PROTOCOL   **Open Access**

The comparative diagnostic accuracy of the Mini Mental State Examination (MMSE) and the General Practitioner assessment of Cognition (GPCOG) for identifying dementia in primary care: a systematic review protocol

Harriet A. Hunt[1]*, Sanne Van Kampen[2], Yemisi Takwoingi[3], David J. Llewellyn[4], Mark Pearson[1] and Christopher J. Hyde[5]

Abstract

Background: Improved dementia identification is a global health priority, and general practitioners (GPs) are often the first point of contact for people with concerns about their cognition. However, GPs often express uncertainty in using assessment tools and the evidence based on which tests are most accurate in identifying dementia is unclear. In particular, there is little certainty around how the accuracy of available brief cognitive assessments compares within a clinical family practice setting.

Grounded in existing brief cognitive assessment evidence, we will compare the diagnostic test accuracy of the Mini Mental State Examination (MMSE) to the General Practitioner Assessment of Cognition (GPCOG) against the best available reference standard when used within a family practice setting.

Methods: We will employ robust systematic review methods to assess studies of diagnostic accuracy where both the MMSE and GPCOG have been evaluated as direct comparisons, i.e. within the same study population. This approach will enable us to minimise between-study heterogeneity, to eliminate the risk of bias due to confounding and increase the opportunity to make clinically useful and useable comparisons of diagnostic accuracy across both the MMSE and GPCOG. This systematic review will be conducted using a pragmatic search strategy, refining searches that build upon studies identified as part of our overview of systematic reviews of the diagnostic accuracy of brief cognitive assessments for identifying dementia in primary care.

Discussion: Through this systematic review, we aim to improve existing evidence on how the diagnostic accuracy of MMSE and GPCOG compares when used to identify dementia within the family practice setting. We also aim to make clinical practice recommendations based upon the variations in diagnostic accuracy identified between the MMSE and GPCOG.

Keywords: Dementia, Brief cognitive assessment, Diagnostic accuracy, Primary care, Systematic review

---

* Correspondence: h.a.hunt@exeter.ac.uk
[1]National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South West Peninsula, Institute of Health Research, University of Exeter Medical School, St Luke's Campus, Exeter, Devon EX1 1TE, UK
Full list of author information is available at the end of the article

Hunt *et al. Diagnostic and Prognostic Research* (2017) 1:14

Page 2 of 6

Background

Improved dementia diagnosis is a global health priority of international bodies such as the World Health Organization [1] and the G8 [2]. Existing delays in the diagnostic pathway have led to debates around case finding and targeted screening within primary care [3–6]. General practitioners (GPs) are often the first point of contact for people with concerns about their cognition, yet GPs often express uncertainty in using assessment tools alongside concerns around the consequences of misdiagnosing dementia [5, 7, 8]. In established healthcare systems, guidelines on the most accurate brief cognitive assessment for identifying dementia in primary care are inconsistent and variable in their specific recommendations. Whilst there is variation in guidance on thresholds, accuracy and suitability of test within different populations, guidelines often feature the same subset of brief cognitive assessments. Examples are available in the UK from the National Institute for Health and Care Excellence [9] and the Royal College of Psychiatrists [10], and in the Netherlands from the Huisartsen Genootschap (GP Society) [11]. These all include the Mini Mental State Examination (MMSE) and the General Practitioner Assessment of Cognition (GPCOG).

