ORIGINAL ARTICLE

Accuracy of pulse oximetry in screening for congenital heart disease in asymptomatic newborns: a systematic review

Shakila Thangaratinam, Jane Daniels, Andrew K Ewer, Javier Zamora, Khalid S Khan

Arch Dis Child Fetal Neonatal Ed 2007;**92**:F176–F180. doi: 10.1136/adc.2006.107656

A supplementary appendix is available at http://adc.bmj.com/supplemental

See end of article for authors' affiliations

Correspondence to: Shakila Thangaratinam, s.thangaratinam@bham.ac.uk

Accepted 6 February 2007
Published Online First 7 March 2007

Objective: To evaluate the accuracy of pulse oximetry as a screening tool for congenital heart disease in asymptomatic newborns.

Design, data sources and methods: Systematic review of relevant studies identified through MEDLINE, EMBASE, *Cochrane Library*, MEDION, and bibliographies of retrieved primary and review articles. Two reviewers independently extracted data on study characteristics, quality and results to construct 2×2 tables with congenital heart disease as the reference standard. A random-effects bivariate model was used to meta-analyse estimates of sensitivity and specificity. Logit pairs of sensitivity and specificity of each study were analysed in a single model, accounting for their correlation due to differences in threshold between studies.

Results: Eight studies were included with a total of 35 960 newborns. Pulse oximetry was performed on asymptomatic newborns in all studies; three studies excluding newborns with an antenatal diagnosis of congenital heart disease. Either functional or fractional oxygen saturation was measured by pulse oximetry with oxygen saturation below 95% as the cut-off level in most studies. On the basis of the eight studies, the summary estimates of sensitivity and specificity were 63% (95% CI 39% to 83%) and 99.8% (95% CI 99% to 100%), respectively, yielding a false positive rate of 0.2% (95% CI 0% to 1%).

Conclusion: Pulse oximetry was found to be highly specific tool with very low false positive rates to detect congenital heart disease. Large, well-conducted prospective studies are needed to assess its sensitivity with higher precision.

Congenital heart disease is the commonest group of congenital malformations and affects 7–8/1000 newborns.[1 2] It contributes to 3% of all infant mortality and 46% of deaths from congenital malformations, with most deaths occurring in the first year of life.[1] A large proportion of these children require surgery in the first year.

One of the major contributors to increased infant mortality and morbidity is clinical deterioration and collapse prior to diagnosis and treatment.[3–6] Early detection of congenital heart disease in the asymptomatic period immediately after birth will reduce clinical deterioration by instigation of appropriate, timely management. Currently infants are screened to detect congenital heart disease by physical examination after birth and another examination at six to eight weeks.[7] However, this method of screening fails to detect up to 50% of congenital heart defects at birth.[8]

Pulse oximetry has been proposed as an alternative screening method for the detection of congenital heart defects. It is a simple, non-invasive investigation which measures the percentage of haemoglobin in blood that is saturated with oxygen. It is proposed that the measurement of oxygen saturation identifies infants with mild cyanosis who do not have an audible murmur or other signs of cardiac abnormality and are not detected by routine clinical examination.[9] Several studies have reported the use of pulse oximetry as a screening tool for the detection of congenital heart disease.[10–17] Although the reported false positive rates have been low, individual studies have had only few cases with cardiac disease leading to imprecision in the estimation of true positive rates (sensitivity).

A recent *Health Technology Assessment* report reviewed the available evidence on the screening strategies for detection of congenital heart disease in newborns with the view to assist in policy making.[1] The review identified four studies in which pulse oximetry was used as a screening test for congenital heart disease in asymptomatic newborns. It did not pool the results statistically and suggested the need for further evaluation of pulse oximetry as a screening method. Since the publication of this report, the results of more, large primary studies have become available, which may potentially alter the report's conclusion.

We therefore conducted a systematic review to collate all results and to update information on accuracy of pulse oximetry to detect congenital heart disease in asymptomatic newborns.

