Appendix A: List of included studies

1. Arbyn M, Schenck U, Ellison E, et al. Metaanalysis of the accuracy of rapid prescreening relative to full screening of pap smears. Cancer 2003 Feb 25;99(1):9-16.
2. Bachmann LM, Kolb E, Koller MT, et al. Accuracy of Ottawa ankle rules to exclude fractures of the ankle and mid-foot: systematic review. BMJ 2003 Feb 22;326(7386):417.
3. Badgett RG, Lucey CR, Mulrow CD. Can the clinical examination diagnose left-sided heart failure in adults? JAMA 1997 Jun 4;277(21):1712-9.
4. Bafounta ML, Beauchet A, Aegerter P, et al. Is dermoscopy (epiluminescence microscopy) useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests. Arch Dermatol 2001 Oct;137(10):1343-50.
5. Bailey JJ, Berson AS, Handelsman H, et al. Utility of current risk stratification tests for predicting major arrhythmic events after myocardial infarction. J Am Coll Cardiol 2001 Dec;38(7):1902-11.
6. Balk EM, Ioannidis JP, Salem D, et al. Accuracy of biomarkers to diagnose acute cardiac ischemia in the emergency department: a meta-analysis. Ann Emerg Med 2001 May;37(5):478-94.
7. Barnes CJ, Pietrobon R, Higgins LD. Does the pulse examination in patients with traumatic knee dislocation predict a surgical arterial injury? A meta-analysis. J Trauma 2002 Dec;53(6):1109-14.
8. Bax JJ, Wijns W, Cornel JH, et al. Accuracy of currently available techniques for prediction of functional recovery after revascularization in patients with left ventricular dysfunction due to chronic coronary artery disease: comparison of pooled data. J Am Coll Cardiol 1997 Nov 15;30(6):1451-60.
9. Bax JJ, Poldermans D, Elhendy A, et al. Sensitivity, specificity, and predictive accuracies of various noninvasive techniques for detecting hibernating myocardium. Curr Probl Cardiol 2001 Feb;26(2):147-86.
10. Berger MY, van d, V, Lijmer JG, et al. Abdominal symptoms: do they predict gallstones? A systematic review. Scand J Gastroenterol 2000 Jan;35(1):70-6.
11. Bonis PA, Ioannidis JP, Cappelleri JC, et al. Correlation of biochemical response to interferon alfa with histological improvement in hepatitis C: a meta-analysis of diagnostic test characteristics. Hepatology 1997 Oct;26(4):1035-44.
12. Brown DL, Doubilet PM. Transvaginal sonography for diagnosing ectopic pregnancy: positivity criteria and performance characteristics. J Ultrasound Med 1994 Apr;13(4):259-66.
13. Brown MD, Rowe BH, Reeves MJ, et al. The accuracy of the enzyme-linked immunosorbent assay D-dimer test in the diagnosis of pulmonary embolism: a meta-analysis. Ann Emerg Med 2002 Aug;40(2):133-44.
14. Brown MD, Lau J, Nelson RD, et al. Turbidimetric D-dimer test in the diagnosis of pulmonary embolism: a metaanalysis. Clin Chem 2003 Nov;49(11):1846-53.
15. Carlson KJ, Skates SJ, Singer DE. Screening for ovarian cancer. Ann Intern Med 1994 Jul 15;121(2):124-32.
16. Carter BG, Butt W. Review of the use of somatosensory evoked potentials in the prediction of outcome after severe brain injury. Crit Care Med 2001 Jan;29(1):178-86.
17. Chappell ET, Moure FC, Good MC. Comparison of computed tomographic angiography with digital subtraction angiography in the diagnosis of cerebral aneurysms: a meta-analysis. Neurosurgery 2003 Mar;52(3):624-31.
18. Cher DJ, Conwell JA, Mandel JS. MRI for detecting silicone breast implant rupture: meta-analysis and implications. Ann Plast Surg 2001 Oct;47(4):367-80.
19. Chien PF, Khan KS, Ogston S, et al. The diagnostic accuracy of cervico-vaginal fetal fibronectin in predicting preterm delivery: an overview. Br J Obstet Gynaecol 1997 Apr;104(4):436-44.
20. Chien PF, Arnott N, Gordon A, et al. How useful is uterine artery Doppler flow velocimetry in the prediction of pre-eclampsia, intrauterine growth retardation and perinatal death? An overview. BJOG 2000 Feb;107(2):196-208.

