Pulse oximetry as a screening test for congenital heart defects in newborn infants: a test accuracy study with evaluation of acceptability and cost-effectiveness

AK Ewer, AT Furmston, LJ Middleton, JJ Deeks,
JP Daniels, HM Pattison, R Powell, TE Roberts,
P Barton, P Auguste, A Bhoyar, S Thangaratinam,
AM Tonks, P Satodia, S Deshpande,
B Kumararatne, S Sivakumar, R Mupanemunda
and KS Khan

HTA

How to obtain copies of this and other HTA programme reports

An electronic version of this title, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (www.hta.ac.uk). A fully searchable DVD is also available (see below).

Printed copies of HTA journal series issues cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our despatch agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per issue and for the rest of the world £3 per issue.

How to order:

– fax (with **credit card details**)
– post (with **credit card details** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you to either print out your order or download a blank order form.

Contact details are as follows:

Synergie UK (HTA Department)
Digital House, The Loddon Centre
Wade Road
Basingstoke
Hants RG24 8QW

Email: orders@hta.ac.uk

Tel: 0845 812 4000 – ask for 'HTA Payment Services'
(out-of-hours answer-phone service)

Fax: 0845 812 4001 – put 'HTA Order' on the fax header

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *University of Southampton* and drawn on a bank with a UK address.

Paying by credit card
You can order using your credit card by phone, fax or post.

Subscriptions

NHS libraries can subscribe free of charge. Public libraries can subscribe at a reduced cost of £100 for each volume (normally comprising 40–50 titles). The commercial subscription rate is £400 per volume (addresses within the UK) and £600 per volume (addresses outside the UK). Please see our website for details. Subscriptions can be purchased only for the current or forthcoming volume.

How do I get a copy of *HTA on DVD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd/index.shtml). *HTA on DVD* is currently free of charge worldwide.

The website also provides information about the HTA programme and lists the membership of the various committees.

Pulse oximetry as a screening test for congenital heart defects in newborn infants: a test accuracy study with evaluation of acceptability and cost-effectiveness

AK Ewer,[1,2]* AT Furmston,[3] LJ Middleton,[3] JJ Deeks,[4]
JP Daniels,[1,3] HM Pattison,[5] R Powell,[5] TE Roberts,[6]
P Barton,[6] P Auguste,[6] A Bhoyar,[7] S Thangaratinam,[1,2]
AM Tonks,[8] P Satodia,[9] S Deshpande,[10]
B Kumararatne,[11] S Sivakumar,[12] R Mupanemunda[13]
and KS Khan[14]

[1]University of Birmingham, School of Clinical and Experimental Medicine, Birmingham, UK
[2]Birmingham Women's Healthcare NHS Foundation Trust, Birmingham, UK
[3]University of Birmingham, Birmingham Clinical Trials Unit, Birmingham, UK
[4]University of Birmingham, Public Health, Epidemiology and Biostatistics, Birmingham, UK
[5]Aston University, School of Life and Health Sciences, Birmingham, UK
[6]University of Birmingham, Health Economics Unit, School of Health and Population Sciences, Birmingham, UK
[7]Birmingham Children's Hospital NHS Foundation Trust, Birmingham, UK
[8]West Midlands Perinatal Institute, Birmingham, UK
[9]University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK
[10]The Shrewsbury and Telford Hospital NHS Trust, Shrewsbury, UK
[11]The Royal Wolverhampton Hospitals NHS Trust, Wolverhampton, UK
[12]Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK
[13]Heart of England NHS Foundation Trust, Birmingham, UK
[14]Barts and the London School of Medicine and Dentistry, London, UK

*Corresponding author

Declared competing interests of authors: none

This report should be referenced as follows:

NIHR Health Technology Assessment programme

The Health Technology Assessment (HTA) programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

Abstract

Pulse oximetry as a screening test for congenital heart defects in newborn infants: a test accuracy study with evaluation of acceptability and cost-effectiveness

AK Ewer,[1,2]* AT Furmston,[3] LJ Middleton,[3] JJ Deeks,[4] JP Daniels,[1,3]
HM Pattison,[5] R Powell,[5] TE Roberts,[6] P Barton,[6] P Auguste,[6] A Bhoyar,[7]
S Thangaratinam,[1,2] AM Tonks,[8] P Satodia,[9] S Deshpande,[10]
B Kumararatne,[11] S Sivakumar,[12] R Mupanemunda[13] and KS Khan[14]

[1]University of Birmingham, School of Clinical and Experimental Medicine, Birmingham, UK
[2]Birmingham Women's Healthcare NHS Foundation Trust, Birmingham, UK
[3]University of Birmingham, Birmingham Clinical Trials Unit, Birmingham, UK
[4]University of Birmingham, Public Health, Epidemiology and Biostatistics, Birmingham, UK
[5]Aston University, School of Life and Health Sciences, Birmingham, UK
[6]University of Birmingham, Health Economics Unit, School of Health and Population Sciences, Birmingham, UK
[7]Birmingham Children's Hospital NHS Foundation Trust, Birmingham, UK
[8]West Midlands Perinatal Institute, Birmingham, UK
[9]University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK
[10]The Shrewsbury and Telford Hospital NHS Trust, Shrewsbury, UK
[11]The Royal Wolverhampton Hospitals NHS Trust, Wolverhampton, UK
[12]Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK
[13]Heart of England NHS Foundation Trust, Birmingham, UK
[14]Barts and the London School of Medicine and Dentistry, London, UK

*Corresponding author

Background: Screening for congenital heart defects (CHDs) relies on antenatal ultrasound and postnatal clinical examination; however, life-threatening defects often go undetected.
Objective: To determine the accuracy, acceptability and cost-effectiveness of pulse oximetry as a screening test for CHDs in newborn infants.
Design: A test accuracy study determined the accuracy of pulse oximetry. Acceptability of testing to parents was evaluated through a questionnaire, and to staff through focus groups. A decision-analytic model was constructed to assess cost-effectiveness.
Setting: Six UK maternity units.
Participants: These were 20,055 asymptomatic newborns at ≥35 weeks' gestation, their mothers and health-care staff.
Interventions: Pulse oximetry was performed prior to discharge from hospital and the results of this index test were compared with a composite reference standard (echocardiography, clinical follow-up and follow-up through interrogation of clinical databases).
Main outcome measures: Detection of *major CHDs* – defined as causing death or requiring invasive intervention up to 12 months of age (subdivided into *critical CHDs* causing death or intervention before 28 days, and *serious CHDs* causing death or

intervention between 1 and 12 months of age); acceptability of testing to parents and staff; and the cost-effectiveness in terms of cost per timely diagnosis.

Results: Fifty-three of the 20,055 babies screened had a major CHD (24 critical and 29 serious), a prevalence of 2.6 per 1000 live births. Pulse oximetry had a sensitivity of 75.0% [95% confidence interval (CI) 53.3% to 90.2%] for critical cases and 49.1% (95% CI 35.1% to 63.2%) for all major CHDs. When 23 cases were excluded, in which a CHD was already suspected following antenatal ultrasound, pulse oximetry had a sensitivity of 58.3% (95% CI 27.7% to 84.8%) for critical cases (12 babies) and 28.6% (95% CI 14.6% to 46.3%) for all major CHDs (35 babies). False-positive (FP) results occurred in 1 in 119 babies (0.84%) without major CHDs (specificity 99.2%, 95% CI 99.0% to 99.3%). However, of the 169 FPs, there were six cases of significant but not major CHDs and 40 cases of respiratory or infective illness requiring medical intervention. The prevalence of major CHDs in babies with normal pulse oximetry was 1.4 (95% CI 0.9 to 2.0) per 1000 live births, as 27 babies with major CHDs (6 critical and 21 serious) were missed. Parent and staff participants were predominantly satisfied with screening, perceiving it as an important test to detect ill babies. There was no evidence that mothers given FP results were more anxious after participating than those given true-negative results, although they were less satisfied with the test. White British/Irish mothers were more likely to participate in the study, and were less anxious and more satisfied than those of other ethnicities. The incremental cost-effectiveness ratio of pulse oximetry plus clinical examination compared with examination alone is approximately £24,900 per timely diagnosis in a population in which antenatal screening for CHDs already exists.

Conclusions: Pulse oximetry is a simple, safe, feasible test that is acceptable to parents and staff and adds value to existing screening. It is likely to identify cases of critical CHDs that would otherwise go undetected. It is also likely to be cost-effective given current acceptable thresholds. The detection of other pathologies, such as significant CHDs and respiratory and infective illnesses, is an additional advantage. Other pulse oximetry techniques, such as perfusion index, may enhance detection of aortic obstructive lesions.

Funding: The National Institute for Health Research Health Technology programme.

Contents

List of abbreviations

ANNP	advanced neonatal nurse practitioner
ANOVA	analysis of variance
AS	aortic stenosis
ASD	atrial septal defect
AVSD	atrioventricular septal defect
CAR	congenital anomalies register
CEAC	cost-effectiveness acceptability curve
CHD	congenital heart defect
CI	confidence interval
CN	consultant neonatologist
CoA	coarctation of the aorta
ECE	expedited clinical examination
FN	false-negative
FP	false-positive
HCHS	Health and Community Health Services
HLH	hypoplastic left heart
HLHS	hypoplastic left heart syndrome
HTA	health technology assessment
IAA	interrupted aortic arch
ICER	incremental cost-effectiveness ratio
IQR	interquartile range
NETSCC	NIHR Evaluation, Trials and Studies Coordinating Centre
NICE	National Institute for Health and Clinical Excellence
NIHR	National Institute for Health Research
PDA	patent ductus arteriosus
PS	pulmonary stenosis
PSA	probabilistic sensitivity analysis
QALY	quality-adjusted life-year
RCOG	Royal College of Obstetrics and Gynaecology
SHO	senior house officer
SpR	specialist registrar
STAI	State-Trait Anxiety Inventory
STARD	Standards for the Reporting of Diagnostic Accuracy Studies
TGA	transposition of the great arteries
TN	true-negative
TOF	tetralogy of Fallot
TP	true-positive
VSD	ventricular septal defect
WTP	willingness to pay

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

Executive summary

Background

Congenital heart defects (CHDs) are the most common group of congenital malformations and one of the leading causes of infant death in the developed world. Early detection of major CHDs – which cause death or require invasive intervention before 1 year of age – may improve outcome. Current routine screening for CHDs relies on a mid-trimester anomaly ultrasound scan in pregnant women, involving imaging of the heart chambers, and a postnatal clinical examination involving assessment of the cardiovascular system. Both of these have a relatively low detection rate and a number of babies are discharged from hospital before a CHD is diagnosed. A proportion of these may die or present in such a poor clinical condition that the outcome, despite treatment, is compromised.

This report assesses the accuracy, acceptability to both parents and clinical staff, and the cost-effectiveness of an alternative approach, based on pulse oximetry screening, to determine the value of this method in diagnosing those CHD lesions that are potentially life-threatening.

Objectives

This health technology assessment completed three distinct pieces of work, the objectives of which were to determine:

- the accuracy of pulse oximetry against the composite reference standard of echocardiography, clinical follow-up and interrogation of regional and national clinical databases, and to determine the added value of pulse oximetry over routine antenatal ultrasound screening
- the acceptability of pulse oximetry testing to parents and health-care staff
- the cost-effectiveness of pulse oximetry testing compared with existing strategies for CHD screening.

Methods

A test accuracy study was performed in six large maternity units with delayed verification in test negatives. The index test of pulse oximetry testing was performed in 20,055 eligible newborns prior to discharge from hospital. Those not achieving predetermined oxygen saturation thresholds underwent the reference standard of echocardiography. All other infants were followed up to 12 months of age through the interrogation of regional and national congenital anomaly and cardiac registries and clinical follow-up. The study compared the accuracy of the index test in detecting major CHDs subdivided into critical (causing death or requiring invasive intervention before 28 days) and serious (causing death or requiring invasive intervention between 1 and 12 months of age).

Acceptability of testing to participants was evaluated through a structured questionnaire completed after testing and again at 1 year in a subsample. The characteristics of those who declined to take part in the study when first approached were also analysed. Acceptability of

testing to clinical staff was evaluated through focus groups and e-mail surveys with those who had taken part in the study.

For the economic evaluation, resource usage data were collected alongside the test accuracy study to establish the cost of pulse oximetry testing. A decision-analytic model was constructed to estimate the cost-effectiveness of various screening and prevention strategies, using an NHS perspective and an outcome of cost per case of timely diagnosis of major CHDs.

Results

Main findings of test accuracy study

There were 53 cases of major CHDs (24 critical and 29 serious) within the study cohort of 20,055 babies. Of those with an abnormal result following pulse oximetry testing, 26/195 [13.33%; 95% confidence interval (CI) 8.9% to 18.92%] had major CHDs (18 critical and 8 serious). Of the babies who had a normal pulse oximetry result, 27/19,860 (0.14%; 95% CI 0.09% to 0.20%) had major CHDs (6 critical and 21 serious).

For the full cohort, pulse oximetry had a sensitivity of 75.0% (95% CI 53.3% to 90.2%) for critical cases and 49.1% (95% CI 35.1% to 63.2%) for critical plus serious cases combined. For the cohort in which CHDs had not been suspected following antenatal ultrasound, pulse oximetry had a sensitivity of 58.3% (95% CI 27.7% to 84.8%) for critical cases (12 babies) and 28.6% (95% CI 14.6% to 46.3%) for critical plus serious cases combined (35 babies). One in 119 babies (0.84%) without serious or critical CHDs had a false-positive (FP) result (specificity 99.16%, 95% CI 99.02% to 99.28%). In addition, in the FP cohort, 6/169 (3.5%) of babies had CHDs defined as significant but not major, and 40/169 (24%) had a respiratory or infective illness requiring hospital treatment. Thus, in the test-positive cohort, in total, 72/195 (37%) had a condition requiring medical intervention.

Main findings of acceptability studies

Across all parts of the study, parent and staff participants were predominantly satisfied with pulse oximetry screening, perceiving it to be an important and valuable test to detect unwell babies. There was no evidence that mothers given FP results were more anxious after taking part in the screening processes than those given true-negative (TN) results, although the former were less satisfied with the test and gave higher depression scores (a small, but statistically significant difference). In multivariate analyses, higher anxiety and depression was predicted by lower optimism, lower overall satisfaction and ethnicity (white British/Irish participants being less anxious). Satisfaction with screening was predicted by higher perception of the treatment's ability to control heart disease, comprehensibility of heart disease, and lower stress, anxiety and depression. White British/Irish participants were more satisfied than those of other ethnicities. Indian mothers were less satisfied overall with screening and Pakistani mothers were more stressed during screening than white British/Irish mothers. Communication problems were indicated as a cause of worry by participants, and staff identified a need for further training in communicating with parents about the study and for giving results, especially where a positive result is found. Malfunctioning equipment could increase anxiety for parents and staff alike.

Despite creating an additional workload in testing babies, the test was seen as reassuring for staff, and positively impacted on the roles of those caring for ill babies as they could be treated while in a less critical condition.

Main findings of health economics

The pulse oximetry test took on average 6.9 minutes (median 5 minutes, range 1 to 30 minutes) to be completed. Taking into account equipment costs, the procedure is estimated to cost £6.24 per test. The additional cost of including pulse oximetry as an adjunct to the current practice of clinical examination of the newborn was estimated to be approximately £24,900 per additional case of timely diagnosis. This estimate was shown to be robust in an extensive sensitivity analysis.

Conclusions

Implications for health care

Pulse oximetry is a safe, simple, non-invasive, feasible and reasonably accurate test that is acceptable to parents and clinical staff, and has a sensitivity that appears to be superior to that of antenatal screening and clinical examination. The use of both pre- and postductal saturations compared with postductal alone appears to be advantageous and, in practice, does not take significantly longer to perform.

Pulse oximetry adds value to existing screening procedures and is likely to identify cases of critical CHD that would otherwise go undetected. The detection of other pathologies, such as significant CHDs and respiratory and infective illnesses, is an additional advantage.

Pulse oximetry as an adjunct to clinical examination is twice as costly, but will detect more babies with CHDs.

Recommendations for research

Pulse oximetry improves detection rates of critical CHDs. The majority of critical cases missed by pulse oximetry (and by other screening methods) are associated with obstruction of the aortic arch as these conditions are often not associated with hypoxaemia. Further investigation of other oximetry techniques, such as perfusion index, may enhance the detection rates for these lesions.

Further research should be conducted with mothers of different ethnicities to gain a greater understanding of factors limiting participation and satisfaction with testing. Research should also be conducted with mothers given FP results at the time of discharge from hospital to verify whether they experience heightened anxiety at this point.

Funding

The National Institute for Health Research Health Technology programme.

Chapter 1

Introduction and background

Congenital heart defects

The term congenital heart defect (CHD) encompasses a variety of lesions with a wide spectrum of clinical importance, ranging from those of no functional or clinical significance to potentially life-threatening 'critical' lesions. If undiagnosed, infants with a critical lesion are at risk of acute cardiovascular collapse or death. At least 18 distinct types of CHD are recognised, with many anatomical variations.[1]

Congenital heart defects are the most common group of congenital malformations, with a reported incidence of between 4 and 10 per thousand live-born infants.[1–3] CHDs are one of the leading causes of infant death in the developed world, accounting for more deaths than any other type of malformation; up to 40% of all deaths from congenital abnormalities,[1,4] and 3–7.5% of infant deaths.[5,6] Apparent increases in the prevalence of CHDs[7,8] are likely to be explained predominantly by the increased detection of relatively minor defects (which may never become clinically apparent), as echocardiography has become more sophisticated and widely available.[8,9]

Most newborns with a CHD can be diagnosed using echocardiography and, if necessary, stabilised with prostaglandin infusion and treated with surgery or transcatheter intervention.[10] Surgery has resulted in marked improvements in survival, particularly for those infants with potentially life-threatening conditions. However, if such defects are not detected sufficiently early then severe hypoxaemia, shock, acidosis and death are potential sequelae. Such cardiovascular compromise, if not lethal, can have significant long-term effects as a consequence of significant multiorgan insults, including hypoxic–ischaemic brain injury. Timely recognition of these conditions is likely to improve outcome and therefore the evaluation of screening strategies to enhance early detection is of great importance.

Definitions

There is considerable variation in the definitions of severity of CHDs within the published literature. Terms such as major, critical, severe, complex, serious and significant are frequently used, but the lack of agreed definitions makes comparisons across studies difficult.

Because of the wide spectrum of severity, it is vital when evaluating the potential benefit of screening to be clear about the nature of CHDs that is considered important from the screening perspective. For the purposes of this report, and in keeping with previously published studies,[8] we have defined a congenital heart defect as 'a gross structural abnormality of the heart or intrathoracic great vessels that is actually or potentially of functional significance'.[11] We have further categorised CHDs according to echocardiographic findings and clinical course, in order to indicate the clinical importance of the lesion (*Table 1*).

Major CHDs resulting in death or requiring invasive intervention (surgery or cardiac catheterisation) during infancy are the lesions in which early detection by screening is most likely to improve outcome. We have further divided this group into two subcategories: critical

TABLE 1 Study definitions of congenital heart disease

Normal	No echocardiographic abnormalities
Non-significant	Presence of any one of the following at birth and findings no longer detected at 6 months: small PDA; small interatrial communication (PFO/ASD); muscular VSD; mildly abnormal turbulence at branch pulmonary artery
Significant	Presence of any one of the following at birth and findings persist for longer than 6 months of age: small PDA/PFO/muscular VSD; mildly abnormal turbulence at branch pulmonary artery. Also any cardiac lesion that requires regular monitoring beyond 6 months or requiring drug treatment, but not categorised as serious or critical
Serious	Any cardiac lesion not defined as critical, which requires intervention (cardiac catheterisation or surgery) or results in death between 1 month and 1 year of age
Critical	All infants with hypoplastic left heart, pulmonary atresia with intact ventricular septum, simple transposition of the great arteries or interruption of the aortic arch
	All infants dying or requiring surgery within the first 28 days of life with the following conditions: CoA; aortic valve stenosis; pulmonary valve stenosis; TOF; pulmonary atresia with VSD; total anomalous pulmonary venous connection

ASD, atrial septal defect; CoA, coarctation of the aorta; PDA, patent ductus arteriosus; PFO, patent foramen ovale; TOF, tetralogy of Fallot; VSD, ventricular septal defect.
Non-significant and significant CHDs are classified as 'minor' CHDs, and serious and critical as 'major' CHDs.

and serious. Critical lesions are most likely to present in the first few days or weeks of life, usually as a result of closure of the ductus arteriosus. During fetal life the ductus is an important shunt between the aorta and the pulmonary artery, which allows blood to bypass the uninflated lungs in utero. After birth, once the lungs are inflated, alteration of blood flow to the lungs and an increase in blood oxygen levels causes the ductus to contract and close. This event usually takes place over the first few hours of life, but patency may persist for longer than this. In a CHD, in which normal blood flow to the lungs or to the systemic circulation is restricted, blood flow through the ductus can be vital. In these 'duct-dependent' circulations, closure of the ductus can be catastrophic and cause acute cardiac decompensation as previously described.

For the purposes of this study, we have defined critical CHDs in accordance with a previously published UK categorisation[2] to include all potentially life-threatening duct-dependent conditions plus infants dying or undergoing invasive procedures (surgery or cardiac catheterisation) within the first 28 days of life, although it is accepted that death from undiagnosed CHDs can occur after that age. Serious CHDs are defined as those defects not classified as critical that result in death or invasive intervention within 12 months of age (see *Table 1*).

The prevalence of major defects remains essentially unchanged at around 2.5 per 1000 live births.[1,7,8] Critical lesions have an estimated incidence of 1–1.8 per 1000 live births[2,12,13] and this group accounts for between 15% and 25% of all CHDs, depending on the definitions used.[2,10]

Prognosis for congenital heart defects

In recent years there have been significant improvements in the survival of infants born with CHDs, mainly as a result of improved imaging and advances in surgical techniques. It is likely that the survival rates for infants operated upon in the 21st century will be even better than those for infants who received surgery 15–20 years ago. In the USA, between 1995 and 2005, the mortality rate from CHDs declined by 42%, and the recent aggregated postoperative mortality rate is estimated at 3.7%.[1]

Most deaths from a CHD occur within the first year and these are commonly as a result of extracardiac anomalies (e.g. trisomies 13 and 18), cardiovascular collapse in the neonatal period or perioperative mortality. Mortality from specific cardiac defects varies depending on prevalence, association with extracardiac defects, likelihood of acute decompensation in the neonatal period and amenability to surgical intervention. For example, ventricular septal defect (VSD), the most common CHD, accounting for 37% of all defects[8] and the highest proportion of deaths in infancy due to a CHD (17%), is more frequently associated with extracardiac defects, and most infants dying with VSD in the first year of life have non-cardiac cause of death.[14] By contrast, transposition of the great arteries (TGA) and hypoplastic left heart (HLH) account for 4% and 2% of CHDs, respectively, but account for 7% and 15% of all deaths in infancy; 82% of deaths with TGA and 79% of deaths from HLH are due to the cardiac condition alone.[14]

The majority of children with major CHDs will require surgical intervention within the first year of life. The majority of CHDs are amenable to surgery and, in general, for infants with an isolated CHD surgical mortality rates are low. Some surgery, but by no means all, is corrective, resulting in an anatomically normal heart.

Whatever the mode of detection of CHDs, the definitive surgical procedure will be almost certainly the same. However, circulatory collapse prior to surgery, resulting in shock and acidosis, can have a significant effect on outcome. Poor clinical status at the time of operation increases surgical mortality and adversely affects outcome.[15] Evidence for worse surgical outcome associated with perioperative instability has also been shown for specific critical lesions such as TGA,[16] coarctation of the aorta (CoA)[17] and HLH.[18]

At present, it is estimated that over 80% of babies born with a CHD will survive to 16 years of age,[6] and this survival rate is likely to increase with increasing surgical expertise. Data on long-term sequelae, particularly neurodevelopmental outcome, are lacking, but it is estimated that severe neurological deficit occurs in 5–10% of patients and milder neurological problems occur in up to 25%.[6] Thus, for the majority of patients with CHDs the outcome is favourable, with survival to adulthood the norm, and quality of life is likely to be reasonable.[6] Detection of critical CHDs in a timely manner, i.e. a preoperative diagnosis prior to death or cardiovascular collapse, is highly likely to further improve survival and long-term outcome.

Detection of congenital heart defects

Primary prevention is not possible; therefore, detection of CHDs prior to the onset of symptoms allows the possibility of interventions that may influence the natural history of the condition and subsequent outcome. As previously stated, this is of particular importance in infants with potentially life-threatening, critical CHDs, most of whom are asymptomatic at birth[2] and in whom deterioration or death can occur before the condition is recognised.[19]

An undetected CHD is a significant cause of unexpected neonatal or infant death. In one UK study, 15% of infants with CHDs who died before 12 months of age had a CHD that was undiagnosed prior to death.[19] Failure to diagnose a critical CHD prior to discharge from hospital occurred in up to 26% of infants in a Swedish study over an 8-year period, with an increase in infants discharged without diagnosis over the study period.[20] In the same study, 20% of critical lesions were undetected prior to discharge.[20] In UK studies, 25–30% of infants with potentially life-threatening conditions[2,15] and almost 80% of infants with obstructive left heart defects (the main causes of death from an undiagnosed CHD after discharge and before diagnosis) left

hospital undiagnosed.[21] Similar data have been reported in the USA; 1 in 10 infants with a CHD dying in the first year of life did not have the malformation diagnosed before death and, of the infants who died in the first week of life, one-quarter did not have a diagnosis before death.[22] Death at home or in hospital emergency rooms occurred in 50% of infants with undetected critical CHDs.[23]

Gold standard

Postnatal echocardiography is well established as the gold standard for diagnosing CHDs. However, it has to be remembered that, as studies of prevalence show, echocardiography may also contribute to an apparent rising incidence of CHDs mainly as a result of the detection of abnormalities which are of no functional or clinical significance.[7,8] As a result, echocardiography is likely to have significant limitations as a screening tool, mainly because of the high false-positive (FP) rate,[6,10] but also as a result of cost and lack of availability of trained personnel to perform the examinations.

Screening procedures

A screening programme is directed at a population that may be at risk of a disease or its complications and offers one or more tests to identify those who need further investigation or treatment. Screening targets apparently healthy fetuses and newborn babies, and provides parents and health professionals with information with which to make informed choices about their health. It has the potential to reduce morbidity and improve quality of life through early diagnosis, but there can be disadvantages, and any screening programme should be systematically evaluated before implementation as a public health policy. Current screening strategies to detect CHDs in the UK include physical examination and antenatal ultrasound.

Physical examination

Current UK guidance recommends examination of the cardiovascular system in all infants shortly after birth and again at 6–8 weeks of age,[24] and such practice has been in place for decades.[25] The cardiovascular component of the newborn screening clinical examination involves auscultation of the heart for murmurs and additional sounds, palpation of the peripheral pulses, particularly the femoral pulses, and observation for cyanosis.

A UK retrospective study of all infants with a CHD diagnosed in the first 12 months of life showed that routine neonatal examination failed to detect over half of the babies with a CHD and the later examination at 6 weeks missed one-third of babies.[26] Some CHDs are even more difficult to detect by examination; in a further UK study of 120 babies with CHDs causing obstruction to the systemic circulation – HLH, CoA, aortic stenosis (AS) and interrupted aortic arch (IAA) – 94 (78%) were discharged home without a diagnosis and, although an initial examination was noted in 34 babies, only 8 were referred.[21] The difficulties encountered when determining cardiovascular problems by physical examination are well known; many of the physical signs which may alert the clinician to CHDs are unreliable. Mild cyanosis is difficult to detect with the naked eye even if the examiner is very experienced,[27] and peripheral pulses may still be palpable as a result of blood flow through the patent ductus arteriosus. The prevalence of murmurs detected during the neonatal examination is estimated to be between 0.6% and 4.2%.[10,28] The presence of a murmur has been shown to be associated with a 54% chance of a CHD.[28] However, they often do not occur in critical lesions such as valve atresias and TGA.[10] Murmurs associated with minor CHDs are common, and those associated with more complex lesions may become apparent only after the pulmonary vascular resistance has fallen, which may occur after discharge from hospital.

One of the main additional factors that is likely to influence early detection by physical examination is length of hospital stay following delivery. The sooner a baby is discharged, the

lower the likelihood of the development of clinical signs or symptoms which could be detected by examination. Kuehl *et al.*[22] identified an increased risk of missing critical CHD if babies were discharged before 48 hours of age. However, there is an increasing trend towards earlier discharge in North America and Scandinavia.[29–32] In the UK, an increasing proportion of babies are being sent home within 12 hours after delivery, and this trend is likely to continue.[6] This is liable to have a further impact on the number of babies discharged from hospital before a CHD has been diagnosed.

Antenatal ultrasound (mid-trimester or anomaly scan)

The opportunity to perform a four-chamber ultrasound view of the fetal heart in order to diagnose fetal cardiac anomalies was recognised in the early 1980s.[33–35] Ultrasound is now established as a routine screening procedure, with 97% of all UK units offering second-trimester anomaly scans to all pregnant women between 18 and 22 weeks' gestation.[36] Both the Royal College of Obstetrics and Gynaecology (RCOG) and the National Institute for Health and Clinical Excellence (NICE) have issued guidelines relating to the nature of antenatal anomaly screening and have recommended that, in addition to the four-chamber view of the heart, the outflow tracts should also be visualised.[37,38] However, in the UK in 2002, only 57% of units reported routinely examining the outflow tracts and the range of units offering this service within health regions was 25–84%.[36]

The identification of a cardiac anomaly at this stage of pregnancy allows time for appropriate counselling regarding the nature of the lesion and the management options that are available, for example termination of pregnancy or planned delivery and timely medical and surgical intervention. Cardiac defects diagnosed antenatally are frequently complex and associated with extracardiac abnormalities, such as chromosome disorders, genetic syndromes or are part of a multiple malformation disorder that may be associated with a poor outcome. However, antenatal diagnosis can improve the outcome for isolated critical conditions.[39] Antenatal detection of HLH,[18] CoA[17] and TGA[40] has been reported as resulting in better surgical survival. This is likely to be as a result of better clinical state prior to surgery with reduction in preoperative acidosis, cardiovascular compromise and end-organ dysfunction as previously discussed.[15,41]

The effectiveness of antenatal ultrasound screening to detect CHDs in low-risk populations at 18–23 weeks' gestation is extremely variable with sensitivities ranging from 0% to 81%.[38,42] In the UK in 1999, the average national detection rate was reported as approximately 23%[33] and similar figures have been reported more recently in the UK and in other countries.[18,40,43–49] Data from the north-east of England show that over the 20-year period from 1985 to 2004 an average of only 8% of life-threatening CHDs were diagnosed antenatally, and although there was an increase in the proportion of infants diagnosed antenatally over the time period, at best only 20% were picked up in this way.[2]

In addition, it has to be remembered that a significant proportion of those babies detected antenatally have associated non-cardiac abnormalities that may themselves be life-threatening. The antenatal detection rate in isolated CHDs is lower. Prenatal detection also offers the opportunity for termination of pregnancy, which will reduce the prevalence of CHDs in the live-born population.

With intensive training, improvements in the detection rates of CHDs during the routine anomaly scan can be demonstrated; however, in the UK the proportion of CHDs detected antenatally still varies between 20% and 55%.[35]

National standards for the anomaly scan published in 2010 by the NHS Fetal Anomaly Screening Programme state that a four-chamber view of the heart and views of the right and left ventricular

outflow tracts should be performed. The use of colour-flow Doppler to improve detection is encouraged, but not a requirement.[50] The expected detection rate for serious CHDs is stated as 50%.[50]

Fetal echocardiography

It is possible to detect most forms of CHDs even in the first trimester,[51] using antenatal echocardiography; the main exceptions to this being milder obstructive lesions of the great arteries [AS, pulmonary stenosis (PS) and CoA], persistent patent ductus arteriosus (PDA), secundum atrial septal defects (ASDs) and some forms of VSD.[52] However, it is not a universal screening test. Tertiary obstetric centres will often offer fetal echocardiography to women at a high risk of a pregnancy complicated by a CHD – risk factors such as family history of CHDs, maternal diabetes and exposure to teratogens, such as lithium or phenytoin, would usually lead to fetal echocardiography by a paediatric cardiologist or individuals trained in the technique. However, despite the high risk, the absolute numbers of cardiac abnormalities in these groups are low and the majority of cases of CHDs occur in low-risk groups, as outlined in the previous section. Therefore, the majority of cardiac abnormalities detected antenatally will be identified in low-risk groups after a routine anomaly scan.[35] Even in tertiary centres the ability to detect specific defects varies with the nature of the defect. Four-chamber abnormalities are more readily detected than outflow tract lesions. For example, antenatal diagnosis has been reported as 81% for HLH in one tertiary centre, whereas TGA was diagnosed antenatally in only 25% of cases.[53]

A systematic review of five studies comprising over 60,000 unselected and low-risk patients found that among the low-risk populations frequency of correct diagnosis ranged from 35% to 86%. However, the potential for ascertainment bias and the choice of reference standard that limited the validity of these findings was also reported, and the variation in sensitivity across studies was not explainable by clinical factors, such as scanning regime, operator skill and equipment. The review concluded that the evidence about the accuracy of fetal echocardiography did not lend support to its routine use among unselected and low-risk populations during the second trimester to detect congenital heart disease.[54] Additionally, expanding fetal echocardiography screening to include the low-risk population also has significant training and resource implications.[6,53,55]

Pulse oximetry

The rationale for pulse oximetry screening is based on the fact that hypoxaemia is present, to some degree, in the majority of CHDs. This may result in obvious cyanosis; however, mild degrees of hypoxaemia cannot be detected by clinical observation, even by experienced clinicians.[27] The difficulty is exacerbated in infants with pigmented skin.[56]

Pulse oximetry was developed as a non-invasive method to determine arterial oxygen saturations (SpO_2) and has been widely used in intensive care, operating theatres and emergency units for over 30 years. The ability to detect the different absorption spectra oxygenated and deoxygenated haemoglobin allows pulse oximeters to measure the amount of oxygen-saturated haemoglobin in the capillaries of an extremity, such as a finger or an ear lobe in an adult, or a hand or foot in a baby. Pulse oximetry thus allows the detection of hypoxaemia that would not necessarily produce visible cyanosis; the technique correlates well with arterial blood gas measurements[56] and does not require calibration.