A number of systematic reviews [12–18] have explored the individual diagnostic accuracy of brief cognitive assessments for dementia in isolation, and across a range of populations and settings. In an overview of systematic reviews of the diagnostic accuracy of brief cognitive assessments for identifying dementia in primary care, we identified two brief cognitive assessments that can be compared to identify the test with better diagnostic performance. These tests were two of the three most frequently assessed brief cognitive assessments within the 13 included systematic reviews [12, 13, 19–29] with the MMSE featuring in 8 reviews and the GPCOG featuring in 4 reviews. The clock drawing test (CDT) was the third most frequently assessed tool, featuring in 4 reviews. We judged this to be less comparable to the MMSE in terms of administration complexity, timing and domains assessed, compared to the GPCOG. As the most frequently assessed test within our overview, the MMSE is also included as one of the index tests within this review as, whilst copyright restrictions are now enforced, it remains one of the most popular brief cognitive assessments employed in practice [27, 30]. The MMSE is based on a 30-point scale of 11 questions testing five domains of cognitive function (orientation, registration, attention and calculation, recall and language) [31]. The GPCOG was the second most frequently assessed index test within our overview. The GPCOG is a publicly available test that has two sections: a patient examination (GPCOG-patient) with a maximum score of 9 (optimum performance) covering time orientation, clock drawing, reporting recent events and a word-recall task, and an optional informant questionnaire (GPCOG-informant) with a maximum score of 6 with questions assessing the patient's memory of recent events and their executive function [32]. In comparison, the CDT is a standard assessment where the patient is asked to draw a clock face marking the hours and then draw the hour and minute hands to correctly indicate a specific time (e.g. quarter past 3). There are a number of scoring approaches, but the Shulman method uses a 6-point scoring system [33] whilst the Sunderland method uses a 10-point system [34]. Taking into consideration the ubiquity of the index tests, the comparability of the tests mentioned above and their common use within guidelines, we have chosen to compare the MMSE against the GPCOG as index tests within this systematic review. Therefore, the aim of this systematic review is to compare the diagnostic accuracy evidence of the MMSE and the GPCOG for identifying dementia, particularly within a primary care setting and using direct (within study) comparisons.

This use of direct comparisons should reduce between-study heterogeneity and allow us to draw firm conclusions about the comparative accuracy of these brief cognitive assessments within the same or similar populations [35, 36]. To our knowledge, this type of systematic review has not previously been conducted to compare the accuracy of brief cognitive assessments for identifying dementia.

This evidence will contribute strongly to clinical practice and policy making by demonstrating the presence or absence of superiority in the diagnostic accuracy of GPCOG relative to that of MMSE for identifying dementia in primary care.

Methods

The primary outcome is the comparative accuracy of the two tests assessed via direct comparisons, i.e. the diagnostic accuracy of the two tests are compared within the same population in a study (comparative study).

The secondary outcome of the review is to identify other common test-related factors identified by included studies, such as ease of administration or administration time. Whilst beyond our primary focus of test accuracy, these other factors may contribute to the overall usefulness of the tests when applied in a primary care setting, and we will incorporate them in our findings in order to make useful research and clinical recommendations.

This systematic review will be conducted using a pragmatic search strategy, refining searches that build upon studies identified as part of our overview of systematic reviews of the diagnostic accuracy of brief cognitive assessments for identifying dementia in primary care. Further details are given below (PROSPERO reference 42015022078).

Hunt *et al. Diagnostic and Prognostic Research* (2017) 1:14

Page 3 of 6

Overview search methods

To build the search database for the overview of systematic reviews of the diagnostic accuracy of brief cognitive assessments for identifying dementia in primary care, we searched the Cochrane Database of Systematic Reviews, EMBASE, MEDLINE and PsychINFO for systematic reviews from inception until August 2015. Search strategies are shown in the Additional file 1. According to best searching practice for diagnostic accuracy reviews, we applied no date or language restrictions, and where reviews were updated, we used the latest version available. Additional papers were identified through Zetoc alerts and incorporated at the title and abstract screening phase. We ran updated searches on the Cochrane Database of Systematic Reviews in February 2016.

Eligibility criteria

Adults aged 18 years or over recruited from a primary care or general practice population were included, and we did not exclude patients who were selected on the basis of an existing diagnosis or condition which might reasonably be expected to feature in primary care (e.g. stroke).

The target condition was all-cause (non-differentiated) dementia. We also included reviews that focused specifically on differentiated forms of dementia such as Alzheimer's disease, vascular dementia and dementia with Lewy bodies. We excluded reviews that focused on mild cognitive impairment (MCI). Where reviews investigated both dementia and MCI, we extracted data referring to dementia and excluded data referring solely to MCI.