21. Chin AS, Goldman LE, Eisenberg MJ. Functional testing after coronary artery bypass graft surgery: a meta-analysis. Can J Cardiol 2003 Jun;19(7):802-8.
22. Clark TJ, Mann CH, Shah N, et al. Accuracy of outpatient endometrial biopsy in the diagnosis of endometrial hyperplasia. Acta Obstet Gynecol Scand 2001 Sep;80(9):784-93.
23. Clarke CE, Davies P. Systematic review of acute levodopa and apomorphine challenge tests in the diagnosis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 2000 Nov;69(5):590-4.
24. Colin C, Lanoir D, Touzet S, et al. Sensitivity and specificity of third-generation hepatitis C virus antibody detection assays: an analysis of the literature. J Viral Hepat 2001 Mar;8(2):87-95.
25. Conde-Agudelo A, Kafury-Goeta AC. Triple-marker test as screening for Down syndrome: a meta-analysis. Obstet Gynecol Surv 1998 Jun;53(6):369-76.
26. Contopoulos-Ioannidis DG, Ioannidis JP. Maternal cell-free viremia in the natural history of perinatal HIV-1 transmission: a meta-analysis. J Acquir Immune Defic Syndr Hum Retrovirol 1998 Jun 1;18(2):126-35.
27. Cook DJ, Fitzgerald JM, Guyatt GH, et al. Evaluation of the protected brush catheter and bronchoalveolar lavage in the diagnosis of nosocomial pneumonia. J Intensive Care Med 1991 Jul;6(4):196-205.
28. Dales RE, Stark RM, Raman S. Computed tomography to stage lung cancer. Approaching a controversy using meta-analysis. Am Rev Respir Dis 1990 May;141(5 Pt 1):1096-101.
29. de Albuquerque FL, Picano E. Comparison of dipyridamole and exercise stress echocardiography for detection of coronary artery disease (a meta-analysis). Am J Cardiol 2001 May 15;87(10):1193-6.
30. de Bernardinis, Violi V, Roncoroni L, et al. Discriminant power and information content of Ranson's prognostic signs in acute pancreatitis: a meta-analytic study. Crit Care Med 1999 Oct;27(10):2272-83.
31. de Jaeger, Litalien C, Lacroix J, et al. Protected specimen brush or bronchoalveolar lavage to diagnose bacterial nosocomial pneumonia in ventilated adults: a meta-analysis. Crit Care Med 1999 Nov;27(11):2548-60.
32. de Vries SO, Hunink MG, Polak JF. Summary receiver operating characteristic curves as a technique for meta-analysis of the diagnostic performance of duplex ultrasonography in peripheral arterial disease. Acad Radiol 1996 Apr;3(4):361-9.
33. Delgado-Bolton RC, Fernandez-Perez C, Gonzalez-Mate A, et al. Meta-analysis of the performance of 18F-FDG PET in primary tumor detection in unknown primary tumors. J Nucl Med 2003 Aug;44(8):1301-14.
34. Detrano R, Janosi A, Lyons KP, et al. Factors affecting sensitivity and specificity of a diagnostic test: the exercise thallium scintigram. Am J Med 1988 Apr;84(4):699-710.
35. Deville WL, van der Windt DA, Dzaferagic A, et al. The test of Lasegue: systematic review of the accuracy in diagnosing herniated discs. Spine 2000 May 1;25(9):1140-7.
36. Dijkhuizen FP, Mol BW, Brolmann HA, et al. The accuracy of endometrial sampling in the diagnosis of patients with endometrial carcinoma and hyperplasia: a meta-analysis. Cancer 2000 Oct 15;89(8):1765-72.
37. Dove SB. Radiographic diagnosis of dental caries. J Dent Educ 2001 Oct;65(10):985-90.
38. Dwamena BA, Sonnad SS, Angobaldo JO, et al. Metastases from non-small cell lung cancer: mediastinal staging in the 1990s--meta-analytic comparison of PET and CT. Radiology 1999 Nov;213(2):530-6.
39. Ebell MH, Flewelling D, Flynn CA. A systematic review of troponin T and I for diagnosing acute myocardial infarction. J Fam Pract 2000 Jun;49(6):550-6.
40. Ebell MH, White LL, Weismantel D. A systematic review of troponin T and I values as a prognostic tool for patients with chest pain. J Fam Pract 2000 Aug;49(8):746-53.
41. Engels EA, Terrin N, Barza M, et al. Meta-analysis of diagnostic tests for acute sinusitis. J Clin Epidemiol 2000 Aug;53(8):852-62.
42. Engman ML. An update on EBCT (Ultrafast CT) scans for coronary artery disease. J Insur Med 1998;30(3):175-9.
43. Fahey MT, Irwig L, Macaskill P. Meta-analysis of Pap test accuracy. Am J Epidemiol 1995 Apr 1;141(7):680-9.
44. Faron G, Boulvain M, Irion O, et al. Prediction of preterm delivery by fetal fibronectin: a meta-analysis. Obstet Gynecol 1998 Jul;92(1):153-8.