Oximeters can measure either functional or fractional oxygen saturations. 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.[13,57,58]

The fetus is hypoxic in utero, with oxygen saturations of 30–60%.[59] The events that follow delivery – clamping of the umbilical cord, initial inflation of the lungs allowing pulmonary gas exchange and an increase in pulmonary circulation – result in a rapid rise in the arterial oxygen tension. The mean preductal (right hand) and postductal (foot) saturations rise to 73% and 67%, respectively, within the first 2 minutes after birth, and to 92% and 89% by 10 minutes.[60] The difference between right hand and lower limb saturations reflects right-to-left shunting across the ductus arteriosus and this generally diminishes with time in most infants, with both pre- and postductal saturations reaching 95% by 1 hour of age.[60] Saturations are generally stable in the first 24 hours with a mean of 98% (similar to values obtained from older neonates); however, periods of desaturation may occur,[61] and larger studies have shown that many healthy newborns (up to 5%) may have episodes of saturations of < 95% within the first 24 hours.[13]

These findings have led to exploration of the possibility that pulse oximetry may be useful in detecting hypoxaemia associated with CHDs in apparently healthy newborns, and a number of studies have been published that have used pulse oximetry as a screen for CHDs in this group.[13,62–71]

In a systematic review published in 2007, we reviewed the eight published studies available at that time.[72] We highlighted a number of problems when assessing the accuracy of this test as there was a wide variation in methodology, particularly patient selection, timing of saturation measurement, cut-off levels for a positive test result, nature of CHD screened for, rigour of follow-up and type of oximeters used (*Table 2*). In addition, most of the studies were relatively small and as a consequence the prevalence of CHDs was relatively low, particularly in the studies in which patients with an antenatally suspected CHD were excluded. None of the studies calculated power a priori and sample size was often inadequate to estimate sensitivity precisely. Since the publication of our systematic review, three more studies have been published,[68–70] including two large Scandinavian studies[69,70] (see *Table 2*).

To date, only four studies have recruited more than 10,000 patients,[64,68–70] although all of these excluded patients with antenatally diagnosed CHDs. In one study, this led to a significant reduction in the prevalence of CHDs with only 4/31 patients with critical CHDs receiving screening.[68] Sixteen out of 31 patients with critical CHDs (52%) were identified antenatally in this study cohort and therefore not screened, and 11 babies with critical CHDs were not screened for other reasons.[68] In the two largest studies, the antenatal diagnosis rate was comparatively low – only 3.3% of critical lesions detected in the Swedish study[70] and 7% of total CHDs detected antenatally in the Norwegian study.[69]

These two Scandinavian studies have significantly increased the number of patients screened by pulse oximetry[69,70] and demonstrate sensitivity rates for critical CHDs in large cohorts of 62% and 77%, respectively. The Norwegian study also reported the sensitivity rate for all CHDs, but this was low at only 10%. When combined with physical examination both studies demonstrated an increase in sensitivity for pulse oximetry.

The methodological differences are important when comparing the results of studies and assessing test accuracy. There were various thresholds for abnormality of test results. Studies that screened for all CHDs generally had much lower sensitivity than those that screened specifically for critical or cyanotic lesions. Some studies with very high sensitivity rates[66,71] tended to recruit lower numbers and therefore had a lower disease prevalence, which makes interpretation of these results more difficult (see *Table 2*). Ascertainment of missed cases (either those dying outside hospital or those presenting to other cardiac units) in order to identify false-negatives (FNs) was complete in only three studies.[13,64,70]

TABLE 2 Details of pulse oximetry screening studies

Author	Threshold	Functional/ fractional	Single/ repeat	Timing of testing	No. of patients	Sensitivity (%)	Specificity (%)	FP rate	Antenatal diagnosis included	Respiratory problems included	Follow-up post discharge completed
All CHDs											
Richmond, 2002[13]	<95% foot	FRAC	R	At 2 hours and at discharge	5626	25 (12.7–41.2)	99.6 (99.4–99.7)	0.4%	Yes	Yes	Yes
Arlettaz, 2005[71]	<95% foot	FUNC	S	6–12 hours Repeat 4–6 hours Median 8 hours	3262	96.8 (73.6–100)	99.7 (99.5–99.9)	0.3%	Yes	Five PPHN only	No
Bakr, 2005[62]	<94% hand and foot	FRAC	R	Before discharge, three readings Average 32 hours	5211	30.8 (9.1–61.4)	100 (99.9–100)	1/5006	No	No	Yes Not mortality
Meberg, 2008[69]	<95% foot	FUNC	R	First day Median 6 hours	50,008	9.9 (7.3–13.2)	99.4 (99.3–99.5)	0.6%	No	Yes	Yes Not mortality
Specified CHDs: cyanotic or critical											
Hoke, 2002[63]	<92% in leg >6% diff leg <arm	FUNC	R	<6 hours 24 hours discharge	2876 + 32	75 (57.8–87.9)	87.9 (86.6–89.1)	1.8%	Yes	1 PPHN Nil else	No
Koppel, 2003[64]	≤95% foot	FUNC	S	>24 hours	11,281	60 (14.7–94.7)	100 (100–100)	1/11,281	No	1 PPHN Nil else	Yes

Author	Threshold	Functional/ fractional	Single/ repeat	Timing of testing	No. of patients	Sensitivity (%)	Specificity (%)	FP rate	Antenatal diagnosis included	Respiratory problems included	Follow-up post discharge completed
Reich, 2003[65]	<95% foot or hand ≤90% or >3% difference	FUNC	R	Before discharge	2114	66.7 (9.4–99.2)	99.9 (99.7–100)	0.1%	No	No	Yes 150 days
de Wahl Granelli, 2005[66]	<95% hand and foot or >3% difference	FUNC	S	12–48 hours Median 24 hours	266	89.4 (79.4–95.6)	99 (96.4–99.9)	1%	Yes	No	N/A
Rosati, 2005[67]	<96% foot	FUNC	S	>24 hours Median 72 hours	5292	66.7 (9.4–99.2)	100 (99.9–100)	0%	No	No	Yes Not mortality
Sendelbach, 2008[68]	<96% foot	FUNC	R	4 hours Repeat at discharge	15,233	75 (19.4 –99.4)	94.4 (94–94.7)	1/15, 233	No	Yes indirectly	No
de Wahl Granelli, 2009[70]	<95% hand and foot or >3% difference	FUNC	R	Median 38 hours	39,821	62.1 (42.3–79.3)	99.8 (99.8– 99.9)	0.17%	No	Yes	Yes
Meberg, 2008[69]	<95% foot	FUNC	R	First day Median 6 hours	50,008	77.1 (59.4–89.0)	99.4 (99.3–99.5)	0.6%	No	Yes	Yes Not mortality

FRAC, fractional; FUNC, functional; N/A, not applicable; PPHN, persistent pulmonary hypertension of the newborn; R, repeat; S, single.

In general, studies that screened later (24–72 hours)[62,64,65,67,70] had lower FP rates; however, this has to be tempered by the fact that critical CHDs, if undiagnosed, may present with clinical deterioration in the first 48 hours of life; therefore, earlier testing may help to identify these patients before such an event can take place. The increasing trend towards earlier discharge for uncomplicated births also has to be taken into consideration when determining optimal timing. Some studies also included the detection of additional diagnoses such as respiratory problems and infections,[13,69,70] which, although classified as FPs, were also potentially serious illnesses. These non-cardiac problems tended to present earlier and suggest a potential additional advantage to earlier screening.

Also important is the threshold for an abnormal saturation; the higher the percentage saturation thresholds the greater the sensitivity but the lower the specificity, and the converse is true for a lower cut-off. Although a number of levels have been proposed as the lower saturation threshold, the most frequently used in the published studies is around ≤ 95%. Most studies have used postductal saturations as a single measurement, although some studies used both pre- and postductal saturations and included a difference between the two in the definition of abnormality (see *Table 2*). Use of a differential criterion may increase the likelihood of detecting those obstructive left heart lesions, such as CoA, where the ductus supplies some of the systemic circulation and may result in a difference in saturation between the upper and lower limbs.

The two recent, large prospective studies discussed above[69,70] demonstrate encouraging results, but were mostly limited to reporting the accuracy of pulse oximetry in detecting CHDs. Furthermore, they did not take into account the added influence of antenatal screening on the results, parents' and health-care professionals' acceptability of the test, psychosocial impact of FP results or identification of a non-critical CHD, and provided only limited analysis of the cost-effectiveness of such a screening programme.

Acceptability of pulse oximetry screening
When new antenatal or neonatal screening programmes are introduced, consideration should be given to the acceptability of screening to parents and to the psychological impact of the screening procedure. Any screening procedure is likely to raise anxiety as it raises the possibility, previously not considered by most parents, that a child may have a serious health condition. Our systematic review of pulse oximetry as a screening test identified no studies that evaluated the acceptability of screening to parents or the psychosocial impact of FP results.[72] It is important to address this as the literature on antenatal and neonatal screening more generally suggests that acceptability of screening has an impact on uptake and the effect of inaccurate results may extend over a considerable time period.[73–76] An earlier health technology assessment (HTA) review of newborn screening for congenital heart defects[6] reported focus groups with parents in the results of the review. Although welcoming pulse oximetry as a screening method, these parents raised similar concerns to those found in studies of other types of neonatal and antenatal screening, particularly in relation to test accuracy and poor communication by health-care professionals. If screening is to be implemented as a standard procedure, it must also be acceptable to health-care staff; it is important to assess the impact of screening on the roles of staff involved.

Cost-effectiveness of pulse oximetry screening
Adding the pulse oximetry test as an adjunct to the current practice of using clinical examination to screen for CHDs in newborn infants has the potential to improve the number of cases that receive a timely diagnosis. But such benefits need to be assessed against the additional costs that would be incurred and consideration of whether or not this is an appropriate use of limited health-care resources.

Knowles *et al.*[6] undertook a systematic review of the clinical and economic literature and used the data obtained to undertake a model-based economic evaluation of pulse oximetry to diagnose CHDs in the newborn. The authors developed a comprehensive model that suggested that pulse oximetry offered what could be considered a cost-effective screening test for detecting CHDs early in newborn infants. But the evaluation was based on secondary data from existing published sources and it was acknowledged that pulse oximetry required testing in a UK clinical setting to obtain primary data on its accuracy and associated costs. Against this background, we developed a new decision-analytic model based on the care pathways assessed in the current study and carried out an economic evaluation using new primary data on test accuracy and costs. The objective of the economic evaluation for this study was to assess the cost-effectiveness of pulse oximetry as an adjunct to current practice compared with current practice alone to ensure that decision-makers use available resources effectively.

Aims of the Health Technology Assessment project

The project was commissioned by the National Institute for Health Research (NIHR) HTA programme in 2007. The overall aim of the study was to evaluate the use of pulse oximetry as a screening procedure for the detection of serious and critical CHDs by assessing:

- the accuracy of pulse oximetry against the reference standards of echocardiography, and interrogation of regional and national congenital anomaly and cardiac defect databases
- the added value of pulse oximetry over routine antenatal ultrasound screening
- the acceptability to parents and clinical staff of pulse oximetry testing
- the cost-effectiveness of pulse oximetry testing compared with existing strategies of screening for CHDs.

The key objectives were:

1. to establish the feasibility of pulse oximetry as a screening test for CHDs among newborn infants
2. to determine the accuracy (sensitivity and specificity) of pulse oximetry using echocardiography and interrogation of regional and national congenital anomaly and cardiac defect databases as reference standards
3. to determine the acceptability of pulse oximetry testing among mothers whose babies tested positive and a selection of true-negative (TN) tests and also among health-care staff
4. to determine the cost and cost-effectiveness of pulse oximetry testing for CHDs and to compare this with other strategies for screening.

The accuracy study was designed and managed in such a way as to meet the Standards for the Reporting of Diagnostic Accuracy Studies (STARD) criteria for methodological quality.[77] Acceptability was assessed using a questionnaire survey and qualitative interviews. Cost-effectiveness analysis involved decision-analytic model-based economic evaluation.

Chapter 2

Diagnostic accuracy of pulse oximetry as a screen for congenital heart defects in asymptomatic newborn infants

Introduction

This chapter describes a test accuracy study (the PulseOx study) to determine the accuracy of pulse oximetry screening for serious and critical CHDs in asymptomatic newborns:

- *population* apparently healthy asymptomatic newborn infants with a gestation ≥ 35 weeks prior to discharge from hospital
- *index test* oxygen saturations from pulse oximetry readings from the right hand (preductal) and either foot (postductal)
- *reference standard* echocardiography, clinical follow-up and follow-up through interrogation of regional and national clinical databases
- *target condition* major CHDs (see *Table 1*).

Methods

Using recommended methods for diagnostic accuracy evaluation[77] (see *Appendix 1*), a study protocol was developed for a prospective accuracy study using delayed verification for test negatives.[78,79]

Research Ethics Committee (Trent Research Ethics Committee and local research ethics committees) and NHS Trust research governance approval was obtained for recruitment in six large obstetric units based in the UK: Birmingham Women's Hospital; Birmingham Heartlands Hospital; City Hospital, Birmingham; New Cross Hospital, Wolverhampton; Royal Shrewsbury Hospital; and University Hospital of Coventry and Warwickshire – all serving large, socioeconomically and ethnically diverse populations. These hospitals represented a spectrum of activity from specialised tertiary referral centre to busy district general hospital.

Study sample

Recruitment was organised and supported by research midwives, who worked with local community and antenatal clinics, midwifery teams and obstetrics leads. A dedicated co-ordinating midwife was employed at each centre with responsibility for overseeing staff training, consent, testing and data collection in her centre. In addition, a lead research midwife at Birmingham Women's Hospital liaised with all the co-ordinating midwives at each centre, providing training for consent procedures, trouble-shooting of recruitment and midwifery problems.

All pregnant women booked for delivery at one of the six participating units were approached for consent to be recruited into the study. Study information leaflets were given to mothers at the time of their antenatal booking visit or at the mid-trimester scanning visit. The leaflet emphasised
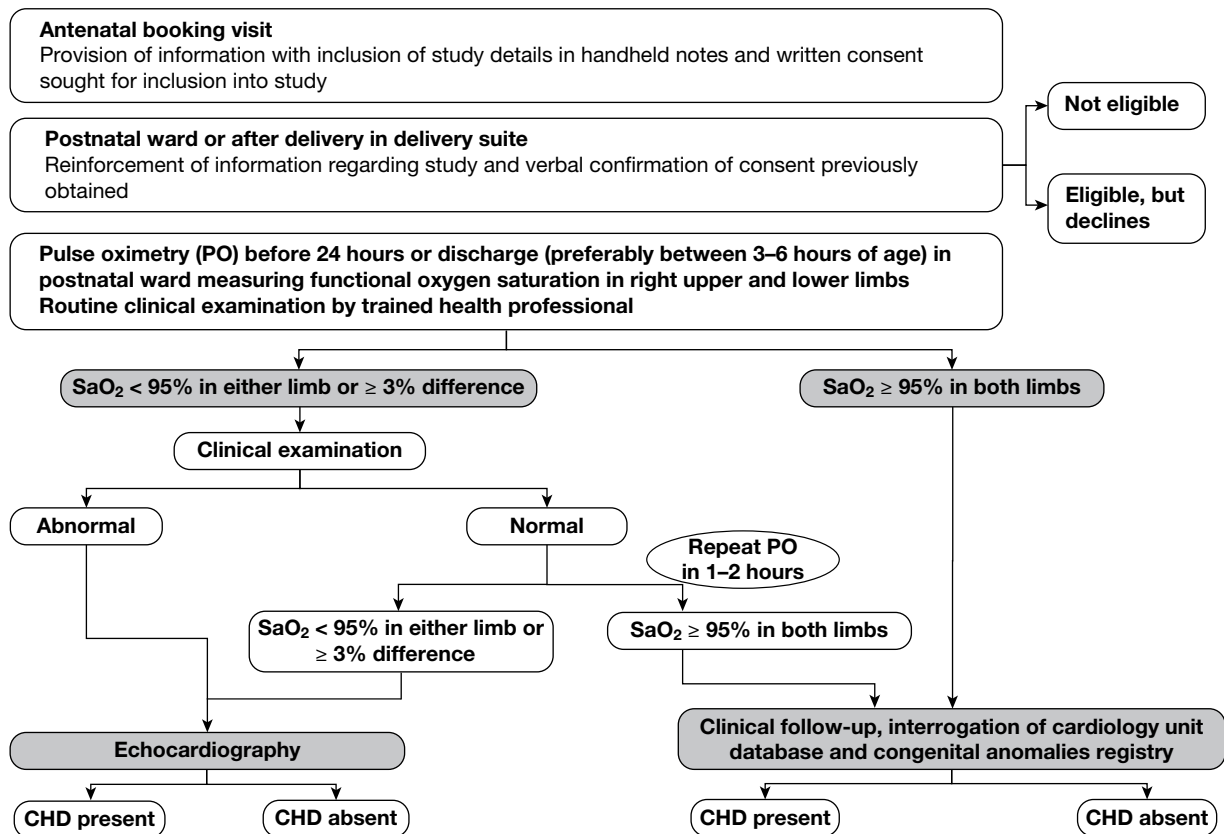
FIGURE 1 Flow chart of study organisation.

that pulse oximetry was introduced for routine use in the participating hospitals for the duration of the study, but that parents are entitled to choose whether or not they want their baby screened. Whenever possible, the written consent of the women was sought between 20 and 24 weeks or at least prior to labour and reconfirmed postnatally prior to testing eligible babies. The information leaflet and consent forms were also provided in minority languages (Arabic, Bengali, Chinese, French, Kurdish, Polish, Punjabi, Somali and Urdu) and also as audio-recordings on CD. Consent was supported by interpreters whenever possible.

All consecutive asymptomatic newborns with a gestation of 35 weeks or more were eligible for inclusion in the study. This included babies who were suspected antenatally of having a CHD. We excluded mothers who were unable to give consent through incapacity or lack of interpreter. Any newborn who showed symptoms of cardiovascular abnormalities at birth (such as overt cyanosis, dyspnoea or tachypnoea) was excluded from the study.

Index test

Pulse oximetry was performed after delivery and prior to discharge in all eligible, consented cases and was embedded in routine postnatal care provided by the midwife on the postnatal ward. We aimed to test early (within 3–6 hours of age), but in practice expected a range of testing times prior to discharge from hospital.

Clinical examination was not performed independently of the pulse oximetry result as ethically midwives would be expected to alert the medical staff of low saturations. This precluded a comparison of clinical examination with pulse oximetry on account of work-up bias.

Standardisation of pulse oximetry and feasibility pilot study

The Radical-7® pulse oximeter (Masimo, Irvine, CA, USA) was used in the study with the reusable probe LNOP YI. These have been shown to produce recordings free from motion artefact and to be capable of achieving stable, accurate readings of functional saturations in an active subject and also when perfusion is low with extremely low intra- and interobserver variability.[66] Five oximeters were available for each unit in order to ensure that readings could be taken at all times and babies were not missed because of faulty or misplaced machines. Oximeters were available on both the postnatal ward and delivery suite.

An initial technical feasibility study was carried out to establish the most practicable methods of testing and reporting results in the clinical setting. This allowed us to develop programmes for quality assurance and training. Testing was carried out by midwives and midwifery assistants depending on the availability of staff and working pattern preferences within each hospital. All staff involved in obtaining pulse oximetry readings were trained and their technique observed by the research midwife to ensure standard proficiency in testing. The importance of obtaining a high-quality, stable reading before documenting saturations was emphasised. If the saturation fluctuated between two or more readings and the signal quality was acceptable then the lower of the readings was documented.

Pulse oximetry was performed at a time deemed appropriate to clinical staff and after written informed consent was obtained. If consent had been obtained antenatally, then verbal confirmation of consent was obtained at the time of testing. An observational assessment of the baby was made prior to testing to ensure that the baby was asymptomatic. Either the oximeter was taken to the baby's cot on a trolley or the detachable front section of the oximeter was removed to allow handheld testing. The reusable probe was used on the right hand and either foot in a non-specified order and remained on the limb until a satisfactory trace and consistent saturation reading was obtained. The date and time of the test and the results – expressed as percentage functional saturation for each limb – were recorded on the patient data sheet. The reusable probe was cleaned with an alcohol swab before and after each patient test.

A saturation of < 95% in either limb or a difference of > 2% between the limb readings was considered to be abnormal. This threshold was chosen in an attempt to increase the sensitivity to detect left heart obstructive lesions, such as AS, CoA and IAA – treatable conditions that were most commonly missed in earlier studies with different thresholds.

If the oxygen saturations were considered abnormal, an expedited clinical examination (ECE) was performed by an individual trained in examination of the newborn. If this clinical examination was unremarkable, pulse oximetry was repeated 1–2 hours later. If the saturations remained abnormal on this second occasion, or if abnormalities of the cardiovascular system were identified on the ECE, these babies were classified as test positive and underwent echocardiography (*Figure 1*).

Other screening tests

Antenatal screening ultrasound

It was routine practice in each participating centre to perform screening mid-trimester anomaly ultrasound at between 18 and 22 weeks, including a four-chamber assessment of the fetal heart. Any pregnant woman in whom a suspected fetal cardiac anomaly was identified at this stage was referred to a fetal medicine centre for fetal echocardiography. In addition to this, in one of the centres (BWH), women at high risk of fetal cardiac anomaly [i.e. those with a family history of CHDs, maternal diabetes, increased nuchal translucency, taking medication known to increase the risk of CHDs (e.g. antiepileptic drugs), fetal dysrhythmias] were also tested by fetal echocardiography. In order to identify the results of antenatal anomaly ultrasound we

reviewed all fetal echocardiography results of study participants in order to identify screening scans that detected a suspicion of a CHD. We assumed that every woman received an anomaly scan and that all fetuses, in whom there was concern about cardiac abnormality, underwent fetal echocardiography.

Clinical examination

Patients who tested negative on the initial pulse oximetry test underwent a routine clinical examination, usually within the first 24–36 hours or prior to discharge, whichever was sooner. Patients who tested positive on initial testing underwent the ECE of the cardiorespiratory systems as described above. All patients receiving this examination also had the routine neonatal screening examination prior to discharge. No examination was performed without knowledge of the saturation result as we felt this would be unethical. It was not possible, therefore, to compare directly the accuracy of oximetry testing with examination. Only details of the expedited examination were recorded on the patient data sheet. Details of the routine clinical examination were collected retrospectively only if a CHD was suspected as a result.

Reference standard

A composite reference standard combining echocardiography and follow-up was used to identify immediate- and late-presenting cases of CHDs. Echocardiography was performed in all test-positive cases by trained individuals and the timing of the procedure was decided by the senior clinician responsible for the care of the baby; whenever possible, this was on the same day and always no longer than 72 hours after a positive pulse oximetry test result. In a small number of cases and at the discretion of the responsible clinician, infants perceived to be at low risk of a CHD despite a positive test were allowed to go home prior to echocardiography.

All echocardiograms were stored on videotape, collated, reviewed and ratified by a paediatric cardiologist. If there was disagreement then a second echocardiogram was produced.

The echocardiography result was categorised into one of the following groups: critical, serious, significant, non-significant or normal (see *Table 1*). We were particularly interested in the accuracy of diagnosing critical (potentially life-threatening conditions requiring invasive intervention within 1 month of life) and serious (requiring invasive intervention within 12 months of life) CHDs. Any baby with significant echocardiographic findings was followed up by a cardiologist in a routine manner. Those with non-significant findings were followed up until 6 months to establish non-significance.

All babies were also followed up through information obtained from regional and national registries. At 12 months after the end of recruitment, the regional congenital anomalies registers (CARs) and mortality registers for the West Midlands and surrounding regions [East Midlands and South Yorkshire (EMSYCAR), Oxfordshire, Berkshire and Buckinghamshire (CAROBB), South West (SWCAR) and Wales (CARIS)] were interrogated using diagnostic codes for cardiac abnormalities. Birmingham Children's Hospital paediatric cardiology database (HeartSuite) and the national Central Cardiac Audit Database (CCAD) were also interrogated to identify all CHD cases requiring intervention within 1 year (i.e. critical and serious cases).

Sample size and power consideration

The sample size of at least 20,000 neonates was calculated on the basis that the lower limits of the confidence intervals (CIs) for both sensitivity and specificity would exceed certain values for an estimated prevalence of 2 per 1000. For an assumed sensitivity of 75% and specificity of 99.5%, the study had 80% power to prove that the sensitivity was at least 52% and over 90% power to prove that the specificity was above 99.3% (both using a one-tailed significance level of 2.5%). Estimates were calculated by simulation (using 10,000 repetitions) to account for both the

sampling variability in the observed number of cases of CHDs and also the observed sensitivity and specificity. Binomial exact methods were incorporated to estimate CIs and a range of higher and lower prevalence rates were used to explore the effect on power and size of CIs.

An independent monitoring committee confidentially reviewed interim analyses to determine whether or not our assumptions for sample size were borne out at 3–4 months after commencement of recruitment in the accuracy study. It determined if the principal question on index test accuracy had been answered and monitored any adverse scenarios. The monitoring committee also reviewed reports of recruitment, compliance and estimated completeness of verification at 6-monthly intervals.[80] They met on two occasions and recommended continuing recruitment until the target recruitment of 20,000 babies was met.

Data analysis

For the analysis, the subjects were divided into two main cohorts. The first included all recruited babies, with the second excluding those individuals with an antenatal suspicion of a CHD from anomaly ultrasound screening that was subsequently confirmed by fetal echocardiography. This was undertaken in order to identify those in whom a positive pulse oximetry test result could make a difference to subsequent testing (most likely using echocardiography) and contingent health care. Those who underwent fetal echocardiography for a reason other than an abnormal anomaly scan were not excluded from the full cohort, as fetal echocardiography screening in this manner is not currently available as a routine test across the UK.

The diagnostic accuracy for the cohorts detailed above was evaluated by calculating sensitivity, specificity and predictive values for critical cases alone and for critical plus serious cases combined (see *Table 1* for definitions). The 95% CIs for each estimate were calculated using binomial exact methods.[81] The accuracy of anomaly scan alone was evaluated in a similar fashion. The accuracy of pulse oximetry according to the timing of test was evaluated using a logistic regression model allowing for time from birth to the first stage of the pulse oximetry testing as a continuous variable.

Results

Between February 2008 and January 2009, there were 26,513 deliveries within the participating hospitals. Mothers of 20,055 eligible babies (75.64%) consented to the study; 685 infants were ineligible for inclusion (2.58%); 2005 infants were not included because consent was declined (7.56%); and in 3768 the opportunity for screening was missed (14.21%). The main reason babies were missed was insufficient staffing levels to obtain informed consent and/or to undertake the screening. The total percentage recruitment of eligible babies whose parents consented to the study was 84%. Demographic details of the mothers included in the study are described in *Table 3*.

Seven hundred and seventy-two babies (3.85%) had an abnormal result with the first test, with 696 proceeding to an ECE (*Figure 2*). Seventy-six babies were wrongly classified at the time as test negative and therefore the expedited examination did not take place; these babies were followed up as per reference standard for a normal result. Of the 76 babies, 70 had both hand and foot saturation readings of > 94% and a difference of 3%; one had both saturation readings of > 94% and a difference of 4%. A further five cases correctly classified as test positive did not have the expedited clinical examination; in two cases the mothers declined the examination and for the remaining three the examination was not done for unknown reasons. None of these 76 babies

TABLE 3 Characteristics of mothers and babies included in the PulseOx study

Total consented, 2005[5]	
Characteristic	***n* (%)**
Age (years)	
< 20	1474 (7)
20–24	4541 (23)
25–29	5833 (29)
30–34	4784 (24)
35–39	2766 (14)
≥ 40	657 (3)
Ethnicity	
White	11,158 (56)
Asian	4956 (25)
Black	1323 (7)
Other	1550 (8)
Not known	1068 (5)
Gestational age (weeks)	
35–36	731 (4)
37–38	4396 (22)
39–40	11,029 (55)
> 40	3899 (19)
Gravida[a]	
1	7500 (37)
2	5708 (28)
3	3273 (16)
4	1707 (9)
5+	1866 (9)
Missing	1
Parity	
1	9197 (46)
2	5800 (29)
3	2834 (14)
4+	2224 (11)
Pregnancy type	
Singleton	19,537 (97)
Twins	518 (3)
Baby gender	
Female	9874 (49)
Male	10,181 (51)
Birthweight (kg)	
Mean (SD)	3.33 (0.52)
No. recruited in each centre	
Birmingham Women's	5708
Birmingham Heartlands	3378
City Hospital	1845
New Cross	2460
UHCW	3300
Royal Shrewsbury	3364

SD, standard deviation; UHCW, University Hospital of Coventry and Warwickshire.
a   Including current pregnancy.
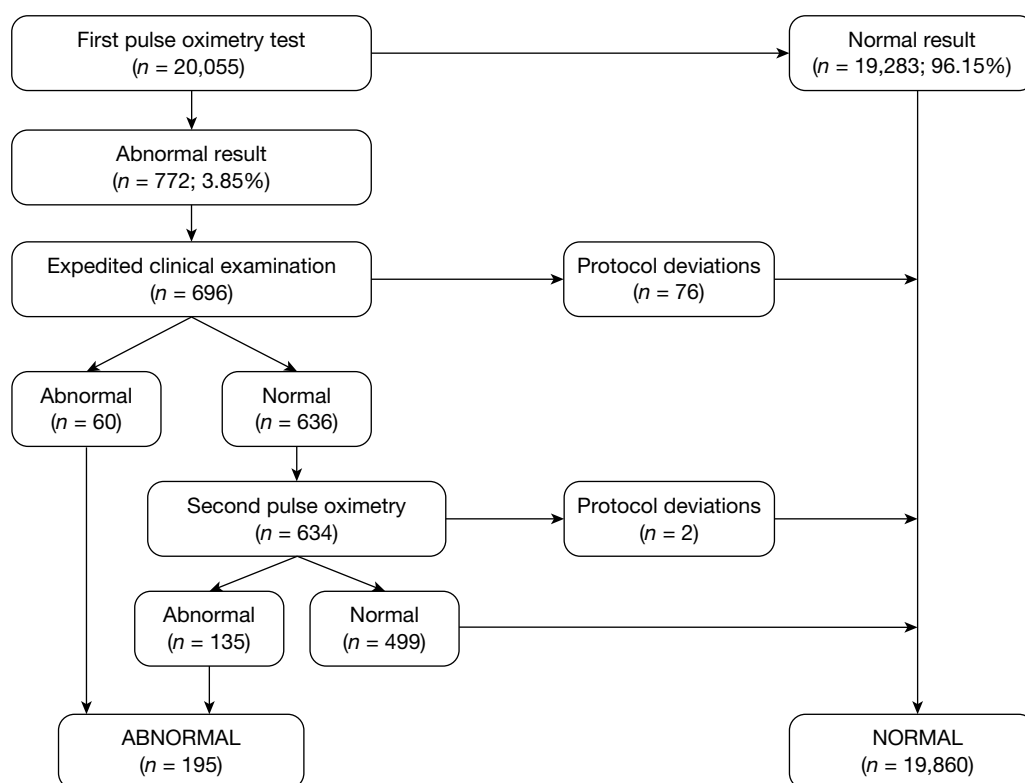
FIGURE 2 Simplified flow of patients through the PulseOx study.

was found to have any evidence of a CHD and they were classified as TN on an intention-to-test basis given that they did not complete testing.

The ECE raised a suspicion of cardiac abnormality in 60 out of 696 (8.62%) babies. Of the remaining 636 infants who were apparently normal on examination, 634 went on to have a second pulse oximetry test, and an abnormal result occurred on this occasion in 135 (21.23%). Two babies from this group did not have a second test because of parental refusal.

Of the resulting 195 babies (0.97% of the total cohort) who were test positive, 193 proceeded to reference standard of echocardiography (*Figure 3*). Two babies missed any echocardiography testing, despite repeated attempts to arrange this, and in one case the tape of the echocardiography result was lost and so the provisional normal echocardiographic diagnosis could not be ratified. These three babies were followed up via clinical databases and were not found to have evidence of CHDs.