Identification of studies for this systematic review

To identify eligible studies for this systematic review, we will first assess the 13 systematic reviews included within our overview review (methods described above) and identify included reviews that contained studies including direct comparisons of the diagnostic accuracy of MMSE and GPCOG for identifying dementia in primary care. Once we have identified these studies, we will carry out citation tracking via Google Scholar, i.e. clicking on the appropriate link (e.g. "cited by 15") to view details on the articles that have cited the original study. We will also use these initial studies to conduct snowball searching, i.e. checking the bibliographies for relevant original studies for possible inclusion within this systematic review. We will use Zetoc alerts to proactively identify recent studies published that meet our criteria (using the terms "MMSE", "GPCOG", "test accuracy" and "dementia"). Finally, when we have identified studies using the above methods, we will conduct a traditional search taking a start date 1 year prior to the most recently published identified study up to the current day, using MEDLINE, EMBASE and PsychINFO databases. The rationale is that this search will cover the maximum period of time not covered in the overview review searches with some date overlap to ensure all potential sources are included, using the most efficient means to identify the most recent evidence. This will also enable us to confirm whether we identified all relevant studies via the overview searches.

Index tests

The index tests are the MMSE [31] and the GPCOG [32]. The MMSE is one of the most widely used brief cognitive assessments currently used, and development of the GPCOG has been independent to the development of the MMSE.

The conventional threshold for the MMSE is 24 (also shown as <24), where out of a maximum possible 30 points, scores below 24 indicate impairment [22]. The GPCOG comprises of two sections: the section completed by the individual being assessed, known as GPCOG-patient, and an optional section for a relative or friend to complete (if present) known as GPCOG-informant. GPCOG-patient has 9 items with possible total scores of between 0 (indicating severe impairment) and 9 (indicating no impairment). GPCOG-informant has 6 items with possible total scores of between 0 (indicating severe impairment) and 6 (indicating no impairment). GPCOG-patient can be conducted by itself, with a conventional threshold of 8 out of 9 (<8). If informants are available, a score of GPCOG-patient between 5 and 8 precipitates the GPCOG-informant and the scores are combined ("GPCOG-total") with a conventional threshold of 11 out of a maximum 15 (<11). If no informant is available, the conventional threshold of 8 stands. It is also possible to conduct a staged GPCOG assessment where GPCOG-informant is only required if GPCOG-patient is scored between 5 and 8 out of 9. This is known as "GPCOG Two stage".

For our assessment, we will stratify GPCOG into 3 types of test: GPCOG-patient with a threshold of <8, GPCOG-total with a threshold of <11 and GPCOG Two stage [37].

Reference standard

There is currently no gold standard test for identifying dementia in primary care. We will accept reference standards consisting of the following tools alone, clinical diagnosis alone or clinical diagnosis combined with one or a combination of the following assessment tools:

- Diagnostic and Statistical Manual (DSM) III/III-R/ IV/IV-R,
- Clinical Dementia Rating (CDR),
- International Classification of Diseases (ICD) 10,
- Geriatric Mental State–Automated Geriatric Examination for Computer Assisted Taxonomy (GMS-AGECAT),

Hunt *et al. Diagnostic and Prognostic Research* (2017) 1:14

Page 4 of 6

- Cambridge Mental Disorders of the Elderly Examination (CAMDEX),
- International Psychogeriatric Association World Health Organization (IPA-WHO) criteria.

Reference standards are selected on the basis of many variables such as common practice within individual clinics, practitioner preference, specialisation and experience of healthcare professionals and practice managers and are subject to changes in cost and fashion. Many of the globally accepted reference standards such as the World Health Organization-supported ICD and the DSM produced by the American Psychiatric Association are updated regularly; the DSM-5 (sometimes referred to as DSM-V) was released in 2013 [38], and the ICD-11 is due for release by 2018 [39].

Data extraction, selection and coding

All sources will be managed using the latest version of EndNote software. Two reviewers will pilot the screening for titles and abstracts on the first 15 sources, and we will write screening notes to help with title/abstract and full-text screening. Title/abstract and full-text screening will be conducted by the same two reviewers, and a third reviewer will resolve any disagreements.