45. Farquhar C, Ekeroma A, Furness S, et al. A systematic review of transvaginal ultrasonography, sonohysterography and hysteroscopy for the investigation of abnormal uterine bleeding in premenopausal women. Acta Obstet Gynecol Scand 2003 Jun;82(6):493-504.
46. Fleischmann KE, Hunink MG, Kuntz KM, et al. Exercise echocardiography or exercise SPECT imaging? A meta-analysis of diagnostic test performance. JAMA 1998 Sep 9;280(10):913-20.
47. Fraile M, Rull M, Julian FJ, et al. Sentinel node biopsy as a practical alternative to axillary lymph node dissection in breast cancer patients: an approach to its validity. Ann Oncol 2000 Jun;11(6):701-5.
48. Gianrossi R, Detrano R, Colombo A, et al. Cardiac fluoroscopy for the diagnosis of coronary artery disease: a meta analytic review. Am Heart J 1990 Nov;120(5):1179-88.
49. Gill CJ, Lau J, Gorbach SL, et al. Diagnostic accuracy of stool assays for inflammatory bacterial gastroenteritis in developed and resource-poor countries. Clin Infect Dis 2003 Aug 1;37(3):365-75.
50. Gisbert JP, Pajares JM. Diagnosis of Helicobacter pylori infection by stool antigen determination: a systematic review. Am J Gastroenterol 2001 Oct;96(10):2829-38.
51. Glas AS, Roos D, Deutekom M, et al. Tumor markers in the diagnosis of primary bladder cancer. A systematic review. J Urol 2003 Jun;169(6):1975-82.
52. Goodman CM, Cohen V, Thornby J, et al. The life span of silicone gel breast implants and a comparison of mammography, ultrasonography, and magnetic resonance imaging in detecting implant rupture: a meta-analysis. Ann Plast Surg 1998 Dec;41(6):577-85.
53. Gordon I, Barkovics M, Pindoria S, et al. Primary vesicoureteric reflux as a predictor of renal damage in children hospitalized with urinary tract infection: a systematic review and meta-analysis. J Am Soc Nephrol 2003 Mar;14(3):739-44.
54. Gould MK, Maclean CC, Kuschner WG, et al. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. JAMA 2001 Feb 21;285(7):914-24.
55. Gould MK, Kuschner WG, Rydzak CE, et al. Test performance of positron emission tomography and computed tomography for mediastinal staging in patients with non-small-cell lung cancer: a meta-analysis. Ann Intern Med 2003 Dec 2;139(11):879-92.
56. Greco S, Girardi E, Masciangelo R, et al. Adenosine deaminase and interferon gamma measurements for the diagnosis of tuberculous pleurisy: a meta-analysis. Int J Tuberc Lung Dis 2003 Aug;7(8):777-86.
57. Hallan S, Asberg A. The accuracy of C-reactive protein in diagnosing acute appendicitis--a meta-analysis. Scand J Clin Lab Invest 1997 Aug;57(5):373-80.
58. Harvey RT, Gefter WB, Hrung JM, et al. Accuracy of CT angiography versus pulmonary angiography in the diagnosis of acute pulmonary embolism: evaluation of the literature with summary ROC curve analysis. Acad Radiol 2000 Oct;7(10):786-97.
59. Hobby JL, Tom BD, Bearcroft PW, et al. Magnetic resonance imaging of the wrist: diagnostic performance statistics. Clin Radiol 2001 Jan;56(1):50-7.
60. Hoffman RM, Clanon DL, Littenberg B, et al. Using the free-to-total prostate-specific antigen ratio to detect prostate cancer in men with nonspecific elevations of prostate-specific antigen levels. J Gen Intern Med 2000 Oct;15(10):739-48.
61. Honest H, Bachmann LM, Gupta JK, et al. Accuracy of cervicovaginal fetal fibronectin test in predicting risk of spontaneous preterm birth: systematic review. BMJ 2002 Aug 10;325(7359):301.
62. Hoogendam A, Buntinx F, de Vet HC. The diagnostic value of digital rectal examination in primary care screening for prostate cancer: a meta-analysis. Fam Pract 1999 Dec;16(6):621-6.
63. Hrung JM, Sonnad SS, Schwartz JS, et al. Accuracy of MR imaging in the work-up of suspicious breast lesions: a diagnostic meta-analysis. Acad Radiol 1999 Jul;6(7):387-97.
64. Huicho L, Campos M, Rivera J, et al. Fecal screening tests in the approach to acute infectious diarrhea: a scientific overview. Pediatr Infect Dis J 1996 Jun;15(6):486-94.
65. Huicho L, Campos-Sanchez M, Alamo C. Metaanalysis of urine screening tests for determining the risk of urinary tract infection in children. Pediatr Infect Dis J 2002 Jan;21(1):1-11, 88.