In the total cohort, 24 babies were ultimately diagnosed via echocardiography or clinical database follow-up as having a critical CHD (0.12%), and 29 babies were diagnosed as having a serious CHD, making 53 in total (2.6 per 1000 live births) (see *Figure 3* and *Appendix 2*). In addition, 20 babies were found to have a significant CHD, although for the purposes of our main analysis these were classified as six FPs (see *Figure 3* and *Appendix 2*) and 14 TNs (see *Figure 3*).
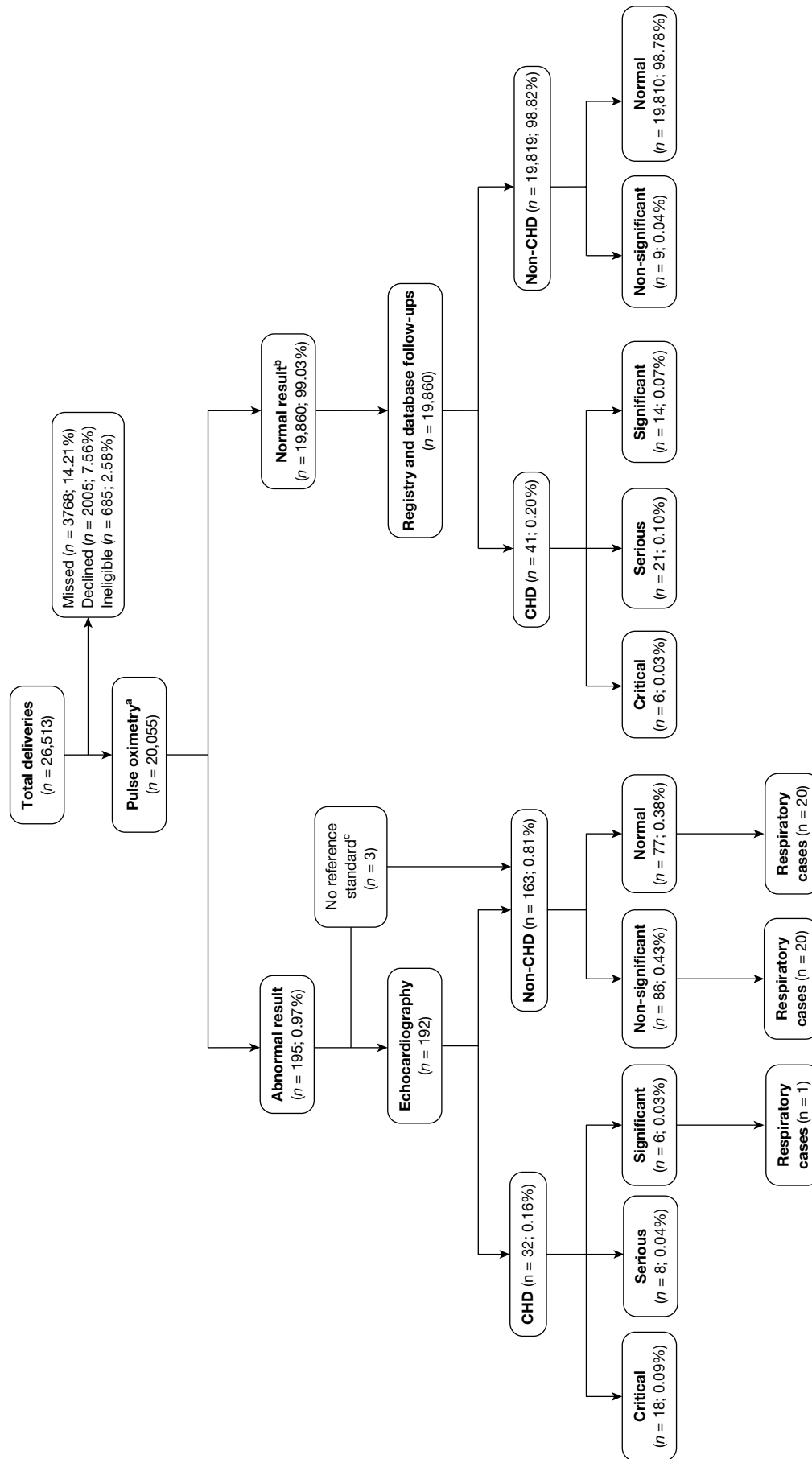
FIGURE 3 Flow of participants through the study. Percentages quoted are as of a proportion of the total recruited (*n* = 20,555). a, See test methods for details of index text; b, this includes 78 babies who missed some or all of the stages of the index test after the first pulse oximetry stage; these were followed up as per reference standard for normal result and have been confirmed as non-CHD; c, these were followed up as per reference standard for normal result and have been confirmed as non-CHD.

Assessment of test accuracy

Accuracy of pulse oximetry

Of the 20,055 babies screened, 53 had major CHDs (24 critical and 29 serious) – a prevalence of 2.6 per 1000 live births (see *Figure 3*). As stated previously, we examined the test accuracy both for the full cohort and then for the cohort excluding those suspected of having a CHD after routine anomaly scan and fetal echocardiography in order to assess the added value of pulse oximetry. Details of testing outcomes for these two cohorts are shown in *Figure 4* (cohorts F and G).

Full cohort

Of the 195 babies who had an abnormal result following pulse oximetry testing, 26 had major CHDs, with 18 critical and 8 serious cases. For the full cohort, pulse oximetry had a sensitivity of 75.00% (95% CI 53.29% to 90.23%) for critical cases and 49.06% (95% CI 35.06% to 63.16%) for critical plus serious cases combined (*Table 4*). Six critical and 21 serious cases were not identified by pulse oximetry (FNs).

Cohort in which congenital heart disease was not suspected antenatally

Of the 20,055 babies who took part in the study, following retrospective review of all cases of CHDs identified, we found that 12 out of 24 with critical CHDs (50.00%) had already been suspected by prenatal screening as having a high likelihood of anomaly (sensitivity 50.00%, 95% CI 29.12% to 70.88%) (*Table 5*), whereas 19 out of 53 critical and serious cases combined were detected (sensitivity 35.85%, 95% CI 23.14% to 50.20%). Sixteen out of 20,002 babies without critical or serious CHDs were incorrectly identified at this stage (specificity 99.92%, 95% CI 99.87% to 99.95%), although this was reduced to five cases after fetal echocardiography examination of these babies was undertaken (cohort B; see *Figures 4* and *5*). It was noted that one serious case was incorrectly diagnosed as a non-CHD after fetal echocardiography (cohort D; see *Figure 4*).

Therefore, for the cohort in whom pulse oximetry results could affect postnatal management as they had not been suspected prenatally, pulse oximetry had a sensitivity of 58.33% (95% CI 27.67% to 84.83%) for critical cases (12 babies) and 28.57% (95% CI 14.64% to 46.30%) for critical plus serious cases combined (35 babies) (*Table 6*). One in 119 babies (0.84%) without a serious or critical CHD had a FP pulse oximetry result (specificity 99.16%, 95% CI 99.02% to 99.28%); this was similar for both the full cohort and those not diagnosed prenatally.

TABLE 4 Accuracy of pulse oximetry in full cohort (G in *Figure 4*, *n* = 20,055)

	Critical cases alone	Critical plus serious cases
No. of TPs	18	26
No. of FNs	6	27
No. of FPs	177	169
No. of TNs	19,854	19,833
% Sensitivity (95% CI)	75.00 (53.29 to 90.23)	49.06 (35.06 to 63.16)
% Specificity (95% CI)	99.12 (98.98 to 99.24)	99.16 (99.02 to 99.28)
% Positive predictive value (95% CI)	9.23 (5.56 to 14.20)	13.33 (8.90 to 18.92)
% Negative predictive value (95% CI)	99.97 (99.93 to 99.99)	99.86 (99.80 to 99.91)

FN, false-negative; FP, false-positive; TN, true-negative; TP, true-positive.

A

Anomaly scan/DS screen	Reference standard CHD diagnosis			
	Critical	Serious	No CHD	
Suspicion	12	7	16	35
Clear	12	22	19,986	20,020
	24	29	20,002	20,055
	0.12%	0.14%		

B

Fetal echo	Reference standard CHD diagnosis			
	Critical	Serious	No CHD	
Confirmed	12	6[a]	5[a]	23
Clear	0	1	11	12
	12	7	16	35
	34.29%	20.00%		

C

Pulse oximetry	Reference standard CHD diagnosis			
	Critical	Serious	No CHD	
Fail	11	5	2	18
Pass	1	1	3	5
	12	6	5	23
	52.17%	26.09%		

D

Pulse oximetry	Reference standard CHD diagnosis			
	Critical	Serious	No CHD	
Fail	0	0	0	0
Pass	0	1	11	12
	0	1	11	12
	0.00%	8.33%		

E

Pulse oximetry	Reference standard CHD diagnosis			
	Critical	Serious	No CHD	
Fail	7	3	167	177
Pass	5	19	19,819	19,843
	12	22	19,986	20,020
	0.06%	0.11%		

F

Pulse oximetry	Reference standard CHD diagnosis			
	Critical	Serious	No CHD	
Fail	7	3	167	177
Pass	5	20	19,830	19,855
	12	23	19,997	20,032
	0.06%	0.11%		

G

Pulse oximetry	Reference standard CHD diagnosis			
	Critical	Serious	No CHD	
Fail	18	8	169	195
Pass	6	21	19,833	19,860
	24	29	20,002	20,055
	0.12%	0.14%		

FIGURE 4 Results of antenatal and pulse oximetry testing all participants recruited to the study are in cohort A. Thirty-five individuals (cohort B) were identified by prenatal screening for anomalies and Down syndrome to have suspicion or high likelihood of a CHD (all 35 were in fact based on anomaly scans with zero prenatally diagnosed Down syndrome cases). CHDs were confirmed in 23 of these 35 by fetal echocardiography (cohort C) – echocardiography would be undertaken after birth regardless of pulse oximetry results. Combining the 12 who were clear on fetal echocardiography (cohort D) with the 20,020 who had no suspicion on prenatal screening (cohort E) defines the group (cohort F) in whom a positive pulse oximetry test result could make a difference to subsequent testing (most likely using echocardiography) and contingent health care. In settings without antenatal testing the whole cohort (cohort G) would benefit from pulse oximetry testing. In settings with antenatal testing, but without fetal echocardiography those in cohort E are the appropriate group. a, two babies (one a serious case, the other without a CHD) had missing fetal echocardiography results; we have made the assumption that these babies would undergo echocardiography after birth.

TABLE 5 Accuracy of antenatal ultrasound anomaly scan (cohort A in *Figure 4*, *n* = 20,055)

	Critical cases alone	Critical plus serious cases
No. of TPs	12	19
No. of FNs	12	34
No. of FPs	23	16
No. of TNs	20,008	19,986
% Sensitivity (95% CI)	50.00 (29.12 to 70.88)	35.85 (23.14 to 50.20)
% Specificity (95% CI)	99.89 (99.83 to 99.93)	99.92 (99.87 to 99.95)
% Positive predictive value (95% CI)	34.29 (19.13 to 52.21)	54.29 (36.65 to 71.17)
% Negative predictive value (95% CI)	99.94 (99.90 to 99.97)	99.83 (99.76 to 99.88)

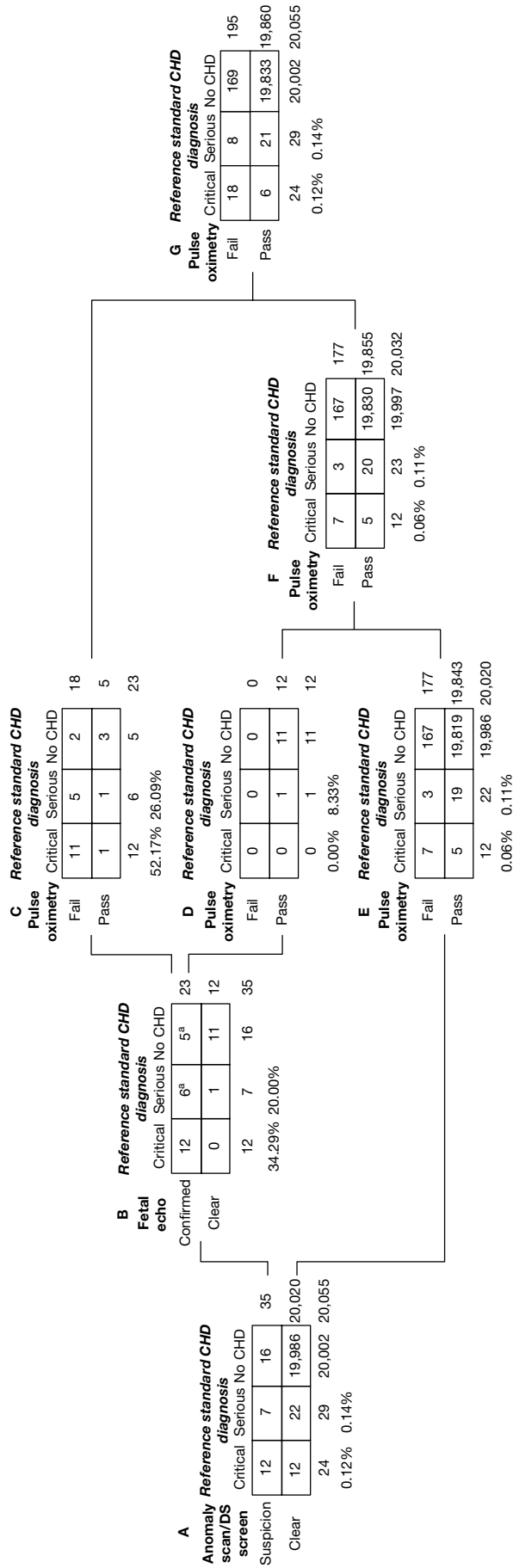FN, false-negative; FP, false-positive; TN, true-negative; TP, true-positive.

TABLE 6 Accuracy of pulse oximetry in cohort where a positive test would influence subsequent management (cohort F in *Figure 4*, *n* = 20,032)

	Critical cases alone	Critical plus serious cases
No. of TPs	7	10
No. of FNs	5	25
No. of FPs	170	167
No. of TNs	19,850	19,830
% Sensitivity (95% CI)	58.33 (27.67 to 84.83)	28.57 (14.64 to 46.30)
% Specificity (95% CI)	99.15 (99.01 to 99.27)	99.16 (99.03 to 99.29)
% Positive predictive value (95% CI)	3.95 (1.60 to 7.98)	5.65 (2.74 to 10.14)
% Negative predictive value (95% CI)	99.97 (99.94 to 99.99)	99.87 (99.81 to 99.92)

FN, false-negative; FP, false-positive; TN, true-negative; TP, true-positive.
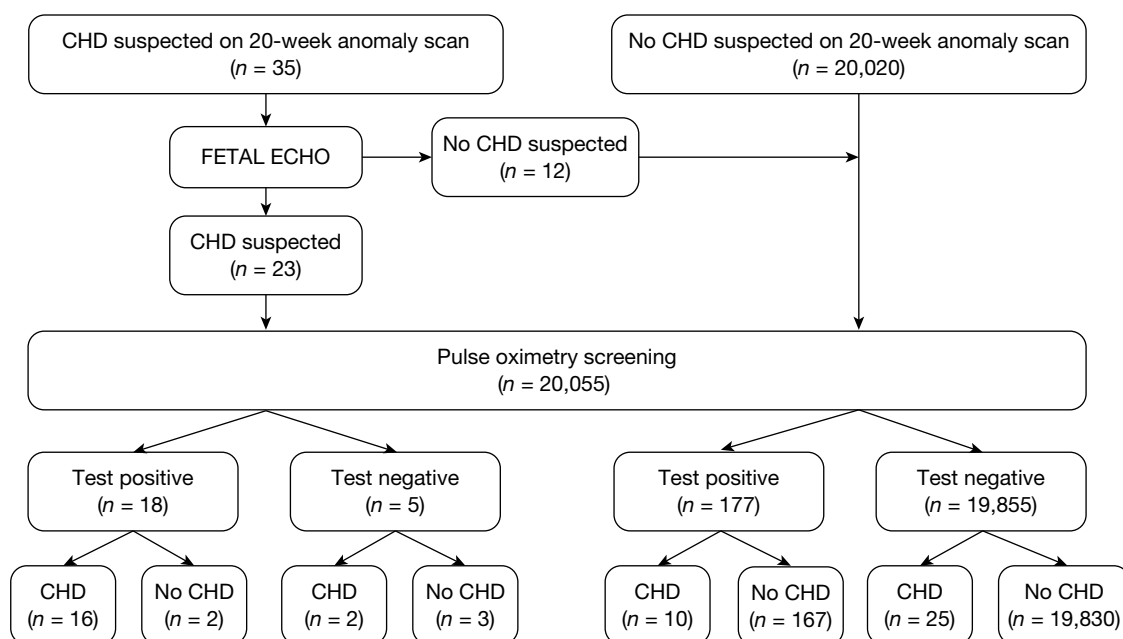
FIGURE 5 Simplified results of antenatal and pulse oximetry testing.

Timing of pulse oximetry

For the full cohort (cohort G; see *Figure 4*), earlier testing showed a strong association with increased sensitivity [odds ratio of true-positive (TP) to FN with hours to testing as the explanatory variable 0.93 (95% CI 0.88 to 0.98; $p = 0.008$)], although when those babies suspected antenatally of having a CHD were not included (cohort F; see *Figure 4*) this association was weaker and non-significant (odds ratio 0.97, 95% CI 0.93 to 1.02; $p = 0.2$) (*Table 7*). There was evidence of an increased rate of FP outcomes from earlier timings [odds ratio FP to TN 0.988 (95% CI 0.977 to 0.998; $p = 0.02$)]. This result did not change when babies suspected of having a CHD antenatally were excluded.

False-positives

There were 169 babies in the full cohort (see *Table 4*) who tested positive, but did not have a critical or serious CHD. Of these, six babies had a significant CHD and a further 40 cases had respiratory or infective conditions that required medical intervention (antibiotic treatment, oxygen therapy or respiratory support). Thus, the total number of test-positive infants in whom there was neither a significant CHD nor intercurrent illness requiring treatment was 123 [63.08% of those who were test positive (195) or 0.61% of the total cohort].

The study by de Wahl Granelli *et al.*[70] used both pre- and postductal saturations < 95% and a difference of > 3% as the test-positive threshold rather than the single measurement < 95% and the > 2% difference used in this study. If de Wahl Granelli *et al.*'s threshold[70] had been used as an alternative in our study, this would have reduced the number of FPs by 61; however, one critical case [hypoplastic left heart syndrome (HLHS), which had been suspected antenatally], one serious case (truncus arteriosus) and one significant case (Ebstein's anomaly) – neither suspected antenatally – and 13 respiratory conditions would have been missed.

Similarly, using a threshold of < 95% in the foot only – the most frequently used test in previous studies[13,64,67–69,71] – would have reduced the number of FPs by 84, but three critical cases (two

TABLE 7 Performance of pulse oximetry in the detection of serious and critical CHD cases by time of testing

Cohort where positive test would influence subsequent testing (n = 20,032, F in *Figure 4*)

	0–6 hours (n = 4937, 25%)	6–12 hours (n = 4822, 24%)	12–24 hours (n = 5320, 27%)	> 24 hours (n = 4953, 25%)
No. of TPs	3	3	3	1
No. of FNs	6	1	7	11
No. of FPs	59	40	37	31
No. of TNs	4869	4778	5273	4910
% Sensitivity (95% CI)	33.33 (7.49 to 70.07)	75.00 (19.41 to 99.37)	30.00 (6.67 to 65.25)	8.33 (0.21 to 38.48)
% Specificity (95% CI)	98.80 (98.46 to 99.09)	99.17 (98.87 to 99.41)	99.30 (99.04 to 99.51)	99.37 (99.11 to 99.57)

Full cohort (n = 20,055, G in *Figure 4*)

	0–6 hours (n = 4956, 25%)	6–12 hours (n = 4823, 24%)	12–24 hours (n = 5323, 27%)	> 24 hours (n = 4953, 25%)
No. of TPs	18	3	4	1
No. of FNs	7	1	8	11
No. of FPs	60	40	38	31
No. of TNs	4871	4779	5273	4910
% Sensitivity (95% CI)	72.00 (50.61 to 87.93)	75.00 (19.41 to 99.37)	33.33 (9.92 to 65.11)	8.33 (0.21 to 38.48)
% Specificity (95% CI)	98.78 (98.44 to 99.07)	99.17 (98.87 to 99.41)	99.28 (99.02 to 99.49)	99.37 (99.11 to 99.57)

with HLH identified antenatally and one with CoA not diagnosed antenatally) and one serious case (truncus arteriosus not diagnosed antenatally) would have been missed. Two babies with significant CHDs and nine with respiratory conditions would also have been missed.

False-negatives

Six babies from the full cohort with critical CHDs were not detected by pulse oximetry testing (see *Figure 3* and *Appendix 2*). One of these (a baby with congenitally corrected TGA and PS) had a suspected diagnosis of CHD following antenatal anomaly screening, and in a further three babies a CHD was identified prior to discharge from hospital because of an abnormal routine clinical examination. Two babies (one with TGA, VSD and a coarctation and one with a hypoplastic aortic arch and coarctation) were discharged home after both the pulse oximetry screen and postnatal examination were normal. Both presented with clinical symptoms relating to the CHD and one of the babies (hypoplastic aortic arch and coarctation) presented in a collapsed state. Both went on to have a surgical correction of the lesion. No baby died with an undiagnosed CHD in our study cohort.

A further 21 babies with normal pulse oximetry screening were diagnosed with serious CHDs (see *Figure 3* and *Appendix 2*). In this cohort, two babies had CoA, one had a hypoplastic aortic arch with a PDA, two had tetralogy of Fallot (TOF), five had PS (one with a VSD) and one had an atrioventricular septal defect (AVSD). In addition, seven babies had an isolated VSD; one had persistent PDA, one had anomalous coronary artery and one had aortopulmonary window, all of whom received invasive intervention within 12 months. One baby from this group was identified by abnormal clinical examination prior to discharge (one PS); the remaining 20 were identified by interrogation of the relevant databases.

Discussion

Summary of main findings

This is the largest UK test accuracy study using pulse oximetry in the detection of major CHDs. The study has demonstrated the feasibility and accuracy of pulse oximetry testing in a cross-section of UK maternity units. In asymptomatic infants, pulse oximetry screening had a sensitivity of 75.00% (95% CI 53.29% to 90.23%) for critical lesions and a sensitivity of 49.06% (95% CI 35.06% to 63.16%) for critical plus serious lesions.

For the cohort in which the test results could affect postnatal management as a CHD had not been suspected prenatally, pulse oximetry had a sensitivity of 58.33% (95% CI 27.67% to 84.83%) for critical cases and 28.57% (95% CI 14.64% to 46.30%) for critical plus serious cases combined.

When combined with the standard screening of antenatal ultrasound and physical examination 22/24 cases of critical CHDs were detected prior to discharge from hospital, with pulse oximetry alone identifying seven cases. FP results occurred in 1 in 119 babies (0.84%) without serious or critical CHDs (specificity 99.16%, 95% CI 99.02% to 99.28%); however, out of the FP cohort of 169 babies, 46 (27%) had additional medical problems requiring intervention (specifically significant CHDs, respiratory disorders and infections).

The timing of pulse oximetry testing varied, with the majority of testing occurring in the first 24 hours. There was a significant association of increased sensitivity with earlier timing, but this is likely to be owing to the fact that those with an antenatal suspicion of a CHD were tested much earlier. Earlier testing was associated with a higher FP rate.

Strengths and limitations of methods

The validity of our findings relies on the quality of the study. We complied with and reported all criteria for a high-quality test accuracy evaluation[77] (see *Appendix 1)*. The study population and recruitment criteria were well defined. Recruitment of consecutive eligible subjects was representative of a spectrum of maternity activity. A sample size calculation was performed to ensure that the study had sufficient power to exclude clinically unacceptable accuracy and the study recruited to that target. The rationale for the index test was clarified and it was performed by trained staff. The initial reference standard of echocardiography was performed by independent, trained individuals. The robustness of the data has been facilitated by the rigorous follow-up to 1 year of age of all recruited babies.

This study used a composite reference standard of echocardiography and database interrogation, therefore, not all babies entered into the study underwent the gold standard reference of echocardiography. This would have been impracticable and we are confident that the database follow-up was sufficiently rigorous to account for all potential FNs. Because only those testing positive underwent echocardiography, it was not possible to blind the echocardiographers to the index test result; however, all echocardiograms were ratified by an independent cardiologist. Echocardiography is an objective test and unlikely to be influenced by bias; in addition, the independent cardiologists allocated the final echocardiography category after clinical data on outcomes (such as surgery) were available.

In general, pulse oximetry screening took place before routine clinical examination and we felt that it was unethical for the results of pulse oximetry readings to be withheld from clinical staff. This lack of blinding meant that we were unable to compare the accuracy of the physical examination with that of pulse oximetry. As we described in *Chapter 1*, the published evidence relating to the accuracy of the physical examination shows that it has a relatively low sensitivity.[2] A large Swedish study compared CHD detection in different regions and showed that the rate of critical CHDs diagnosed after discharge in those regions which did not use pulse oximetry was significantly higher than in those that used examination alone.[70]

We included those infants who had been suspected of having a CHD following antenatal ultrasound; this increased the prevalence of CHDs in our study cohort and simplified study recruitment. We tested those infants who fulfilled the eligibility criteria in order to ascertain if pulse oximetry would have identified them had the condition not been suspected and we undertook further sensitivity analysis after excluding these patients. The antenatal detection rate in our cohort was relatively high compared with other published studies, for example in de Wahl Granelli *et al.*'s study[70] of almost 40,000 babies undergoing pulse oximetry testing, only 3.3% of critical lesions were detected antenatally compared with 50% in our study. We feel that our findings represent the value of pulse oximetry when antenatal detection rates are high and this implies that in regions where antenatal detection rates are lower, the added value of pulse oximetry is likely to be even greater.

We were unable to access directly the anomaly scan data for all study participants and we made two assumptions – that all women recruited into the study had undergone an anomaly scan and all anomaly scans in which there was a suspicion of a CHD were referred for fetal echocardiography. This is standard practice in the region where the study was performed and we believe that these assumptions are reasonable. With this additional information we were able to calculate the added value of pulse oximetry over anomaly scan screening.

There were a number of protocol violations in our study which are described in detail. The likelihood that all protocol violations would have resulted in a negative test and the robustness

of the subsequent clinical follow-up makes it extremely unlikely that these violations would have had any impact on the outcome of the test accuracy study.

The FP rate for the test accuracy study was relatively high compared with some previous studies, and there are a number of explanations for this. The initial test was performed in the first 24 hours in a significant proportion of babies, and this has been shown in previous studies to be associated with a higher FP rate. However, as an increasing number of newborn babies in the UK and elsewhere are discharged from hospital in the first 24 hours, we felt that it was important to collect data in this crucial time window in order to inform test accuracy within current clinical practice. In addition, the threshold for a test-positive case in our study using pre- and postductal saturations and a > 2% difference between the two was more conservative than in any previous published work. This was an attempt to increase the sensitivity of the test with particular reference to those lesions that are commonly missed, such as obstruction of the aortic arch. In previously published pulse oximetry studies these lesions have been the most commonly reported lesions undetected by saturation testing. In our study, 3/7 (43%) babies with critical CoA or IAA were detected by pulse oximetry, which compares with 3/14 (21%) in de Wahl Granelli et al.'s study,[70] which used a difference of > 3%. Analysis showed that a threshold of > 2% resulted in the detection of one additional critical case, but at the expense of a higher FP rate.

The failure to detect a significant number of serious CHD lesions can partly be explained by the fact that the majority of babies with these conditions were likely to be acyanotic at the time of testing and, therefore, would not be expected to test positive. Increasingly early intervention for VSD – which accounts for 14–16% of CHDs requiring intervention within 12 months[1] – also increased the number categorised as serious.

Some technical problems were identified with the pulse oximeters (see *Chapter 3*), particularly relating to battery life. The same oximeters were used in all centres; however, it became apparent that some experienced problems, whereas others did not. On investigation, it became clear that in some units testing was undertaken with the oximeter as a full unit connected to the mains and in others the detachable front of the oximeter was used as a portable device taken to the baby. This had not been specified in the training and either was felt to be acceptable at the start of the study. The detachable front relies on a rechargeable internal battery and the centres using this method encountered problems with battery life, which occasionally resulted in inadequate saturation readings and a delay in testing. This situation was resolved at the time by using another oximeter within the hospital. The likely cause of this problem was felt to be that the machines were used so frequently, in a way for which they were not designed, that they were not able to fully charge and over time the battery gradually became discharged. Once this problem was identified further recommendations and training significantly reduced the frequency of these events. Occasionally, saturation readings were unobtainable because of 'wear and tear' to the reusable oximeter probe; this was easily identifiable and the probes were replaced. In general, the reusable probes lasted approximately 6 months. The problem with both the batteries and probes resulted in an unobtainable saturation result and so no (potentially unreliable) data were recorded in these situations; a functioning, alternative oximeter was used in each case.

Interpretation of findings

When combined with the routine anomaly scan and newborn physical examination screening, 92% of critical CHDs were detected in our study cohort and no baby died with an unidentified CHD. The detection rate of critical CHDs by pulse oximetry was 75% in the full cohort, which is comparable with other large studies.[69,70] The detection rate for serious lesions is lower; however, it has to be remembered that a significant proportion of the serious lesions not identified by screening were non life-threatening acyanotic conditions (VSD, PDA, etc.), which would not

usually be associated with hypoxaemia and therefore would not be expected to be detected by low oxygen saturations. The consequences of missing such lesions are important, but not as potentially devastating as missing the life-threatening critical lesions. The critical lesions most likely to be missed by pulse oximetry screening were those causing obstruction to the aortic arch (such as CoA and IAA), which is consistent with findings in other studies.[69,70]

If the results from this study were applied to a population of 100,000 babies, approximately 264 babies would have critical or serious CHDs. Of these, pulse oximetry testing would identify 130. Approximately 120 babies would have life-threatening critical lesions and pulse oximetry testing would detect 90 of these.

In our cohort, 50% of critical lesions were detected by antenatal anomaly scan. In a cohort of 100,000 babies with a similar antenatal detection rate pulse oximetry testing could, on average, detect an additional 35 cases of critical CHD. This figure is likely to be higher in areas where the rates of detection by antenatal ultrasound are lower.

Furthermore, pulse oximetry is also likely to detect 30 cases of significant CHD and 199 cases of respiratory or infective illness requiring medical intervention. These cases can be considered as 'secondary targets' for the screening procedure. The pulse oximetry screen detects hypoxaemia and, therefore, any condition other than a major CHD which presents with hypoxaemia alone is likely to be identified. This includes potentially lethal conditions such as group B streptococcal infection and pulmonary hypertension. The management of such illnesses is routine in maternity units and, therefore, identification of such babies can be considered an additional benefit of the screening procedure.

In the population of 100,000, 843 babies with an abnormal test would not have critical or serious CHDs (FPs); however, only 614 would be completely normal (i.e. no significant CHDs or other illness).

Implications for the economic model

The decision-analytic model needs to consider the sensitivities obtained in this study. One of the key issues concerning the screening test in this project is that the intervention is relatively simple, inexpensive and without a high risk of harm to the babies. Pulse oximetry is highly specific and reasonably sensitive, particularly for detecting critical lesions, but the acceptability needs to be evaluated (see *Chapter 3*). A decision-tree analysis model will be required to investigate the cost-effectiveness of pulse oximetry as an adjunct to routine care. Because the accuracy of routine care (i.e. clinical examination) was not evaluated in the test accuracy study, data for clinical examination will have to be obtained from an alternative source.

Implications for practice

Pulse oximetry is a safe, non-invasive, feasible, reasonably accurate test that has a sensitivity that is superior to that of antenatal screening and clinical examination. The use of both pre- and postductal saturations compared with postductal alone appears to be advantageous and in practice does not take significantly longer to perform.

Pulse oximetry adds value to existing screening procedures and is likely to identify cases of critical CHDs that would otherwise go undetected. The detection of other pathologies such as significant CHDs and respiratory and infective illnesses is an additional advantage. Pulse oximetry does not detect all major CHDs. Those lesions missed include the critical lesions associated with obstruction of the aortic arch and the serious lesions, such as VSDs, which are acyanotic.

Earlier testing and the use of more conservative cut-off thresholds increase the identification of babies with pathology, but at the expense of slightly higher FP rates.

Equipment malfunctions were uncommon and mainly related to inadequate charging of the oximeters' rechargeable battery or damage to the reusable probe. These problems were eliminated with further training. Handheld oximeters, which use disposable batteries, are also available and offer an alternative solution.

Recommendations for research

Pulse oximetry improves detection rates of critical CHD, but does not detect all cases. The majority of critical cases missed by pulse oximetry (and by other screening methods) are associated with obstruction of the aortic arch, as these conditions are often not associated with hypoxaemia. Because these lesions are associated with reduced systemic blood flow, further investigation of other oximetry techniques, such as perfusion index, may enhance the detection rates for these lesions.