A bespoke data abstraction form will be piloted by two reviewers using two included studies. Key data extracted will include characteristics of included systematic reviews (references and author details, overall goal of review, date review conducted, date published, participant details), included study details (such as authors, year of study, date of publication, country of study, outcomes reported, test timings) and general review limitations as well as components of the 2×2 table (TP, FP, TN, FN) or other accuracy data such as sensitivity, specificity and disease prevalence if raw numbers are not available. The data abstraction form will be accompanied by a briefing document explaining how it should be used. Data will be abstracted by one reviewer, spot-checked by a second, with a third reviewer acting as moderator if necessary.

Assessment of methodological quality

We will use the QUADAS-2 [40] tool to assess methodological quality of diagnostic accuracy studies for systematic reviews. Whilst this tool is developed for studies focussing on a single index test, we will assess the suitability of using the tool for studies that focus on direct comparisons of two index tests by piloting the QUADAS-2 tool on one of the included studies. We will tailor QUADAS-2 in line with suitability in assessing quality of studies using direct comparisons, for example assessing the reference standard against MMSE and then the reference standard against GPCOG.

Data synthesis and analysis

Study of specific estimates of the sensitivity and specificity (and their 95% confidence intervals) of GPCOG and MMSE will be presented graphically on a forest plot. We will also use these forest plots and summary receiver operating characteristic (SROC) plots to visually explore heterogeneity.

We will consider possible sub-group analyses investigating, for example tests using lower and higher thresholds. Other aspects that may be suitable for investigating through sub-grouping could include variations in population details such as prevalence, and variations in cases and control groups (e.g. confirmed dementia, probably dementia, people with memory problems, healthy people).

We will perform meta-analysis if the quantity and nature of the included studies permit. Again, if data allow, we will use a hierarchical meta-regression model with test type as a covariate to estimate and compare SROC curves or summary points [36]. A priori uncertainty about thresholds for determining test positivity and the likelihood of implicit thresholds suggests estimation of SROC curves using a hierarchical SROC (HSROC) meta-regression model may be preferable [41]. However, we will consider using a bivariate meta-regression model to estimate and compare summary points [42, 43] if studies use a common threshold.

We will create a summary of result table with additional summary tables of subgroup results (potential subgroups listed above) if relevant. If feasible and appropriate, we will consider translating any summary results into natural frequencies and other metrics such as predictive values to help improve understanding by readers.

We will not assess reporting bias because its impact on diagnostic accuracy is unclear, and the tools for investigating it are in the early stages of development [44].

Discussion

We do not foresee any practical or operational issues with the conduct of this systematic review. All differences between the protocol and systematic review will be reported in the full systematic review.

Additional file

Additional file 1: Search strategy, formatted for EMBASE (OVID). (DOCX 24 kb)

Hunt *et al. Diagnostic and Prognostic Research* (2017) 1:14

Page 5 of 6

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions
HH created the first draft, and all authors contributed to, read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South West Peninsula, Institute of Health Research, University of Exeter Medical School, St Luke's Campus, Exeter, Devon EX1 1TE, UK. [2]National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South West Peninsula, Plymouth University Peninsula Schools of Medicine and Dentistry, ITTC building, Tamar Science Park, Drake Circus, Plymouth, Devon PL4 8AA, UK. [3]Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [4]University of Exeter Medical School, St Luke's Campus, Exeter, Devon EX1 1TE, UK. [5]Institute of Health Research, University of Exeter Medical School, St Luke's Campus, Exeter, Devon EX1 1TE, UK.