66. Hurley JC. Concordance of endotoxemia with gram-negative bacteremia. A meta-analysis using receiver operating characteristic curves. Arch Pathol Lab Med 2000 Aug;124(8):1157-64.
67. Imran MB, Khan MA, Aslam MN, et al. Diagnosis of coronary artery disease by stress echocardiography and perfusion scintigraphy. J Coll Physicians Surg Pak 2003 Aug;13(8):465-70.
68. Ioannidis JP, Salem D, Chew PW, et al. Accuracy and clinical effect of out-of-hospital electrocardiography in the diagnosis of acute cardiac ischemia: a meta-analysis. Ann Emerg Med 2001 May;37(5):461-70.
69. Jensen JE, Nielsen SH, Foged L, et al. The MICRAL test for diabetic microalbuminuria: predictive values as a function of prevalence. Scand J Clin Lab Invest 1996 Apr;56(2):117-22.
70. Jorgensen K, Andersen TJ, Dam H. The diagnostic efficiency of the Rorschach Depression Index and the Schizophrenia Index: a review. Assessment 2000 Sep;7(3):259-80.
71. Jorm AF. Methods of screening for dementia: a meta-analysis of studies comparing an informant questionnaire with a brief cognitive test. Alzheimer Dis Assoc Disord 1997 Sep;11(3):158-62.
72. Kardaun JW, Kardaun OJ. Comparative diagnostic performance of three radiological procedures for the detection of lumbar disk herniation. Methods Inf Med 1990 Jan;29(1):12-22.
73. Kearon C, Julian JA, Newman TE, et al. Noninvasive diagnosis of deep venous thrombosis. McMaster Diagnostic Imaging Practice Guidelines Initiative. Ann Intern Med 1998 Apr 15;128(8):663-77.
74. Kellen M, Aronson S, Roizen MF, et al. Predictive and diagnostic tests of renal failure: a review. Anesth Analg 1994 Jan;78(1):134-42.
75. Kertai MD, Boersma E, Bax JJ, et al. A meta-analysis comparing the prognostic accuracy of six diagnostic tests for predicting perioperative cardiac risk in patients undergoing major vascular surgery. Heart 2003 Nov;89(11):1327-34.
76. Kinkel K, Hricak H, Lu Y, et al. US characterization of ovarian masses: a meta-analysis. Radiology 2000 Dec;217(3):803-11.
77. Kittler H, Pehamberger H, Wolff K, et al. Diagnostic accuracy of dermoscopy. Lancet Oncol 2002 Mar;3(3):159-65.
78. Klassen TP, Rowe PC. Selecting diagnostic tests to identify febrile infants less than 3 months of age as being at low risk for serious bacterial infection: a scientific overview. J Pediatr 1992 Nov;121(5 Pt 1):671-6.
79. Koelemay MJ, Lijmer JG, Stoker J, et al. Magnetic resonance angiography for the evaluation of lower extremity arterial disease: a meta-analysis. JAMA 2001 Mar 14;285(10):1338-45.
80. Koumans EH, Johnson RE, Knapp JS, et al. Laboratory testing for Neisseria gonorrhoeae by recently introduced nonculture tests: a performance review with clinical and public health considerations
15. Clin Infect Dis 1998 Nov;27(5):1171-80.
81. Kowalski J, Tu XM, Jia G, et al. A comparative meta-analysis on the variability in test performance among FDA-licensed enzyme immunosorbent assays for HIV antibody testing. J Clin Epidemiol 2001 May;54(5):448-61.
82. Kwok Y, Kim C, Grady D, et al. Meta-analysis of exercise testing to detect coronary artery disease in women. Am J Cardiol 1999 Mar 1;83(5):660-6.
83. Lensing AW, Hirsh J. 125I-fibrinogen leg scanning: reassessment of its role for the diagnosis of venous thrombosis in post-operative patients. Thromb Haemost 1993 Jan 11;69(1):2-7.
84. LeSar CJ, Meier GH, DeMasi RJ, et al. The utility of color duplex ultrasonography in the diagnosis of temporal arteritis. J Vasc Surg 2002 Dec;36(6):1154-60.
85. Lewis JD, Ng K, Hung KE, et al. Detection of proximal adenomatous polyps with screening sigmoidoscopy: a systematic review and meta-analysis of screening colonoscopy. Arch Intern Med 2003 Feb 24;163(4):413-20.
86. Li J. Capnography alone is imperfect for endotracheal tube placement confirmation during emergency intubation
16. J Emerg Med 2001 Apr;20(3):223-9.

87. Liberman M, Sampalis F, Mulder DS, et al. Breast cancer diagnosis by scintimammography: a meta-analysis and review of the literature. Breast Cancer Res Treat 2003 Jul;80(1):115-26.
88. Littenberg B, Mushlin AI. Technetium bone scanning in the diagnosis of osteomyelitis: a meta-analysis of test performance. Diagnostic Technology Assessment Consortium. J Gen Intern Med 1992 Mar;7(2):158-64.
89. Maas JW, Evers JL, ter RG, et al. Pregnancy rate following normal versus abnormal hysterosalpingography findings: a meta-analysis. Gynecol Obstet Invest 1997;43(2):79-83.
90. Mackenzie R, Palmer CR, Lomas DJ, et al. Magnetic resonance imaging of the knee: diagnostic performance studies. Clin Radiol 1996 Apr;51(4):251-7.
91. Makrydimas G, Sotiriadis A, Ioannidis JP. Screening performance of first-trimester nuchal translucency for major cardiac defects: a meta-analysis. Am J Obstet Gynecol 2003 Nov;189(5):1330-5.
92. Mantha S, Roizen MF, Barnard J, et al. Relative effectiveness of four preoperative tests for predicting adverse cardiac outcomes after vascular surgery: a meta-analysis. Anesth Analg 1994 Sep;79(3):422-33.
93. Mijnhout GS, Hoekstra OS, van Tulder MW, et al. Systematic review of the diagnostic accuracy of (18)F-fluorodeoxyglucose positron emission tomography in melanoma patients. Cancer 2001 Apr 15;91(8):1530-42.
94. Mirvis SE, Shanmuganathan K, Miller BH, et al. Traumatic aortic injury: diagnosis with contrast-enhanced thoracic CT--five-year experience at a major trauma center. Radiology 1996 Aug;200(2):413-22.
95. Mitchell MF, Schottenfeld D, Tortolero-Luna G, et al. Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis. Obstet Gynecol 1998 Apr;91(4):626-31.
96. Mitchell MF, Cantor SB, Ramanujam N, et al. Fluorescence spectroscopy for diagnosis of squamous intraepithelial lesions of the cervix. Obstet Gynecol 1999 Mar;93(3):462-70.
97. Mol BW, Dijkman B, Wertheim P, et al. The accuracy of serum chlamydial antibodies in the diagnosis of tubal pathology: a meta-analysis. Fertil Steril 1997 Jun;67(6):1031-7.
98. Mol BW, Bayram N, Lijmer JG, et al. The performance of CA-125 measurement in the detection of endometriosis: a meta-analysis. Fertil Steril 1998 Dec;70(6):1101-8.
99. Mol BW, Lijmer JG, Ankum WM, et al. The accuracy of single serum progesterone measurement in the diagnosis of ectopic pregnancy: a meta-analysis. Hum Reprod 1998 Nov;13(11):3220-7.
100. Mol BW, Meijer S, Yuppa S et al. Sperm penetration assay in predicting successful in vitro fertilization. A meta-analysis. J Reprod Med 1998;43:503-8.
101. Nallamothu BK, Saint S, Bielak LF, et al. Electron-beam computed tomography in the diagnosis of coronary artery disease: a meta-analysis. Arch Intern Med 2001 Mar 26;161(6):833-8.
102. Nelemans PJ, Leiner T, de Vet HC, et al. Peripheral arterial disease: meta-analysis of the diagnostic performance of MR angiography. Radiology 2000 Oct;217(1):105-14.
103. Ng PC, Dear PR. The predictive value of a normal ultrasound scan in the preterm baby--a meta-analysis. Acta Paediatr Scand 1990 Mar;79(3):286-91.
104. Oei EH, Nikken JJ, Verstijnen AC, et al. MR imaging of the menisci and cruciate ligaments: a systematic review. Radiology 2003 Mar;226(3):837-48.
105. Oei SG, Helmerhorst FM, Keirse MJ. When is the post-coital test normal? A critical appraisal. Hum Reprod 1995 Jul;10(7):1711-4.
106. Olaniyan OB. Validity of colposcopy in the diagnosis of early cervical neoplasia--a review. Afr J Reprod Health 2002 Dec;6(3):59-69.
107. Olatidoye AG, Wu AH, Feng YJ, et al. Prognostic role of troponin T versus troponin I in unstable angina pectoris for cardiac events with meta-analysis comparing published studies. Am J Cardiol 1998 Jun 15;81(12):1405-10.
108. Orr RK, Porter D, Hartman D. Ultrasonography to evaluate adults for appendicitis: decision making based on meta-analysis and probabilistic reasoning. Acad Emerg Med 1995 Jul;2(7):644-50.
109. Pai M, Flores LL, Pai N, et al. Diagnostic accuracy of nucleic acid amplification tests for tuberculous meningitis: a systematic review and meta-analysis. Lancet Infect Dis 2003 Oct;3(10):633-43.

