

# Should systematic reviews of diagnostic tests go beyond test accuracy?

Marie Westwood<sup>1</sup> and Penny Whiting<sup>2</sup>

<sup>1</sup> Centre for Reviews and Dissemination, University of York (mew3@york.ac.uk); <sup>2</sup> Department of Social Medicine, University of Bristol

## Background

Systematic reviews of diagnostic tests often focus exclusively on test accuracy studies, as this is the dominant study design in primary research. However, reviews that fail to adequately consider the clinical context of testing and its impact on patient outcome may result in misleading conclusions and failure to note significant gaps in the evidence base. We consider how broader evidence surrounding test accuracy has been dealt with, using the published systematic review literature and our own experience.

## Objectives

To examine how issues beyond test accuracy have been considered in published diagnostic systematic reviews. To discuss our own experience in conducting these reviews and assess whether a restriction to test accuracy studies would have changed findings.

## Methods

We examined a convenience sample of diagnostic (diagnosis, screening, monitoring, prognosis) abstracts from the DARE database (1998 to 2007) to determine whether they addressed outcomes beyond test accuracy. Based upon the published DARE abstracts, we recorded additional outcomes addressed, whether or not data on these outcomes were identified, and whether or not conclusions incorporated these data or a null finding.

We assessed the contribution of information other than test accuracy data to reviews of diagnostic tests that we have conducted.

## Results

### Sample of published systematic reviews

Our DARE sample included 234 abstracts: 13 from 2007; 25 per year from 2000-2006; 23 per year from 1999 and 1998. Of these reviews, 33 (14%) included outcomes other than measures of test accuracy, 18 used additional data identified to inform their conclusions and one included a null finding in its conclusions; seven reported data which were not used in their conclusions and seven reported a null finding which did not form part of their conclusions. Additional outcomes (see Figure 1) included: adverse events (six abstracts); changes to patient management (14 abstracts); changes to patient outcome (17 abstracts); patient preferences (five abstracts); test failure (seven abstracts). Samples were too small to adequately assess any trend over time, but the highest proportion of reviews to include outcomes beyond test accuracy occurred in the 2006 sample (36%).