METHODS

We carried out a systematic review with a prospective protocol using widely recommended methods.[18–21]

Identification of studies

We searched MEDLINE (1996–2006), EMBASE (1996–2006), *Cochrane Library* (2006) and MEDION (a database of diagnostic test reviews set up by Dutch and Belgian researchers) for relevant citations using the search terms ''pulse NEAR oximetry'', ''infant-newborn'', ''neonate'', ''newborn'', ''infant'', ''congenital heart disease''. The reference lists of all known primary and review articles were examined to identify cited articles not captured by electronic searches. There were no language restrictions. A comprehensive database of potentially relevant citations was constructed.

Study selection and data extraction

Studies which evaluated the accuracy of pulse oximetry in asymptomatic newborns for the detection of congenital heart disease were selected by a two-stage process. First, the electronic searches were scrutinised and full manuscripts of all citations that were likely to meet the predefined selection criteria were obtained by two independent reviewers (ST and JD). Second, final inclusion or exclusion decisions were made by the reviewers (ST and JD) after examination of these

manuscripts. Studies which met the predefined and explicit criteria regarding population, tests, outcomes and study design were selected for the review (see appendix 1, available at http://adc.bmj.com/supplemental). Where disagreements occurred, they were resolved by consensus. In cases of duplicate publications, the most recent or complete versions were selected.

From each selected article, we extracted information on the study population which included age, test characteristics along with frequency and method of testing, and methodological quality, including verification of diagnosis of congenital heart disease by echocardiography. Accuracy data were used to construct 2×2 tables of pulse oximetry results (test positive if pulse oximetry values were below a threshold as defined in the primary study, and test negative if these were above the threshold) and presence or absence of congenital heart disease diagnosed by echocardiography (wherever employed). Where accuracy data were not extractable, we contacted the corresponding author by letter or email to seek his or her assistance in data extraction.

Assessment of methodological quality

All manuscripts meeting the selection criteria were assessed for their methodological quality. Quality was defined as the confidence that the study design, conduct and analysis minimised assessment bias in the estimation of test accuracy. On the basis of existing checklists,[18 21–23] quality assessment involved scrutinising study design and relevant features of the population, test and outcomes of the study. A study was considered to be of good quality if it used a prospective design, consecutive enrolment, full verification of the test result with reference standard, and contained an adequate description of the test.[18 21–23]

Data synthesis

From individual studies, measures of accuracy such as sensitivity and specificity were calculated with the 95% confidence interval. The true positive rate and false positive rate for various test thresholds were plotted in the receiver operating characteristics (ROC) space. We used a random-effects bivariate model to meta-analyse estimates of sensitivity and specificity. Rather than using a single outcome measure per study, such as the diagnostic odds ratio in the summary ROC approach, the bivariate model preserves the two-dimensional nature of diagnostic data by directly analysing the logit transformed sensitivity log (sensitivity/(1−sensitivity)) and specificity log (specificity/(1−specificity)) of each study in a single model.[24] This model estimates and incorporates the correlation that might exist between logit sensitivity and specificity within studies due to possible differences in threshold between studies.

We applied a standard correction of adding 0.5 to all four cells of the 2×2 table when either sensitivity or specificity was 100%. The model produces the following results: a random effect estimate of the mean sensitivity and specificity with the corresponding 95% confidence interval, the amount of between-study variation for sensitivity and specificity separately, and the strength and shape of the correlation between sensitivity and specificity. Using the parameters of the bivariate distribution we calculated a confidence ellipse around the summary estimates of sensitivity and specificity. All analyses were performed using the MetaDisc statistical package[25] except to fit the bivariate model for which the Proc Mixed procedure in SAS (version 8.2 for Windows) was used.