Chapter 3

Evaluating the acceptability of pulse oximetry screening to parents and health-care staff

Introduction

Assessment of acceptability to parents

As indicated in *Chapter 1*, the acceptability of perinatal screening to parents is likely to have an impact on uptake, and screening may have a negative impact on the psychological health of parents, even those for whom screening suggests that there is no problem. It is therefore important to assess the mother's experience of testing and psychological sequelae. In particular, we should seek to identify whether or not anxiety is increased by a FP result and whether or not any heightened anxiety persists over time. Raised anxiety may be due, to some extent, to the quality of information provided to parents.[82] Some research suggests that diagnostic tests showing the baby to be healthy dispel high levels of anxiety created by an initial positive result.[74] In this chapter we examine the extent to which mothers are satisfied with the testing process across three groups (TP, FP and TN results). We also examine the emotional state (anxiety and depression) of mothers and, particularly, compare those given FP results with those given TN results.

Across all groups, we can also identify any demographic or psychological factors that predict satisfaction with testing or negative psychological outcomes. Evidence of the relationship between demographic factors and the acceptability of antenatal and perinatal screening is limited.[83] Willingness to participate and acceptability of perinatal screening for group B *Streptococcus* during labour was found to vary with age and ethnicity.[80] However, this procedure involved taking samples from the mother rather than non-invasive measurements from the baby. The same study found that satisfaction with aspects of the test process varies with maternal anxiety and perceptions of the screened illness. Pulse oximetry detects not only congenital heart disease, but also other health problems, such as respiratory disorders and infections. Satisfaction will, therefore, be examined by parental perceptions of test findings, as well as the strict study criteria for a 'TP' result. Understanding factors underlying satisfaction and negative emotions allows vulnerable groups to receive any extra support needed, as appropriate, in making decisions about taking part in screening.

Assessment of acceptability to health-care staff

If screening is to be implemented as a standard procedure, it must be acceptable to health-care staff. The response of staff is likely to be affected by the impact of screening in terms of time and effort required, the perceived benefits of testing, and professional views on the impact of screening on parents.

Aims

- To identify factors predictive of participation or non-participation in pulse oximetry screening.

- To assess satisfaction with pulse oximetry as a screening method and levels of distress across result groups, in particular whether or not FP and TN participants have different levels of distress (anxiety and depression) and satisfaction.
- To identify factors predicting negative outcomes (distress, low satisfaction with testing).
- To assess acceptability of pulse oximetry screening to health-care staff and impact of screening on staff roles.
- To explore professional views on perceived benefits of testing and the impact of screening on parents.

ACCEPTABILITY TO PARENTS

Methods

Participation in the study

Data were collected from the records of the mothers who declined to take part in the study. This consisted of basic demographic details. The likelihood of declining entry into the study with respect to the information collected on all participants and decliners (age, parity, ethnicity) was examined using a multivariate logistic regression model. Covariates were first considered individually and then in combination if statistically important ($p < 0.1$). Adjusted estimates of effect size from the model including any important variables identified are presented as odds ratios with 95% CIs. Effects within groupings were also examined to see if any interaction effects of identified variables were statistically important. Analysis was performed in SAS version 9.2 (SAS Institute Inc., Cary, NC, USA) using PROC LOGISTIC.

Acceptability of pulse oximetry for routine screening

For mothers who participated in the study, a cross-sectional questionnaire survey was administered, with a 1-year follow-up postal questionnaire for mothers of babies with FP test results.

Participants and procedures

For FP cases, questionnaires were either given to mothers by a midwife before discharge from hospital or posted to mothers at home. The median time to questionnaire completion after the baby's birth was 30 days [interquartile range (IQR) 12 to 58 days]. Of 169 mothers of babies with FP results in the whole study, 148 were approached to complete an acceptability questionnaire; of these, 119 (80.4%) responded.

A follow-up questionnaire was sent to the mothers of babies who screened FP and who returned a baseline questionnaire. Follow-up data were obtained from 51 participants (42.9%); questionnaires were completed at a median of 385 days after birth (IQR 355 to 385 days).

Mothers of babies who received TP results were approached at a time when health-care staff perceived them to be ready to respond to the questions. These questionnaires were administered face to face by a member of the research team or given to mothers to complete on their own before discharge if the baby was relatively well. The median time from birth to questionnaire completion was 20 days (IQR 3 to 29 days). Of 26 mothers of babies with TP results in the study, 21 were approached to complete a questionnaire; of these, 15 (71.4%) participated.

Mothers of babies who received FN results were not routinely approached to take part in this aspect of the study. Three mothers whose babies' problems were identified very quickly did

complete the questionnaire in face-to-face interviews, but this sample was not sufficient for even descriptive analysis.

A sample of mothers whose babies received TN results was approached to complete the acceptability questionnaire at the time of discharge from hospital. Midwifery staff were asked to hand questionnaires to mothers who had taken part in the study from October 2008 to the end of the study in mid-January 2009. The median number of days to questionnaire completion was 1 day after the baby's birth (IQR 0 to 2 days). Two thousand questionnaires were provided to participating hospitals, in proportional with the number of births, but no record was kept of how many women were approached to complete them. Of the 6297 mothers of babies with TN results recruited into the study during this period, 679 completed questionnaires. Given that not all questionnaires supplied may have been distributed the minimum response rate was 33.95% (679/2000).

Measures

We define acceptability broadly in this study to incorporate the psychological impact of screening on parents. The acceptability of pulse oximetry was assessed using a questionnaire designed for the study, but incorporating standardised measures. The questionnaire was designed on the basis of the literature on antenatal and neonatal screening and diagnosis. In order to maximise face and content validity of the questionnaire it was developed by one of the investigators (HMP), a health psychologist, in consultation with the research team, which included experienced midwives, paediatricians and representatives of paediatric cardiology support groups.

Acceptability questionnaire: time 1
Satisfaction

Measures of acceptability and patient satisfaction are often problematic because of a lack of variability in what are largely very positive evaluations.[84,85] Measurement of separate components relevant to participants enables investigation of variability between components and by participant characteristics. Questions were included on information provision and communication, and on the testing procedure itself, with responses indicated on five-point scales.

Eight items asked participants to indicate satisfaction with information and communication:

- the information provided about pulse oximetry before testing (items 1–3)
- satisfaction with opportunities to discuss the test (item 4)
- opportunities to change their mind about having the test (item 5)
- satisfaction with presentation of the results, information given on what happened next and opportunities to discuss test results (items 11–13).

Five items asked about the conduct of the test:

- how happy the mother was with the way the test was done (item 6)
- how comfortable the baby was (item 7)
- how stressed the mother felt while the test was carried out (item 8)
- confidence with how the test was done and confidence with test results (items 9 and 10).

Subscale scores and an 'overall satisfaction' score were devised. The internal consistency of each subscale was assessed using Cronbach's α. An α value of ≥ 0.8 was taken to indicate that items were measuring aspects of the same variable and therefore could be combined. The single item 'Would you recommend the test to someone else?' (item 17) was also included to assess the construct validity of the overall satisfaction measure. This item has been shown to demonstrate satisfaction reliably.[84]

- Items 1–3 were combined to form the subscale score for 'satisfaction with information' ($\alpha = 0.87$).
- Items 4 and 5 were analysed separately.
- To assess satisfaction with test procedure, α was calculated for items 6–8 at 0.51. Removing item 8 improved the internal consistency, so items 6 and 7 ($r_s = 0.74$, $p < 0.001$) were combined for the subscale 'happiness with test'. Item 8 ('stress') was analysed separately.
- Items 9 and 10 ($r_s = 0.77$, $p < 0.001$) were combined to form the subscale 'confidence in test'.
- Items 11–13 ($\alpha = 0.88$, time 1; $\alpha = 0.95$, time 2) were combined to form the subscale 'post-test communication satisfaction'.

To create the overall satisfaction scale, Cronbach's α for the scale was calculated for all subscales related to satisfaction with the test procedure (items 4 and 5, mean items 6 and 7, item 8, mean items 9 and 10, mean items 11–13). Including item 8, $\alpha = 0.75$; without item 8 $\alpha = 0.87$. The overall satisfaction scale was therefore created as the sum of all subscales with the exception of item 8, which was analysed separately. The overall satisfaction scale score significantly correlated with item 17, 'Would you recommend the test to someone else?' ($r_s = 0.41$, $p < 0.001$), indicating good validity. Satisfaction, happiness and confidence scales were then scored so that higher scores indicated higher satisfaction, happiness and confidence; item 8 ('stress') was scored so that a higher score indicated higher feelings of stress while the test was being conducted.

General feelings about the pulse oximetry test

Single items asked each mother to indicate whether or not she thought it was important for her baby to have the test, for all babies to have the test and whether or not heart problems would have been found without the test. Higher scores indicate more positive perceptions of the test.

Perception of test results

Participants were given seven options as to what they thought the pulse oximetry test had shown for their baby, i.e. whether or not they thought there was:

1. no problem with the baby's health
2. a minor heart condition
3. a minor health condition, but not a heart condition
4. a serious heart condition
5. a serious health condition, but not a heart condition
6. 'do not know'
7. 'other'.

A binary scale was created with 0 = no problem with the baby's health or 'do not know'; 1 = all other responses.

Psychological variables

The rest of the questionnaire measured psychological variables that may be associated with the impact or reported acceptability of screening:

Anxiety The Spielberger State-Trait Anxiety Inventory (STAI)[86] is a frequently used measure of anxiety. The six-item short-form measure of current anxiety levels has been demonstrated to yield results that are comparable with the full form.[87] Higher scores indicate higher anxiety.

Depression The depression subscale of the Hospital Anxiety and Depression Scale (HADS)[88] has been widely used and validated in clinical studies as a measure of mood, including in previous screening research.[89] Higher scores indicate higher depression.

Illness perceptions Items assessing representations of illness consequences, timeline, treatment control and illness comprehensibility were adapted for the context of heart disease in babies from the Brief Illness Perception Questionnaire (IPQ).[90] Higher scores indicated perceptions of higher severity, longer time line, less treatment control and lower illness comprehensibility.

Optimism A short (two-item) version of the Life Orientation Test[91,92] was used to measure dispositional optimism. With the present data set, the correlation between these item scores was low ($r_s = 0.19$, $p < 0.001$). The items were therefore not combined and were instead analysed separately.

Comments
The questionnaire ended with a blank box for free responses. Participants were invited to add ideas for improving the test procedure or any other comments.

Demographic and clinical information
Data on mothers' age, parity and ethnicity, hospital and location of test, person who tested and pulse oximetry status (TN, FP, TP, FN) were obtained from other study data (see *Table 3*).

Acceptability questionnaire: time 2 (false-positive group only)
Satisfaction
Satisfaction items 1–13 and item 17 asking whether or not participants would recommend the test to others were repeated.

General feelings about the pulse oximetry test
Items were presented as in the Acceptability Questionnaire time 1.

Anxiety
Short form of the STAI (as above).

Comments
This questionnaire also ended with a blank box. Participants were asked to give ideas on how the test could have been improved or to make any other comments.

Missing data
Where a scale or subscale contained at least three items, if a single item was missing then the value was substituted with that individual's mean score for the other items in the scale. Where more than one item was missing, or if a scale contained only two items, data were not substituted.

Analysis

Differences between the two larger groups, TN and FP, were examined using the Mann–Whitney *U*-test. Differences in the three main outcome measures – anxiety, depression and overall satisfaction – were also examined by maternal perceptions of the test results positive/negative using the Mann–Whitney *U*-test. Wilcoxon matched pairs was used to detect any change in anxiety between time 1 and time 2 for participants in the FP group. Variable distributions were inspected and transformed using log transformations where appropriate; where variables were negatively skewed the transformed scores were multiplied by –1 to maintain non-transformed directions of effect. One-way analyses of variance (ANOVAs) were computed to identify differences in the three main outcome variables (anxiety, depression and overall satisfaction) by ethnic group.

Hierarchical multiple regression equations were calculated to identify characteristics that would predict negative outcomes following testing (dependent variables: 'anxiety' and 'overall satisfaction'). The first block of variables contained test result (TN or FP), parent-defined status (no problem or problem), parity, mother's age, optimism, illness perceptions and ethnicity [where 0 = white (British and Irish); 1 = other ethnic group]. The second block contained 'anxiety', 'depression', 'overall satisfaction' and reported 'stress' during testing (excluding the dependent variable for that equation). Mahalanobis distance values were calculated to check for multivariate outliers among the independent variable blocks; 11 participants were identified as outliers [χ^2 (11) > 31.26 for block 1, χ^2 (4) > 18.47 for block 2, $p < 0.001$, following recommendations of Tabachnick and Fidell[93]]. Regression equations were conducted both with and without these participants. As some differences were seen in terms of significance at $p < 0.05$, regressions excluding these 11 participants are presented.

The free-text comments were thematically analysed using a framework approach.[94]

Results

Participation in pulse oximetry testing

The number of mothers who participated in the PulseOx study was 20,055 and 2005 declined. Both parity ($p < 0.0001$) and ethnicity ($p < 0.0001$) had an effect on the likelihood of declining, but age of mother did not ($p < 0.2$) (*Table 8*). Compared with the largest ethnic group (white British and Irish), which had the lowest rate of declining (5%), all of the other major ethnic groups had an increased likelihood of declining (*Table 9*); the largest increase was for black African mothers (21%, odds ratio 4.6, 95% CI 3.8 to 5.5, $p < 0.0001$). Compared with mothers with a first baby, mothers with more than one baby were more likely to decline; this was particularly evident in mothers with four or more babies (14% vs 7%, odds ratio 1.5, 95% CI 1.3 to 1.7, $p < 0.0001$).

Acceptability to parents

Data describing participants by age, parity and ethnicity and details of test results are provided in *Table 10*.

Table 11 shows what participants thought the test showed. It can be seen that the majority of participants with TN results thought that the pulse oximetry test had shown that there was no problem with their babies' health. For those with FP results, just under half also believed the test showed no problem. So there was some mismatch between parental perceptions and research-defined test results. However, most TP participants believed that the test had shown there to be a serious problem.

Emotional state and satisfaction by test result

Table 12 summarises the scores for the three main outcome measures (anxiety, depression and overall satisfaction), as well as the satisfaction subscales and participants' general feelings about the test. The mean anxiety scores for mothers of babies with TN and FP results were 31.37 (95% CI 30.50 to 32.24) and 33.30 (95% CI 31.01 to 35.59), respectively. These scores are slightly lower than those values obtained in previous studies of pregnant women[87] and screening for group B *Streptococcus* perinatally,[80] but still within in the lower part of the normal range as indicated in the normative values for the STAI Manual (means 35.2 for working adult females, 36.17 for 19- to 39-year-old females).

Mothers of babies with TP results had high anxiety and depression scores compared with the other two groups, but similar scores on satisfaction measures. This was not tested statistically, however, because of the small size of this group. The two larger groups (TNs and FPs) were tested

TABLE 8 Proportions of women in the age, ethnicity and parity groups who declined to take part in the study

	Proportion in sample (*n* = 22,060) *n* (%)	No. who declined	No. as proportion (%) of total who declined (*n* = 2050)
Age group (years)			
< 20	1586 (7.2)	112	5.5
20–24	5027 (22.8)	486	23.7
25–29	6448 (29.2)	615	30.0
30–34	5240 (23.8)	456	22.2
35–39	2988 (13.5)	222	10.8
≥ 40	727 (3.3)	70	3.4
Missing	44		
Ethnicity			
White (British and Irish)	11,223 (50.9)	605	29.5
Asian (Indian)	1374 (6.2)	152	7.4
Asian (Pakistani)	3361 (15.2)	553	27.0
Asian (Bangladeshi)	569 (2.6)	108	5.3
Black (Caribbean)	625 (2.8)	76	3.7
Black (African)	854 (3.9)	182	8.9
Other	2960 (13.4)		
Not given	459		
Missing	635		
Parity			
1	9906 (44.9)	709	34.6
2	6387 (29.0)	587	28.6
3	3168 (14.4)	334	16.3
4+	2588 (11.7)	364	17.8
Missing	11		

TABLE 9 Odds ratios of declining by ethnicity

Comparison with white (British and Irish)	Odds ratio (95% CI)	*p*-value
Asian (Indian)	2.2 (1.8 to 2.7)	< 0.0001
Asian (Pakistani)	3.3 (2.9 to 3.7)	< 0.0001
Asian (Bangladeshi)	3.9 (3.1 to 4.9)	< 0.0001
Black (Caribbean)	2.4 (1.9 to 3.1)	< 0.0001
Black (African)	4.6 (3.8 to 5.5)	< 0.0001

for differences on these measures. Significant differences were seen on depression and most satisfaction measures, with FP participants scoring as more depressed and less satisfied. Anxiety was not significantly elevated in the FP group ($z = -1.71$, $p = 0.09$) and, although statistically significant, the difference in median depression between FP and TN participants is only one point, which is unlikely to be clinically significant. However, the median time to completion of the questionnaires by FP mothers was much longer than for TN mothers (30 days and 1 day, respectively) and greater differences may have been detected if the FP mothers had responded immediately. Differences were not seen in whether participants would recommend the test to others or in the perceived importance of the test. Median scores indicate that most people were satisfied with the test procedures: on satisfaction subscale measures, a score of 1 indicates 'very satisfied', 2 'satisfied' and 3 'neither'. No subscale score for any group yielded a median of 4, 'dissatisfied', or 5, 'very dissatisfied'.

TABLE 10 Characteristics of participants by test result group

	TN (*n* = 679)	FP (*n* = 119)	TP (*n* = 15)	Total (*n* = 813)
Age [median, (IQR)]	28 (24, 32)	29 (25, 31)	27 (24, 34)	28 (24,32)
Parity [median, (IQR)]	2 (1,2)	2 (1,2)	2 (1,2)	2 (1,2)
	Frequency (%)	Frequency (%)	Frequency (%)	Frequency (%)
Ethnicity				
White (British and Irish)	386 (56.8)	56 (47.1)	3 (20.0)	445 (54.7)
Asian (Indian)	34 (5.0)	6 (5.04)	1 (6.6)	41 (5.0)
Asian (Pakistani)	83 (12.2)	24 (20.2)	4 (26.7)	111 (13.7)
Asian (Bangladeshi)	13 (1.9)	5 (4.2)	3 (20.0)	21 (2.6)
Black (Caribbean origin)	20 (2.9)	5 (4.2)	0 (0)	25 (3.1)
Black (African origin)	20 (2.9)	7 (5.9)	0 (0)	27 (3.3)
Other	82 (12.1)	11 (9.2)	0 (0)	93 (11.4)
Not given	21 (3.1)	1 (0.8)	1 (6.6)	23 (2.8)
Missing data	20 (2.9)	4 (3.4)	3 (20.0)	27 (3.3)
First test location				
Delivery suite	53 (7.8)	11 (9.2)	3 (20.0)	67 (8.2)
Postnatal unit	624 (91.9)	108 (90.8)	5 (33.3)	737 (90.7)
Neonatal unit	2 (0.3)	0 (0)	7 (46.7)	9 (1.1)
Person who performed first test				
Midwife	256 (37.7)	64 (53.8)	4 (26.7)	324 (39.9)
Midwifery assistant	182 (26.8)	38 (31.9)	3 (20.0)	223 (27.4)
Care assistant	126 (18.6)	7 (5.9)	0 (0)	133 (16.4)
Neonatal nurse	2 (0.3)	1 (0.8)	5 (33.3)	8 (1.0)
Doctor	0 (0)	0 (0)	2 (13.3)	2 (0.2)
Nurse	113 (16.6)	9 (7.6)	1 (6.7)	123 (15.1)

'Not given' = ethnicity was recorded as 'not given'. If no response was given, data were treated as missing.

TABLE 11 Frequencies of parent perceptions of test results for each group

What parents thought pulse oximetry showed	TN, frequency (%)	FP, frequency (%)	TP, frequency (%)
No problem	599 (88.2)	56 (47.1)	1 (6.7)
Minor heart condition	5 (0.7)	29 (24.4)	1 (6.7)
Minor health condition, not heart condition	2 (0.3)	11 (9.2)	0 (0)
Serious heart condition	8 (1.2)	2 (1.7)	11 (73.3)
Serious health condition, not heart condition	0 (0)	2 (1.7)	0 (0)
Do not know	16 (2.4)	7 (5.9)	1 (6.7)
Other	34 (5.0)	11 (9.2)	0 (0)
Missing data	15 (2.2)	1 (0.8)	1 (6.7)

TABLE 12a Median scores (IQR) for each test result group on anxiety, depression and 'general feelings about the test' measures

	TP		TN		FP (time 1)		Difference between TN and FP, *p*-value
	Median (IQR)	*n*	Median (IQR)	*n*	Median (IQR)	*n*	
Anxiety (possible range: 20–80)	48.3 (25 to 67.7)	14	30 (20 to 36.7)	594	32.7 (23.3 to 40)	102	0.09
Depression (possible range: 0–21)	7.5 (5.6 to 12)	14	3 (2 to 6)	644	4 (2 to 7)	115	0.008
Important for your baby to have test? 1 = 'definitely not'; 5 = 'yes, definitely'	5 (5 to 5)	14	5 (5 to 5)	678	5 (5 to 5)	118	0.95
Important for all babies to have test? 1 = 'definitely not'; 5 = 'yes, definitely'	5 (5 to 5)	14	5 (5 to 5)	676	5 (5 to 5)	117	0.63
Would have found heart problem without test? 1 = 'yes, definitely'; 5 = 'definitely not'	3 (2 to 4)	14	3 (2 to 3)	674	3 (3 to 4)	116	0.003

TABLE 12b Median scores (IQR) for each test result group on satisfaction outcome measures

	TP		TN		FP		Difference between TN and FP, *p*-value
	Median (IQR)	*n*	Median (IQR)	*n*	Median (IQR)	*n*	
Overall satisfaction Possible scale range: 1–30	20.5 (18.3 to 22.8)	12	23 (21 to 25)	637	20 (18.5 to 23.1)	109	< 0.001
Stress at time of testing (item 8) 1 = 'strongly disagree'; 5 = 'strongly agree'	2 (2 to 4)	13	1 (1 to 2)	667	3 (2 to 4)	114	< 0.001
Satisfaction with information 1 = 'very dissatisfied'; 5 = 'very satisfied'	4 (4 to 4.8)	12	4.67 (4 to 5)	661	4 (4 to 5)	108	< 0.001
Opportunities to discuss (Item 4) 1 = 'very dissatisfied'; 5 = 'very satisfied'	4 (3.5 to 4)	13	5 (4 to 5)	667	4 (4 to 5)	116	< 0.001
Opportunities to change mind (Item 5) 1 = 'very dissatisfied'; 5 = 'very satisfied'	4 (3.5 to 4)	13	5 (4 to 5)	659	4 (4 to 5)	111	< 0.001
Happiness with test 1 = 'strongly disagree'; 5 = 'strongly agree'	4.5 (4 to 5)	13	5 (4 to 5)	674	4 (4 to 5)	115	< 0.001
Confidence in test 1 = 'strongly disagree'; 5 = 'strongly agree'	4 (4.5)	13	5 (4 to 5)	668	4 (3.5 to 5)	115	< 0.001
Post-test communication satisfaction 1 = 'strongly disagree'; 5 = 'strongly agree'	4 (3.9 to 5)	14	4.67 (4 to 5)	661	4 (3.33 to 5)	117	< 0.001
Would recommend test? 1 = 'definitely not'; 5 = 'yes, definitely'	5 (4.75 to 5)	14	5 (5 to 5)	674	5 (5 to 5)	117	0.52

For the FP group there was no significant change in anxiety ($z = -0.24$, $p = 0.81$) from time 1 (median = 32.7) to time 2 (median = 30). There was also no change in post-test communication satisfaction ($z = -0.93$, $p = 0.36$), whether or not they would recommend the test ($z = -0.03$, $p = 0.98$), the importance of their and all babies being tested ($z = -1.51$, $p = 0.13$; $z = -1.58$, $p = 0.12$, respectively), and whether or not they believed a heart problem would be found without the test ($z = -0.34$, $p = 0.73$). Further results include only TN and FP participants.

Table 13 shows that parents who perceived the test to show a problem with the baby's health were less satisfied with the testing process than those who thought the test showed there to be no problem ($z = -2.23$, $p = 0.03$) and felt more stressed while the test was being done ($z = -4.25$, $p < 0.001$). However, the scores for anxiety and depression did not differ significantly ($z = -0.55$, $p = 0.59$; $z = -1.7$, $p = 0.09$, respectively). Thus, testing positive for problems (whether as defined by the research team or as perceived by parents) would appear to negatively impact on satisfaction with the test process and retrospectively reported stress experienced during the test procedure.

One-way ANOVAs were performed to test differences in anxiety, depression, overall satisfaction and stress experienced by ethnicity. Significant differences by ethnicity were found on all these measurements: $F_{(5,574)} = 8.89$ ($p < 0.001$), $F_{(5,624)} = 5.63$ ($p < 0.001$), $F_{(5,611)} = 5.35$ ($p < 0.001$) and $F_{(5,640)} = 6.88$ ($p < 0.001$), respectively. Post hoc analyses indicated that Asian (Indian), Asian (Pakistani) and Asian (Bangladeshi) participants were more anxious than white (British and Irish) mothers; Asian (Indian) and Asian (Pakistani) mothers were more depressed than white (British and Irish) mothers; Asian (Indian) participants were less satisfied (on overall satisfaction measure) than white (British and Irish) participants, and Asian (Pakistani) participants reported more stress (higher item 8 scores) than white (British and Irish) participants.

Factors associated with emotional state and satisfaction in multivariate analyses

Multiple regression equations to identify predictors of the outcome measures in a multivariate context were conducted. Bivariate measures of parental perceptions of test result (no problem/problem) and ethnicity [white (British and Irish)/all others] were used in the multiple regression equations. *Table 14a* shows that optimism continued to predict anxiety, as did ethnicity, overall satisfaction and depression. More optimistic and less depressed participants were less anxious; white (British and Irish) participants were less anxious than other participants. Being more satisfied overall was associated with being less anxious.

TABLE 13 Median scores (IQRs) for mothers who perceived no problem and those who perceived a problem following pulse oximetry testing

	No problem		Problem		
	Median (IQR)	*n*	Median (IQR)	*n*	*p*-value
Anxiety (time 1)	30 (20 to 36.7)	622	30 (20 to 40)	57	0.59
Depression	3 (2 to 6)	674	4 (2 to 7)	65	0.09
Satisfaction					
Overall satisfaction	23 (20.3 to 25)	664	22.5 (19.3 to 24.5)	63	0.03
Stress (item 8)	2 (1 to 2)	693	2 (1,4)	66	< 0.001

Overall satisfaction was predicted by test result: FP participants reported lower satisfaction than TN participants (*Table 14b*). Perceptions of treatment control and comprehensibility predicted overall satisfaction: perceiving treatment to be more helpful was associated with being more satisfied and perceiving a higher understanding of heart disease in babies was also associated with being more satisfied. White (British and Irish) participants were more satisfied overall than were people of other ethnicities. People who found the testing process more stressful (item 8), were more anxious, more depressed and less satisfied overall.

Parental comments

Participants

Free-response comments were given on 139 questionnaires: 7 in the TP group, 67 in the TN group, 41 from the FP group for the first questionnaire (FP1) and 24 for the second (FP-only) questionnaire (FP2). Fifteen participants wrote comments at both time 1 and time 2 in the FP group, therefore, 124 participants in total gave comments.

TABLE 14a Multiple regression equation: dependent variable – anxiety

Model	Standardised β	p-value	R^2 change	F change	p-value (F change)
Model 1					
Test result	0.00	0.94	0.13	7.90	<0.001
Parent's perception of test result	0.03	0.60			
Parity	0.00	0.97			
Mother's age	−0.05	0.29			
Optimism LOT 1	−0.18	<0.001			
Optimism LOT 2	−0.17	<0.001			
Consequences	−0.02	0.57			
Timeline	0.01	0.81			
Treatment control	0.07	0.11			
Comprehensibility	0.08	0.05			
Ethnicity	0.20	<0.001			
Model 2					
Test result	−0.08	0.11	0.11	26.96	<0.001
Parent's perception of test result	0.05	0.28			
Parity	−0.02	0.69			
Mother's age	−0.04	0.33			
Optimism LOT 1	−0.16	<0.001			
Optimism LOT 2	−0.11	0.01			
Consequences	−0.01	0.73			
Timeline	0.02	0.55			
Treatment control	0.05	0.22			
Comprehensibility	0.05	0.24			
Ethnicity	0.13	0.001			
Overall Satisfaction	−0.14	0.001			
Stress (item 8)	0.07	0.09			
Depression	0.27	<0.001			

For overall model (2), $F = 12.82$, $n = 586$. Adjusted $R^2 = 0.12$ (model 1) and 0.22 (model 2).

TABLE 14b Multiple regression equation: dependent variable – overall satisfaction

Model	Standardised β	p-value	R^2 change	F change	p-value (F change)
Model 1					
Test result	−0.30	<0.001	0.14	8.14	<0.001
Parent's perception of test result	0.09	0.7			
Parity	0.02	0.61			
Mother's age	−0.02	0.69			
Optimism LOT 1	0.01	0.9			
Optimism LOT 2	0.02	0.62			
Consequences	0.04	0.36			
Timeline	0.00	0.94			
Treatment control	−0.12	0.003			
Comprehensibility	−0.11	0.01			
Ethnicity	−.016	<0.001			
Model 2					
Test result	−0.21	<0.01	0.07	17.47	<0.01
Parent's perception of test result	0.08	0.08			
Parity	0.03	0.46			
Mother's age	−0.03	0.50			
Optimism LOT 1	−0.03	0.51			
Optimism LOT 2	−0.04	0.31			
Consequences	0.03	0.48			
Timeline	−0.00	0.93			
Treatment control	−0.11	0.01			
Comprehensibility	−0.10	0.01			
Ethnicity	−0.10	0.02			
Stress (item 8)	−0.17	<0.001			
Anxiety	−0.14	0.001			
Depression	−0.11	0.01			

For overall model (2), $F = 10.69$, $n = 586$. Adjusted $R^2 = 0.12$ (model 1) and 0.19 (model 2).

Perceptions of testing

Across all groups, perceptions of pulse oximetry were predominantly positive. It was perceived to be useful, quick, safe, non-invasive, painless and non-distressing for the baby, and reassuring for parents. Of high importance was its potential to detect problems early, before discharge, allowing treatment to be started and lives to be saved. Many reported feeling pleased or glad to have had the opportunity to take the test, and reported feeling reassured. Some parents reported gratitude – this was particularly the case for the parents of the FP group for which pulse oximetry had identified a health condition. Participants from all groups felt that the test should be a routine part of standard care; many also said he or she would recommend the test to other parents.

Communication and process issues

Positive comments were made about hospital staff by participants in all groups:

> All the staff were friendly and professional.
>
> (FP1, respondent 1)

However, some concerns were raised, particularly on receiving a positive result within the FP group:

> We did get the impression that the midwives and baby doctor did not fully understand the protocol when the test failed … this led to increased anxiety.
>
> (FP1, 20)

Before testing

At the time of testing, participants (mainly those in the TN group) reported that it would be useful to have more information on how the test is conducted, what it does and about what happens after testing, and pictures/diagrams of equipment were suggested. One participant also suggested people should be:

> … informed more verbally that it is not compulsory.
>
> (TN44)

During and after testing

A key concern for participants in the TN and FP groups was communication when their babies 'failed' a test. Some parents felt that staff could have been more reassuring when informing them about the test result or would have found it helpful if they had been given more information about what this meant:

> I was very worried about my baby failing the test but wasn't quickly given an explanation or reasons why this could happen in a healthy baby.
>
> (FP2, 24)

For participants in these groups, babies would eventually be declared healthy or else relatively minor problems would be identified and reassurance might seem to be appropriate. However, the staff member giving the results would be aware that it would be possible that the condition could be serious at the time of testing and would need to avoid giving false reassurance. Positive comments were also received about communication; nevertheless, where communication problems were perceived, worry and anxiety were clearly exacerbated.

Participants also commented on aspects of the test procedure: some experienced repeated testing late at night; one participant was absent when the test was conducted; one participant reported an extended stay in hospital waiting for the test, whereas another felt that the test was given too soon after birth:

> … don't feel the body of the baby has had time to adjust after birth.
>
> (TP3)

Equipment problems

Some participants reported experiencing problems because of faulty equipment:

> My daughter was tested every hour during the night. It was found that some of the equipment used was faulty.
>
> <div align="right">(FP2, 7)</div>

Reported consequences of equipment failure included distress and unnecessary scanning. There were some problems with faulty equipment; however, it is interesting that none of the participants in the TN group reported problems of this type. It may be that, where positive results were found but then negative results were later received, the problem could have been attributed to the equipment rather than changes in the baby's body:

> The nurse could not get a significant reading.
>
> <div align="right">(FP2, 23)</div>

This experience could have been attributed to the baby's state, the nurse's ability to use the equipment or to the equipment itself, but the reliability of the equipment was key:

> The equipment needs to be adequate for this job as this is unfair on parents, babies and the midwives who have to give worrying results.
>
> <div align="right">(TP22)</div>

ACCEPTABILITY TO MIDWIVES, MIDWIFERY ASSISTANTS, NURSERY NURSES AND HEALTH-CARE ASSISTANTS

Methods

Design

Four focus groups were conducted at two hospital sites involved in the study. Interviews were conducted in March and April 2009.

