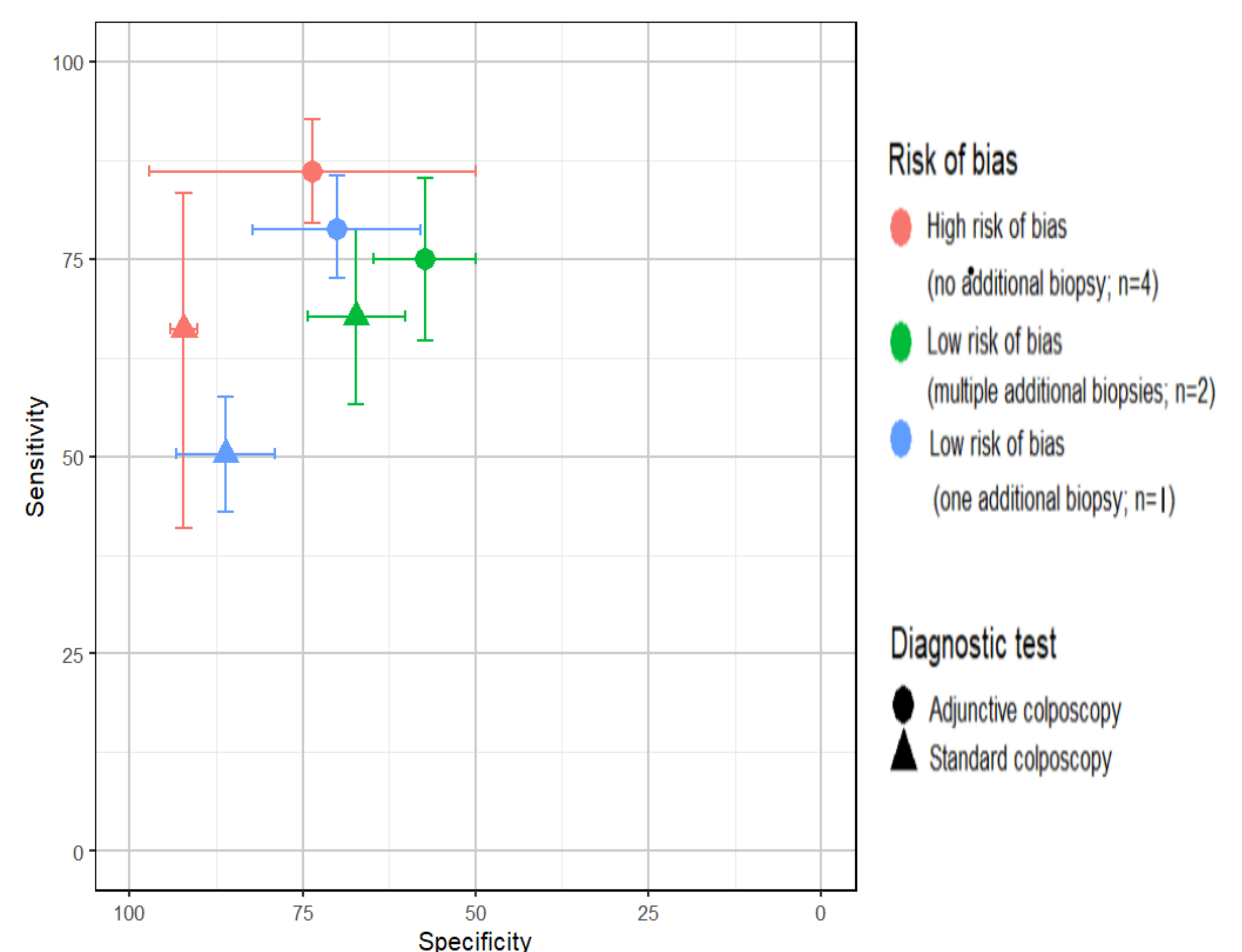
Systematic reviews of diagnostic accuracy should consider assessing the magnitude of verification bias.

Sensitivity analyses and simulation can help quantify the possible magnitude of this bias.

Verification bias should be weighed against the benefits of limiting unnecessary testing.

Where appropriate, direction and magnitude of bias issues should be considered in QUADAS-2 domains.