RESULTS

Literature identification and quality of the studies

Figure 1 summarises the process of literature identification and selection. A total of 558 citations were identified by electronic searches. After a detailed assessment of the papers, eight primary articles[10–17] met the selection criteria, including a total of 35 960 newborns. Salient features of each study according to the population subgroups, test characteristics and reference standards are provided in appendix 1 (available at http://adc.bmj.com/supplemental). Pulse oximetry was performed on asymptomatic newborns in all studies, with three studies excluding newborns who were antenatally diagnosed as having congenital heart disease.[14 15 17]

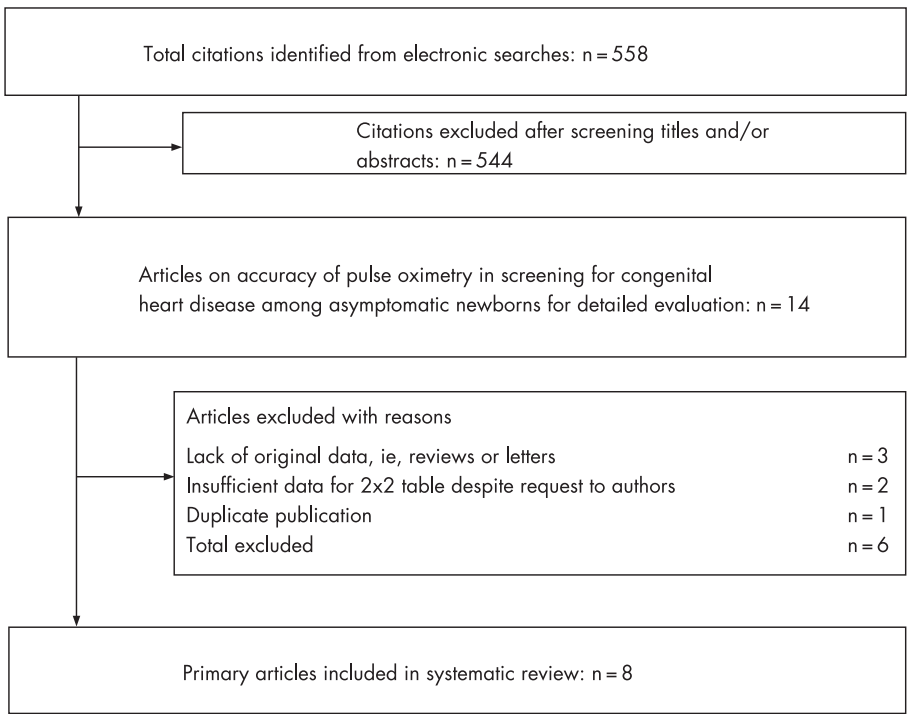

Figure 1 Study selection process for systematic review of pulse oximetry as a screening test for congenital heart disease.

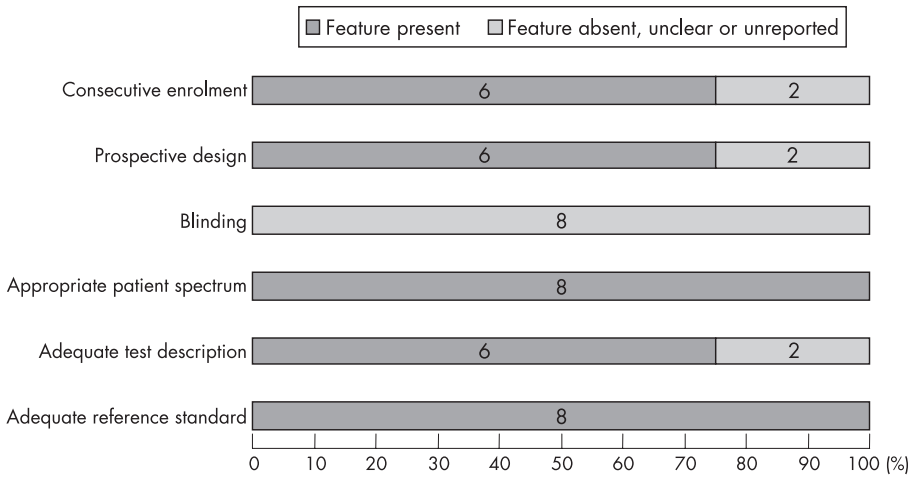

Figure 2 Quality of the studies included in the review.