Participants

The focus groups involved the staff groups who had carried out pulse oximetry testing and/or took consent as part of the PulseOx study. Focus groups were set up with the support of ward managers, at times when staff were most likely to be able to take time out of ward duties. All staff in the relevant group who were available at the interview times were invited to take part.

The six hospitals involved in the study chose to involve people in 'assistant' roles to a varying degree in screening, as fitted their current practice. The two hospital sites were selected for their differing mixture of staff who were involved in carrying out testing. At hospital 1, midwives, health-care assistants and nursery nurses carried out testing; at hospital 2, only midwives and midwifery assistants were involved. At each site, two focus groups were conducted: one with midwives ('MW' groups) and one with people in 'assistant' roles ('A' groups – midwifery assistants, health-care assistants and nursery nurses). Focus groups were conducted separately for midwives and 'assistants' because of the different roles and different levels of training and status. It was felt that, with the interviews taking place at participants' place of work, staff might feel inhibited from voicing their thoughts in the presence of other staff groups. In total, 24 staff members participated. At hospital 1 the two focus groups participants were (1) two midwives

(group H1MW) and (2) four nursery nurses and one health-care assistant (group H1A). At hospital 2 the two focus participants were (1) six midwives (group H2MW) and (2) 11 midwifery assistants (group H2A).

Procedures and method

Focus groups were conducted in quiet rooms off the hospital wards and were recorded on two digital recorders and transcribed verbatim. Two researchers (RP and HMP) facilitated each group. A topic guide was used to ensure that the areas covered in the discussion included the experience of using pulse oximetry, the costs and benefits of testing, perceptions of how the mothers felt about taking part and reasons for declining, practical problems, implications for roles in carrying out pulse oximetry testing, changes in perception of testing over the course of the study and perceptions of neonatal monitoring more generally.

Analysis

A thematic analysis was conducted using the framework approach.[94] A list of superordinate and subthemes was devised and used to code the manuscript line by line. The coded manuscripts enabled the creation of charts indexing extracts belonging to each theme for each participant, without losing sight of how extracts were embedded in the data. The detailed records of the analysis made by RP allowed another researcher (HMP) to undertake independent analysis of scripts to ensure agreement on the analysis process, including the themes identified and interpretations of the data.

Results

Perceptions of the test

In all groups, staff were predominantly positive about the pulse oximetry test, viewing it as being highly important. Many knew of babies who had had conditions detected by the test, even if they had not themselves detected a case; it seems that, for many, these examples justified the test:

> If it helps one baby, even if you only have one baby a year, if it's your baby its very very important.
>
> (H2A)

The detection of both heart problems and respiratory problems was seen as beneficial, as was the reassurance for healthy babies.

Pulse oximetry as part of the PulseOx study

Several of the themes which emerged from the focus groups related to the conduct of the PulseOx study rather than pulse oximetry as a method for screening.

Staff involvement

In H2MW, a discussion arose about the process by which staff were involved in the study. The study was an addition to an already 'stressful' workload, and being research rather than standard practice was viewed by some as less important:

> I think it lost its validity in the fact that it was research, just research and it wasn't like a test we're doing to like everybody.
>
> (H2MW)

It appears that staff did not initially feel involved in the study and resentment towards the study may have been reduced by a higher level of consultation:

> We're sensible people we know things have to happen but I think just sometimes being consulted on stuff and feel you are part of the process will just remove some of that feeling of we're being put upon.
>
> (H2MW)

Consent process

It had been intended that parents would give initial informed consent antenatally. In practice, this often did not happen and many parents only gave consent postnatally. Assistants at Hospital 2 felt that the original process would have been helpful, rather than parents being given a lot of information when on the ward, and they noted that interpreters would have been more available antenatally. It was also felt that post labour was not an ideal time for the consent process:

> If they'd had a section … or a long labour, they'd just be really shattered.
>
> (H2MW)

A time delay could mean that consent and testing was passed to the next shift, increasing the likelihood that the test could be missed. With ward stays being short for apparently well mothers and babies, staff could have limited time in which to perform the test. In contrast, examples were noted where antenatal information increased interest in the study from parents:

> You've still got some coming through who've had the information and they're like 'Oh we know about this'.
>
> (H1A)

Influences on participation

The consent process was raised as a barrier to participation in all focus groups. First, the quantity of paperwork could be off-putting. Some staff felt that some parents who did not participate had not understood what the study was about. Some mothers seemed to associate 'screening' with more invasive procedures that could affect decision-making about keeping a baby:

> I suppose they saw it in a similar way to you know, the screening in pregnancy for Down's syndrome.
>
> (H1MW)

Thus, not only could the elaborate consent process have seemed strange, but it may also have implied the study procedure to be more dangerous for babies than it actually was. Paradoxically, the formal consent procedure seemed to increase anxiety in parents rather than inform and reassure them.

Some parents seemed to have been negatively influenced by the test being part of a study rather than routine care:

> 'You're going to use my baby as a guinea pig' type of thing.
>
> (H2A)

On the other hand, the size of the study, drawing participants across the Midlands, was found to be reassuring.

Cultural differences in the various ethnic groups were noted as having an impact on participation at both sites. At Hospital 1, both focus groups identified Somalian parents as being more wary of taking part, though most did consent:

> They're very much wanting to leave things to nature and not wanting to interfere so that's obviously somewhere within that culture.
>
> (H1MW)

In focus groups at hospital 2, members of the South Asian community were perceived as being less likely to take part. Staff felt that there may have been comprehension problems although written documentation was provided in their own languages, and South Asian mothers were perceived to be more likely to want to wait for their husbands before deciding whether or not to take part. Across focus groups, the majority of staff who voiced an opinion felt that the participation rate had increased substantially since the end of the study (testing continued post study at these sites), which suggests that the study procedures were a key obstacle in restricting access to participation.

The experience of carrying out the test

Staff in the 'assistant' groups reported finding the test 'fiddly' at first, but easy to carry out once they had practised a little. If the test needed to be carried out quickly, for example when parents were keen to take a baby home, the test could be more difficult as the baby needed to be calm for the test to be effective. Midwives at both hospitals seemed to be less confident because the majority of testing had been carried out by assistant staff and the expertise of midwifery assistants was recognised by midwives:

> I go with the MA and watch them.
>
> (H2MW)

Some practical issues were encountered with the testing equipment. For example, problems were encountered in keeping the machines sufficiently charged; at hospital 1, strategies such as keeping machines plugged into the power supply were tried (taking babies to the trolley rather than the machine to the babies); hospital 2 changed to machines that took batteries, eliminating this problem. Problems with the equipment could be frustrating and stressful for staff and were seen to lead to anxiety for parents:

> That's a worry for the parents if you do not pick it up straight away because they think 'Oh there must be a problem'.
>
> (H2A)

Staff reported positive feelings about carrying out the test and pride on detecting ill babies. This was particularly true for 'assistant' staff, although the extent of this differed between individuals:

> It was something we'd never done, so it was interesting to do it.
>
> (H2A)

However, testing could also cause anxiety for staff:

> I was always dreading finding a horrible [unwell] one.
>
> (H1A)

Nevertheless, staff generally seemed happy to carry out the test, apart from the problem of time pressure:

> If we've got the time then I quite enjoy … working my way round.

> (H1A)

Both of these hospital sites continued with pulse oximetry testing after the end of the study, and staff reported the reduction in paperwork resulting from the test being part of usual care as being helpful.

The impact of carrying out the test on staff

It was generally felt that the impact of carrying out testing after the study, when the test became part of routine care, was different from that during the study. Once the test entered routine care, written information at booking and only verbal consent were required, as for any other procedure. Without formal informed consent being required, staff found it easier to fit pulse oximetry into a routine. However, on a busy ward (with well babies and mothers), even without the process of formal consent, the test could require a substantial amount of time:

> You can do at least 15 a day.

> (H1A)

The change in midwives' involvement in testing varied with the institution. At both sites, there was consensus that the majority of testing was carried out by those in assistant roles. This appears to have caused considerable friction at both sites, as all staff (midwives and assistants) were trained in testing for the study and the expectation had been that everybody would carry out testing. Although some assistants seemed to have accepted testing as part of their role, others were less accepting and encouraged midwives to share the role:

> I was like you know 'it is not our job'.

> (H1A)

In contrast, there was some suggestion that midwives perceived assistants to have chosen to take on the testing:

> They've made it part of their role.

> (H2MW)

The idea that testing should be a shared role seems to be accepted by both midwives and assistants, but the test was also viewed as being part of the baby check process by midwives – a task mainly carried out by assistants at both sites.

Assistants felt that workload per se was not an explanation for midwives not testing as assistants were also very busy. They tended to attribute lack of testing to inclination:

> They just do not want to do it, do they?

> (H2A)

Or belief that the midwife's role was more important:

> Their role obviously is more, it appears more important than our role, so they think they've got to get their job done and we can perhaps leave things.

> (H1A)

Midwives could initially lack confidence with testing, as they had less practice. However, when lack of confidence had been given by midwives as a reason for not testing, it was not well received by assistants, who felt that they could easily gain confidence by practice, especially with their higher level of training:

> I can deliver a baby but I cannot put a thing on a hand!
>
> (H2A)

There was a sense that it had become an assistant role because assistants had taken it on board:

> It's just easier if you do it, you do it properly … then it's just another job that's been given to you.
>
> (H1A)

At hospital 2, the resentment of assistants was compounded by their not being allowed to consent people for the study:

> We're good enough to do the test but not good enough to consent them.
>
> (H2A)

However, despite the friction, midwives at both hospitals spoke positively and sympathetically of the role taken by assistants:

> They're fantastic in doing them.
>
> (H1MW)

It seemed that clarity as to whose responsibility testing is would have been helpful, and may have reduced resentment on the part of assistants:

> If … they'd said 'Right, this is going to be a nursery nurse's job or a health-carer's job' then fine, you get on with it.
>
> (H1A)

However, when asked how they would have felt if testing was presented as a job for assistants, responses from midwifery assistants at hospital 2 included:

> It would have been just another thing on the list. The list is growing.
>
> (H2A)

Overall, it would seem that nominating testing as an assistant role may have reduced friction between the staff groups, but would increase the already busy workload of assistants.

Perceptions of parent experiences

Participants were generally perceived to be happy to take part. It was noticed that parents tended to be happy to take part if they saw the test being conducted on another baby (H2A). Parents seemed interested in the test results (H2A) and reassured by test results (H2MW). However, staff were aware that waiting for test results could cause anxiety. The initial problems with oximeters encountered at hospital 2 were felt to exacerbate the anxiety of parents:

> They think there's something wrong with the baby.
>
> (H2A)

> We'd say it's probably the machine … get another machine … hopefully it'll work that time … the majority of the time it did. When it did not then, they'd be panicking because it took you such a long time to get the results.
>
> (H2A)

A disadvantage to observing testing in others was that if the other babies had passed, but a mother's own baby was referred for further testing:

> It scares the life out of the mum.
>
> (H1A)

Where babies had a positive test result, anxiety was perceived to increase while waiting for an echocardiogram, although it was anticipated that a clear echocardiogram would provide reassurance.

Communication

Staff seemed to appreciate being able to inform parents about the test 'in simple terms' (H2A) since the end of the study, treating the test in the same way as any other test that was carried out. One midwife observed that people were not as informed, but seemed to imply this in a positive way:

> It's not this big rigmarole about can you read the information leaflet, can you sign there.
>
> (H2MW)

This change in communication strategy was seen to result in fewer refusals. During the study, the information provided to parents as required by the ethics committee could make it difficult for staff to respond to parents' questions:

> And then they start thinking … 'well, what can go wrong?' And then you're like 'Well I do not know!'
>
> (H1A)

When babies failed the oximetry test, the midwife or assistant would have to break the news to the parent(s). Understandably, staff disliked this situation: 'You do feel a bit like the baddie' (H1A) and would try to reassure them. As assistants are not trained to the same level as midwives, they were asked about how they felt about any training they had received in communicating test results. Although they had received training in carrying out the test, they were not trained in talking to parents about the results. Some felt their experience was beneficial in this situation, but training may have been useful.

ACCEPTABILITY TO MEDICAL STAFF AND ADVANCED NEONATAL NURSE PRACTITIONERS

Methods

Design

The fast turnover of junior medical staff and the wide geographical distribution of these and more senior staff meant that it was not feasible to run focus groups with medical and nurse practitioner staff. Instead, open-ended format questionnaires were administered by e-mail. A summary of each staff group's responses was e-mailed to participants to allow them to respond to the thoughts of their peers.

Participants

The principal investigator at each of the six study sites was asked to provide e-mail addresses of staff who had acted as senior house officers (SHOs), specialist registrars/clinical research fellows/staff grade doctors (SpRs), consultant neonatologists (CNs) and advanced neonatal nurse practitioners (ANNPs) during the time of the PulseOx study. In total, 119 staff members were contacted by e-mail; 26 (22%) completed and returned e-mail surveys. Eight participants replied a second time after receiving the response summary; four of these responded that they had nothing further to add. The respondents consisted of six SHOs from five hospital sites, nine SpR group doctors from five hospital sites (including two clinical research fellows; no staff grade doctors responded), six consultant neonatologists from four sites and five ANNPs from four hospital sites. Thus, although the response rate was only 22%, responses were well distributed across staff groups and hospital sites.

Measures and procedure

The questionnaire e-mailed to staff included open-ended questions about the same issues raised with the focus groups above. In addition, examples were requested of cases where there were problems in connection with the test or cases where the test proved beneficial. These questions were sent by e-mail on 19 June 2009; reminders to non-respondents were sent 29 June 2009.

Within each group's responses, a summary of responses was developed. The summaries were sent to respondents on 13 July 2009 and a reminder was sent on 21 July 2009. Participants were invited to comment on the contents of the summary document, for example where an issue had been raised that the participant had not considered or disagreed with.

Analysis

A thematic analysis of the returned surveys and responses to summaries was conducted using the framework approach.[94] The same procedures were followed as for the focus groups reported earlier.

Results

Impact on roles

The staff groups had different levels of involvement with study participants. SHOs were predominantly involved with examining babies who had failed the initial test; one reported telephoning to arrange echocardiograms (SHO6); another being involved with carrying out echocardiography under supervision on three babies (SHO4). SpR group members reported reviewing babies where there was concern, and reassuring parents (SpR1, SpR4, SpR7), making judgements on admitting babies (SpR6), and carrying out or arranging echocardiograms (SpR4, SpR5, SpR6, SpR7). The one consultant who reported carrying out echocardiography reported an increased echocardiography role (CN1); the other consultants did not feel that pulse oximetry had greatly affected their roles. ANNP group members also reported a low level of involvement, reviewing some babies who had failed the initial test (ANNP4, ANNP5).

Some SHO and SpR group participants reported that reviewing babies who failed their first test could be time-consuming, particularly on a busy shift (SHO1, SpR1, SpR4):

> … only became a concern when alone on nightshifts.
>
> (SHO2)

However, some felt that the impact on their role was minimal (e.g. SHO4, SHO5, SpR9) and a positive impact was also reported:

> … a safety net in case we missed subtle signs of CHD at baby check.
>
> (SHO5)

> … adds to confidence when examining baby.
>
> (ANNP1)

> … reassurance for junior staff.
>
> (SpR9)

A number of staff, across groups, noted the positive impact of pulse oximetry on their roles in that babies with problems were identified early, before collapse. Other positive impacts noted were increasing the awareness of cardiac problems and facilitating decision-making about the seriousness of heart murmurs.

Some respondents noted the impact on other staff groups:

> My main concern was the additional workload for an already understaffed and overworked postnatal ward.
>
> (SHO5)

On the other hand, it was also suggested that it was 'empowering':

> … empowering for midwives to be involved in the 'cardio-respiratory monitoring of the newborn baby.
>
> (CN2)

The majority of participants were very positive about pulse oximetry testing becoming routine. Some concerns were reported about resources to do this:

> The extra work involved needs to be recognised and must be factored in, with appropriate resources and support provided.
>
> (ANNP1, response to summary)

However, it was also noted that not to continue would be a waste of costs already paid:

> I would hate to see the midwifery training investment to go to waste.
>
> (CN2)

Benefits and costs of pulse oximetry testing

Across all staff groups, pulse oximetry testing was widely regarded as both worthwhile and effective:

> Highly effective … definitely worthwhile.
>
> (CN3)

Some participants were more reserved in their commendation but even where reservations were noted, participants tended still to be positive about the test's use:

> … a little variable, depending on when and how it is done, but overall a very useful tool … very worthwhile – I would like to see it adopted as standard.
>
> (ANNP5)

Early detection of CHDs and some respiratory conditions was cited as a key benefit of carrying out pulse oximetry across all staff groups. The benefits of identifying ill babies while they were stable, rather than later when they could be in an unstable condition, and of helping to reduce the risk of collapse were noted. ANNPs also felt that the testing could reassure staff and increase awareness of oxygen saturation monitoring and other cardiac issues with midwifery and neonatal unit staff, and another identified parental reassurance as a benefit:

> … adds to confidence when examining baby.
>
> (ANNP1)

Some participants interpreted the question about the costs of testing as requiring a monetary response, with estimates from 'many thousands for each trust' (SHO2), to 'minimal' (SpR3). However, most respondents took a broader view of the costs of pulse oximetry. Most commented on the costs of staff time – for training, for midwives in carrying out the test and doctors who reviewed babies who failed the test; this was particularly emphasised by the consultants. Equipment, rather than staff time, could be viewed as a cost:

> I'd imagine that the paperwork and extra monitors have been expensive but the midwives do all the work at no extra cost.
>
> (SHO3)

Respondents from most staff groups noted the anxiety caused by FP results:

> I have concerns about the added stress involved for the families of babies that failed the screening once or twice but subsequently proved to be healthy.
>
> (SHO5)

When asked about effectiveness and how worthwhile the test was, some participants considered the cost-effectiveness of testing. To some, the benefits were such that testing had to be worth it:

> Perhaps not incredibly cost-effective but worth it for the benefits of detecting ill babies.
>
> (SHO3)

Others noted the importance of investigating the cost-effectiveness of pulse oximetry.

Changes in evaluation of the test over time

Some staff reported changes in their thoughts about the test through the experience of being involved in the study. Change was particularly seen in the SHO and ANNP groups. Within the ANNP group there was some concern at the start that the study would be a waste of time, creating more work for staff; these staff had changed opinion to see the test as worthwhile, increasing confidence when examining babies and detecting some life-threatening lesions. Within the SHO group, initial thoughts about the test had tended to be either not knowing much, being 'interested' or, in one case, expecting it to 'involve lots of work' (SHO6). Most participants reported their opinions as having changed to being impressed or at least seeing the test as useful. Within the consultant group, opinions were largely unchanged since the start of the study; staff reported viewing pulse oximetry positively even at the start of the study. Some felt that the study had reinforced their positive opinion and the importance of seeing the study's results were noted:

… more confidence on its effectiveness.

(CN1)

To weigh up the benefits … against the increased parental anxiety.

(CN2)

Problems encountered in testing

Some problems were encountered with the process of testing and some suggested that a dedicated screening team would be needed if screening became a national programme:

Often delays in midwives in undertaking the test.

(SHO6)

Additional workload for an already understaffed and overworked postnatal ward.

(SHO5)

The feeling was also expressed that some mothers had not understood the purpose of the test because of language issues. Other procedural problems noted were not being contacted with an abnormal result (SpR4), some midwives needing support in the placement of probes (ANNP5) and errors in fixing tape to babies, leading to unreliable readings (SpR8).

Other problems reported were associated with communication with parents and their understanding of the test. Conflict was noted by two SpR participants (SpR6, SpR7) when an echocardiogram was not immediately available and the family wished to go home (it is possible that both participants were referring to the same case). Discussion enabled a compromise to be reached, with the family going home and returning to hospital for the echocardiogram:

I think proper communication and the choices/alternatives given for echo scan would solve most of these problems and it only intensifies if we are rigid.

(SpR6)

A consultant at another hospital reported a case where, after positive pulse oximetry results, it was difficult for staff to persuade parents that the baby could have a significant problem.

(CN3)

Staff perspectives of parents' perceptions of the test

Staff in all groups perceived parents to have mixed feelings about undergoing testing. Most staff felt that the majority of parents viewed the test positively, with most parents being reassured by negative results. However, staff were also aware that the test could cause anxiety and anxiety was particularly likely if a baby had failed the first test and were aware of the extent to which a normal repeat test was reassuring after failing an initial test:

Several thought it was an additional thing to worry about.

(SHO5)

Requiring an echocardiogram was seen to provoke anxiety; this anxiety was particularly perceived as unfortunate where the echocardiogram revealed a baby to be well:

Some parents with babies requiring echo that was subsequently normal will have had their first few days of joy and celebration clouded by anxiety.

(CN2)

Failing the test could also cause anger or frustration, particularly if they had to wait for an echocardiogram:

> One father was extremely cross that his baby had failed the test and was found to have a serious condition.
>
> (ANNP5)

However, it was largely felt that parents whose babies did have a cardiac or respiratory problem were pleased to have had the test:

> Most parents I had contact with were thankful that the oximetry picked their baby up with either respiratory or cardiac condition.
>
> (CN1)

Senior house officers seemed to be less likely than other groups to comment on the positive impact for parents with ill babies compared with SpR, CN or ANNP group participants. This is likely to reflect the stage of the testing pathway at which parents encountered the different staff members, with SHOs mostly involved at the early stage of reviewing babies who had failed the first test.

Communication

The extent to which parents understood the purpose of the test, particularly in the context of the study, was questioned:

> A small number did not really understand what the test was about and its limitations.
>
> (ANNP5)

Some participants thought that parents saw the test as one more health check rather than thinking about it in the context of a study (SpR4, CN3, CN4). These consultants (CN3 and CN4) thought that parents did not understand about the study because of the way that midwifery staff took consent:

> Most considered it part of normal care and an expectation that they would want to take part was often assumed by those explaining the study to them … those seeking parents' acceptance need to be properly trained.
>
> (CN3)

This seems to contradict findings that the paperwork of the information and consent processes took a great deal of time for midwifery staff and assistants, a process that eased when the test did become part of routine care and the test could be treated like any other routine test.

Discussion

Summary of the findings on acceptability

White British and Irish mothers were more likely to participate in the study. This was also found in an earlier study of perinatal screening, for group B *Streptococcus*, in which ethnic minority groups were under-represented in the study sample.[80]

Across all parts of the study, parent and staff participants were predominantly satisfied with pulse oximetry screening, perceiving it to be an important and valued test to detect ill babies. Despite

creating an additional workload, the test was seen as reassuring for staff, and positively impacted on the roles of those caring for babies as they could be treated before the condition deteriorated.

There was no evidence that mothers given FP results were more anxious after taking part in the screening processes than those given TN results, though they were less satisfied with the test and gave higher depression scores (a small, but statistically significant difference). However, questionnaires tended to be returned later by the FP group, so anxiety and depression may have reduced by the time they recorded their responses. In multivariate analyses, higher anxiety and depression was predicted by lower optimism, lower overall satisfaction and ethnicity (white British/Irish participants being less anxious). Satisfaction with screening was predicted by higher perception of the treatment's ability to control heart disease, comprehensibility of heart disease, and lower stress, anxiety and depression. White (British/Irish) participants were more satisfied than those of other ethnicities. Participants given FP results recalled the testing process as more stressful, probably reflecting the need for a second test and communication about the results. Lower comprehensibility of the condition and lower satisfaction with testing also predicted higher reported stress during the test. Of those with FP results, under half believed that the test showed no problem. This mismatch between research-defined test results and parental perceptions may indicate that communication was inadequate. However, as most parents with TP results believed the test had shown there to be a serious problem, it is more likely that parents are not distinguishing between the pulse oximetry test and the further investigations that followed a positive result.

Indian mothers were less satisfied overall with screening and Pakistani mothers were more stressed during screening than white British/Irish mothers. It would be useful to identify the factors leading to these effects to ensure that women receive sufficient information in a format which they understand. Participants who believed that heart disease can be controlled by treatment were more satisfied with screening – people who believed something could be done about the illness were perhaps more willing to put up with the testing procedure. Those who felt they understood heart disease both were more satisfied with screening and found the test less stressful. It may be helpful to focus more information on heart disease itself to increase parents' understanding of the condition.

Communication problems were indicated as a cause of worry by participants, and staff identified a need for further training in communicating with parents about the study and for giving results, especially where a positive result is found. Malfunctioning equipment could increase anxiety for parents and staff alike. The problems relating to equipment and the reasons for them are described in *Chapter 2*.

Staff who carried out testing found the formal consent procedure required in research to be the major factor impacting on their time, and they also perceived this to be an impediment to participation, with mothers being given a large amount of information to read at a time not long after having given birth. This is supported by a parent comment that the standard information sheet was off-putting because of required wording about the risks of taking part in research. Uptake of testing was observed by focus group members to increase when the study was completed and screening was continued as standard care in those hospitals. A lack of clarification as to whose role it was to perform testing on the wards could result in friction between busy staff groups, but this seemed to be less of a problem when the procedure was carried out as standard care than when it carried the paperwork of a research study. Paperwork was also cited as a cost of testing in group B *Streptococcus* screening during labour.[80] On the other hand, it is important that parents understand what is involved even in routine screening and are given the chance to make an informed choice. This raises broader issues of how informed consent is obtained in clinical practice in the perinatal period.

Strengths of the study

This study is the first to address the acceptability of pulse oximetry screening from the perspectives of both parents and health-care professionals. In so doing, a convergence of opinions was seen on the benefits and problems associated with screening, strengthening the findings. Factors have been identified that may enable more parents to understand and accept screening in the future, and that will facilitate the testing process on hospital wards.

Limitations of the study

Because of research ethics constraints, other than basic demographic data, no information was collected from those who declined to take part. So the study does not provide an insight into reasons why parents found either the study or the testing procedure so unacceptable that they would not allow their baby to be screened. Two hospitals continued testing as routine practice after the end of the research study. The incidence of parents declining testing (after receiving an information sheet at booking, and verbal explanation and consent at the time of testing) is reported as extremely low.

There was no follow-up TN group for comparison with the FP follow-up group, again because of constraints imposed by the Research Ethics Committee. However, as there was no difference in anxiety at baseline between TN and FP participants, and anxiety did not change over time for FP participants, this omission is less problematic than if anxiety had been elevated for mothers receiving FP results.

An untested variable that may have affected emotional state or satisfaction with screening was the time between birth and questionnaire completion. It could be hypothesised that someone completing a questionnaire on discharge from hospital would be more anxious and the test would be more salient than in case of someone completing the questionnaire later. However, this variable was heavily confounded by test result group, with TN participants receiving the questionnaire on discharge, although the questionnaire was often sent by post to FP participants.

The analysis of parent comments was limited to those who both completed a questionnaire and those who also chose to give additional information – as such, this is a small subsample of mothers and views may not be representative. However, it might be expected that mothers who had particularly negative experiences or perceptions might be more likely to give comments so, although data may be missing on 'ordinary' test experiences, we may have gathered reasonable data on participants who had exceptionally poor (or good) experiences.

There are also sampling issues with the staff studies. The first contained representative samples of staff on duty within each of the two hospitals, but only two hospitals were included; each hospital has different managerial structures on wards, so roles may have been differently affected at other hospitals. The second only included a small proportion of neonatal staff involved with pulse oximetry screening, but the small sample was well distributed across staff groups and hospitals.

Findings in light of the limitations

Although ethnic minority groups were under-represented in the acceptability evaluation, the larger groups in the population were still present in sufficient numbers to allow comparison of their responses.

It is not clear whether or not the absence of a finding of elevated state anxiety among mothers receiving FP results is because of a lack of increased anxiety in this group or because anxiety had reduced by the time they completed questionnaires. At the least, the findings indicate that there is no need for concern about persisting increases in anxiety for these women. However, it would be

useful to identify whether or not there is an elevation in anxiety at the time of discharge so that support can be put in place if necessary.

Although some issues raised in focus groups may differ from those raised in other hospitals, for example issues surrounding workforce organisation, other aspects of the findings seem to be robust as they are supported by the data from parent and e-mail group responses, i.e. general high satisfaction with screening, impact of paperwork, importance of communication training and problems caused by malfunctioning equipment.

Ramifications for the economic model

The results show that having clearly defined testing roles and functioning machines facilitated screening among ward staff, particularly when screening was part of standard care rather than within a study. The time costs of study paperwork, particularly taking informed consent, were important within the context of a study, but would be of less importance should testing be widely adopted as standard care. However, the findings indicate the need for training for staff not only in conducting the test but also in communicating the results to parents. Some staff suggested that screening should be carried out by a dedicated workforce. Both of these factors have cost implications.

Recommendations for practice

Many of the problematic issues raised by staff related to the research study rather than screening per se, for example taking informed consent, and so would be of less concern if it was adopted as standard practice. As suggested above, the findings clearly show the need for staff carrying out screening to receive training in communicating results to parents in a way that minimises both parental anxiety and false reassurance. Just under half of those with FP results believed the test showed no problem. This may be an indication that mothers confused the pulse oximetry test with later investigations, but it could also suggest inadequate communication of results. However, it is clear that staff from all professional groups were positive about testing and saw the benefits of identifying babies with problems early. Parents may also need information on heart disease in order to understand better both the screened conditions and the importance of testing. The timing of the giving of that information is important and probably needs to be during pregnancy when options for birth are discussed with midwives. The screening role should be clearly defined as a responsibility for a particular staff group or groups to avoid conflict between these groups.

Recommendations for research

Further research should be conducted with mothers of different ethnicities to gain a greater understanding of factors limiting participation and satisfaction with testing. Research should also be conducted with mothers given FP results at the time of discharge from hospital to verify whether or not they experience heightened anxiety at this point.

A key factor limiting participation in this study, and one which impacted greatly on ward staff, was the complex information and consent procedure required by Research Ethics Committees. Although it is extremely important to ensure that research participants fully understand what research is about and participate only after giving full informed consent, when this process both limits participation among particular groups and greatly impacts on the roles of NHS staff, the process raises other ethical questions: to reduce health inequalities participation is needed from all parent groups; excessive time costs for staff running the study has implications in terms of funding required – reducing these costs enables funds to be put to use elsewhere. Conversely, routine screening should not be carried out without full information provision. Further research is needed on whether or not parents whose babies undergo routine screening have been given information and guidance to give informed consent.

Chapter 4

Economic evaluation

Introduction

This chapter reports the model-based economic evaluation that was carried out alongside this primary study. The objective of the economic evaluation for this study was to compare the cost and cost-effectiveness of pulse oximetry as an adjunct to current practice with current practice alone, based on an outcome of cost per timely diagnosis. The choice of outcome follows the recommendations of Knowles et al.[6] and is explained later in this section.

Background

The funding call from the NIHR Evaluation, Trials and Studies Coordinating Centre (NETSCC) for this project was initiated in part by the publication of the HTA report by Knowles et al.[6] As part of that project, the authors had undertaken a systematic review of the clinical and economic literature and used the data obtained to undertake a model-based economic evaluation of pulse oximetry to diagnose CHDs in the newborn. They also assessed the role of screening using echocardiography. The authors had developed a comprehensive model that strongly suggested that pulse oximetry offered what could be considered a cost-effective screening test for detecting CHDs early in newborn infants. As the evaluation was based on published data, it was acknowledged that pulse oximetry required testing in a UK clinical setting to obtain primary data on its accuracy and associated costs. Against this background, we agreed that a pragmatic approach to the economic evaluation was required which would build on the model structure developed by Knowles et al. as far as possible, but adapting it where necessary to fit the pathways and new data that would be gained from the primary study. The majority of data required for this economic evaluation had already been obtained by Knowles et al.[6] and repeating these efforts would not be an efficient use of health-care research resources. The principal new data that were required from the current primary study were the accuracy and costs associated with a pulse oximetry test in the newborn. Therefore, the current economic analysis requested limited funding from the NETSCC as a consequence of adopting this pragmatic approach. As a result we draw heavily on much of the secondary information presented in Knowles et al.[6] We are also grateful to Knowles et al.,[6] who provided us with access to their model and the data used in their model-based economic evaluation.

Methods

The details of the accuracy study are described in *Chapter 2*. For the economic analysis we developed decision-analytic model, using as a starting point the model developed by Knowles et al.[6]

Model structure and screening strategies

The primary study and the model differ in that, in the study, all infants enrolled received pulse oximetry, but in the model for the purposes of the base-case economic evaluation, we compare

the relative cost-effectiveness of a strategy in which pulse oximetry is an adjunct to routine practice with routine practice alone.

- *Strategy 1* In the main model structure, the strategy referred to as *clinical examination* represents 'routine practice'. Thus, it refers to the clinical pathway that newborn infants would have followed if pulse oximetry had not been carried out. Typically, all infants would receive a routine clinical examination by a trained professional as described in *Chapter 1* before they are discharged from the maternity unit.

If the clinical examination result is abnormal the infant will go on to receive further tests, depending on the abnormality or concern raised. Those in whom a CHD is suspected will usually require diagnostic echocardiography.

- *Strategy 2* The alternative strategy in the model is referred to as *pulse oximetry as an adjunct to clinical examination.* Infants for whom the pulse oximetry test is considered normal (test negative) will proceed to receive the same clinical examination as in Strategy 1, and, in the model, will follow the same pathway as they would if they had not received the pulse oximetry test. For instance, if the clinical examination provides an abnormal result that is suggestive of a CHD, the infant will be sent for diagnostic echocardiography. If a non-CHD related abnormality is suspected they will be sent for other appropriate tests.

