2019

Performance of commercial tests for molecular detection of Shiga toxin-producing Escherichia coli (STEC): A systematic review and meta-analysis protocol

Gillian A.M. Tarr
University of Calgary

Chu Yang Lin
University of Alberta

Diane Lorezetti
University of Calgary

Linda Chui
University of Alberta

Phillip I. Tarr
Washington University School of Medicine in St. Louis

See next page for additional authors

BMJ Open

Performance of commercial tests for molecular detection of Shiga toxin-producing *Escherichia coli* (STEC): a systematic review and meta-analysis protocol

Gillian A M Tarr,[1] Chu Yang Lin,[2] Diane Lorenzetti,[3,4] Linda Chui,[5,6] Phillip I Tarr,[7] Lisa Hartling,[8] Ben Vandermeer,[8] Stephen B Freedman[9]

To cite: Tarr GAM, Lin CY, Lorenzetti D, *et al*. Performance of commercial tests for molecular detection of Shiga toxin-producing *Escherichia coli* (STEC): a systematic review and meta-analysis protocol. *BMJ Open* 2019;**9**:e025950. doi:10.1136/bmjopen-2018-025950

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2018-025950).

Received 9 August 2018
Revised 22 December 2018
Accepted 4 January 2019

For numbered affiliations see end of article.

Correspondence to
Dr Gillian A M Tarr;
Gillian.Tarr@albertahealthservices.ca

ABSTRACT

Introduction Rapid detection of Shiga toxin-producing *Escherichia coli* (STEC) enables appropriate treatment. Numerous commercially available molecular tests exist, but they vary in clinical performance. This systematic review aims to synthesise available evidence to compare the clinical performance of enzyme immunoassay (EIA) and nucleic acid amplification tests (NAATs) for the detection of STEC.

Methods and analysis The following databases will be searched employing a standardised search strategy: Medline, Embase, Cochrane CENTRAL Register of Controlled Trials, Cochrane Database of Systematic Reviews, PubMed, Scopus and Web of Science. Grey literature will be searched under advice from a medical librarian. Independent reviewers will screen titles, abstracts and full texts of retrieved studies for relevant studies. Data will be extracted independently by two reviewers, using a piloted template. Quality Assessment of Diagnostic Accuracy Studies-2 will be employed to assess the risk of bias of individual studies, and the quality of evidence will be assessed with the Grading of Recommendations Assessment, Development and Evaluation approach. A bivariate random-effects model will be used to meta-analyse the sensitivity and specificity of commercial STEC diagnostic tests, and a hierarchical summary receiver operator characteristic curve will be constructed. Studies of single test accuracy of EIA and NAATs and studies of comparative accuracy will be analysed separately.

Ethics and dissemination Ethics approval was not required for this systematic review and meta-analysis. Findings will be disseminated in conferences, through a peer-reviewed journal and via personal interactions with relevant stakeholders.

PROSPERO registration number CRD42018099119.

Strengths and limitations of this study

- There is little evidence reviewing the relative clinical performance of commercially available tests for Shiga toxin-producing *Escherichia coli* (STEC).
- A key strength of this study is the comprehensive comparison of enzyme immunoassays and nucleic acid amplification tests to inform clinical practice.
- A limitation is the lack of a common gold standard for STEC identification, which may introduce heterogeneity into our analysis.
- Another limitation is that the finding of a Shiga toxin (Stx) 1 producing STEC that does not also produce Stx2, especially in the absence of bloody diarrhoea, is of unclear clinical and epidemiological value.

INTRODUCTION

Shiga toxin-producing *Escherichia coli* (STEC) cause significant disease. Although prototypical *E. coli* O157:H7 is the leading cause of haemolytic uraemic syndrome (HUS), other STEC serotypes have been associated with severe disease and large outbreaks.[1–4] Multiple serotypes have now been linked to disease. Unlike the O157 serotype, detection of non-O157 serotypes has increased significantly in the past decade, though likely because of dissemination of technology to detect these organisms.[5] Patients infected with STEC often seek care through emergency departments (EDs), especially if they have bloody diarrhoea. Strong evidence suggests that antibiotics may increase the risk of developing HUS if administered to people infected with STEC,[6–8] and a recent meta-analysis demonstrated that the early administration of fluids is associated with improved outcomes.[9] Therefore, it is important that healthcare providers have a means of detecting STEC that is both rapid and applicable to any serotype.