110. Palomaki GE, Neveux LM, Haddow JE. Can reliable Down's syndrome detection rates be determined from prenatal screening intervention trials? J Med Screen 1996;3(1):12-7.
111. Patel SR, Wiese W, Patel SC, et al. Systematic review of diagnostic tests for vaginal trichomoniasis. Infect Dis Obstet Gynecol 2000;8(5-6):248-57.
112. Patrick DL, Cheadle A, Thompson DC, et al. The validity of self-reported smoking: a review and meta-analysis. Am J Public Health 1994 Jul;84(7):1086-93.
113. Paulson WD, Ram SJ, Birk CG, et al. Does blood flow accurately predict thrombosis or failure of hemodialysis synthetic grafts? A meta-analysis. Am J Kidney Dis 1999 Sep;34(3):478-85.
114. Petersen JR, Smith E, Okorodudu AO, et al. Comparison of four methods (L/S ratio, TDx FLM, lamellar bodies, PG) for fetal lung maturity using meta-analysis. Clin Lab Manage Rev 1996 Mar;10(2):169-75.
115. Picano E, Bedetti G, Varga A, et al. The comparable diagnostic accuracies of dobutamine-stress and dipyridamole-stress echocardiographies: a meta-analysis. Coron Artery Dis 2000 Mar;11(2):151-9.
116. Rao JK, Weinberger M, Oddone EZ, et al. The role of antineutrophil cytoplasmic antibody (c-ANCA) testing in the diagnosis of Wegener granulomatosis. A literature review and meta-analysis. Ann Intern Med 1995 Dec 15;123(12):925-32.
117. Reed JF, III. Meta-analysis of the reliability of noninvasive carotid studies. Biomed Instrum Technol 1991 Nov;25(6):465-71.
118. Revah A, Hannah ME, Sue AQA. Fetal fibronectin as a predictor of preterm birth: an overview. Am J Perinatol 1998;15(11):613-21.
119. Romagnuolo J, Bardou M, Rahme E, et al. Magnetic resonance cholangiopancreatography: a meta-analysis of test performance in suspected biliary disease. Ann Intern Med 2003 Oct 7;139(7):547-57.
120. Rosado B, Menzies S, Harbauer A, et al. Accuracy of computer diagnosis of melanoma: a quantitative meta-analysis. Arch Dermatol 2003 Mar;139(3):361-7.
121. Rosenberg D, Cretin S. Use of meta-analysis to evaluate tolonium chloride in oral cancer screening. Oral Surg Oral Med Oral Pathol 1989 May;67(5):621-7.
122. Samson DJ, Flamm CR, Pisano ED, et al. Should FDG PET be used to decide whether a patient with an abnormal mammogram or breast finding at physical examination should undergo biopsy? Acad Radiol 2002 Jul;9(7):773-83.
123. Sarmiento OL, Weigle KA, Alexander J, et al. Assessment by meta-analysis of PCR for diagnosis of smear-negative pulmonary tuberculosis. J Clin Microbiol 2003 Jul;41(7):3233-40.
124. Scheidler J, Hricak H, Yu KK, et al. Radiological evaluation of lymph node metastases in patients with cervical cancer. A meta-analysis. JAMA 1997 Oct 1;278(13):1096-101.
125. Schinkel AF, Bax JJ, Geleijnse ML, et al. Noninvasive evaluation of ischaemic heart disease: myocardial perfusion imaging or stress echocardiography? Eur Heart J 2003 May;24(9):789-800.
126. Scholten RJ, Deville WL, Opstelten W, et al. The accuracy of physical diagnostic tests for assessing meniscal lesions of the knee: a meta-analysis. J Fam Pract 2001 Nov;50(11):938-44.
127. Scholten RJ, Opstelten W, van der Plas CG, et al. Accuracy of physical diagnostic tests for assessing ruptures of the anterior cruciate ligament: a meta-analysis. J Fam Pract 2003 Sep;52(9):689-94.
128. Schreiber G, McCrory DC. Performance characteristics of different modalities for diagnosis of suspected lung cancer: summary of published evidence. Chest 2003 Jan;123(1 Suppl):115S-28S.
129. Schwimmer J, Essner R, Patel A, et al. A review of the literature for whole-body FDG PET in the management of patients with melanoma. Q J Nucl Med 2000 Jun;44(2):153-67.
130. Siegman-Igra Y, Anglim AM, Shapiro DE, et al. Diagnosis of vascular catheter-related bloodstream infection: a meta-analysis. J Clin Microbiol 1997 Apr;35(4):928-36.
131. Silvestri GA, Littenberg B, Colice GL. The clinical evaluation for detecting metastatic lung cancer. A meta-analysis. Am J Respir Crit Care Med 1995 Jul;152(1):225-30.
132. Skupski DW, Rosenberg CR, Eglinton GS. Intrapartum fetal stimulation tests: a meta-analysis. Obstet Gynecol 2002 Jan;99(1):129-34.