If the infants in Strategy 2 have an abnormal pulse oximetry test result (test positive), they will receive an expedited clinical examination (ECE), which is primarily checking for CHDs, and is undertaken sooner than the routine clinical examination. If the ECE shows an abnormality (test positive) the infant will be sent directly for diagnostic echocardiography. If the ECE suggests no CHD-related abnormality (test negative) the infant will undergo a second pulse oximetry test. If this is still abnormal, the infant will proceed to diagnostic echocardiography.

If the second pulse oximetry test is considered normal, the infant will receive the remaining component of the routine clinical examination that was not part of the ECE, i.e. the non-CHD aspects (such as checking the eyes, hips, etc.) and continue to follow routine care as described in Strategy 1, receiving non-CHD-related interventions if appropriate.

In both Strategies 1 and 2, the diagnostic echocardiogram is confirmatory and any infant who has a CHD identified by echocardiography will receive appropriate treatment. Other abnormalities that are unrelated to CHDs will follow other diagnostic procedures that are beyond the focus of the current study.

Model structure

A decision tree model was developed in TreeAge 2009 software (TreeAge Software, Inc., Williamstown, MA, USA) to represent the alternative strategies. The pathways of the model represent, as far as possible, the clinical procedures carried out in the study. The decision tree structure is presented in *Figure 6*. The branches to the right of the decision node (square symbol) represent the strategies being compared. Babies receiving either strategy have a possibility of being in one of the categories of CHD to the right of the chance node (circle symbol). The pathways combine the probability of an infant following a particular path and the associated costs. *Table 15* shows the prefixes used in the decision tree, for example in *Figure 6* 'prev_critical', represents the prevalence of newborns with critical CHDs.

The decision tree was modelled from the disease severity category, followed by the probability of an infant with or without a CHD testing either positive or negative, respectively. However, in
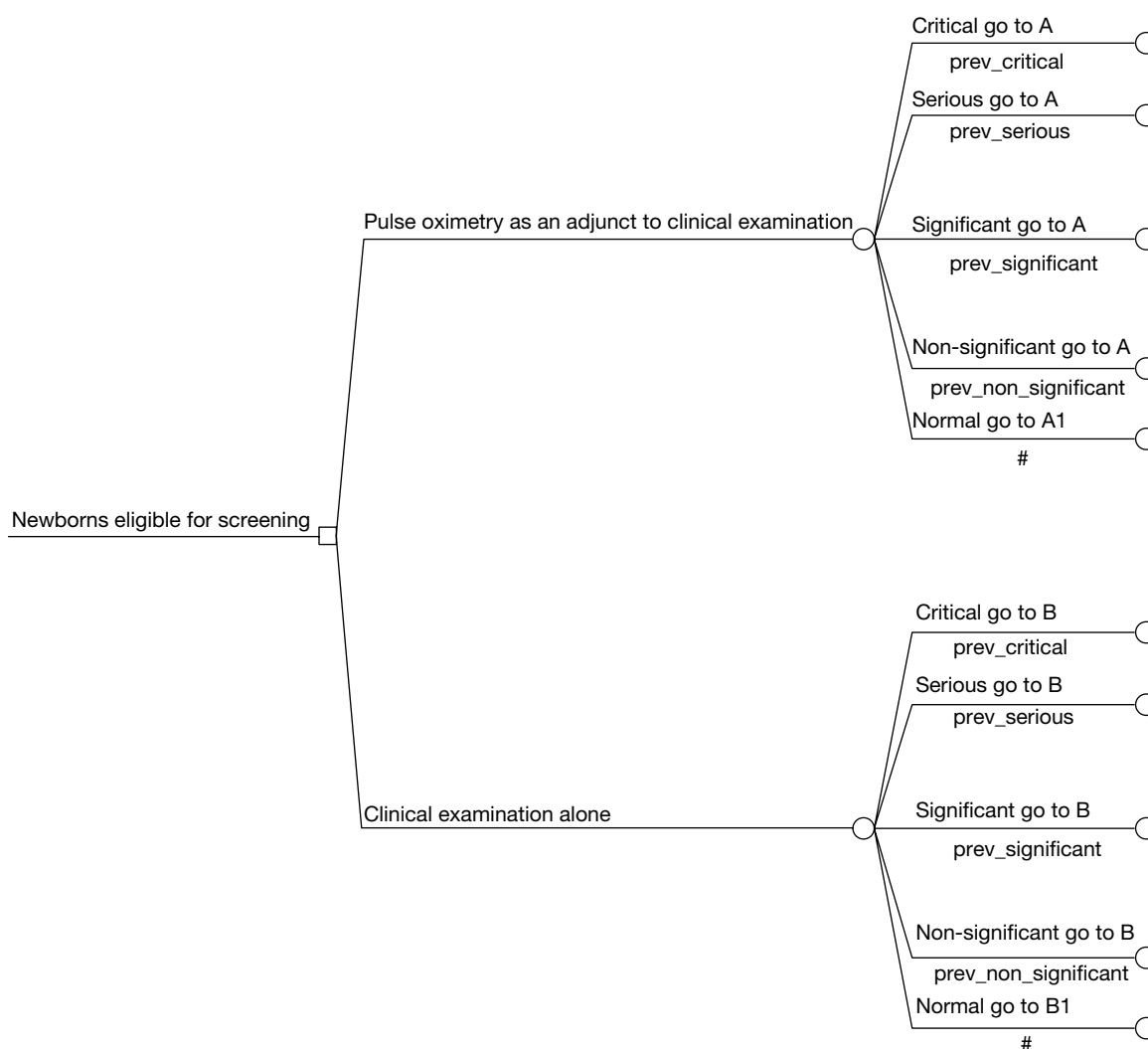
FIGURE 6 Basic decision tree structure.

clinical practice the test result is known before the specific type of CHD is diagnosed. Modelling the test result first followed by the category of CHD or vice versa makes no mathematical difference in terms of the expected values calculated.[6,95] It should also be noted that the pathway followed is based only on the test results, not on information that would not be known to the clinicians at the time.

Newborns with a CHD are assumed to have the CHD confirmed by echocardiographic assessment. If the diagnosis of a CHD is missed in the neonatal period, it could be confirmed by follow-up on the congenital anomaly or other clinical registers. For both these cases the probability of 1 is attached to this event as represented by branches A and B (*Figures 7* and *8*). Newborns without a CHD are not expected to have a confirmed CHD from the diagnostic assessment or follow-up of the registries and a probability of '1' is also attached to this event, but in corresponding diagrams represented by branches A1 and B1 (*Figures 9* and *10*).

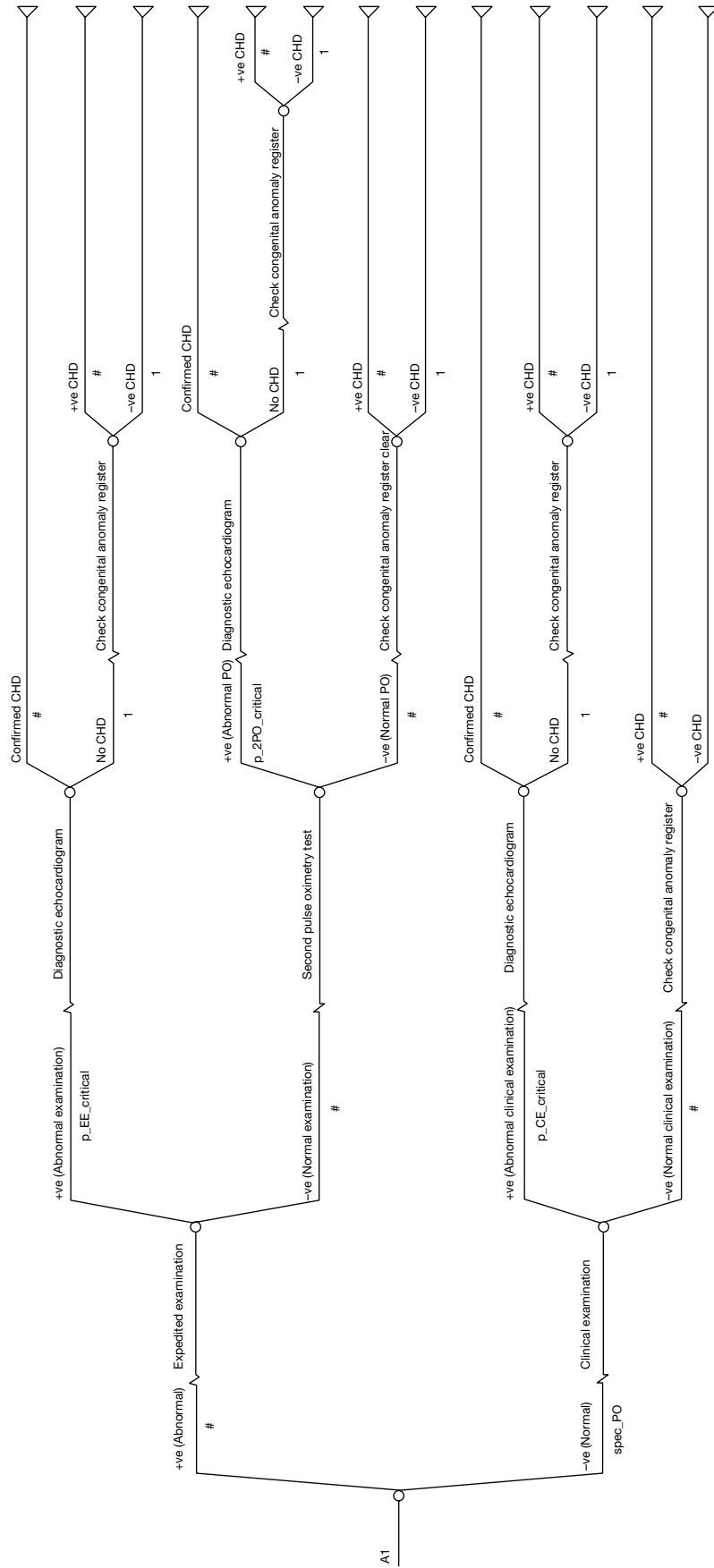The model runs for a period of 1 year, which is the year of follow-up in the study.

FIGURE 7 Pathway of the pulse oximetry (PO) as an adjunct to clinical examination strategy for newborns with a CHD.
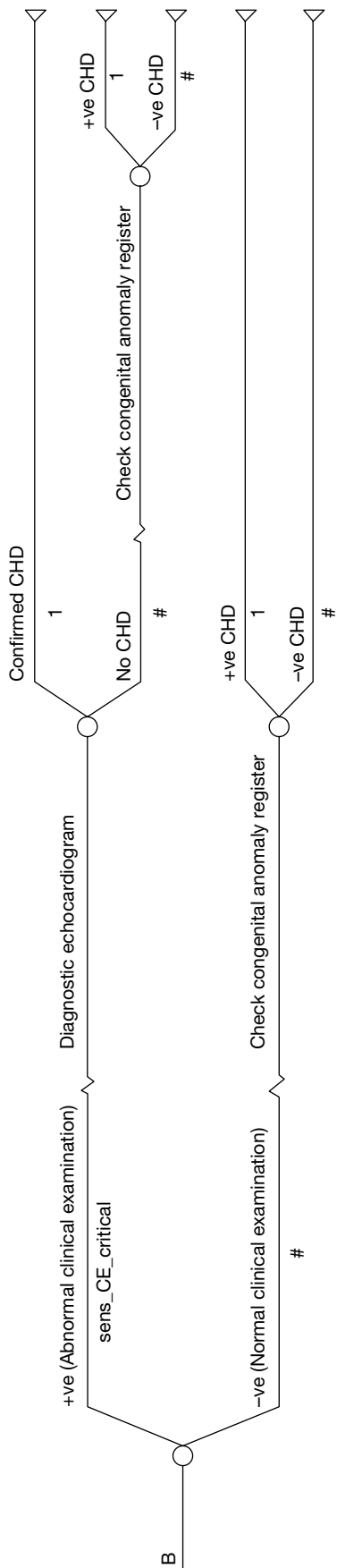
FIGURE 8 Pathway of the clinical examination alone clinical strategy for newborns with a CHD.

FIGURE 9 Pathway of the pulse oximetry (PO) as an adjunct to clinical examination strategy for newborns without a CHD.
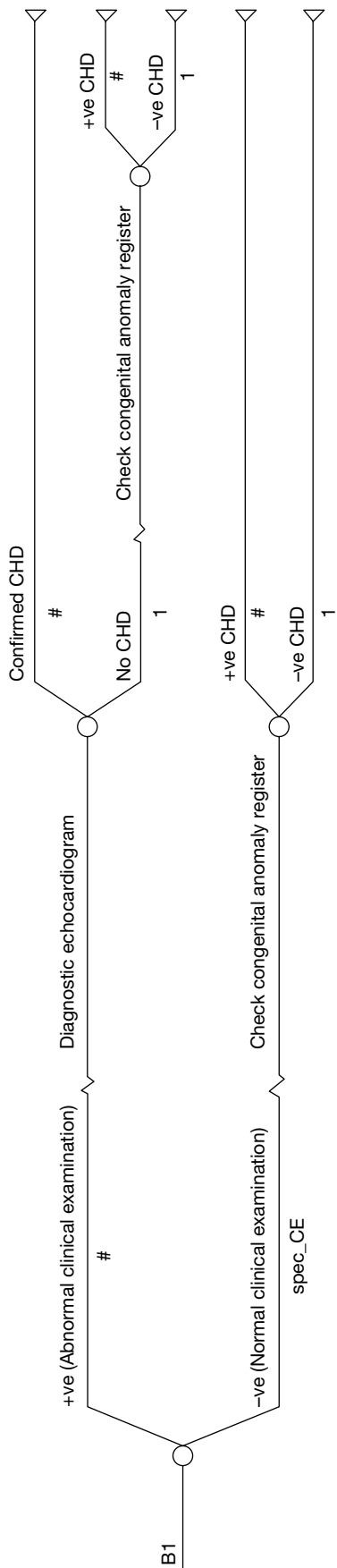
FIGURE 10 Pathway of the clinical examination-alone strategy for newborns without a CHD.

TABLE 15 Description of the prefixes used in the decision tree

Prefix	Description
c	Cost of a screening test or an assessment
prev	Prevalence of a newborn with a categorised CHD
p_PO	Proportion of newborns with an abnormal pulse oximetry result by categorised CHD
p_CE	Probability of newborn with an abnormal clinical examination result by categorised CHD
p_EE	Proportion of newborns with an abnormal expedited examination result
p_2PO	Proportion of newborns with an abnormal second pulse oximetry result
sens_PO1	Sensitivity of first pulse oximetry by categorised CHD
spec_PO1	Specificity of first pulse oximetry
sens	Sensitivity of screening test by categorised CHD
spec	Specificity of screening test

Clinical data used in the model

The decision model was populated with prevalence data from the current study. The population prevalence is presented in subgroups according to the severity of the CHD. Two separate population prevalence groups are considered. Group 1 defines the population prevalence for newborns with a CHD, but excludes the infants in whom a CHD was suspected in the antenatal period by antenatal ultrasound screening and subsequent fetal echocardiography. These are the babies for whom the addition of pulse oximetry testing could influence management of CHDs in the neonatal period. Group 2 (described as the full cohort in *Chapter 2*) includes all babies and assumes that antenatal screening is unavailable or the anomaly was missed on antenatal screening. These prevalence data are presented in *Table 16*.

The test accuracy data used for the model were drawn as far as possible from the current study, but are supplemented where necessary by Knowles *et al.*[6] Test accuracy data vary according to both the specific test being undertaken and the population being screened. Thus, test accuracy data are estimated for each strategy being compared and for both populations. The test accuracy data for the strategies of 'pulse oximetry as an adjunct to clinical examination' and 'clinical examination' for population Group 1 are presented in *Table 17*. The test accuracy for these two strategies and an additional strategy of 'pulse oximetry alone' are presented for population Group 2 in *Table 18*. The sensitivity and specificity of the use of 'pulse oximetry as an adjunct to clinical examination' to detect CHDs came directly from the current study. However, the accuracy of 'clinical examination' strategy was not estimated in the primary study for reasons that are explained in *Chapter 2*. This information was available to some extent in Knowles *et al.*[6] The primary study estimated the accuracy of the clinical examination *only* in conjunction with a preceding pulse oximetry test (which if negative would result in a routine clinical examination or if positive would be followed by an expedited clinical examination). Thus, in the primary study the clinical examination was only ever performed in the knowledge of the pulse oximetry result. Thus, the accuracy of *routine* 'clinical examination' had to be estimated in the following way: the probability of a newborn having an abnormal clinical examination following normal pulse oximetry result was taken from the current study for both critical and serious cases. However, the analogous probability was not collected in the current study for the significant and non-significant subgroups – although this is required for our model to be consistent and comparable with that of Knowles *et al.*[6] Therefore, we followed the approach of Knowles *et al.*[6] and used appropriate data for these subgroups from Wren *et al.*[26] From this source, we categorised the individual CHD cases based on the definitions described in *Table 1* for significant and non-significant cases and calculated the probability of an abnormal clinical examination following a normal pulse oximetry test as 0.4520. The specificity of clinical examination was taken from Knowles *et al.*[6] The probability of a newborn having an abnormal ECE following an

TABLE 16 Prevalence data by disease category used to populate the models

Group	Critical	Serious	Significant	Non-significant	Normal
Group 1[a]	12/20,032	23/20,032	18/20,032	95/20,032	19,884/20,032
Group 2[b]	24//20,055	29/20,055	20/20,055	95/20,055	19,887/20,055

a   Babies suspected of having a CHD through antenatal ultrasound screening were excluded from this group.
b   All babies in the study included in this group.

abnormal pulse oximetry result was also recorded in the current study and for all subgroups, which includes critical and serious cases and significant, non-significant and normal cases. Thus, the sensitivity and specificity data for the strategy of clinical examination branch (not estimated directly in the current primary study directly, but only following a normal pulse oximetry test) were taken as the weighted average of the ECE (which followed an abnormal pulse oximetry test) and the routine clinical examination (which followed a normal pulse oximetry test). Wherever we have used the current study to inform the probability, we used the actual observed fraction to give the point estimate and to determine the variability. In some cases the observed fraction is not based on the whole population. We have taken account of this in the variance of the distributions used in probabilistic sensitivity analysis (PSA) and assume that the observed fraction is an unbiased estimate of the true population.

Costs and resource-use data

The majority of cost data used in the model were based on secondary costs presented in Knowles *et al.*[6] All cost data that are used, from published sources or the current study, are reported in 2009 UK prices, having been appropriately inflated if necessary using the Health and Community Health Services (HCHS) pay and price index.[96] However, the primary study presented an opportunity to estimate the true cost of carrying out a pulse oximetry test on a newborn. This cost had only been approximately estimated in the study by Knowles *et al.*[6]

We carried out a time-and-motion study, and all staff across the study sites asked to record the time it took them to carry out the test on the infant and record the result. A member of the economic research team (TR) observed some tests being carried out. The time-and-motion study was carried out for a week in January 2009, very close to the end of the study, in the hope that by this time any initial difficulties with equipment and/or technique would have been resolved. All centres were asked to record the duration of the test. Only one centre failed to return results, because a key staff member in the centre had recently left. In total, 312 forms were returned. Overall, pulse oximetry was assumed to be carried out by a midwife and the average duration of the test was 6.9 minutes. The minimum recorded time to carry out the test was 1 minute and the maximum amount of time was 30 minutes (median 5 minutes). The time recorded to complete the pulse oximetry test relates only to the time taken to undertake the test, as consent had been given prior to the birth. It was noted that there was no variation between the observed time and the reported time to complete the test. The main cost components of carrying out the test are the equipment itself and the staff time. Use of disposable oximetry probes was also included. These cost £150 and last approximately 6 months. Using the salary data and costs presented in Curtis 2009,[96] the cost of staff time for carrying out the test was £5.64. We also estimated a cost of £2.42, assuming that the pulse oximetry test was carried out by a midwifery assistant. The equipment used for the test is a pulse oximetry machine that costs £1100. The machine is assumed to have a life of 5 years. The annuitised cost of the equipment is estimated following the method of Drummond *et al.* using a discount rate of 3.5%.[97] An annual maintenance cost of 10% is then added, which is the usual maintenance cost we apply to all technical equipment. There were six centres in total: one had six machines, four centres had five machines and one centre had four

TABLE 17 Proportions required for the model for Group 1

Proportions and probabilities required	Critical	Serious	Significant	Non-significant	Normal
Pulse oximetry as an adjunct to clinical examination strategy					
Newborn having an abnormal (positive) pulse oximetry screen result (p_PO)	7/12	4/23	4/18	87/95	652/19,884
Newborn having a negative screen (specificity/detection rate) (1–p_PO)	N/A	N/A	N/A	N/A	19,232/19,884
Proportion of newborns having an abnormal expedited examination given an abnormal pulse oximetry result (p_EE)	6/7	1/4	1/4	21/87	20/576[a]
Proportion of newborns with an abnormal repeat pulse oximetry result having passed the expedited examination given an initial abnormal pulse oximetry result (p_2PO)	1/1	2/3	3/3	65/66	57/554[b]
Proportion of newborns having an abnormal (positive) clinical examination given normal (negative) pulse oximetry result (p_CE)	3/5	1/19	0.4520[c]	0.4520[c]	0.0030[d]
Clinical examination alone screening strategy					
Proportion of newborns having an abnormal clinical examination (sensitivity) (sens_CE)[e]	3/4	2/23	0.4070	0.2590	N/A
Proportion of newborns having a normal clinical examination (specificity) (spec_CE)[f]	N/A	N/A	N/A	N/A	0.9960

N/A, not applicable.
a   Seventy-six missing results here.
b   Two missing results here.
c   Probability calculated from the Wren *et al.*[26] study based on the echocardiography definitions used in the study.
d   Probability calculated from the Knowles *et al.*[6] study based on the specificity of clinical examination (0.9970).
e   Proportion and probabilities calculated based on the weighted average of newborns with abnormal expedited and clinical examinations.
f   Probability calculated based on the weighted average of newborns with normal expedited and clinical examinations.

machines. For each centre, the total equipment cost was divided by the number of infants who had the test, to achieve an average cost per infant for the use of the pulse oximetry machine. This cost was estimated to be approximately £0.57. We added £0.03 per infant to cover costs of disposables. We also estimated the cost of equipment using the same method, but assuming a 3-year lifespan for the machine and this cost was £0.78.

The total cost of carrying out the pulse oximetry test, including staff time, equipment and disposables based on a 5-year lifespan and assuming that a midwife carried out all tests, was approximately £6.24. *Table 19* shows the unit costs used in the model compared with the costs used in the Knowles study[6] appropriately inflated.

We assumed that the second pulse oximetry test, if required, will be carried out by a staff member of similar grade (midwife) and will require the same average time for completion as the first pulse oximetry test.

Clinical examination/expedited examination

The time taken for the ECE to be carried out was recorded in the study. The cost of the clinical examination was estimated at £5.43. This cost includes only staff cost, as no equipment was required. We assumed that the routine clinical examination would be carried out by a specialty trainee 2 (ST2) and will take the same average time of 8.57 minutes to complete as the ECE.

TABLE 18 Proportions required for the model for Group 2

Proportions and probabilities required	Critical	Serious	Significant	Non-significant	Normal
Pulse oximetry as an adjunct to clinical examination screening strategy					
Newborn having an abnormal (positive) pulse oximetry screen result (p_PO)	18/24	9/29	6/20	87/95	652/19,887
Newborn having a negative screen (specificity/detection rate) (1−p_PO)	N/A	N/A	N/A	N/A	19,235/19,887
Proportion of newborns having an abnormal expedited examination given an abnormal pulse oximetry result (p_EE)	13/18	5/9	1/6	21/87	20/576[a]
Proportion of newborns with an abnormal repeat pulse oximetry result having a normal expedited examination given an initial abnormal pulse oximetry result (p_2PO)	5/5	3/4	5/5	65/66	57/554[b]
Proportion of newborns having an abnormal (positive) clinical examination given normal (negative) pulse oximetry result (p_CE)	4/6	1/21	0.4520[c]	0.4520[c]	0.0030[d]
Pulse oximetry alone screening strategy					
Proportion of newborns having an abnormal pulse oximetry examination (sensitivity) (sens_PO1)	18/24	9/29	6/20	87/95	652/19,887
Proportion of newborns having a normal pulse oximetry examination (specificity) (spec_PO1)	N/A	N/A	N/A	N/A	19,235/19,887
Clinical examination alone screening strategy					
Proportion of newborns having an abnormal clinical examination (sensitivity) (sens_CE)[e]	17/24	6/29	0.3660	0.2590	N/A
Proportion of newborns having a normal clinical examination (specificity) (spec_CE)[f]	N/A	N/A	N/A	N/A	0.9960

N/A, not applicable.
a Seventy-six missing results here.
b Two missing results here.
c Probability calculated from the Wren *et al.*[26] study based on the echocardiography definitions used in the study.
d Probability calculated from the Knowles *et al.*[6] study based on the specificity of clinical examination (0.9970).
e Proportion and probabilities calculated based on the weighted average of newborns with abnormal expedited and clinical examinations.
f Probability calculated based on the weighted average of newborns with normal expedited and clinical examinations.

Diagnostic echocardiogram

We calculated the cost of diagnostic echocardiographic assessment as £115.57. This cost includes diagnostic echocardiography and the staff required to carry out the procedure. The cost of echocardiography was based on the resource use reported in Knowles *et al.*[6] with prices for 2009 applied using of the HCHS price and pay index.[96] We followed the same assumption that the procedure would take 30 minutes to carry out[6] by a paediatric cardiologist.[96]

Analysis

The model was constructed to investigate the cost-effectiveness of the screening strategies: pulse oximetry as an adjunct to clinical examination versus clinical examination alone. The analyses were carried out from an NHS perspective and based on an outcome of cost per timely diagnosis. The chosen outcome for the analysis was highly influenced by Knowles *et al.*[6] Evidence from their review suggested that the higher postoperative mortality and morbidity may be a result of preoperative collapse. Given the paucity of evidence relating to the possible outcomes (adverse neurological sequelae, especially cognitive, speech and language and motor deficits) caused by

TABLE 19 Unit costs and resource use for current model based analysis and Knowles *et al*.[6]

Unit cost data for tests	Current study (2009 prices) (£)	Knowles *et al*.[6] inflated to 2009 prices (£)[a]	Knowles *et al*. 2005 (2001 prices) (£)	Duration of procedures (minutes)		Difference between unit costs	Source
				Current study	Knowles *et al*.[6]		
Pulse oximetry test (first)	6.24	2.04	N/A[b]	6.9	2.0	Unit cost differ owing to the average time taken to complete the pulse oximetry test	Knowles *et al*. 2005,[6] Curtis[96] and BWH[c]
Pulse oximetry test (second)	6.24	N/A	N/A	6.9	N/A	A second pulse oximetry test was not modelled in the Knowles *et al*. study[6]	Curtis[96] and BWH[c]
Expedited examination	5.43	N/A	N/A	8.57	N/A	This procedure was not modelled in the Knowles *et al*. study[6]	Knowles *et al*. 2005,[6] Curtis[96] and BWH[c]
Clinical examination	5.43	1.59	1.17	8.57	2.0	Unit costs differ owing to the time taken to complete clinical examination (expedited examination used as a proxy)	Knowles *et al*. 2005,[6] Curtis[96] and BWH[c]
Clinical examination (non-cardiac examination)	5.43	N/A	N/A	8.57	N/A	Assumed that the non-cardiac aspect of the clinical examination will take the same time as an expedited examination	Knowles *et al*. 2005,[6] Curtis[96] and BWH[c]
Diagnostic echocardiography	115.57	114.37	84.17	30	30	Assumed the diagnostic echocardiography will take 30 minutes as assumed by Knowles *et al*.[6]	Knowles *et al*. 2005,[6] Curtis[96] and BWH[c]

N/A, not applicable.

a   Inflated by HCHS price and pay index 2008/9.[96]
b   Newborns received one pulse oximetry test in the Knowles *et al*.[6] study. Cost of pulse oximetry test was not stated in the study, but was calculated at £1.65.
c   Birmingham Women's Hospital.

preoperative collapse, the authors suggested that a suitable end point for an evaluation is 'timely' diagnosis. Like Knowles *et al*.,[6] we also assumed that management of CHDs and prevention of preoperative collapse commence after a confirmed diagnosis made by an echocardiographic assessment. Hence, we base our evaluation on an outcome of 'timely diagnosis' of a CHD as confirmed by echocardiography before preoperative collapse or death of the infant. The results are presented in terms of the incremental cost-effectiveness ratio (ICER), namely the additional cost per additional case of timely diagnosis of clinically significant CHDs.

Two main analyses were carried out. Analysis 1 is the base-case analysis and was carried out on the population defined by Group 1 (see *Table 16*). Thus, in this analysis all infants who were antenatally screened and subsequently received a suspected diagnosis via fetal echocardiography are removed from the population. Analysis 2 is carried out on the entire population of babies defined by Group 2 (see *Table 16*). Thus, Analysis 2 also includes those babies with a CHD that was detected through antenatal ultrasound screening. The rationale for this analysis is that the accuracy and availability of antenatal screening process vary significantly across the UK and the number of babies suspected antenatally in the present study was relatively high compared with other published studies both in the UK and in other developed countries (see *Chapter 1*). In Analysis 2, we also include a third strategy of pulse oximetry alone, to

compare with the use of pulse oximetry as an adjunct to clinical examination and the clinical examination-alone strategies.

Deterministic sensitivity analysis and PSA was used to assess the uncertainty in the model input parameters for both Analyses 1 and 2. In deterministic sensitivity analysis there is no randomness, and during each calculation each model parameter uses its specified point value. If an analysis is repeated using the same parameters, then the results will be unchanged. In contrast, the PSA process assigns a particular distribution to model parameters, which represents the amount and variation pattern and generates cost-effectiveness results based on random sampling of mean cost and mean effectiveness estimates. This process is repeated 10,000 times in a Monte Carlo simulation of the model and shows that the variation of the ICER is a result of the variation of model input parameters. Beta and Dirichlet distributions were used for the proportions and prevalence, respectively. These are standard, and theoretically correct, distributions for representing the uncertainty in binomial and multinomial data, respectively.[98] Uncertainty in cost parameters reflects variation in practice between different centres. This is handled through deterministic sensitivity analysis.

Deterministic sensitivity analysis

Deterministic sensitivity analyses were conducted on Analysis 1 only (base-case analysis). We carried out two sensitivity analyses in which the threshold for an abnormal pulse oximetry was changed and four analyses in which an aspect of the input costs was changed.

1. Changing the thresholds:
    i. *Changing the threshold to < 95% in either limb or a differential of > 3%* This analysis was considered to determine the impact on the ICER of the base-case results by changing the threshold for an abnormal pulse oximetry test result from a differential of > 2% to determine an abnormal pulse oximetry test result to a differential of > 3%.[65,66,70]
    ii. *Changing the threshold to < 95% for foot saturations alone* This analysis was carried out to determine the impact on the costs and the effects of using foot saturations alone. In the base-case analysis we used saturations of < 95% in either limb or a differential of > 2% to determine an abnormal pulse oximetry test result. However, other published studies have used foot saturations alone to determine the accuracy of pulse oximetry test, hence, the rationale for this analysis.[13,64,67–69,71]
2. Changing an aspect of the costs:
    i. *Changing the cost to include hospitals without available echocardiography* In all of the analyses the cost used for the echocardiography assessment was £115.57. This cost included cost of the echocardiogram and the paediatric cardiologist's time to complete the assessment.[6] We varied the cost of the echocardiography, the rationale being to account for those hospitals without available echocardiography and, thus, an additional cost of travel. The most pragmatic way to assess the impact of changing this cost was to double it.
    ii. *Changing the cost of the pulse oximetry test* In this analysis we changed the cost from £6.24, which was estimated on the assumption that the equipment had a lifespan of 5 years, to a cost of £6.45, assuming that equipment had a lifespan of only 3 years. In this analysis we used the median length of time taken to carry out the pulse oximetry test as opposed to the mean. The median duration for carrying out the test was 5 minutes, which changed the total cost of carrying out the test from £6.24 to £4.68. The mean duration was 6.9 minutes and was used to estimate the cost of pulse oximetry test for the base case. For economic evaluation use of the mean is considered appropriate in order to capture the outliers. As outliers exist, if they are ignored the true costs will be underestimated.[99,100]

Results

Analysis 1

The results of Analysis 1 are presented in *Table 20*. Clinical examination alone was the least costly intervention, with an average cost of the strategy estimated at £614,000 and 91.5 cases detected of clinically significant CHD per 100,000 live births. Pulse oximetry as an adjunct to clinical examination had an expected total cost for the strategy of £1,358,800 and 121.4 cases of CHD detected per 100,000 live births. The additional cost of the strategy compared with clinical examination alone was estimated at £744,700 and approximately 30 additional cases (29.9 cases in *Table 20*) received a timely diagnosis per 100,000 live births. The estimated ICER was calculated as £24,900 per timely diagnosis. This ICER indicates that each additional case of significant CHD timely diagnosed by the strategy of 'pulse oximetry as an adjunct to clinical examination' will cost an additional £24,900 compared with the strategy of clinical examination alone.