Historically, STEC testing has focused on the O157 serogroup using culture on sorbitol-MacConkey agar, leveraging its inability to ferment sorbitol.[10] This attribute is not

shared by other STEC serogroups, so they are overlooked if sorbitol-MacConkey agar culture is the only detection method employed. Further, culture can take days to yield results, delaying informed management.[11] In light of the limitations of culture, enzyme immunoassay (EIA) and nucleic acid amplification tests (NAATs) have been developed to detect STEC irrespective of serogroup. Reflecting their popularity, the US Council of State and Territorial Epidemiologists has recently revised the probable STEC case definition to include laboratory evidence from EIA and NAAT.[12]

Numerous tests to detect STEC are commercially available.[13 14] The EIAs detect Shiga toxin (Stx), and most NAATs detect the Stx genes Stx1 and Stx2, and some additionally seek a locus that is specific to the O157 serogroup. For NAAT, STEC is often one of several enteropathogens detected by the assay. EIA has suboptimal sensitivity, particularly if a time-consuming enrichment step is not conducted.[15–18] Commercial NAATs appear to be more sensitive, but results vary by study and test.[19–21] NAATs are more costly than traditional microbiological techniques owing to the equipment and consumables required to perform them. However, the higher cost may be compensated by increased ascertainment[21] and/or improved patient outcomes.[22] As laboratories consider NAATs, it is crucial to identify the best testing strategy to support time-sensitive, cost-effective treatment decisions. Thus, we will conduct a systematic review of commercial EIA and NAAT for STEC detection to determine if and how their performance differs in terms of diagnostic test accuracy (DTA).

METHODS AND ANALYSIS

This systematic review and meta-analysis will be conducted in accordance with reporting requirements for Preferred Reporting Items for Systematic Reviews and Meta-analyses statement (PRISMA). This protocol was prepared according to PRISMA-Protocol and PRISMA-DTA guidelines.[23 24]

Research question

What is the accuracy of commercially available EIA and NAAT for the detection of STEC and how do they differ?

Eligibility criteria

- Participants: study participants with acute diarrhoea, who provide a stool specimen or rectal swab for diagnostic testing; any age or subpopulation.
- Setting: healthcare systems or medical facilities, including outpatient clinics, EDs, hospitals, long-term care centres and similar, without geographical limitation.
- Index tests: any commercially available EIA or NAAT for the detection of Stx, or Stx1 and Stx2; NAAT for the identification of the O157 serogroup, if available. Included studies may assess the accuracy of commercially available EIA, NAAT or both, including comparative accuracy studies.
- Reference standard: at least one of the following: enhanced protocols, real-time PCR, sequencing and/or other NAAT.
- Target condition: acute diarrhoea associated with STEC infection.
- Study designs: cross-sectional diagnostic accuracy studies, encompassing all studies with both index and reference tests conducted on stool samples/swabs collected at a single point of time during the acute diarrhoea illness, including both single test and comparative accuracy studies.
- Report characteristics: years 2005 to present (2015 to present for conference abstracts), published or unpublished, in any language.

Literature searches

The following databases will be searched from 2005: MEDLINE, Cochrane CENTRAL Register of Controlled Trials, Cochrane Database of Systematic Reviews, EMBASE, PubMed, SCOPUS and Web of Science. Clinical trial databases (ClinicalTrials.gov), Food and Drug Administration applications, package inserts for commercial assays, company product websites and literature, government/non-governmental organization reports and conference abstracts will also be searched under the advice of STEC subject experts and a medical librarian. The reference lists of included studies will be scanned to identify additional studies of relevance to this review. The specific search strategy can be found in online supplementary appendix I.

Study records

Data management

Records retrieved will be uploaded into EndNote V.8 (Philadelphia, Pennsylvania, USA), and deduplicated using EndNote V.8 and Rayyan for Systematic Reviews (Qatar, 2018).

Selection process

Two reviewers (GAMT, CYL) will independently screen all titles and abstracts in duplicate, and a third reviewer (SBF) will adjudicate any disagreements. Studies will be included if the title and abstract indicate that the manuscript may contain data related to the evaluation of EIA and/or NAAT for the detection of STEC. The full text of all potentially relevant citations will then be obtained and reviewed by two independent reviewers (GAMT, CYL) using the predefined eligibility criteria outlined above, with the involvement of a third reviewer (SBF) in case consensus cannot be reached. Reasons for inclusion and exclusion will be documented. A tool to document the selection process will be developed, piloted with the first 25 search results and modified as necessary.