Fig. 1 Sensitivity analysis: high vs. low risk of verification bias



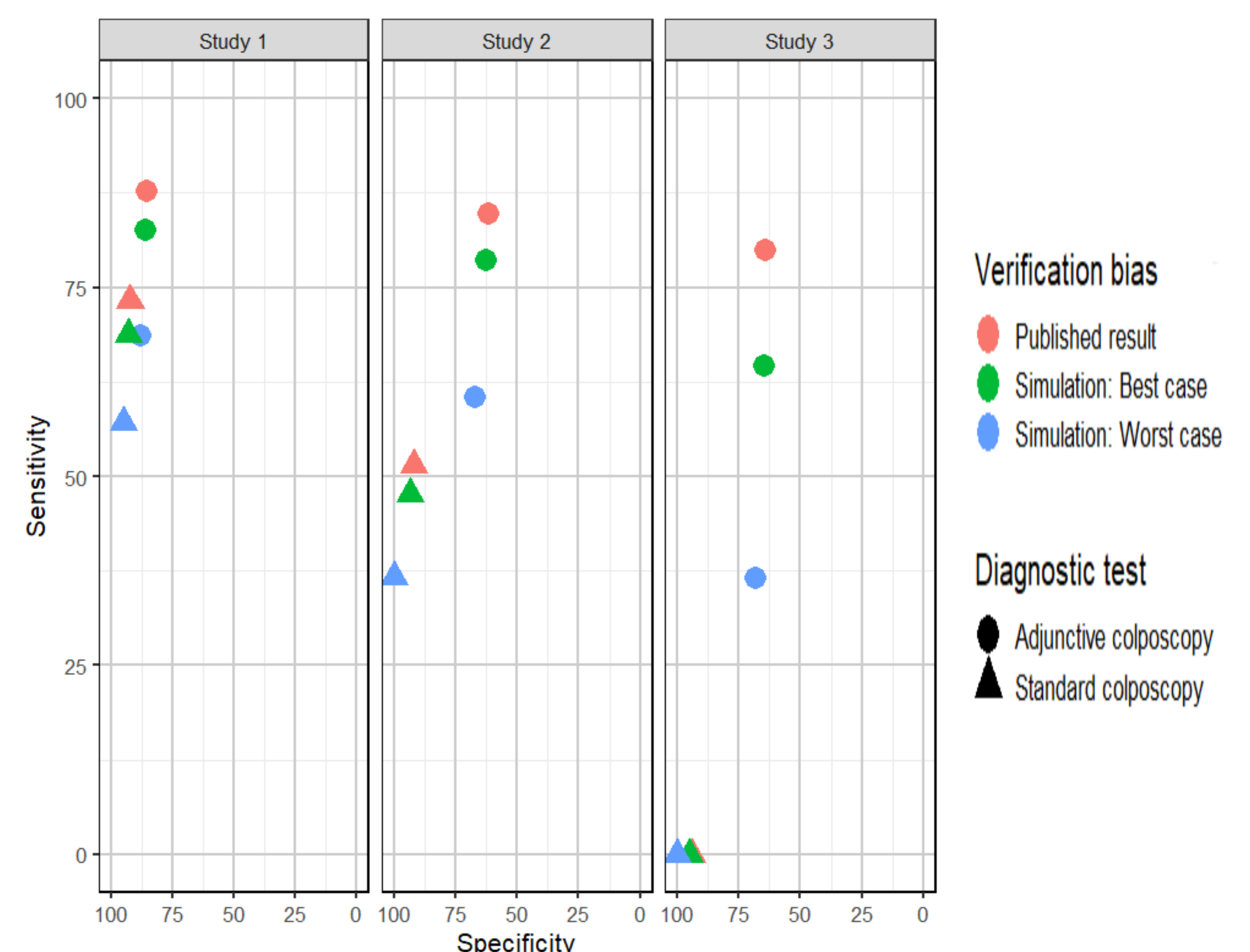
## Results

8 out of 11 included studies did not use biopsy as the gold standard in all patients and were at high risk of verification bias. 3 studies used either single or multiple biopsies in all patients, and were at low risk of bias.

Fig.1 shows that studies at high risk of bias produced larger estimates of sensitivity and specificity.

Fig.2 shows that when simulations were performed sensitivity could be significantly overestimated, but specificity was largely unchanged. This may have better face validity as verification bias was unlikely to significantly affect specificity.

Fig. 2 Simulation: impact of varying rate of misclassified disease in studies at high risk of verification bias



<sup>1</sup> Peron M, Llewellyn A, Moe-Byrne T, Walker S, Walton M, Harden M, et al. Adjunctive colposcopy technologies for assessing suspected cervical abnormalities: systematic reviews and economic evaluation. Health Technol Assess [in press]. HTA project no 16/30/01. PROSPERO registered: CRD42017054515