For comparison, and in an attempt to validate our model, we used the input costs [pulse oximetry test (£1.65), clinical examination (£1.17) and echocardiographic assessment (£84.17)] reported in the study by Knowles *et al.*[6] in our model. The ICER reported by Knowles *et al.*[6] was £4894 (in 2001 prices). When we use their costs in our model structure, the corresponding ICER for the Analysis 1 is £7800 per timely diagnosis of a CHD. These results are presented in *Appendix 4*, *Table 25*. Factors that have contributed to this difference include the assumption made by Knowles *et al.*[6] that the pulse oximetry test will take the same amount of time as the clinical examination, which were both assumed to be of a much shorter duration (2 minutes) than estimated in our time-and-motion study (see *Table 18*); additionally, our model incorporates a second pulse oximetry test if the first test is positive but the clinical examination is negative. In contrast, the analysis carried out by Knowles *et al.*[6] assumed only one pulse oximetry test that adds to our costs.

Figure 11a shows the Monte Carlo simulation for Analysis 1. For the 10,000 runs of the Monte Carlo simulation, the scatterplot shows little variation in the incremental cost. This is not surprising, as all newborns received the first pulse oximetry test and a very small proportion received any further tests apart from the clinical examination.

The results for Analysis 1 (pulse oximetry as an adjunct to clinical examination vs clinical examination alone) are presented in the form of a cost-effectiveness acceptability curve (CEAC). The CEAC informs us of the probability that a screening strategy is cost-effective at society's willingness to pay (WTP) for a timely diagnosis of a clinically significant CHD. Interpretation of results presented in quality-adjusted life-years (QALYs) is straightforward for decision-makers because the acceptable threshold used by NICE is presented in an outcome of cost per QALY. To interpret this result in the absence of the use of QALYs, we consider the following. The threshold used by NICE is £20,000 per QALY; this means, that below and up to this threshold society is assumed to be willing to pay £20,000 per QALY for 1 year of life in full health. Thus, for society to be willing to pay £100,000 a newborn with timely diagnosis of a CHD would need to gain just five QALYs. The technologies used to treat this condition are new and advancing all the time; high-quality multicentre follow-up studies are lacking, but the indications are that the majority of these children reach early adulthood in good health.[6] From *Figure 11a*, we can see that at the £100,000 WTP threshold, the probability that 'pulse oximetry as an adjunct to clinical examination' would be cost-effective is considerably greater than 90%. So an infant who receives a timely diagnosis as a result of the strategy of 'pulse oximetry as an adjunct to clinical examination' would only have to achieve five QALYs to reach a 90% chance of the strategy being considered cost-effective. However, treatment costs have not been included in this model-based analysis. If we assume that

TABLE 20 Summary of the deterministic results for Analyses 1 and 2 cost per timely diagnosis of a CHD present/100,000 live births (2008–9 prices)

Strategy	Screening strategy	Expected total cost per screening strategy (£)	Difference in costs	Effectiveness CHD detected	Incremental CHD detected	ICER per timely diagnosis of a significant CHD (£)
Analysis 1: base-case results						
A	Clinical examination alone	614,100	–	91.5	–	–
B	Pulse oximetry as an adjunct to clinical examination (Strategy B vs A)	1,358,800	744,700	121.4	29.9	24,900
Analysis 2: changing the population to include babies detected through antenatal ultrasound screening and the addition of the pulse oximetry-alone strategy						
A	Clinical examination alone	621,000	–	151.2	–	–
B	Pulse oximetry alone (Strategy B vs A)	1,068,900	447,900	164.5	13.3	33,600
C	Pulse oximetry as an adjunct to clinical examination (Strategy C vs B)	1,370,200	301,300	216.0	51.5	5900
Exclusion of the pulse oximetry-alone screening strategy due to extended dominance						
A	Clinical examination alone	621,000	–	151.2	–	–
C	Pulse oximetry as an adjunct to clinical examination (Strategy C vs A)	1,370,200	749,200	216.0	64.8	11,600

treatment of a CHD costs on average £50,000, for example, then for an accurate timely diagnosis alone, society's WTP would be £50,000 (£100,000 minus £50,000). If this is the case, then *Figure 11b* shows that the probability that 'pulse oximetry as an adjunct to clinical examination' is cost-effective still remains > 90%.

Analysis 2

The deterministic results of Analysis 2 are also presented in *Table 20*. This analysis is the same as that carried out in Analysis 1, but on a different population. In this analysis, the population includes infants who may have been diagnosed antenatally and also includes the additional strategy of pulse oximetry alone. The deterministic results for Analysis 2 show that the clinical examination-alone strategy is the least costly screening strategy, with an expected total cost of £621,000, and detects 151 cases of CHD per 100,000 live births. The pulse oximetry alone-strategy is shown to be considerably more expensive (£1,068,900), but detects an average of 14 additional cases of CHDs per 100,000 live births compared with clinical examination alone. In contrast, the strategy of 'pulse oximetry as an adjunct to clinical examination' is more expensive again than the pulse oximetry-alone strategy (£1,370,200), but detects 52 additional cases of CHDs per 100,000 live births compared with pulse oximetry alone.

The estimated ICER reported for the comparison of pulse oximetry alone versus clinical examination alone was £33,600 per timely diagnosis. The strategy of 'pulse oximetry as an adjunct to clinical examination' is, more costly and more effective than pulse oximetry alone, with an estimated ICER of £5900. This is an example of what is known as *extended dominance*: at any threshold for which society is willing to pay the extra cost of pulse oximetry alone compared with clinical examination alone, society must also be willing to pay the additional cost of pulse oximetry as an adjunct to clinical examination compared with pulse oximetry alone. Thus, the option of pulse oximetry alone can be excluded and pulse oximetry as an adjunct to clinical examination can be compared directly with clinical examination alone.

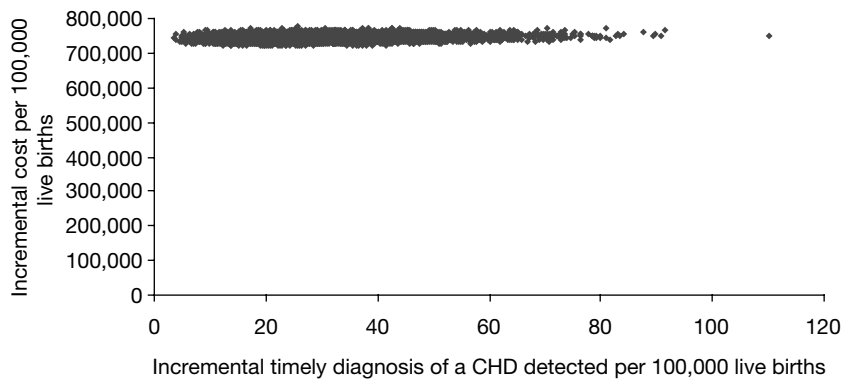## Probability sensitivity analysis results for the base case

FIGURE 11a Scatterplot using distributions around the accuracy data; incremental cost-effectiveness scatterplot of pulse oximetry as an adjunct to clinical examination vs clinical examination alone.
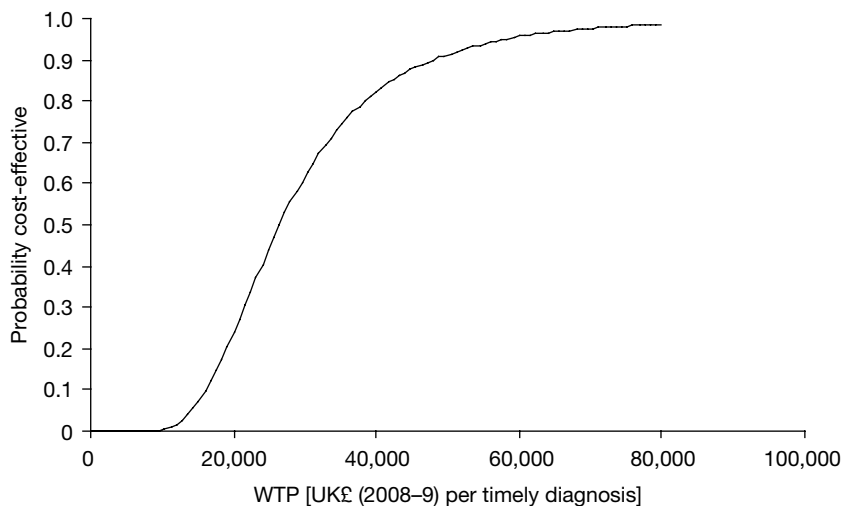

FIGURE 11b Cost-effectiveness acceptability curve using distributions around the accuracy data.

Table 20 presents the deterministic results from Analysis 2 when this comparison is made. Here, 'pulse oximetry as an adjunct to clinical examination' is more costly and more effective than clinical examination alone and has an estimated ICER of £11,600 per case of timely diagnosis of a CHD diagnosed. This result is in line with intuition. Given the much higher prevalence of CHDs in this population (Group 2), a relatively lower ICER for the strategy of pulse oximetry as an adjunct to clinical examination' for this analysis (Analysis 2) compared with Analysis 1 is expected.

Figures 12a and *12b* are analogous to *Figures 11a* and *11b* as presented above. *Figure 12a* presents the results of the Monte Carlo simulation for pulse oximetry as an adjunct to clinical examination compared with clinical examination alone. As before, we observe that there is relatively little variation in the incremental cost. *Figure 12b* presents the CEAC. As stated before, the CEAC informs us of the probability that a screening strategy is cost-effective at any value of society's WTP for a timely diagnosis. In the absence of antenatal ultrasound screening, which is assumed for this analysis, it is clear that the addition of pulse oximetry as an adjunct to routine

Probabilistic sensitivity analysis for pulse oximetry as an adjunct to clinical examination versus clinical examination alone
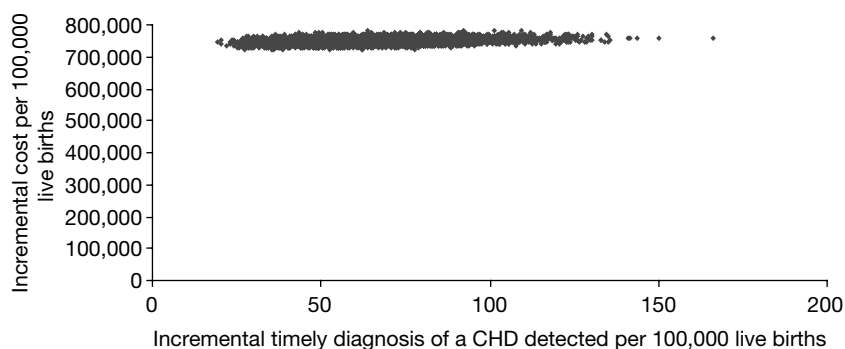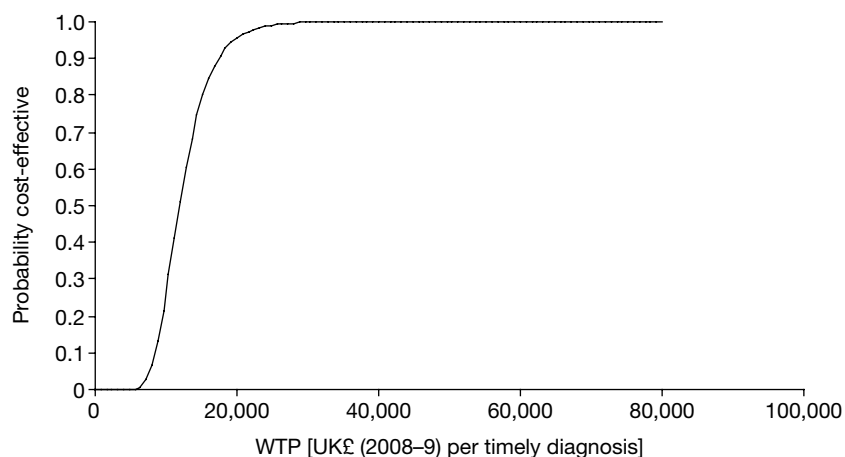

FIGURE 12a Scatterplot using distributions around the accuracy data; incremental cost-effectiveness scatterplot of pulse oximetry as an adjunct to clinical examination vs clinical examination alone.

FIGURE 12b Cost-effectiveness acceptability curve using distributions around the accuracy data.

examination would be considered cost-effective with almost certainty (probability almost = 1), even at WTP thresholds of £20,000 if a newborn infant could achieve just one additional QALY as a result of a timely diagnosis of confirming a CHD.

Sensitivity analyses
Changing the threshold to <95% in either limb or a differential of >3%
In this deterministic sensitivity analysis, the base-case analysis (Analysis 1) was repeated, but using a differential in either limb of >3% as opposed to the differential of >2% that was used in the base case. Before undertaking this sensitivity analysis the population prevalence for each severity category had to be re-estimated using this threshold (or cut-off). With this threshold, in the strategy 'pulse oximetry as an adjunct to clinical examination', more babies would now 'pass' the pulse oximetry test and proceed to routine clinical examination, and fewer babies would proceed to an expedited examination. Thus, the proportion of babies defined as normal, for example, would be expected to change.

However, it must be emphasised that this less conservative threshold was not used in the primary study and so primary data on the diagnosis of the additional babies that proceed to routine clinical examination would not be known with certainty for this cut-off. That is, we do not know from the study whether or not the routine examination would identify them as cases of a suspected CHD and, in particular, what degree of severity, if any, would be identified. Probabilities that exist in the literature refer to babies have not had a pulse oximetry test. Ideally, we want to use the probabilities from the clinical examination strategy and then infer the probabilities that apply to a (preceding) negative pulse oximetry test to give a correct weighted average. But because of the small number of babies in the study who have a positive pulse oximetry test (at this threshold) we do not have enough data to estimate a weighted average that we can use with a high degree of confidence. Nonetheless, the analysis was considered relatively important to explore the impact of a change in the cut-off for the test.

The table summarising the revised proportions is presented in *Appendix 4, Table 23*. The results for this sensitivity analysis are summarised and presented in *Table 21*. The results suggest that pulse oximetry as an adjunct to clinical examination is more costly and more effective than clinical examination alone, with an estimated total cost of £1,311,100, and approximate number of cases detected of 121.4 per 100,000 live births compared with the estimated total cost of £613,700 and approximately 91.5 cases detected per 100,000 live births for the strategy of clinical examination alone. The calculated ICER was approximately £23,300, which suggests that each additional accurate case of timely diagnosis of a clinically significant CHD will cost an additional £23,300. Because of the caveat explained above for this analysis, the different costs and different effects presented in *Table 21* as a result of the clinical examination strategy (at this cut-off) will be different to those presented in the base case and other sensitivity analyses. The implication of the results is that it may be possible to reduce the FP rate without causing any great problems and the ICER does not change dramatically from the base case. But no inference should be made from this sensitivity analysis about what the correct cut-off should be and any screening policy should certainly not be made based on the results of this sensitivity analysis.

Probabilistic sensitivity analysis was carried out for this analysis using different thresholds, but is not reported because both the result and implication of the PSA are very similar to the base case.

Sensitivity analysis using threshold of <95% for foot saturations alone

By changing the test threshold from the difference between hand and foot saturation to a threshold based on foot saturation alone, there will be a difference in each population. The summary presenting the revised proportions is presented in *Appendix 4, Table 24*. The deterministic results for this analysis are also presented in *Table 21*. The total cost for clinical examination alone was estimated at £611,000 compared with a cost for pulse oximetry as an adjunct to clinical examination of £1,278,200. The calculated ICER was £26,700 per case of timely diagnosis of a clinically significant CHD.

One-way deterministic sensitivity analysis (doubling the cost for the echocardiogram assessment)

In this analysis the cost of echocardiography was doubled. The rationale for increasing this cost could be the inclusion of travel costs for infants referred for echocardiography. Some hospitals may not have echocardiography on site and so infants would need to be transported to another site for this investigation. This additional cost is borne by the NHS. The results show some sensitivity to the increased cost of echocardiography but in line with intuition given the increase (see *Table 21*).

TABLE 21 Summary of the deterministic sensitivity analysis results

Strategy	Screening strategy	Expected total cost per screening strategy (£)	Difference in costs	Effectiveness CHD detected	Incremental CHD detected	ICER per timely diagnosis of a significant CHD (£)
Changing the threshold to <95% in either limb or a differential of >3%						
A	Clinical examination alone	613,700	–	91.5	–	–
B	Pulse oximetry as an adjunct to clinical examination (Strategy B vs Strategy A)	1,311,100	697,300	121.4	29.9	23,300
Using a threshold of <95% for foot saturations alone						
A	Clinical examination alone	611,000	–	88.8	–	–
B	Pulse oximetry as an adjunct to clinical examination (Strategy B vs Strategy A)	1,278,200	667,200	113.7	24.9	26,700
Doubling the cost for the echocardiogram assessment (£231.14)						
A	Clinical examination alone	685,200	–	91.5	–	–
B	Pulse oximetry as an adjunct to clinical examination (Strategy B vs Strategy A)	1,508,200	823,000	121.4	29.9	27,500
Changing the cost of the pulse oximetry test						
Using a cost of £6.45 for the pulse oximetry test assuming a lifespan of 3 years (instead of 5 years)						
A	Clinical examination alone	614,100	–	91.5	–	–
B	Pulse oximetry as an adjunct to clinical examination (Strategy B vs Strategy A)	1,380,500	766,400	121.4	29.9	25,600
Using a cost of £4.68 for the pulse oximetry (assuming median duration for carrying out the test as opposed to mean duration)						
A	Clinical examination alone	614,100	–	91.5	–	–
B	Pulse oximetry as an adjunct to clinical examination (Strategy B vs Strategy A)	1,197,300	583,200	121.4	29.9	19,500

One-way deterministic sensitivity analysis (using a cost of £6.45 for pulse oximetry)

(a) In this sensitivity analysis it was assumed that the life of a pulse oximetry machine was 3 years as opposed to 5 years that was used in base case. The resulting ICER shows that this change does not have a dramatic effect on the results (see *Table 21*).

(b) In this sensitivity analysis it was assumed that the cost of the pulse oximetry test was £4.68 estimated using the median length of time taken to carry out the test (5 minutes) as opposed to using the mean duration. This has a favourable effect on the ICER, which is reduced to £19,500, as would be expected. But using the median duration as opposed to the mean duration risks underestimating the true cost of carrying out the test.

Discussion

Principal findings

The results of this analysis suggest that pulse oximetry as an adjunct to clinical examination is twice as costly, but detects almost 30 additional cases of CHD per 100,000 live births compared with clinical examination alone. The ICER of this strategy compared with clinical examination alone is approximately £24,000 per case of timely diagnosis in a population in which antenatal screening for CHDs already exists (Analysis 1). The PSA suggests that at WTP thresholds of £100,000, the probability of 'pulse oximetry as an adjunct to clinical examination' being cost-effective is > 90%. Such a WTP threshold is plausible if a newborn with timely diagnosis of a CHD gained just five QALYs, even when treatment costs are taken into consideration. Additional deterministic sensitivity analyses were carried out, none of which made any impact on the base-case ICER. One of these analyses changed the test cut-off from the base-case value and showed no major impact on the ICER. Careful interpretation of these results is required and this analysis did not allow a cost-effective threshold to be determined.

A separate analysis showed that for a population in which no antenatal screening was assumed, the strategy of pulse oximetry alone was almost twice as costly and detected approximately 15 additional cases of CHD per 100,000 live births compared with a strategy of clinical examination alone, with an ICER of approximately £33,600 per timely diagnosis. However, the analysis also suggested that pulse oximetry as an adjunct to clinical examination would be preferred to pulse oximetry alone, with a corresponding ICER of £5900.

Strengths and limitations of the study

The strength of this model-based economic evaluation is that the accuracy of the strategy of pulse oximetry as an adjunct to clinical examination has been estimated in a primary study. Furthermore, the primary cost data have been collected via a time-and-motion study for carrying out the test and represent the first primary cost data for use of pulse oximetry in the UK setting. In addition, the study represents collaborative work of clinicians, midwives, statisticians and health economists. One weakness is that in the pragmatic design of the study it was not possible to measure the accuracy of 'clinical examination alone', as it was always performed after a negative pulse oximetry test and it was not possible for the clinicians to be blind to the result of the pulse oximetry. However, this limitation has been overcome by using appropriate statistical methods and the accuracy was estimated by using weighted average of the results of the routine clinical examination (which followed a negative pulse oximetry test) and the ECE (which followed a positive pulse oximetry test).

Within the test-positive group there were six babies who had a significant CHD and 40 babies who had respiratory or infective conditions. These were considered FPs, as they did not have a major CHD and the inclusion of these conditions in cost-effectiveness analysis was not included in the original economic assessment protocol and therefore has not been included. However,

the identification of these conditions by pulse oximetry screening is a bonus and likely to reduce distress and health-care costs by identifying them before problems arise.

We did not include the cost of the following in this analysis: (1) training for carrying out the test; and (2) counselling of parents following echocardiography. The reasons were that (1) the procedure is not a lengthy or arduous task and training in the use of new equipment is a regular part of routine for clinical staff so that over time it becomes irrelevant in the overall cost and (2) appropriate counselling is a routine part of echocardiography and is therefore included in the cost.

The costs presented in this analysis are likely to represent the upper limit of carrying out pulse oximetry screening for a number of reasons. First, in our analysis study, the test was always assumed to be carried out by a midwife, although health-care assistants, whose time is relatively less costly, can perform the test. Second, the average cost of the test is influenced by the presence of a few outliers: a test duration of 30 minutes was reported only once, 20 minutes was the next longest duration and, on the whole, most staff took less than 10 minutes to carry out the test and this is likely to improve over time. Finally, our analysis assumes that additional knock-on tests, such as echocardiography, are carried out by a consultant, but a trained technician can also perform these tests. Overall, over time the costs associated with testing are likely to fall.

Strengths and limitations in relation to other studies

Knowles et al.[6] carried out a model-based economic evaluation to assess the cost-effectiveness of pulse oximetry as an adjunct to clinical examination in a UK setting using an outcome of cost per timely diagnosis, based entirely on secondary data for both accuracy and costs. The strength of the current evaluation is that it is based on primary accuracy data in an appropriately powered setting and primary cost and resource data measured through a time-and-motion study. The key difference between the results of the current study and the results of Knowles et al.[6] pertains to the duration of staff time to carry out both pulse oximetry test and the clinical examination, which were both found to be of longer duration than the 2 minutes for each test assumed by Knowles et al.[6] The current study also assumed a second pulse oximetry test in the event of the clinical examination (that followed a negative first pulse oximetry test) being indicative of a CHD. Consequently, the ICERs estimated in the current study are higher than those estimated by Knowles et al.,[6] but this cost is largely accounted for by the different assumptions about cost and resource use.

Meaning of the study

The additional cost of including pulse oximetry as an adjunct to clinical examination to test for CHDs in the newborn was estimated to be approximately £24,900 per additional case of timely diagnosis. This estimate was shown to be robust to extensive sensitivity analysis. The techniques involved in treating this condition are new and so evidence on the life expectancy and quality of life of these infants as they develop through childhood and young adulthood is improving all the time. The early evidence suggests that the majority[1,6] of these infants survive into young adulthood at least, and with good quality of life. If we assume that they achieve as few as just 5 years of full health overall, which could allow for discounting, then the average cost per QALY associated with the timely diagnosis would be approximately £5000. This falls well within the acceptable thresholds used by decision-makers such as NICE. Treatment costs are not included in this ICER. The addition of the treatment cost was not part of the remit of this study, but the addition would obviously make the incremental cost-effectiveness threshold less favourable than that presented. However, the objective of including pulse oximetry as an adjunct to routine examination is to identify babies with potential CHD problems as soon as possible for a timely diagnosis. Treatment costs will necessarily be incurred for the babies identified, but if a CHD was not detected with pulse oximetry in the neonatal period and a CHD became apparent later

then treatment costs would still be necessary. Furthermore, in the absence of a timely diagnosis the risk is that treatment cost are likely to be greater, although survival prospects may risk being poorer. The only clear additional treatment costs associated with a timely diagnosis of a CHD is for those infants who may otherwise have died as a result of an unidentified CHD.

Unanswered questions and future research

The model shows a high relative uncertainty in the effectiveness of pulse oximetry as an adjunct to clinical examination compared with clinical examination alone, but this uncertainty is highly unlikely to affect the decision to use pulse oximetry. Any future research wishing to evaluate the added value of incorporating pulse oximetry to current routine practice of clinical examination would need to carefully consider the ethical dilemma of blinding medical staff to low saturation rates as discussed earlier in this report. Disregarding this issue, estimating the size of sample needed to compare the two tests in combination with clinical examination alone is particularly difficult, as no completely unbiased estimate of clinical examination accuracy exists and estimates vary greatly. A previous HTA evaluation of newborn screening[6] used an estimate of 32% for clinical examination sensitivity, although this was based on a retrospective database review.[8] de Wahl Granelli's study[70] in Sweden estimated this figure to be much higher at 63%, although this figure was estimated from a different population to the remainder of the study and is not necessarily reliable.

Assuming that the true value for clinical examination sensitivity lies somewhere between the above estimates at around 50%, absolute increases in sensitivity of pulse oximetry over clinical examination of 15%, 20% and 25% (assuming independence of the errors for the two tests) would need sample sizes of 90,000, 51,000 or 33,000 to be detected with 90% power at the 5% level, based on the same prevalence rate found in our study for all major CHDs of 2.6 per 1000. These estimates are calculated using the method described by Alonzo *et al.*[101] for sample size calculations in comparative studies of screening tests. The latter estimate of 33,000 (comparing 50% with 75% sensitivity) may seem fairly realistic, considering that the estimate of sensitivity in our study was 75%, although we should point out that the index test used in our study incorporated a targeted examination which may not necessarily reflect the clinical examination used in current practice.

Chapter 5

Discussion

Introduction

This HTA completed three distinct pieces of work to determine:

- the accuracy of pulse oximetry against the reference standards of echocardiography, clinical follow-up and interrogation of regional and national congenital anomaly and cardiac defect databases, and the added value of pulse oximetry over routine antenatal ultrasound screening
- the acceptability of pulse oximetry testing to parents and health-care staff
- the cost-effectiveness of pulse oximetry testing compared with existing strategies for CHD screening.

Each of these has been described in detail, the main findings reported and the conclusions discussed in the light of any limitations identified at the end of each of the three preceding chapters.

This chapter attempts to focus on the key findings and limitations emerging from all of the work undertaken. It is not a comprehensive summary of all of the issues raised, for which the reader is encouraged to consult the previous three chapters.

Congenital heart defects are the commonest group of congenital abnormalities, affecting 6–9 per 1000 live births, and approximately 1–2 per 1000 will have a potentially life-threatening critical defect. CHDs are responsible for more deaths than any other group of abnormalities, either in association with non-cardiac problems or as isolated lesions, and are a leading cause of infant death in the developed world.

Congenital heart defects can be readily diagnosed by echocardiography, yet as a routine screening procedure, echocardiography is likely to be impracticable because of the significant resource implications and the high FP rate.[6]

The current screening for CHDs relies on an antenatal anomaly ultrasound scan in the second trimester and a postnatal physical examination performed shortly after birth and again at 6 weeks of age. Both of these screening procedures are unsatisfactory, and approximately 30–50% of babies with a CHD will be discharged from hospital without the lesion being detected. In the infants with critical lesions, the closure of the ductus arteriosus prior to diagnosis can result in profound circulatory collapse or death, and 15–30% of infants with critical lesions either present in very poor condition or die prior to diagnosis.

The majority of CHDs are amenable to surgical or transcatheter intervention and once the condition is suspected the infant can be stabilised using a prostaglandin infusion to maintain ductal patency prior to intervention. Improvements in surgical techniques have led to a significant increase in survival rates for infants with CHDs, particularly if the diagnosis is made in a timely manner. It is estimated that approximately 80% of infants with CHDs will survive to the age of 16 years, the majority with a good quality of life.

Because a significant proportion of CHDs are associated with varying degrees of hypoxaemia, the use of pulse oximetry (a non-invasive method of measuring blood oxygen saturations) has been explored by a number of research groups using differing methodologies and patient group selection. This project is the first to fully evaluate pulse oximetry in terms of accuracy, feasibility, acceptability and cost-effectiveness, and examines its potential to be introduced as a routine screening following delivery.

Evaluation of pulse oximetry

The accuracy of pulse oximetry testing for detecting major CHDs (using both pre- and postductal measurements with threshold levels of < 95% in either limb and/or a difference of > 2% between the two) was estimated in a primary test accuracy study using echocardiography and clinical and database follow-up as the composite reference standard. A positive result from the index test led to the reference standard of echocardiography and all test-negative patients were followed up to 12 months of age using the composite reference standard of clinical follow-up and interrogation of regional and national clinical databases.

In asymptomatic infants, pulse oximetry screening had a sensitivity of 75.00% (95% CI 53.29% to 90.23%) for critical lesions and a sensitivity of 49.06% (95% CI 35.06 to 63.16%) for critical plus serious lesions. For the cohort in which the test results could affect postnatal management as a CHD had not been suspected prenatally following ultrasound screening, pulse oximetry had a sensitivity of 58.33% (95% CI 27.67% to 84.83%) for critical cases and 28.57% (95% CI 14.64% to 46.30%) for critical plus serious cases combined.

False-positive results occurred in 1 in 119 babies (0.84%) (specificity, 99.16%, 95% CI 99.02% to 99.28%); however, 27% of the test FP cohort had additional problems that required medical intervention (specifically a significant CHD, respiratory disorders or infections).

There have been a number of previous studies investigating the accuracy of pulse oximetry in detecting CHDs (see *Chapter 1*). The majority of studies involved a relatively small number of patients and were underpowered to address test accuracy.[72] Recently, two studies examining test accuracy have been reported from Scandinavia, recruiting large cohorts of 50 000 (in a Norwegian study) and 40,000 subjects (in a Swedish study).[69,70] The studies reported sensitivities of 77.1% (95% CI 59.4% to 89.0%) and 62.07% (95% CI 42.3% to 79.3%) and specificities of 99.4% (95% CI 99.3% to 99.5%) and 99.82% (95% CI 99.77% to 99.86%), respectively. FP rates were 0.6% and 0.17%, respectively. The studies used different pulse oximetry test methods; one used postductal saturations alone[69] and the other used pre- and postductal saturations[70] with similar thresholds to our study, although the pre- and postductal difference threshold was > 3% rather than the > 2% threshold we used. Testing was generally early in the Norwegian study (median age, 6 hours)[69] and later in the Swedish study (median age, 38 hours).[70] The reference standards were also different; the Norwegian study did not perform echocardiography in all test-positive cases and three patients with a major CHD (including two critical cases) who tested positive were sent home without a diagnosis.[69] In addition, the ascertainment of CHD cases dying in the community was incomplete and the relatively low incidence of critical CHDs (0.7 per 1000 live births) may reflect some unidentified missed cases, which could affect the sensitivity analysis.[69] The Swedish study performed echocardiography in all test-positive cases and undertook full ascertainment of CHD cases with a reference standard similar to the one used in our study.[70] The prevalence of critical CHDs in that study was 1.3 per 1000 live births, which is

almost identical to the prevalence we identified in our cohort. Both studies excluded antenatally diagnosed CHDs but in both the number of cases identified by ultrasound screening appeared very low, particularly in the Swedish study, in which the antenatal detection rate of critical CHDs was only 3%[70] compared with 50% in our cohort.

Both studies failed to detect some cases of critical CHDs and, in keeping with our findings, the majority of critical cases that were missed were those with aortic arch obstruction.

The FP rate in the Norwegian study was similar to that noted in our study, and the lower FP rate in de Wahl Granelli *et al.*'s study[66] is likely to reflect the much later timing of testing. Both studies report the additional diagnoses of respiratory problems and infections in the FP cohorts.

Neither study reported acceptability, and although there were some comments on cost–benefit in the Swedish study, this issue was not addressed in a full cost-effectiveness analysis as described in this report.

Our study, therefore, reports a sensitivity that is slightly higher than that of the Swedish study and slightly lower than that of the Norwegian study but with better ascertainment. In addition, our reporting of acceptability and cost-effectiveness gives a more complete understanding regarding the implications of undertaking pulse oximetry as a potential screening policy. CHDs and pulse oximetry screening for the condition fulfil all of the programme appraisal criteria set out by the National Screening Committee (*Table 22* and *Appendix 6*).

Across all parts of the study, parent and staff participants were predominantly satisfied with pulse oximetry screening, perceiving it to be an important and valuable test. However, differences between ethnic groups were identified, with white British and Irish mothers being more likely to participate in the study, and reporting lower anxiety and higher satisfaction with the screening process, than mothers of other ethnic groups. It would be useful to identify the factors leading to these effects, to ensure that women receive appropriate support in making decisions about testing. There was no evidence that mothers given FP results were more anxious after taking part in the screening processes than those given TN results, although the former were less satisfied with the test and gave higher depression scores. Satisfaction with screening was predicted by higher perception of the treatment's ability to control heart disease, comprehensibility of heart disease, and lower stress, anxiety and depression. Focusing information on heart disease itself may increase parents' understanding of the condition and increase satisfaction with screening. Participants given FP results recalled the testing process as more stressful, probably reflecting the need for a second test and communication about the results.