Data extraction

Two reviewers will extract data independently and in duplicate using a structured form. The form will be piloted on the first five included studies and modified as

necessary. Discordances will be resolved through discussions involving the reviewers and subject matter experts. First and last study authors will be contacted if data necessary to calculate sensitivity or specificity are absent from the manuscript. Study characteristics and study outcomes (table 1) will be extracted from included studies.

Risk of bias assessments

To assess the risk of bias in individual studies, we will employ the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2).[25] We will follow the recommended process for tailoring the QUADAS-2 to our systematic review, including iteratively tailoring the QUADAS-2 assessment tool and piloting it on at least five studies until consensus has been reached on a version of the tool.[25] As part of this process, we will review the Standards for Reporting of Diagnostic Accuracy[26] and prior QUADAS-2 modifications for comparative accuracy studies[27] for relevant criteria. For comparative accuracy studies, we will add a signalling question regarding the assessment of EIA and NAATs in the same group of patients. The risk of bias in individual studies (for all outcomes reported) will be rated as low/unclear/high.[28] Assessments will be made independently by two reviewers, and disagreements will be resolved by discussion, or where necessary, by a third reviewer. Risk of bias will be reported for all included studies.

Data synthesis

Separate synthesis will be conducted for EIA and NAAT. For each of test type, data will be quantitatively synthesised if at least four studies have been identified. If the number of included studies for either EIA or NAAT is insufficient, point estimates and CIs from the individual papers will be shown, and the comparison of EIA and NAAT will be based on the range of estimates reported in individual papers.

If four or more studies are included for a given test type, a bivariate random-effects model[29] will be used to calculate summary estimates and confidence intervals of primary outcomes and secondary outcomes, and a hierarchical summary receiver operating characteristic (ROC) curve[30] will be constructed.[31] The summary point for sensitivity and specificity with confidence ellipse and the hierarchical summary ROC curve will be graphed. These analyses take into account the correlation between sensitivity and specificity and potential threshold effects (eg, due to cycle thresholds used in PCR).[31] Meta-analysis packages in R[32] and RevMan[33] will be used to conduct all analyses.

Comparative accuracy

To compare EIA and NAAT, we will meta-analyse only comparative accuracy studies that evaluate both types of the test against the same reference standard. If no comparative accuracy studies are identified, we will graphically compare point estimates and CIs for sensitivity and specificity resulting from the separate meta-analysis of each type of test. If there is adequate consistency in reference standards used to assess single test accuracy, we will pool EIA and NAAT studies in a single meta-analysis and include test type as a covariate to test the difference in accuracy between EIA and NAAT.

Subgroup analysis

To identify study characteristics that may be contributing to heterogeneity, we will conduct subgroup analyses when at least four studies are available per subgroup:

- Funding (industry vs other).
- Data source (published vs unpublished).
- Age (<10 years old and <18 years old).
- Location of care.
- Diarrhoea duration (<7 days, ≥7 days, not specified).
- Presence of bloody diarrhoea.
- Specimen type.
- Test brand.
- Test targets.
- Reference standard.

Other subgroup analyses not prespecified here will be identified as such in all reports. Subgroup analyses will illustrate the magnitude of differences in accuracy, and thus allow readers to interpret whether they are clinically meaningful. We will obtain statistical evidence of whether these factors contribute to heterogeneity in the primary analysis by adding each to the bivariate random-effects model as a predictor.

A sensitivity analysis excluding studies with a high risk of bias will be conducted. Additional sensitivity analyses will be added if other potential biases become apparent during the review.

Quality of evidence assessment

For the quality of evidence for each test type, two reviewers, one with clinical and one with methodological expertise, will independently use the Grading of Recommendations Assessment, Development and Evaluation approach to assess the quality of evidence for sensitivity and specificity.[34 35] The test will be considered in the context of how it relates to patient-important outcomes to assign importance to the consequences of summary sensitivity and specificity findings (eg, frequency of false negatives). The domains of study design, limitations/risk of bias, directness, consistency, precision and publication bias will be assessed and combined into a summary grade for all important outcomes of the test. Publication bias will be assessed based on differences in accuracy reported in industry-funded versus non-industry-funded studies.