In the reviewed studies either functional or fractional oxygen saturation was measured by pulse oximetry, with oxygen saturation below 95% as the cut-off level in most studies.[10–17] Functional saturation is the ratio of oxygenated haemoglobin to all haemoglobin capable of carrying oxygen; fractional saturation is the ratio of oxygenated haemoglobin to all haemoglobin (including that which does not carry oxygen). Fractional saturation is approximately 2% lower than functional saturation. There was variation in the age of first testing, ranging from 2 h to more than 24 h. In four studies, levels of oxygen saturation were measured after 24 h of birth or just before discharge.[11 14 15 17]

The outcomes assessed were congenital heart disease or critical cardiovascular malformation. Six studies were prospective studies and two were case–control studies. There was a lack of blinding for the reference standard assessment in all the studies. Differential verification of the pulse oximetry results for congenital heart disease was done either by echocardiography in neonates with low oxygen saturation or by clinical follow-up in those with normal levels. Figure 2 provides a summary of the methodological quality of the studies included in the review.

Accuracy of pulse oximetry as a screening tool for congenital heart disease

The sensitivity (true positive rate) of pulse oximetry for detection of congenital heart disease varied between 25% (95% CI 13% to 41%)[16] and 98.5% (95% CI 91.8% to 100%).[12] The test had high specificity in seven studies (98% (95% CI 98% to 99%) to 100% (95% CI 99.8% to 100%)) resulting in false positive rates between 0% and 2% (95% CI 1% to 2%).[10–12 14 16 17] Hoke *et al*'s case–control study had relatively low specificity (88% (95% CI 87% to 89%)) and high false positive rate compared with the other studies (12% (95% CI 11% to 13%)) for threshold level of functional saturation less than 95% in the foot.[13] The actual number of infants with false positive results was not certain in this study as not all newborns with abnormal screening tests underwent echocardiography or clinical examination. The highest sensitivity (98.5% (95% CI 91.8% to 100%)) was obtained using threshold levels of functional saturation <95% in both limbs or at least 3% difference in saturation between foot and right hand.[12] The highest specificity (100%) was obtained when the test was performed after 24 h of birth or near discharge.[11 14 17] The sensitivity of pulse oximetry was higher in those studies that screened for critical or cyanotic congenital heart disease compared with studies that screened for congenital heart disease of any type (fig 3).

The threshold used most commonly to detect congenital heart disease was levels of oxygen saturation less than 95% (table 1). Only one study explored the added value of pulse oximetry above the accuracy achieved through clinical examination. The combination of pulse oximetry and clinical examination had a sensitivity of 76.9% (95% CI 46.2% to 95%) and specificity of 99.9% (95% CI 99.8% to 100%).[11] The bivariate summary estimates of sensitivity and specificity were 63% (95% CI 39% to 83%) and 99.8% (95% CI 99% to 100%) respectively, yielding a false positive rate of 0.2% (95% CI 0% to 1%). The summary estimates of the individual studies are