References
1. WHO, ADI. Dementia: a public health priority. Geneva: World Health Organization; 2012. viii + 102 pp.
2. G8 U. G8 dementia summit declaration. In: Office DoHaPMs, editor. online: https://www.gov.uk/government/publications/g8-dementia-summit-agreements: Department of Health; 2013. Accessed 24 Apr 2017.
3. Dowrick A, Southern A. Dementia 2014: opportunity for change. 2014.
4. Le Couteur D, Doust J, Creasey H, Brayne C. Political drive to screen for predementia: not evidence based and ignores the harms of diagnosis. BMJ. 2013;347(7925):f5125.
5. Fox C, Lafortune L, Boustani M, Brayne C. The pros and cons of early diagnosis in dementia. Br J Gen Pract. 2013;63(612):e510-2.
6. Burns A. Alistair Burns and 51 colleagues reply to David Le Couteur and colleagues. BMJ. 2013;347(oct15_6):f6125.
7. Ballard C, Bannister C. Criteria for the diagnosis of dementia. In: David A, ABaJOB, editors. Dementia. 4th ed. London: Hodder; 2010. p. 794.
8. Robinson L, Tang E, Taylor JP. Dementia: timely diagnosis and early intervention. Bmj. 2015;350:1-6.
9. NICE Pathways. Dementia diagnosis and assessment. 2016. p. 10.
10. Barker A, Arya P, Boston P, Fawzi W, Lennon S, Mikashi S, et al. OP86: Individual patient outcome measures recommended for use in older people's mental health. London: Online: Royal College of Psychiatrists; 2012. Contract No.: OP86.
11. Moll van Charante E, Perry M Vernooij-Dassen MJFJ, Boswijk DFR, Stoffels J, L Achthoven, MN L-K. Dementia Summary Map M21 2017 Website: https://www.nhg.org/standaarden/samenvatting/dementie. Accessed 24 May 2017.
12. Arevalo-Rodriguez I, Smailagic N, Roque IFM, Ciapponi A, Sanchez-Perez E, Giannakou A, et al. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev. 2015;(3)1:Cd010783.
13. Creavin ST, Wisniewski S, Noel-Storr AH, Trevelyan CM, Hampton T, Rayment D, Thom VM, Nash KJ, Elhamoui H, Milligan R, Patel AS. Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. The Cochrane Library. 2016;(1):1-183.
14. Davis DH, Creavin ST, Yip JL, Noel-Storr AH, Brayne C, Cullum S. Montreal Cognitive Assessment for the diagnosis of Alzheimer's disease and other dementias. The Cochrane Library. 2015;(1):1-50.
15. Fage BA, Chan CC, Gill SS, Noel-Storr AH, Herrmann N, Smailagic N, Nikolaou V, Seitz DP. Mini-Cog for the diagnosis of Alzheimer's disease dementia and other dementias within a community setting. The Cochrane Library. 2015;(1):1-33.
16. Harrison JK, Fearon P, Noel-Storr AH, McShane R, Stott DJ, Quinn TJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting. Cochrane Database Syst Rev. 2015;(3)1:Cd010772.
17. Hendry K, Lees Rosalind A, McShane R, Noel-Storr Anna H, Stott David J, Quinn Terry J. AD-8 for diagnosis of dementia across a variety of healthcare settings. Cochrane Database Syst Rev. 2014 [cited 1 1]; (5). Available from: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD011121/abstract. http://onlinelibrary.wiley.com/store/10.1002/14651858.CD011121/asset/v=1&t=idlm4c21&s=8dafa435a321ce12c678db51152a73d7cc4fcb37. Accessed 24 Apr 2017.
18. Quinn TJ, Fearon P, Noel-Storr AH, Young C, McShane R, Stott DJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within community dwelling populations. The Cochrane Library. 2014;(1):1-91.
19. Brodaty H, Low L-F, Gibson L, Burns K. What is the best dementia screening instrument for general practitioners to use? Am J Geriatr Psychiatry. 2006; 14(5):391–400.
20. Carnero-Pardo C, Espejo-Martinez B, Lopez-Alcalde S, Espinosa-Garcia M, Saez-Zea C, Vilchez-Carrillo R, et al. Effectiveness and costs of phototest in dementia and cognitive impairment screening. BMC Neurol. 2011;11(1):92.
21. Carnero-Pardo C, Lopez-Alcalde S, Allegri RF, Russo MJ. A systematic review and meta-analysis of the diagnostic accuracy of the Phototest for cognitive impairment and dementia. Dement Neuropsychologia. 2014;8(2):141–7.
22. Mitchell AJ. A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. J Psychiatr Res. 2009;43(4):411–31.
23. Mitchell AJ, Malladi S. Screening and case finding tools for the detection of dementia. Part I: evidence-based meta-analysis of multidomain tests. Am J Geriatr Psychiatry. 2010;18(9):759–82.
24. Mitchell AJ, Malladi S. Screening and case-finding tools for the detection of dementia. Part II: evidence-based meta-analysis of single-domain tests. Am J Geriatr Psychiatry. 2010;18(9):783–800.
25. Harrison JK, Fearon P, Noel-Storr AH, McShane R, Stott DJ, Quinn TJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a general practice (primary care) setting. The Cochrane Library. 2014;(1):1-57.
26. Naqvi RM, Haider S, Tomlinson G, Alibhai S. Cognitive assessments in multicultural populations using the Rowland Universal Dementia Assessment Scale: a systematic review and meta-analysis. Can Med Assoc J. 2015;187(5):E169–75.