133. Smart SC, Sagar KB. Diagnostic and prognostic use of stress echocardiography in stable patients. Echocardiography 2000 Jul;17(5):465-77.
134. Smith-Bindman R, Hosmer W, Feldstein VA, et al. Second-trimester ultrasound to detect fetuses with Down syndrome: a meta-analysis. JAMA 2001 Feb 28;285(8):1044-55.
135. Sonnad SS, Langlotz CP, Schwartz JS. Accuracy of MR imaging for staging prostate cancer: a meta-analysis to examine the effect of technologic change. Acad Radiol 2001 Feb;8(2):149-57.
136. Stengel D, Bauwens K, Sehouli J, et al. Systematic review and meta-analysis of emergency ultrasonography for blunt abdominal trauma. Br J Surg 2001 Jul;88(7):901-12.
137. Swart P, Mol BW, van d, V, et al. The accuracy of hysterosalpingography in the diagnosis of tubal pathology: a meta-analysis. Fertil Steril 1995 Sep;64(3):486-91.
138. Tan KT, van Beek EJ, Brown PW, et al. Magnetic resonance angiography for the diagnosis of renal artery stenosis: a meta-analysis. Clin Radiol 2002 Jul;57(7):617-24.
139. Tebas P, Nease RF, Storch GA. Use of the polymerase chain reaction in the diagnosis of herpes simplex encephalitis: a decision analysis model. Am J Med 1998 Oct;105(4):287-95.
140. Tsao H, Nadiminti U, Sober AJ, et al. A meta-analysis of reverse transcriptase-polymerase chain reaction for tyrosinase mRNA as a marker for circulating tumor cells in cutaneous melanoma. Arch Dermatol 2001 Mar;137(3):325-30.
141. Urbach DR, Khajanchee YS, Jobe BA, et al. Cost-effective management of common bile duct stones: a decision analysis of the use of endoscopic retrograde cholangiopancreatography (ERCP), intraoperative cholangiography, and laparoscopic bile duct exploration. Surg Endosc 2001 Jan;15(1):4-13.
142. van Beek EJ, Brouwers EM, Song B, Bongaerts AH, Oudkerk M. Lung scintigraphy and helical computed tomography for the diagnosis of pulmonary embolism: a meta-analysis. Clin Appl Thromb Hemost 2001;7:87-92.
143. van Gelder JM. Computed tomographic angiography for detecting cerebral aneurysms: implications of aneurysm size distribution for the sensitivity, specificity, and likelihood ratios. Neurosurgery 2003 Sep;53(3):597-605.
144. Varonen H, Makela M, Savolainen S, et al. Comparison of ultrasound, radiography, and clinical examination in the diagnosis of acute maxillary sinusitis: a systematic review. J Clin Epidemiol 2000 Sep;53(9):940-8.
145. Vasbinder GB, Nelemans PJ, Kessels AG, et al. Diagnostic tests for renal artery stenosis in patients suspected of having renovascular hypertension: a meta-analysis. Ann Intern Med 2001 Sep 18;135(6):401-11.
146. Vasquez TE, Rimkus DS, Hass MG, et al. Efficacy of morphine sulfate-augmented hepatobiliary imaging in acute cholecystitis. J Nucl Med Technol 2000 Sep;28(3):153-5.
147. Visser K, Hunink MG. Peripheral arterial disease: gadolinium-enhanced MR angiography versus color-guided duplex US--a meta-analysis. Radiology 2000 Jul;216(1):67-77.
148. Watson EJ, Templeton A, Russell I, et al. The accuracy and efficacy of screening tests for Chlamydia trachomatis: a systematic review. J Med Microbiol 2002 Dec;51(12):1021-31.
149. Wellnitz U, Binder B, Fritz P, et al. Reliability of telepathology for frozen section service. Anal Cell Pathol 2000;21(3-4):213-22.
150. Wells PS, Lensing AW, Davidson BL, et al. Accuracy of ultrasound for the diagnosis of deep venous thrombosis in asymptomatic patients after orthopedic surgery. A meta-analysis. Ann Intern Med 1995 Jan 1;122(1):47-53.
151. White RH, McGahan JP, Daschbach MM, et al. Diagnosis of deep-vein thrombosis using duplex ultrasound. Ann Intern Med 1989 Aug 15;111(4):297-304.
152. Whitsel EA, Boyko EJ, Siscovick DS. Reassessing the role of QTc in the diagnosis of autonomic failure among patients with diabetes: a meta-analysis. Diabetes Care 2000 Feb;23(2):241-7.
153. Wiese W, Patel SR, Patel SC, et al. A meta-analysis of the Papanicolaou smear and wet mount for the diagnosis of vaginal trichomoniasis. Am J Med 2000 Mar;108(4):301-8.