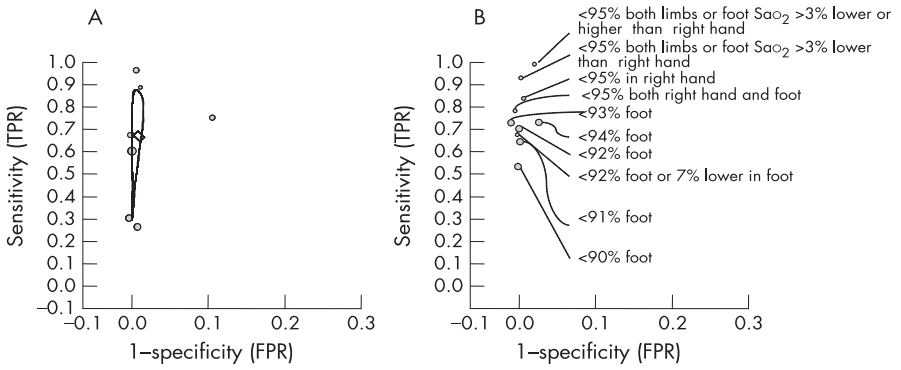

Figure 3 True positive rates (TPR) and false positive rates (FPR) for various threshold levels of oxygen saturation (SaO_2) measured by pulse oximetry for the detection of congenital heart disease in newborns. (A) TPR and FPR of individual studies for the commonest threshold of SaO_2 <95% with summary estimates of sensitivity and specificity and confidence ellipse around the estimates. (B) TPR and FPR for other threshold levels of SaO_2.

Table 1 The accuracy of pulse oximetry for diagnosing congenital heart disease in asymptomatic newborns for the commonest threshold of oxygen saturation <95%

Test	No. of patients	Sensitivity, % (true positive rate) (95% CI)	Specificity, % (95% CI)	False positive rate (%) (95% CI)
Saturation* ⩽95% foot	11 281	60.0 (14.7 to 94.7)	100 (100 to 100)	0
Saturation* ⩽95% foot or hand	2114	66.7 (9.4 to 99.2)	99.9 (99.7 to 100)	0.1 (0 to 0.3)
Saturation* ⩽95% foot	3262	96.8 (73.6 to 100)	99.7 (99.5 to 99.9)	0.3 (0.1 to 0.5)
Saturation† <95% foot	5626	25.0 (12.7 to 41.2)	99.6 (99.4 to 99.7)	0.4 (0.3 to 0.6)
Saturation* ⩽95% foot	5292	66.7 (9.4 to 99.2)	100 (99.9 to 100)	0 (0 to 0.1)
Saturation† <95% hand and foot	5211	30.8 (9.1 to 61.4)	100 (99.9 to 100)	0 (0 to 0.1)
Saturation* <95% foot	2733	75.0 (57.8 to 87.9)	87.9 (86.6 to 89.1)	12.1 (10.9 to 13.4)
Saturation* <95% foot	266	89.4 (79.4 to 95.6)	99.0 (96.4 to 99.9)	1.0 (0.1 to 3.6)
Summary estimate	35 785	63.4 (38.7 to 82.5)	99.8 (99 to 100)	0.2 (0 to 1)

*Functional oxygen saturation; †fractional oxygen saturation.

represented as a confidence ellipse in Figure 3A. The various other reported threshold levels of oxygen saturation and their accuracy measures are given in Figure 3B.

DISCUSSION

Our review has identified pulse oximetry as a potentially useful screening test for congenital heart disease in asymptomatic newborns. Pulse oximetry is a non-invasive, readily available, relatively cheap, well-validated test currently carried out by either a nurse or a doctor.[1] Published individual studies have lacked the large numbers of patients necessary to confidently estimate the accuracy of the test. Our review and meta-analysis have tried to address this issue by collating the results of published studies.

The high specificity reflects the low false positive rate of this test. The highly specific nature of the test also signifies that a low pulse oximetry reading in asymptomatic newborns ''rules in'' congenital heart disease until proved otherwise.[27] The sensitivity of the test, however, is varied, with wide confidence intervals that may be attributed to the low prevalence of the condition. This is best seen in Figure 3A, in which the confidence ellipse displays a narrow specificity axis, a reflection of the consistency in the specificity observed across all studies. The wider sensitivity axis is an indication of the reduced confidence we have in the accuracy of its sensitivity.