For the comparison of EIA and NAAT, we will use a similar approach to grade the quality of evidence, with the same domains as for single test accuracy. Risk of bias will reflect the modifications we make to QUADAS-2 for comparative accuracy studies. Indirectness will be affected by the number of comparative accuracy studies including both EIA and NAAT; if few comparative accuracy studies are identified and the comparison is based on single test accuracy from different studies, quality will be downgraded due to indirectness.

Table 1 Data to be extracted from each included study

Item	Rationale
Study characteristics	
Data source	Peer-reviewed studies will be distinguished from non-peer-reviewed data for potential subgroup analysis
Funding source	Studies funded by diagnostic test companies may be subject to additional bias; potential subgroup analysis
Study design	Cross-sectional studies are expected; other study designs will be noted for potential subgroup analysis
Population	Population restrictions within the study (eg, by age, HUS status, etc) will be noted for potential subgroup analysis
Setting	Country or region; potential subgroup analysis
Clinical data	
Location of care	Primary care versus ED versus hospital, and potentially other; potential subgroup analysis
Diarrhoea definition	Study definition for diarrhoea (eg, ≥3 episodes in 24 hours) will facilitate comparability assessment and interpretation
Diarrhoea duration	Mean/median or restrictions on illness duration at the time of sampling; facilitate comparability assessment and interpretation
Specimen type	Stool specimen or rectal swab; potential subgroup analysis
Bloody diarrhoea	Frequency of bloody diarrhoea; potential subgroup analysis
Test	
Brand name	Ease of reference
Type	EIA or NAAT for main comparison
Enrichment	For EIA tests; potential subgroup analysis
Targets	Toxin versus DNA, STEC-only versus multianalyte; interpretation and potential subgroup analysis
Cycle threshold	Cycle cut-off for positivity; facilitate comparability assessment and interpretation
Comparator/reference standard	Composite standard with component tests, discrepant analysis with confirmatory tests; interpretation and potential source of bias
Specimen comparability	Specimens tested by index and comparator from the same point in time, of the same type; potential source of bias
Outcomes	
Outcome type	For STEC generally, Shiga toxin 1 vs 2 or O157 vs non-O157; distinguish primary and secondary outcomes
No tested	Outcome calculation and interpretation
No confirmatory tested	Outcome calculation and interpretation
No of true positives	Outcome calculation
No of false positives	Outcome calculation
No of true negatives	Outcome calculation
No of false negatives	Outcome calculation
Sensitivity	Primary outcome
Specificity	Primary outcome
Single accuracy measures	For example, AUC, diagnostic accuracy, diagnostic OR; secondary outcome
PPV	Secondary outcome
NPV	Secondary outcome
LR+	Secondary outcome
LR–	Secondary outcome

AUC, area under the curve; ED, emergency department; EIA, enzyme immunoassay; HUS, haemolytic uraemic syndrome; LR, likelihood ratio; NAAT, nucleic acid amplification test; NPV, negative predictive value; PPV, positive predictive value; STEC, Shiga toxin-producing *Escherichia coli*.

Study results will be reported according to the PRISMA-DTA guidelines.[24]

Patient and public involvement

This protocol was designed without patient involvement. Patients were not invited to comment on the systematic review design and were not consulted to develop patient-relevant outcomes. Patients were not invited to contribute to the writing or editing of this protocol for readability or accuracy.

ETHICS AND DISSEMINATION

Findings will be disseminated in conferences, through a peer-reviewed journal and via personal interactions with relevant stakeholders.

Author affiliations
[1]Department of Pediatrics, University of Calgary, Calgary, Alberta, Canada
[2]Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada
[3]Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada
[4]Health Sciences Library, University of Calgary, Calgary, Alberta, Canada
[5]Microbiology Section, Provincial Laboratory for Public Health-Alberta Public Laboratories, Edmonton, Alberta, Canada
[6]Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta, Canada
[7]Department of Pediatrics, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA
[8]Department of Pediatrics, University of Alberta, Edmonton, Alberta, Canada
[9]Sections of Pediatric Emergency Medicine and Gastroenterology, Department of Pediatrics, Alberta Children's Hospital and Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

Contributors GAMT conceived the study, contributed to study design, drafted the protocol and revised the protocol following author comments. CYL contributed to study design, drafted the protocol and provided critical revisions. DL contributed to study design and provided critical revisions. LC contributed to study design and provided critical revisions. PIT contributed to study design and provided critical revisions. LH contributed to study design and provided critical revisions. BV contributed to study design and provided critical revisions. SBF conceived the study, contributed to study design and provided critical revisions. This study was conducted under the umbrella of the Alberta Provincial Pediatric EnTeric Infection TEam (APPETITE), and we would like to acknowledge the contributions of Samina Ali, Bonita Lee, Karen Lowerison, and Kelly Kim to the implementation and/or operations of that study and this systematic review.