Hunt *et al. Diagnostic and Prognostic Research* (2017) 1:14

Page 6 of 6

27. Lin JS, O'Connor E, Rossom RC, Perdue LA, Eckstrom E. Screening for cognitive impairment in older adults: a systematic review for the U.S. preventive services task force. Ann Intern Med. 2013;159(9):601–12.

28. Tsoi KK, Chan JY, Hirai HW, Wong SY, Kwok TC. Cognitive tests to detect dementia: a systematic review and meta-analysis. JAMA internal medicine. 2015;175(9):1450-8.

29. Woodford H, George J. Cognitive assessment in the elderly: a review of clinical methods. QJM. 2007;100(8):469–84.

30. Ismail Z, Rajji TK, Shulman KI. Brief cognitive screening instruments: an update. Int J Geriatr Psychiatry. 2010;25(2):111–20.

31. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189–98.

32. Brodaty H, Pond D, Kemp NM, Luscombe G, Harding L, Berman K, et al. The GPCOG: a new screening test for dementia designed for general practice. J Am Geriatr Soc. 2002;50(3):530–4.

33. Shulman K, Shedletsky R, Silver I. The challenge of time: clock-drawing and cognitive function in the elderly. Int J Geriatr Psychiatry. 1986;1:135–40.

34. Sunderland T, Hill J, Mellow A, Lawlor B, Gundersheimer J, Newhouse P, et al. Clock drawing in Alzheimer's disease: a novel measure of dementia severity. J Am Geriatr Soc. 1989;37:725–9.

35. Takwoingi Y, Leeflang M, Deeks J. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. Ann Intern Med. 2013; 158(7):544–54.

36. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Analysing and presenting results. 2010. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 10 [Internet]. Available from: http://srdta. cochrane.org/handbook-dta-reviews. Accessed 24 May 2017.

37. Brodaty H, Connors MH, Loy C, Teixeira-Pinto A, Stocks N, Gunn J, et al. Screening for dementia in primary care: a comparison of the GPCOG and the MMSE. Dement Geriatr Cogn Disord. 2016;42(5-6):323–30.

38. APA. Diagnostic and Statistical Manual of Mental Disorders (DSM–5) American Psychiatric Association website 2017 Available from: https://www. psychiatry.org/psychiatrists/practice/dsm. Accessed 24 May 2017.

39. WHO. The 11th Revision of the International Classification of Diseases (ICD-11) is due by 2018 2017 [Available from: http://www.who.int/classifications/ icd/revision/en/. Accessed 24 May 2017.

40. Whiting PF, Rutjes AS, Westwood ME, et al. Quadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011; 155(8):529–36.

41. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001;20(19):2865–84.

42. Hamza TH, Reitsma JB, Stijnen T. Meta-analysis of diagnostic studies: a comparison of random intercept, normal-normal and binomial-normal bivariate Summary ROC approaches. Med Decis Making. 2008;28.

43. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982–90.

44. Hamza TH, Reitsma JB, Stijnen T. Meta-analysis of diagnostic studies: a comparison of random intercept, normal-normal, and binomial-normal bivariate summary ROC approaches. Medical Decision Making. 2008;28(5): 639-49.