The validity of the findings of our review is dependent on the methodology of the systematic review and the quality of the individual studies included.[19 22] An extensive literature search was performed using relevant databases without any language restrictions to minimise the possibility of missing any study. The quality of most of the studies was compromised due to the differential verification by either echocardiography (in test positive cases) or clinical follow-up (test negative) cases. Perhaps this is unavoidable. Two of the included studies were case–control studies, a design that biases the results by overestimating the diagnostic odds ratio.[22] Furthermore, the absence of blinding, and absent or poor description of the test or reference standard could have affected the results of the review. The significant heterogeneity observed in the results could be a reflection of the type of saturation chosen for the cut-off level (functional *v* fractional), method of testing and the inclusion or exclusion of newborns diagnosed as having congenital heart disease antenatally, thereby leading to a spectrum of variation. None of the studies evaluated acceptability of ''babies'' testing to parents and the psychosocial impact of false positive results or identification of non-critical congenital heart disease.

We used a bivariate analysis model for meta-analysis with a random effects approach to obtain summary estimates of both sensitivity and specificity. This model accounts for the heterogeneity between studies caused by different threshold settings. In addition, the model acknowledges the difference in precision by which sensitivity and specificity have been measured in each study. This means that studies with larger numbers of patients with the target condition receive more weight in the calculation of the summary estimate of sensitivity, whereas studies with more patients without the target condition are more influential in the pooling of specificity. The model accounts also for the residual heterogeneity due to clinical or methodological differences between studies. Unfortunately, we could not perform an explicit analysis of these potential sources of heterogeneity due to the limited number of studies included in our review.

Given the rarity of the outcome—that is, congenital heart disease in the general population—large, well-conducted and robust studies are essential to confirm the value of pulse oximetry as a screening test, in isolation or in combination with clinical examination to obtain precise estimates of its sensitivity. Further research is needed to evaluate the effect of screening on parents and its acceptability to parents and healthcare professionals,

What is already known on this topic

- Congenital heart diseases in newborns if identified early can be successfully treated.
- Clinical examination of newborns for murmurs and cardiovascular function does not have good detection rate. As a result many babies present with complications after discharge from hospital.
- There is currently no effective screening tool for this condition.
- Pulse oximetry can pick up desaturation of blood objectively.

What this study adds

- Pulse oximetry is highly specific for detecting congenital heart disease.
- Current estimates of sensitivity to detect congenital heart disease are imprecise.
- This is a promising technique that needs large, well-conducted studies to assess the accuracy, effectiveness and feasibility of its use for mass screening of congenital heart disease in newborns.

especially with the possibility of non-significant lesions being detected during echocardiography, and the costs and cost-effectiveness of the screening programme for healthcare services.

CONTRIBUTORS

ST wrote the protocol and the initial draft with input from all authors, performed the initial search and study selection, data extraction and analysis. JD contributed to study selection and data extraction. JZ contributed to statistical analysis and revision of the draft. AKE edited the final manuscript. KSK supervised all aspects of the study including the writing of the manuscript.

Authors' affiliations

Shakila Thangaratinam, Jane Daniels, Khalid S Khan, Academic Unit, Birmingham Women's Hospital, Birmingham, UK
Andrew K Ewer, Neonatal Unit, Birmingham Women's Hospital, Institute of Child Health, University of Birmingham, Birmingham, UK
Javier Zamora, Clinical Biostatistics Unit, Hospital Ramón y Cajal, Dept. Biomathematics, University Complutense of Madrid

Competing interests: None.

Ethical approval was not required for this systematic review.