154. Wijnberger LD, Huisjes AJ, Voorbij HA, et al. The accuracy of lamellar body count and lecithin/sphingomyelin ratio in the prediction of neonatal respiratory distress syndrome: a meta-analysis. BJOG 2001 Jun;108(6):583-8.
155. Williams JW, Jr., Noel PH, Cordes JA, Ramirez G, Pignone M. Is this patient clinically depressed? JAMA
156. Wu AH, Lane PL. Metaanalysis in clinical chemistry: validation of cardiac troponin T as a marker for ischemic heart diseases. Clin Chem 1995 Aug;41(8 Pt 2):1228-33.
157. Zandbergen EG, de Haan RJ, Stoutenbeek CP, et al. Systematic review of early prediction of poor outcome in anoxic-ischaemic coma. Lancet 1998 Dec 5;352(9143):1808-12.

Appendix B: Bayesian model for bivariate meta-analysis of sensitivity and specificity

For each study we used the 2×2 table of test results (positive or negative) and true disease status:

		Disease status	
		Disease	**Healthy**
Test result	Positive	TP	FP
	Negative	FN	TN

Here, TP, FP, FN, and TN are respectively the counts of true positive, false positive, false negative, and true negative findings from each study. The accuracy metrics of interest were the logit-transformed sensitivity and false positive rate (=1-specificity); at the study-level these were estimated as $\hat{\eta} = logit(\frac{TP}{TP+FN})$, and $\hat{\xi} = logit(\frac{FP}{FP+TN})$, respectively.

Throughout the report, we use the logit transformation of a proportion (p): $logit(p) = \ln\left(\frac{p}{1-p}\right)$. The inverse-logit of any number (x) is then defined as $logit^{-1}(x) = \frac{\exp(x)}{1+\exp(x)}$.

Bivariate meta-analysis of sensitivity and specificity

The within studies model is specified as $TP_i \sim Bin(D_i, logit^{-1}(\eta_i))$ and $FP_i \sim Bin(H_i, logit^{-1}(\xi_i))$, where D_i and H_i are the total numbers of affected and unaffected individuals, respectively, for the i-th study. We then model the "true" logit-transformed sensitivity (η_i) and false positive rate (ξ_i) of the i-th study as drawn from a bivariate normal distribution:

$\binom{\eta_i}{\xi_i} \sim N\left(\binom{\bar{\eta}}{\bar{\xi}}, \Omega\right)$. Here, $\bar{\eta}$ and $\bar{\xi}$ are the "average" logit-transformed summary sensitivity and false positive rate (across studies), with covariance matrix $\Omega = \begin{pmatrix} \sigma_\eta^2 & \rho\sigma_\eta\sigma_\xi \\ \rho\sigma_\eta\sigma_\xi & \sigma_\xi^2 \end{pmatrix}$; σ_η^2 and σ_ξ^2 are the

variances of the logit-transformed sensitivity and false positive rate, respectively; ρ is their correlation (at the between-study level); and $\rho\sigma_\eta\sigma_\xi$ is their covariance.

This model was used both for the maximum likelihood and the Bayesian bivariate meta-analyses. The key difference between the two approaches is that in the latter, the model parameters (e.g., the average sensitivity and false positive rate, and the covariance matrix) are treated as random variables and prior distributions need to be specified for them (hyper-priors). For the average logit-transformed sensitivity and false positive rate we used vague normal priors $N(0, 100)$; for the between-study variance we assumed that its square root (i.e., the standard deviation) followed a uniform prior distribution $U(0.01, 5)$. Finally, we assumed that the within study correlation has a uniform prior distribution $U(-1, 1)$. We used alternative prior distributions both for the variances and the correlation. Results were similar to the main analyses, when we used alternative normal priors (more or less informative for the standard deviations) or a Wishart prior for the covariance matrix (as suggested in Verde et al., Statistics in Medicine 2010). For the correlation, we used Fisher's z-transformation and then assumed that the z-transformed correlation follows a normal prior $N(0, 25)$ (as suggested in Paul et al, Statistics in Medicine, 2010).