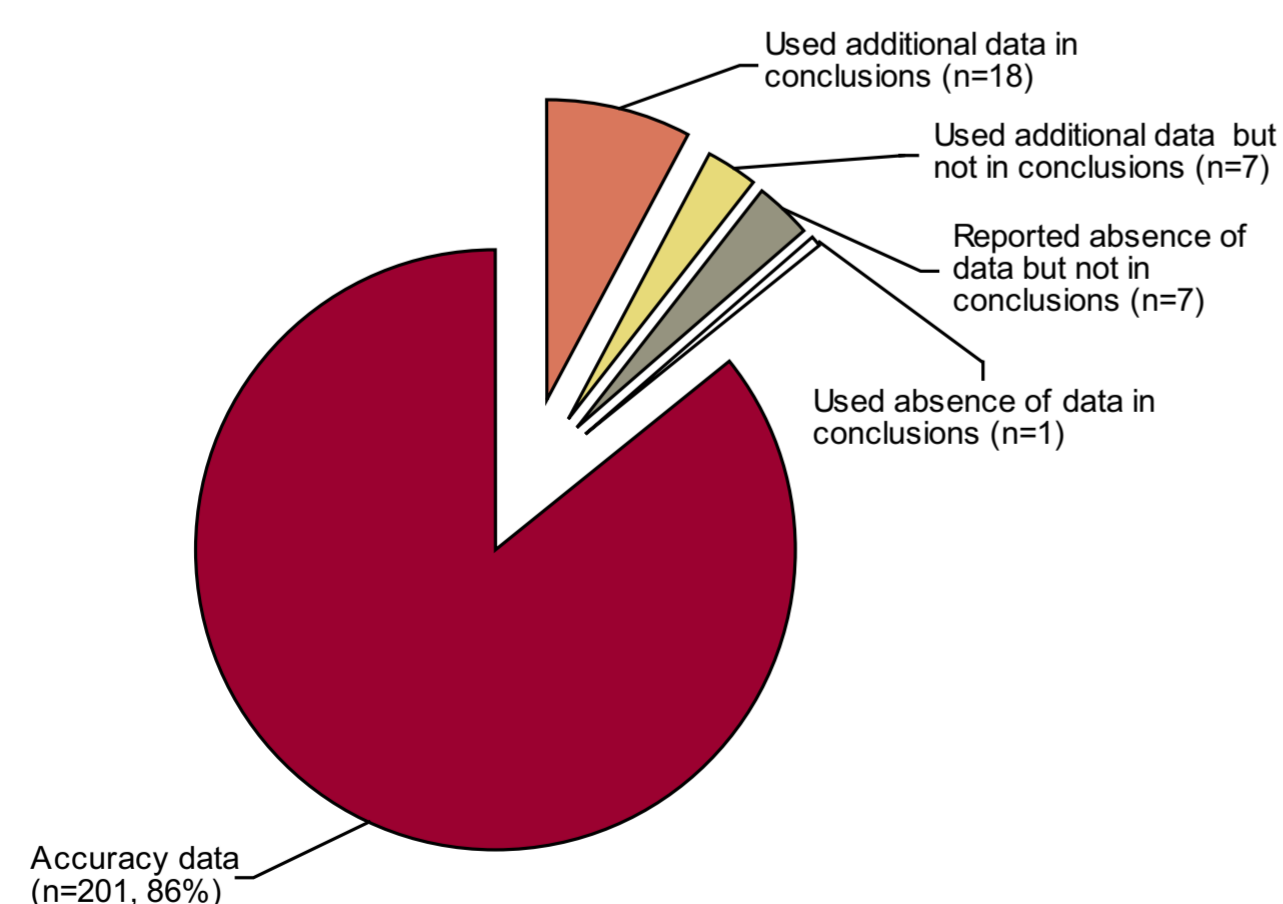


Figure 1: Numbers of reviews addressing outcomes other than test accuracy

## Examples from our experience

Our diagnostic reviews have sought data on the clinical effectiveness (changes to management and/or outcomes) as well as the accuracy of tests, adverse events, patient preferences and resource implications of test introduction. Data were generally sparse and of poor quality, but, in a number of cases, made a significant contribution to overall conclusions.

**Example 1.** An assessment of the accuracy of imaging tests used to investigate confirmed UTI in children.<sup>1</sup>

- Less invasive alternatives were found to have poor accuracy, which could have resulted in a conclusion that universal invasive testing should continue
- One trial was identified which reported that routine testing after a first UTI did not reduce negative outcomes. Studies using follow-up as a reference standard found acute testing to be poorly predictive of negative outcomes
- The review concluded that there was no evidence to support routine imaging at first UTI and current UK guidance now recommends a selective imaging strategy based upon atypical infection and recurrence<sup>2</sup>

**Example 2.** An assessment of the accuracy of imaging tests for the evaluation of peripheral arterial disease.<sup>3</sup>

- Less invasive alternatives (ultrasound, CT, and contrast enhanced MRI) all showed reasonable accuracy when compared with intra-arterial angiography
- Trial data showed that, although initial evaluation using ultrasound or intra-arterial angiography produced similar end outcomes, 22% of patients initially assessed using ultrasound required further angiographic evaluation to formulate a treatment plan
- Data on adverse events and patient preferences lent support to contrast enhanced MRI as a potential replacement test, with CT requiring further research

**Example 3.** An assessment of 'simple' tests (e.g. BMI) for screening for and/or monitoring childhood obesity.<sup>4</sup>

- Accuracy data were limited and poor
- The lack of evidence on the effectiveness of screening or monitoring strategies in reducing obesity in the population or on the effectiveness of weight loss interventions in individuals was highlighted
- Limited data on the attitudes of children, parents and health care professionals and on potential negative effects contributed further to the conclusion that immediate introduction of screening/monitoring would be of doubtful value

## Conclusions

A small proportion of published reviews assess issues beyond test accuracy and those that do find few additional studies. However, we consider that it remains important to frame the objectives and inclusion criteria of reviews of diagnostic tests to address all relevant aspects of test use, including, but not exclusively, accuracy.

## References

1. Westwood ME, Whiting PF, Cooper J, Watt IS, Kleijnen J. Further investigation of confirmed urinary tract infection (UTI) in children under five years: a systematic review. *BMC Pediatr* 2005; 5:2.
2. National Institute for Health and Clinical Excellence. *UTI in children: diagnosis, treatment and long-term management*. NICE clinical guideline 54. Quick reference guide. London: National Institute for Health and Clinical Excellence, August 2007. <http://www.nice.org.uk/nicemedia/pdf/CG54quickrefguide.pdf>
3. Collins R, Burch J, Cranny G, Aguiar-Ibanez R, Craig D, Wright K, Berry E, Gough M, Kleijnen J, Westwood M. Duplex ultrasonography, magnetic resonance angiography and computed tomography angiography for diagnosis and assessment of symptomatic lower limb peripheral arterial disease: systematic review. *BMJ* 2007; 334:1257.
4. Westwood M, Fayter D, Hartley S, Rithalia A, Butler G, Glasziou P, Bland M, Nixon J, Stirk L, Rudolf M. Childhood obesity: should primary school children be routinely screened? A systematic review and discussion of the evidence. *Arch Dis Child* 2007; 92:416.

CRD is part of the National Institute for Health Research and is a department of the University of York. The views expressed in this poster are those of the authors alone.