REFERENCES

1 **Knowles R**, Griebsch I, Dezateux C, *et al*. Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2005;**9**:1–168.
2 **Wren C**, Richardson S, Donaldson L. Temporal variability in the birth prevalence of cardiovascular malformations. *Heart* 2000;**83**:414–9.
3 **Kuehl KS**, Loffredo CA, Ferenez C. Failure to diagnose congenital heart disease in infancy. *Pediatrics* 1999;**103**:743–7.
4 **Malkin JD**, Garber S, Broder MS, *et al*. Infant mortality and early postpartum discharge. *Obstetr Gynecol* 2000;**96**:183–8.
5 **Meberg A**, Otterstad JE, Froland G, *et al*. Early clinical screening for congenital heart disease: the cases we miss. *Cardiol Young* 1999;**9**:169–74.
6 **Bove EL**, Bull C, Stark J, *et al*. Congenital heart disease in the neonate: results of surgical treatment. *Arch Dis Child* 1983;**58**:137–41.
7 **Hall DMB**, Elliman D. *Health for all children*, 4th edn. Oxford: Oxford University Press, 2003.
8 **Wren C**, Richmond S, Donaldson L. Presentation of congenital heart disease in pregnancy: Implications for routine examination. *Arch Dis Child Fetal Neonatal Ed* 1999;**80**:F49–F53.
9 **Salwyer JW**. Neonatal and pediatric pulse oximetry. *Respir Care* 2003;**48**:386–96.
10 **Arlettaz R**, Bauschatz A, Monkhoff M, *et al*. The contribution of pulse oximetry to the early detection of congenital heart disease in newborns. *Eur J Pediatr* 2006;**165**:94–8.
11 **Bakr AF**, Habib HS. Combining pulse oximetry and clinical examination in screening for congenital heart disease. *Pediatr Cardiol* 2005;**26**:832–5.
12 **de Wahl Granelli A**, Mellander M, Sunnegardh J, *et al*. Screening for duct-dependant congenital heart disease with pulse oximetry: a critical evaluation of strategies to maximize sensitivity. *Acta Paediatr* 2005;**94**:1590–6.
13 **Hoke TR**, Donohue PK, Bawa PK, *et al*. Oxygen saturation as a screening test for critical congenital heart disease: a preliminary study. *Pediatr Cardiol* 2002;**23**:403–9.
14 **Koppel R**, I, Druschel C, Carter T, *et al*. Effectiveness of pulse oximetry screening for congenital heart disease in asymptomatic newborns. *Pediatrics* 2003;**111**:451–5.
15 **Reich J**, Miller S, Brogdon B, *et al*. The use of pulse oximetry to detect congenital heart disease. *J Pediatr* 2003;**142**:268–72.
16 **Richmond S**, Reay G, Abu H. Routine pulse oximetry in the asymptomatic newborn. *Arch Dis Child Fetal Neonatal Ed* 2002;**87**:F83–F8.
17 **Rosati E**, Chitano G, Dipaola L, *et al*. Indications and limitations for a neonatal pulse oximetry screening of critical congenital heart disease. *J Perinat Med* 2005;**33**:455–7.
18 **Irwig LM**, Tosteton AN, Gatsonis CA, *et al*. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med* 1994;**120**:667–76.
19 **The Cochrane Collaboration**. Cochrane Methods Working Group on Systematic Reviews of Screening and Diagnostic Tests: Recommended Methods. 6 June, 1996. http://www.cochrane.org/newslett (accessed 24 March 2007).
20 **Deeks JJ**. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001;**323**:157–62.
21 **Khan KS**, Dinnes J, Kleijnen J. Systematic reviews to evaluate diagnostic tests. *J Obstet Gynecol Reprod Biol* 2001;**95**:6–11.
22 **Lijmer JG**, Mol BW, Heisterkamp S, *et al*. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061–6.
23 **Rennie D**. Improving reports of diagnostic tests: the STARD initiave. *JAMA* 2003;**289**:89–90.
24 **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.
25 **Zamora J**, Abraira V, Muriel A, *et al*. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2006;**6**:31.
26 **Sackett DL**, Hayes RB, Guyatt GH, *et al*. *Clinical epidemiology: a basic science for clinical medicine*, 2nd ed. Boston, MA: Little, Brown and Company, 1991.