BUGS language Bayesian model for bivariate meta-analysis of sensitivity and false positive rate

```
#BUGS model: bivariate meta-analysis of test accuracy
#exact binomial likelihood for within-study variability

model {
for( i in 1 : n_studies ) {
tp[i] ~ dbin(tpr[i], n1[i])
fp[i] ~ dbin(fpr[i], n2[i])
logit(tpr[i]) <- m[i,1]
logit(fpr[i]) <- m[i,2]
m[i,1:2] ~ dmnorm(mu0[1:2], sigma.inv[1:2, 1:2])
}
# Priors for means
mu0[1] ~ dnorm(0, 0.01)
mu0[2] ~ dnorm(0, 0.01)

#priors for the elements of the covariance matrix
sigma.inv[1:2, 1:2] <- inverse(sigma[1:2,1:2])
sigma[1,1] <- sigma1*sigma1
sigma[1,2] <- sigma1*sigma2*rho
sigma[2,1] <- sigma1*sigma2*rho
sigma[2,2] <- sigma2*sigma2
sigma1 ~ dunif(0.01,5)
```

```
sigma2 ~ dunif(0.01,5)
rho ~ dunif(-1, 1)

#alternative prior for rho, used in sensitivity analysis
#rho <-  (exp(2*z) - 1)/(1 + exp(2*z))
#z ~ dnorm(0, 0.4)
#for z, alternative precisions gave similar results
# See also, Paul et al. Statistics in Medicine, 2010

# Pooled summaries
x.pool <- mu0[1]
y.pool <- mu0[2]
poolse <- exp(x.pool) / ( 1 + exp(x.pool) )
poolsp <- 1 - exp(y.pool) / ( 1 + exp(y.pool) )

# Variance-covariance matrix for random-effects
sigmaSens <- sigma[1,1]
sigmaSpec <- sigma[2,2]
}
```

Appendix C: Alternative parameterizations of the hierarchical summary receiver operating characteristic curve

We followed Arends 2008 [25] in constructing alternative hierarchical summary receiver operating (hsROC) curves based on the bivariate random effects meta-analysis model. Here, $\bar{\eta}$ is the summary logit-transformed sensitivity, $\bar{\xi}$ is the summary logit-transformed false positive rate (1-specificity), η_i and ξ_i are respectively is the "true" logit-transformed sensitivity and the "true" logit-transformed false positive rate in the i-th study; σ_η^2 and σ_ξ^2 are the variances of the logit-transformed sensitivity and specificity, respectively; $\sigma_{\xi\eta}$ is their covariance.

After fitting the bivariate model, we obtained the following hsROC lines:

Model	Regression line	Intercept	Slope
η ~ ξ regression	$\eta_i = \bar{\eta} + \frac{\sigma_{\xi\eta}}{\sigma_\xi^2}(\xi_i - \bar{\xi})$	$\bar{\eta} - \frac{\sigma_{\xi\eta}}{\sigma_\xi^2}\bar{\xi}$	$\frac{\sigma_{\xi\eta}}{\sigma_\xi^2}$
ξ ~ η regression	$\eta_i = \bar{\eta} + \frac{\sigma_\eta^2}{\sigma_{\xi\eta}}(\xi_i - \bar{\xi})$	$\bar{\eta} - \frac{\sigma_\eta^2}{\sigma_{\xi\eta}}\bar{\xi}$	$\frac{\sigma_\eta^2}{\sigma_{\xi\eta}}$
D ~ S regression	$\eta_i = \bar{\eta} + \frac{\sigma_\eta^2 + \sigma_{\xi\eta}}{\sigma_\xi^2 + \sigma_{\xi\eta}}(\xi_i - \bar{\xi})$	$\bar{\eta} - \frac{\sigma_\eta^2 + \sigma_{\xi\eta}}{\sigma_\xi^2 + \sigma_{\xi\eta}}\bar{\xi}$	$\frac{\sigma_\eta^2 + \sigma_{\xi\eta}}{\sigma_\xi^2 + \sigma_{\xi\eta}}$
Rutter-Gatsonis model	$\eta_i = \bar{\eta} + \frac{\sigma_\eta}{\sigma_\xi}(\xi_i - \bar{\xi})$	$\bar{\eta} - \frac{\sigma_\eta}{\sigma_\xi}\bar{\xi}$	$\frac{\sigma_\eta}{\sigma_\xi}$
MAR	$\eta_i = \bar{\eta} + \frac{\sigma_\eta^2 - \sigma_\xi^2 + \sqrt{(\sigma_\eta^2 - \sigma_\xi^2)^2 + 4\sigma_{\xi\eta}^2}}{2\sigma_{\xi\eta}}(\xi_i - \bar{\xi})$	$\bar{\eta} - \frac{\sigma_\eta^2 - \sigma_\xi^2 + \sqrt{(\sigma_\eta^2 - \sigma_\xi^2)^2 + 4\sigma_{\xi\eta}^2}}{2\sigma_{\xi\eta}}\bar{\xi}$	$\frac{\sigma_\eta^2 - \sigma_\xi^2 + \sqrt{(\sigma_\eta^2 - \sigma_\xi^2)^2 + 4\sigma_{\xi\eta}^2}}{2\sigma_{\xi\eta}}$

D = logit(sensitivity) – logit(1-specificity)
S = logit(sensitivity) + logit(1-specificity)
MAR = major axis regression
η = logit(sensitivity)
ξ = logit(1-specificity).
"Intercept" and "slope" refer to the estimated intercept and slope of the hsROC in the logit space. They describe a line that can be transformed to the hsROC curve. For each model, the corresponding hsROC curve is represented by the following function: sensitivity = invlogit(intercept + slope * logit(specificity)). Estimates of the intercept and slope can be derived from the output of the bivariate regression model in all major statistical packages.