Funding This review is supported by a 2018 Systematic Review Grant from the Alberta Emergency Strategic Clinical Network grant number RES0039208. The Alberta SPOR Support Unit Knowledge Translation Platform is providing in-kind methodological and biostatistical support for the design, conduct and analysis of the review. Dr. Stephen Freedman is supported by the Alberta Children's Hospital Foundation Professorship in Child Health and Wellness. APPETITE is supported by an Alberta Innovates Team Collaborative Research Innovation Opportunity Grant.

Competing interests SBF has previously received in-kind grant support from BioMérieux and Luminex. LC received funding from TechLab for a previous study on SHIGA TOXIN QUICK CHEK and SHIGA TOXIN CHEK. PIT has served as a consultant to BioRad.

Patient consent for publication Not required.

Ethics approval Ethics approval was not required for this systematic review and meta-analysis.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Preußel K, Höhle M, Stark K, *et al*. Shiga toxin-producing Escherichia coli O157 is more likely to lead to hospitalization and death than non-O157 serogroups--except O104. *PLoS One* 2013;8:e78180.
2. Gould LH, Mody RK, Ong KL, *et al*. Increased recognition of non-O157 Shiga toxin-producing Escherichia coli infections in the United States during 2000-2010: epidemiologic features and comparison with E. coli O157 infections. *Foodborne Pathog Dis* 2013;10:453–60.
3. Kuehne A, Bouwknegt M, Havelaar A, *et al*. Estimating true incidence of O157 and non-O157 Shiga toxin-producing Escherichia coli illness in Germany based on notification data of haemolytic uraemic syndrome. *Epidemiol Infect* 2016;144:3305–15.
4. Luna-Gierke RE, Griffin PM, Gould LH, *et al*. Outbreaks of non-O157 Shiga toxin-producing *Escherichia coli* infection: USA. *Epidemiol Infect* 2014;142:2270–80.
5. Tseng M, Sha Q, Rudrik JT, *et al*. Increasing incidence of non-O157 Shiga toxin-producing Escherichia coli (STEC) in Michigan and association with clinical illness. *Epidemiol Infect* 2016;144:1394–405.
6. Wong CS, Mooney JC, Brandt JR, *et al*. Risk factors for the hemolytic uremic syndrome in children infected with *Escherichia coli* O157:H7: a multivariable analysis. *Clin Infect Dis* 2012;55:33–41.
7. Smith KE, Wilker PR, Reiter PL, *et al*. Antibiotic treatment of *Escherichia coli* O157 infection and the risk of hemolytic uremic syndrome, Minnesota. *Pediatr Infect Dis J* 2012;31:37–41.
8. Freedman SB, Xie J, Neufeld MS, *et al*. Shiga Toxin-producing *Escherichia coli* Infection, antibiotics, and risk of developing hemolytic uremic syndrome: a meta-analysis. *Clin Infect Dis* 2016;62:1251–8.
9. Grisaru S, Xie J, Samuel S, *et al*. Associations between hydration status, intravenous fluid administration, and outcomes of patients infected with Shiga toxin-producing escherichia coli: a systematic review and meta-analysis. *JAMA Pediatr* 2017;171:68–76.
10. Tarr PI, Gordon CA, Chandler WL. Shiga-toxin-producing Escherichia coli and haemolytic uraemic syndrome. *Lancet* 2005;365:1073–86.
11. Freedman SB, Vandermeer B, Milne A, *et al*. Diagnosing clinically significant dehydration in children with acute gastroenteritis using noninvasive methods: a meta-analysis. *J Pediatr* 2015;166:908–16.
12. Council of State and Territorial Epidemiologists. *Public health reporting and national notification for Shiga Toxin-Producing Escherichia coli (STEC). In. Vol 17-ID-10*. Atlanta, Georgia, 2017.
13. Health Canada. Medical devices active licenses search. https://health-products.canada.ca/mdall-limh/prepareSearch-preparerRecherche.do?type=active (Accessed 20 Dec 2017).
14. U.S. Food & Drug Administration. Nucleic acid based tests: list of microbial tests. 2018 https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm330711.htm#microbial (Accessed 9 Apr 2018).
15. Grys TE, Sloan LM, Rosenblatt JE, *et al*. Rapid and sensitive detection of Shiga toxin-producing Escherichia coli from nonenriched stool specimens by real-time PCR in comparison to enzyme immunoassay and culture. *J Clin Microbiol* 2009;47:2008–12.
16. Qin X, Klein EJ, Galanakis E, *et al*. Real-Time PCR Assay for detection and differentiation of Shiga toxin-producing Escherichia coli from clinical samples. *J Clin Microbiol* 2015;53:2148–53.
17. Chui L, Patterson-Fortin L, Kuo J, *et al*. Evaluation of enzyme immunoassays and real-time PCR for detecting Shiga toxin-producing Escherichia coli in Southern Alberta, Canada. *J Clin Microbiol* 2015;53:1019–23.
18. Gerritzen A, Wittke JW, Wolff D. Rapid and sensitive detection of Shiga toxin-producing Escherichia coli directly from stool samples by real-time PCR in comparison to culture, enzyme immunoassay and Vero cell cytotoxicity assay. *Clin Lab* 2011;57(11-12):993–8.
19. Buss SN, Leber A, Chapin K, *et al*. Multicenter evaluation of the BioFire FilmArray gastrointestinal panel for etiologic diagnosis of infectious gastroenteritis. *J Clin Microbiol* 2015;53:915–25.
20. Duong VT, Phat VV, Tuyen HT, *et al*. Evaluation of Luminex xTAG Gastrointestinal pathogen panel assay for detection of multiple diarrheal pathogens in fecal samples in Vietnam. *J Clin Microbiol* 2016;54:1094–100.
21. Faron ML, Ledeboer NA, Connolly J, *et al*. Clinical evaluation and cost analysis of Great Basin Shiga toxin direct molecular assay for detection of Shiga Toxin-Producing Escherichia coli in diarrheal stool specimens. *J Clin Microbiol* 2017;55:519–25.

22. Goldenberg SD, Bacelar M, Brazier P, *et al*. A cost benefit analysis of the Luminex xTAG gastrointestinal pathogen panel for detection of infectious gastroenteritis in hospitalised patients. *J Infect* 2015;70:504–11.
23. Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
24. McInnes MDF, Moher D, Thombs BD, *et al*. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018;319:388–96.
25. Whiting PF, Rutjes AW, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
26. Bossuyt PM, Reitsma JB, Bruns DE, *et al*. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* 2003;138:40–4.
27. Wade R, Corbett M, Eastwood A. Quality assessment of comparative diagnostic accuracy studies: our experience using a modified version of the QUADAS-2 tool. *Res Synth Methods* 2013;4:280–6.
28. Santaguida PL, Riley CR, Matchar DB. Assessing risk of bias as a domain of quality in medical test studies. AHRQ Publication No. 12-EHC077-EF. In: *Methods guide for medical test reviews (AHRQ Publication No. 12-EHC017)*. Rockville, MD: Agency for Healthcare Research and Quality, 2012.
29. Reitsma JB, Glas AS, Rutjes AW, *et al*. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982–90.
30. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;20:2865–84.
31. Leeflang MM. Systematic reviews and meta-analyses of diagnostic test accuracy. *Clin Microbiol Infect* 2014;20:105–13.
32. *R: A Language and Environment for Statistical Computing [computer program]*. Vienna, Austria: R Foundation for Statistical Computing, 2017.
33. The Nordic Cochrane Centre, The Cochrane Collaboration. *Review Manager (RevMan) [computer program]. Version 5.3*. Copenhagen, 2014.
34. Schünemann HJ, Oxman AD, Brozek J, *et al*. GRADE: assessing the quality of evidence for diagnostic recommendations. *Evid Based Med* 2008;13:162–3.
35. Schünemann HJ, Schünemann AH, Oxman AD, *et al*. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;336:1106–10.