Despite creating an additional workload in testing babies, the test was seen as reassuring for staff, and had positive impact on staff roles where ill babies were detected and treated while in a less critical condition. Communication problems were indicated as a cause of worry by participants, and staff identified a need for further training in communicating with parents about the study and for giving results, especially where a positive result is found.

The research process was identified as a major factor impacting on staff time for conducting testing, and was also perceived as a barrier to parental participation, with mothers being given a large amount of information to read at a time not long after having given birth. Uptake of testing was observed by focus group members to increase when the study was completed and screening was continued as standard care in those hospitals.

TABLE 22 National Screening Committee programme appraisal criteria for this study

Item	Criteria	Source
The condition		
1	Major CHD is an important health problem. CHD is one of the leading causes of infant death in the developed world and accounts for up to 40% of all deaths from congenital malformations	*Executive summary* and *Chapter 1*
2	The epidemiology and natural history of CHDs is well understood. Babies with critical CHDs are usually asymptomatic at birth and develop symptoms only once the ductus arteriosus closes; during this period most infants will have a degree of hypoxaemia, which is clinically undetectable	*Chapter 1*
3	Primary prevention of CHDs is not possible	*Chapter 1*
4	N/A	
The test		
5	Pulse oximetry is a simple, safe, validated and reasonably precise test	*Abstract* and *Chapter 2*
6	The distribution of test values of pulse oximetry are defined and a suitable threshold has been described	*Chapter 2*
7	Pulse oximetry is acceptable to parents and to health-care staff	*Chapter 3*
8	The further diagnostic intervention following a positive test with pulse oximetry is diagnostic echocardiography	*Chapter 2*
9	N/A	
The treatment		
10	The treatments available for major CHDs include cardiac catheter and surgery. Both are shown to have a significant impact on outcome, with early treatment having a better outcome than later	*Chapter 1*
11	There are agreed evidence-based policies concerning the treatment	*Chapter 1*
12	Antenatal screening detects a proportion of major CHDs before birth. Early screening after birth will identify more cases allowing optimisation of care prior to definitive treatment	*Chapter 1*
The screening programme		
13 and 14	This report provides evidence of the effectiveness and acceptability of pulse oximetry screening	*Chapters 2* and *3*
15	Pulse oximetry does not appear to be associated with harm, either physical or psychological	*Chapters 2* and *3*
16	The cost-effectiveness of pulse oximetry screening is described in detail in this report	*Chapter 4*
17	Other options for detecting CHDs are described and are currently unsatisfactory	*Chapter 1*
18	This report demonstrates how a pulse oximetry screening programme could be managed	*Chapter 2*
19	This report demonstrates that pulse oximetry screening can be achieved within existing staffing levels within existing facilities	*Chapter 2*
20	Evidence-based information can be provided for pulse oximetry screening	*Chapters 1* and *2*
21	The strengths and limitations of the test are known and defined within this report	*Chapters 2*, *3* and *5*

N/A, not applicable.

Strengths and limitations

The robust design and execution of our test accuracy study allows us to be confident that the estimates of accuracy are valid. The study population and recruitment criteria were well defined and recruitment of consecutive eligible subjects was representative of a spectrum of maternity activity. A sample size calculation was performed to ensure that the study had sufficient power to exclude clinically unacceptable accuracy and the study recruited to that target. The rationale for the index test was clarified and the index test was performed by trained staff and the initial reference standard of echocardiography was performed by independent, trained individuals, and the robustness of the data has been facilitated by the rigorous follow-up to 1 year of age of all recruited babies.

This study used a composite reference standard of echocardiography and database interrogation, therefore not all babies entered into the study underwent the gold standard reference of echocardiography; however, we are confident that the database follow-up was sufficiently rigorous to account for all potential FNs.

We included those infants who had been suspected of having a CHD after antenatal ultrasound; this was to increase the prevalence of CHDs in our study cohort and to ease recruitment. We tested those infants who fulfilled the eligibility criteria in order to ascertain if pulse oximetry would have identified them had the condition not been suspected, and we undertook further sensitivity analysis after excluding these patients.

This study is the first to address the acceptability of pulse oximetry screening from the perspectives of both parents and health-care professionals. This has allowed a convergence of opinions to be identified on the benefits and problems associated with screening, strengthening the findings. Factors have been identified that may enable more parents to understand and accept screening in the future.

There was no follow-up TN group for comparison with the FP follow-up group. However, as there was no difference in anxiety at baseline between TN and FP participants, and anxiety did not change over time for FP participants, it is unlikely that anxiety was elevated at follow-up for FP participants compared with TN participants.

The time between birth and questionnaire may have affected emotional state or satisfaction with screening, but this variable was heavily confounded by the test result group: TN participants received the questionnaire on discharge, whereas FP participants often received the questionnaire by post.

Across all parts of the study, parent and staff participants were predominantly satisfied with pulse oximetry screening, perceiving it to be an important and valued test to detect ill babies. Despite creating an additional workload in testing babies, the test was seen as reassuring for staff, and positively impacted upon the roles of those caring for babies, as they could be identified and treated before the condition deteriorated. There was no evidence that mothers given FP results were more anxious after taking part in the screening.

Cost-effectiveness of pulse oximetry testing

Having determined the accuracy of pulse oximetry testing and examined the acceptability of such testing to parents and health-care staff, the study then determined the cost of the testing and the cost-effectiveness of the use of pulse oximetry as an adjunct to routine clinical examination.

A decision-analytic model was used, taking the perspective of the UK NHS and the primary outcome of costs per timely diagnosis. The strategies considered were routine practice (clinical examination only) and pulse oximetry as an adjunct to routine practice. These were then compared for populations in which antenatal screening is similar to that found in our study and a population in which antenatal screening was unavailable.

The results of this analysis suggest that pulse oximetry as an adjunct to clinical examination is twice as costly as clinical examination alone, but detects almost 30 additional cases of critical CHD per 100,000 live births. The ICER of this strategy compared with clinical examination alone is approximately £24,900 per case of timely diagnosis in a population in which antenatal screening for CHDs already exists (Analysis 1). The PSA suggests that at WTP thresholds of

£100,000, the probability of 'pulse oximetry as an adjunct to clinical examination' being cost-effective is >90%. Such a WTP threshold is plausible if a newborn with timely diagnosis of a CHD gained just five QALYs, even when treatment costs are taken into consideration. Additional deterministic sensitivity analyses were carried out, none of which made any impact on the base-case ICER. One of these analyses changed the test cut-off from the base-case value and showed no major impact on the ICER. Careful interpretation of these results is required, and this analysis did not allow a cost-effective threshold to be determined.

A separate analysis showed that for a population in which no antenatal screening was assumed, the strategy of pulse oximetry alone was almost twice as costly, compared with a strategy of clinical examination alone and detected approximately 15 additional cases of critical CHD per 100,000 live births with an ICER of approximately £33,600 per timely diagnosis. However, the analysis also suggested that pulse oximetry as an adjunct to clinical examination would be the preferred strategy to pulse oximetry alone with a corresponding ICER of £5900.

Implications for practice

Pulse oximetry screening for CHDs is simple, safe, feasible and acceptable to both staff and parents. It is reasonably accurate, particularly for life-threatening critical lesions, but it is not sensitive enough to perform as an independent screen and some lesions will not be detected consistently by this method, particularly those causing obstruction to the aortic arch. When combined with the routine screening procedures of antenatal ultrasound and clinical examination, pulse oximetry appears to offer added value in terms of detecting additional cases of critical CHDs in a timely manner. The use of pre- and postductal saturation measurements is likely to offer greater sensitivity than postductal alone and does not take significantly longer to perform. The optimum timing and thresholds for testing need to be decided based on acceptable FP rates and availability of clinical staff to perform echocardiography. Earlier testing may also detect non-cardiac conditions that may be potentially serious and is compatible with the increasing trend to earlier discharge from hospital in the UK; this has to be balanced against the risk of more FP results in healthy babies. It may be more practicable to perform the first test early and repeat the test on more than one occasion if the clinical examination is normal and the saturations are improving. This may prove more cost-effective than performing echocardiography after two positive saturation tests.

The results of our economic analysis demonstrate that the additional cost of using pulse oximetry as an adjunct to current practice is likely to be cost-effective, particularly if the outcomes of neonatal cardiac surgery continue to improve.

Education of staff and particularly parents about CHDs and the screening test is also essential. Parents should be given appropriate and understandable information that describes the nature of CHDs and the presenting features that should alert them to the possibility of the diagnosis. The screening test should also be described, particularly the facts that a negative screening test does not completely exclude the possibility of a major CHD and that a positive test is more likely to result in a FP than a TP.

Recommendations for research

Pulse oximetry improves detection rates of critical CHDs, but does not detect all cases. The majority of critical cases missed by pulse oximetry (and by other screening methods) are associated with obstruction of the aortic arch, as these conditions are often not associated with

hypoxaemia. Further investigation of other oximetry techniques, such as perfusion index, which reflects pulse volume, may enhance the detection rates for these lesions.

Further research should be conducted with mothers of different ethnicities to gain a greater understanding of factors limiting participation and satisfaction with testing. Research should also be conducted with mothers given FP results at the time of discharge from hospital to determine whether or not they experience heightened anxiety at this point.

There is also a clear need for high-quality, multicentre or national data relating to outcome and quality of life for survivors of neonatal cardiac surgery for critical CHDs. This would allow the value for money of any future screening test for these conditions to be evaluated more precisely.

Acknowledgements

Contribution of authors

Andrew K Ewer (Senior Lecturer in Neonatal Paediatrics and Consultant Neonatologist) designed and managed the project, prepared *Chapters 1*, *2* and *5*, and edited the final report.

Lee J Middleton (Statistician) performed the statistical analysis for the test accuracy study and edited *Chapter 2*.

Jon J Deeks (Professor of Health Statistics) designed the test accuracy study and oversaw the analyses of test accuracy.

Jane P Daniels (Deputy Director of Birmingham Clinical Trials Unit) provided management direction, supervising the study co-ordination by Alexandra T Furmston (Trial Co-ordinator).

References

1. Lloyd-Jones D, Adams R, Carnethon M, De Simone G, Ferguson TB, Flegal K, *et al. Heart disease and stroke statistics* 2009 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 2009;**119**:e21–181.

2. Wren C, Reinhardt Z, Khawaja K. Twenty-year trends in diagnosis of life-threatening neonatal cardiovascuar malformations. *Arch Dis Child Fetal Neonatal Ed* 2008;**93**:F33–5.

3. Botto LD, Correa A, Erickson D. Racial and temporal variations in the prevalence of heart defects. *Pediatrics* 2001;**107**:e32.

4. Death Registrations in England and Wales, 2002: causes. *Health Stat Q* 2003;**18**:57–64.

5. Boneva RS, Botton LD, Moore CA, Yang Q, Correa A, Erickson D. Mortality associated with congenital heart defects in the United States: trends and racial disparaties 1979–1997. *Circulation* 2001;**103**:2376–81.

6. Knowles R, Griebsch I, Dezateux C, Brown J, Bull C, Wren C. Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2005;**9**(44).

7. Hoffman JIE, Kaplan S. The incidence of congenital heart disease. *J Am Coll Cardiol* 2002;**39**:1890–900.

8. Wren C, Richmond S, Donaldson L. Temporal variability in birth prevalence of cardiovascular malformations. *Heart* 2000;**83**:414–19.

9. Tikanoja T. Effect of technical development on the apparent incidence of congenital heart disease. *Pediatric Cardiol* 1995;**16**:100–1.

10. Mahle WT, Newburger JW, Matherne GP, Smith FC, Hoke TR, Koppel R, *et al.* Role of pulse oximetry in examining newborns for congenital heart disease: a scientific statement from the AHA and AAP. *Pediatrics* 2009;**124**:823–36.

11. Mitchell SC, Korones SB, Berendes HW. Congenital heart disease in 56,109 births incidence and natural history. *Circulation* 1971;**43**:323–32.

12. Barrington KJ. Neonatal screening for life threatening congenital heart disease. *BMJ* 2009;**338**:117–18.

13. Richmond S, Reay G, Abu Harb M. Routine pulse oximetry in the asymptomatic newborn. *Arch Dis Child Fetal Neonatal Ed* 2002;**87**:F83–8.

14. Samánek M, Goetzová J, Benesová D. Causes of death in neonates born with a heart malformation. *Int J Cardiol* 1986;**11**:63–74.

15. Brown KL, Ridout DA, Hoskote A, Verhulst L, Ricci M, Bull C. Delayed diagnosis of congenital heart disease worsens preoperative condition and outcome of surgery in neonates. *Heart* 2006;**92**:1298–302.

16. Brown JW, Park HJ, Turrentine MW. Arterial switch operation: factors impacting survival in the current era. *Ann Thorac Surg* 2001;**71**:1978–84.

17. Franklin O, Burch M, Manning N, Sleeman K, Gould S, Archer N. Prenatal diagnosis of coarctation of the aorta improves survival and reduces morbidity. *Heart* 2002;**87**:67–9.

18. Tworetsky W, McElhinney DB, Reddy VM, Brook MM, Hanley FL, Silverman NH. Improved surgical outcome after fetal diagnosis of hypoplastic left heart syndrome. *Circulation* 2001;**103**:1269–73.

19. Abu-Harb M, Hey E, Wren C. Death in infancy from unrecognised congenital heart disease. *Arch Dis Child* 1994;**71**:3–7.

20. Mellander M, Sunnegardh J. Failure to diagnose critical heart malformations in newborns before discharge: an increasing problem? *Acta Paediatrica* 2006;**95**:407–13.

21. Abu-Harb M, Wyllie J, Hey E, Richmond S, Wren C. Presentations of obstructive left heart malformations in infancy. *Arch Dis Child* 1994;**71**:F179–83.

22. Kuehl KS, Loffredo CA, Ferencz C. Failure to diagnose congenital heart disease in infancy. *Pediatrics* 1999;**103**:743–7.

23. Chang R-KR, Gurvitz M, Rodriguez S. Missed diagnosis of critical congenital heart disease. *Arch Pediatr Adolesc Med* 2008;**162**:969–74.

24. Hall DMB, Elliman D. *Health for all children*. 4th edn. Oxford: Oxford University Press; 2003.

25. Hall DMB. The role of the routine neonatal examination. *BMJ* 1999;**318**:619–20.

26. Wren C, Richmond S, Donaldson L. Presentation of congenital heart disease in infancy: implications for routine examination. *Arch Dis Child Fetal Neonatal Ed* 1999;**80**:F49–53.

27. O'Donnell CPF, Kamlin COF, Davis PG, Carlin JB, Morley CJ. Clinical assessment of infant colour at delivery. *Arch Dis Child Fetal Neonatal Ed* 2007;**92**:F465–7.

28. Ainsworth SB, Wyllie JP, Wren C. Prevalence and clinical significance of cardiac murmurs in neonates. *Arch Dis Child Fetal Neonatal Ed* 1999;**80**:F43–5.

29. Marbella AM, Chetty VK, Layde PM. Neonatal hospital lengths of stay, readmissions, and charges. *Pediatrics* 1998;**101**:32–6.

30. Kotagal UR, Atherton HD, Eshett R, Schoettker PJ, Perlstein PH. Safety of early discharge for Medicaid newborns. *JAMA* 1999;**282**:1150–6.

31. Wen SW, Liu S, Marcoux S, Fowler D. Trends and variations in length of hospital stay for childbirth in Canada. *Can Med Assoc J* 1998;**158**:875–80.

32. Odlind V, Haglund B, Pakkanen M, Otterblad Olausson P. Deliveries, mothers and newborn infants in Sweden, 1973–2000. Trends in obstetrics as reported to the Swedish Medical Birth Register. *Acta Obstet Gynecol Scand* 2003;**82**:516–28.

33. Bull C. Current and potential impact of fetal diagnosis on prevalence and spectrum of serious congenital heart disease at term in the UK. *Lancet* 1999;**354**:1242–7.

34. Allan LD, Crawford DC, Chita SK, Tynan MJ. Prenatal screening for congenital heart disease. *BMJ* 1986;**292**:1717–19.

35. Sharland G. Fetal cardiac screening; why bother? *Arch Dis Child* 2009;**95**:F64–8.

36. UK National Screening Committee. *Antenatal ultrasound screening ultrasound survey of England 2002*. London: UK National Screening Committee; April 2005.

37. Royal College of Obstetrics and Gynaecology (RCOG). *Ultrasound screening: supplement to ultrasound screening for fetal abnormalities*. London: RCOG; 2000.

38. National Institute for Health and Clinical Excellence (NICE). *Antenatal care: routine care for the healthy pregnant woman*. Clinical Guideline, March 2008. London: NICE; 2008.

39. Khoshnood B, De Vigan C, Vodovar V, Goujard J, Lhomme A, Bonnet D, *et al.* Trends in prenatal diagnosis, pregnancy termination, and perinatal mortality of newborns with

congenital heart disease in France, 1983–2000: a population-based evaluation. *Pediatrics* 2005;**115**:95–101.

40. Bonnet D, Cotri A, Butera G, Fermont L, Le Bidois J, Kachaner J, *et al.* Detection of transposition of the great arteries in fetuses reduces neonatal morbidity and mortality. *Circulation* 1999;**99**:916–18.

41. Verheijen PM, Lisowski LA, Stoutenbeek P, Hitchcock JF, Brenner JI, Copel JA, *et al.* Prenatal diagnosis of congenital heart disease affects preoperative acidosis in the newborn patient. *J Thorac Cardiovasc Surg* 2001;**121**:798–803.

42. Carvalho JS, Mavrides E, Shinebourne EA. Improving the effectiveness of routine prenatal screening for major congenital heart defects. *Heart* 2002;**88**:387–91.

43. Tegnander E, Williams W, Johansen OJ, Blaas H-GK, Eik-Nes SH. Prenatal detection of heart defects in a non-selected population of 30 149 fetuses: detection rates and outcomes. *Ultrasound Obstet Gynecol* 2006;**27**:252–65.

44. Chew C, Halliday JL, Riley MM, Penny DJ. Population-based study of antenatal detection of congenital heart disease by ultrasound examination. *Ultrasound Obstet Gynecol* 2007;**29**:619–24.

45. Garne E, Stoll C, Clementi M, and the Euroscan Group. Evaluation of prenatal diagnosis if congenital heart doseases by ultrasound: experience from 20 European registries. *Ultrasound Obstet Gynecol* 2001;**17**:386–91.

46. Westin M, Saltvedt S, Bergman G, Kublickas M, Almström H, Grunewald C, *et al.* Routine ultrasound examination at 12 or 18 gestational weeks for prenatal detection of major congenital heart malformations? A randomised controlled trial comprising 36 299 fetuses. *BJOG* 2006;**113**:675–82.

47. Jaeggi ET, Sholler GF, Jones ODH, Cooper SG. Comparative analysis of pattern, management and outcome of pre- versus postnatally diagnosed major congenital heart disease: a population-based study. *Ultrasound Obstet Gynecol* 2001;**17**:380–5.

48. Bricker L, Garda J, Henderson J, Mugford M, Neilson J, Roberts T, *et al.* Ultrasound screening pregnancy: a systematic review of the clincial effectiveness, cost effectiveness and women's views. *Health Technol Assess* 2000;**4**(16).

49. Acharya G, Sitras V, Maltau JM, Dahl LB, Kaaresen PI, Hanssen TA, *et al.* Major congenital heart disease in Northern Norway: shortcomings of pre- and postnatal diagnosis. *Acta Obstet Gynecol Scand* 2004;**83**:1124–9.

50. Kirwan D, and the NHS Fetal Anomaly Screening Programme (NHS FASP). *NHS Fetal Anomaly Screening Programme. 18+0 to 20+6 weeks fetal anomaly scan standards and Guidance for England.* NHS Fetal Anomaly Screening Programme; January 2010.

51. Rasiah SV, Publicover M, Ewer AK, Khan KS, Kilby MD, Zamora J. A systematic review of the accuracy of first-trimester ultrasound examination for detecting major congenital heart disease. *Ultrasound Obstet Gynecol* 2006;**28**:110–16.

52. Sharland G. Changing impact of fetal diagnosis of congenital heart disease. *Arch Dis Child Fetal Neonatal Ed* 1997;**77**:F1–3.

53. Sharland G. Fetal cardiac screening: why bother? *Arch Dis Child Fetal Neonatal Ed* 2010;**95**:F64–8.

54. Randall P, Brealey S, Hahn S, Khan KS, Parsons JM. Accuracy of fetal echocardiography in the routine detection of congenital heart disease among unselected and low risk populations: a systematic review. *BJOG* 2005;**112**:24–30.

55. Sharland G. Routine fetal cardiac screening: what are we doing and what should we do? *Prenat Diagn* 2004;**24**:1123–9.

56. Mahle WT. Physical examination and pulse oximetry in newborn infants: out with the old, in with the new? *J Pediatrics* 2008;**152**:747–8.

57. Valmari P. Should pulse oximetry be used to screen for congenital heart disease? *Arch Dis Child Fetal Neonatal Ed* 2007;**92**:F219–24.

58. Poets CF, Southall DP. Noninvasive monitoring of oxygenation in infants and children: practical considerations and areas of concern. *Pediatrics* 1994;**93**:737–46.

59. Seelbach-Göbel B, Heupel M, Kühnert M, Butterwegge M. The prediction of fetal acidosis by means of intrapartum fetal pulse oximetry. *Am J Obstet Gynecol* 1999;**180**:73–81.

60. Toth B, Becker A, Seelbach-Gobel B. Oxygen saturation in healthy newborn infants immediately after birth measured by pulse oximetry. *Arch Gynecol Obstet* 2002;**266**:105–7.

61. Levesque BM, Pollack P, Griffin BE, Nielsen HC. Pulse oximetry: what's normal in the newborn nursery? *Pediatr Pulmonol* 2000;**30**:406–12.

62. Bakr AF, Habib HS. Combining pulse oximetry and clinical examination in screening for congenital heart disease. *Pediatric Cardiol* 2005;**26**:832–5.

63. Hoke TR, Donohue PK, Bawa PK, Mitchell RD, Pathak A, Rowe PC, *et al.* Oxygen saturation as a screening test for critical congenital heart disease: a preliminary study. *Pediatric Cardiol* 2002;**23**:403–9.

64. Koppel RI, Druschel C, Carter T, Goldberg B, Mehta P, Talwar R, *et al.* Effectiveness of pulse oximetry screening for congenital heart disease in asymptomatic newborns. *Pediatrics* 2003;**111**:451–5.

65. Reich J, Miller S, Brogdon B, Casatelli J, Gompf T, Huhta J, *et al.* The use of pulse oximetry to detect congenital heart disease. *J Pediatrics* 2003;**142**:268–72.

66. de Wahl Granelli A, Mellander M, Sunnegardh J, Sandberg K, Ostman-Smith I. Screening for duct-dependant congenital heart disease with pulse oximetry: a critical evaluation of strategies to maximize sensitivity. *Acta Paediatr* 2005;**94**:1590–6.

67. Rosati E, Chitano G, Dipaola L, De Felice C, Latini G. Indications and limitations for a neonatal pulse oximetry screening of critical congenital heart disease. *J Perinat Med* 2005;**33**:455–7.

68. Sendelbach DM, Jackson GL, Lai SS, Fixler DE, Stehel EK, Engle WD. Pulse oximetry screening at 4 hours of age to detect critical congenital heart defects. *Pediatrics* 2008;**122**:e815–20.

69. Meberg A, Brugmann-Pieper S, Reidar D, Jr, Eskedal L, Fagerli I, Farstad T, *et al.* First day of life pulse oximetry screening to detect congenital heart defects. *J Pediatr* 2008;**152**:761–5.

70. de Wahl Granelli A, Wennergren M, Sandberg K, Mellander M, Bejlum C, Inganäs L, *et al.* Impact of pulse oximetry screening on detection of duct dependent congenital heart disease: a swedish prospective screening study in 39 821 newborns. *BMJ* 2009;**338**:A3037.

71. Arlettaz R, Bauschatz A, Mankhoff M, Essers B, Bauersfeld U. The contribution of pulse oximetry to the early detection of congenital heart disease in newborns. *Eur J Pediatrics* 2006;**165**:94–8.

72. Thangaratinam S, Daniels JP, Ewer AK, Zamora J, Khan KS. The accuracy of pulse oximetry in screening for congenital heart disease in asymptomatic newborns: a systematic review. *Arch Dis Child* 2006;**58**:137–41.

73. Green J, Statham H. Testing for fetal abnormality in routine antenatal care. *Midwifery* 1993;**9**:124–35.

74. Pollitt RJ, Green A, McCabe CJ, Booth A, Cooper NJ, Leonard JV. Neonatal screening for inborn errors of metabolism: cost, yield and outcome. *Health Technol Assess* 1997;**1**(7).

75. McCrindle BW, Shaffer KM, Kan JS, Zahka KG, Rowe SA, Kidd L. An evaluation of parental concerns and misperceptions about heart murmurs. *Clin Pediatr* 1995;**34**:25–31.

76. Fyro K, Bodegard G. Four-year follow-up of psychological reactions to false positive screening tests for congenital hypothyroidism. *Acta Paediatr Scand* 1987;**76**:107–14.

77. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;**326**:41–4.

78. Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross-sectional study. *J Clin Epidemiol* 2003;**56**:1118–28.

79. Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross sectional study. In Knottnerus JA, editor. *The evidence base of clinical diagnostics*. London: BMJ Books; 2002. pp. 39–59.

80. Daniels JP, Gray J, Pattison H, Roberts T, Edwards E, Milner P, *et al.* Rapid testing for group B streptococcus during labour: a test accuracy study with evaluation of acceptability and cost-effectiveness. *Health Technology Assess* 2009;**13**(42).

81. Clopper C, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 1934;**26**:404–13.

82. Baillie C, Smith J, Hewison J, Mason G. Ultrasound screening for chromosomal abnormality: women's reactions to false positive results. *Br J Health Psychol* 2000;**5**:377–94.

83. Rowe RE, Garcia J, Davidson LL. Social and ethnic inequalities in the offer and uptake of prenatal screening and diagnosis in the UK: a systematic review. *Publ Health* 2004;**118**:177–89.

84. Ware JEJ, Snyder MK, Wright WR. Defining and measuring patient satisfaction with medical care. *Eval Program Plann* 1983;**6**:247–63.

85. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. *Soc Sci Med* 1997;**45**:1829–43.

86. Spielberger C. *Manual for the state trait anxiety inventory STAI (form Y)*. Palo Alto, CA: Consulting Psychologists Press; 1983.

87. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br J Clin Psychol* 1992;**31**:301–6.

88. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983;**67**:361–70.

89. Michie S, Bobrow M, Marteau TM. Predictive genetic testing in children and adults: a study of emotional impact. *J Med Genet* 2001;**38**:519–26.

90. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res* 2006;**60**:631–7.

91. Scheier MF, Carver CS. Optimism, coping and health: assessment and implications of generalized outcome expectancies. *Health Psychol* 1985;**4**:219–47.

92.   Johnston M, Wright S, Weinman J. *Measures in health psychology: a user's portfolio*. Windsor: NFER-NELSON Publishing Company; 1995.

93.   Tabachnick BG, Fidell LS. *Using multivariate statistics*. Boston, MA: Pearson; Allyn and Bacon; 2007.

94.   Ritchie R, Spencer L. Qualitative data analysis for applied policy research. In Bryman A, Burgess RG, editors. *Analyzing qualitative data*. London: Routledge; 1994.

95.   Hunink M, Glasziou P. *Decision making in health and medicine, integrating evidence and values*. Cambridge: Cambridge University Press; 2001.

96.   Curtis L. *Unit costs of health and social care*. London: Personal Social Services Research Unit; 2009.

97.   Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. 3rd edn. Oxford: Oxford Univesity Press; 2005.

98.   Briggs A, Claxton K, Sculpher M. *Decision modelling for health economic evaluation*. Oxford: Oxford University Press; 2006.

99.   Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ* 1986;**5**:1–30.

100.   Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *BMJ* 2000;**320**:29.

101.   Alonzo TA, Pepe MS, Moskowitz CS. Sample size calculations for comparative studies of medical tests for detecting presence of disease. *Stat Med* 2002;**21**:835–52.

Appendix 1

Standards for the Reporting of Diagnostic Accuracy Studies checklist for the reporting of studies of diagnostic accuracy studies

Section and topic	Item		Source
Title/abstract/key words	1	Identify the article as a study of diagnostic accuracy	Prelims
Introduction	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups	Executive summary and *Chapter 1*
Methods			
Participants	3	Describe the study population: the inclusion and exclusion criteria, setting and locations where the data were collected	*Chapter 2*
	4	Describe participant recruitment: was recruitment based on presenting symptoms, results from previous tests or the fact that the participants had received the index tests or the reference standard?	*Chapter 2*
	5	Describe participant sampling: was the study population a consecutive series of participants defined by the selection criteria in items 3 and 4? If not, specify how participants were further selected	*Chapter 2* (consecutive)
	6	Data collection: was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	*Chapter 2* (prospective)
Test methods	7	Describe the reference standard and its rationale	*Chapter 2*
	8	Describe technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard	*Chapter 2*
	9	Describe definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard	*Chapter 2* and *Table 1*
	10	Describe the number, training and expertise of the persons executing and reading the index tests and the reference standard	*Chapter 2*
	11	Describe whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers	*Chapter 2* (not masked)
Statistical methods	12	Describe methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% CIs)	*Chapter 2*
	13	Describe methods for calculating test reproducibility, if done	Not done
Results			
Participants	14	Report when study was performed, including beginning and end dates of recruitment	*Chapter 2*
	15	Report clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms)	*Chapter 2*
	16	Report the number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended)	*Chapter 2* and *Figure 3*

Section and topic	Item		Source
Test results	17	Report time interval between the index tests and the reference standard, and any treatment administered in between	*Chapter 2*
	18	Report distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition	*Table 1* and *Figure 3*
	19	Report a cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard	*Figure 4*
	20	Report any adverse events from performing the index tests or the reference standard	None
Estimates	21	Report estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% CIs)	*Chapter 2*
	22	Report how indeterminate results, missing data and outliers of the index tests were handled	*Chapter 2*
	23	Report estimates of variability of diagnostic accuracy between subgroups of participants, readers or centres, if done	*Chapter 2*
	24	Report estimates of test reproducibility, if done	Not done
Discussion	25	Discuss the clinical applicability of the study findings	*Chapter 2*

Version: January 2003.

Appendix 2

Clinical details of babies with major congenital heart defects

Lesion	Timing of test (hours)	Antenatal diagnosis	Pulse oximetry		Examination	Pulse oximetry 2		CHD category	Result
			Hand	Foot		Hand	Foot		
TGA	3–6	Yes	81	85	Abnormal	–	–	Critical	TP
TAPVD	12–24	No	73	77	Abnormal	–	–	Critical	TP
CoA, hypoplastic AA, VSD	0–3	Yes	98	72	Abnormal	–	–	Critical	TP
PA, AVSD, TGA	3–6	Yes	83	79	Normal	84	74	Critical	TP
HLHS	0–3	Yes	100	96	Abnormal	–	–	Critical	TP
TGA	3–6	No	53	56	Abnormal	–	–	Critical	TP
HLHS	0–3	Yes	88	86	Normal	90	90	Critical	TP
CoA	12–24	No	95	84	Abnormal	–	–	Critical	TP
PA, DILV	0–3	Yes	93	94	Normal	93	84	Critical	TP
HLHS	0–3	Yes	92	97	Normal	100	97	Critical	TP
HLHS	0–3	Yes	91	84	Abnormal	–	–	Critical	TP
HLHS	0–3	Yes	94	94	Abnormal	–	–	Critical	TP
TGA	3–6	No	79	73	Normal	69	79	Critical	TP
TGA, VSD	3–6	Yes	83	91	Abnormal	–	–	Critical	TP
TGA	24+	No	96	91	Abnormal	–	–	Critical	TP
TGA	6–12	No	83	74	Abnormal	–	–	Critical	TP
PA, VSD	0–3	Yes	79	86	Abnormal	–	–	Critical	TP
CoA	6–12	No	100	96	Abnormal	–	–	Critical	TP
CoA, hypoplastic AA, VSD	3–6	No	98	97	Abnormal	–	–	Critical	FN
Congenitally corrected TGA, PS	0–3	Yes	95	97	Abnormal	–	–	Critical	FN
TGA, VSD, CoA	3–6	No	95	97	Normal	–	–	Critical	FN
Hypoplastic arch, AS, VSD	24+	No	99	99	Abnormal	–	–	Critical	FN
TA, AS	12–24	No	98	96	Abnormal	–	–	Critical	FN
Hypoplastic AA, CoA, VSD	12–24	No	98	99	Normal	–	–	Critical	FN
Complete AVSD	3–6	Yes	92	81	Normal	93	77	Serious	TP
TA, AVSD	12–24	Yes	84	84	Abnormal	–	–	Serious	TP
Tricuspid atresia, VSD	0–3	Yes	87	90	Abnormal	–	–	Serious	TP
TOF	3–6	Yes	87	92	Abnormal	–	–	Serious	TP
TA	6–12	No	93	97	Normal	94	97	Serious	TP
Coronary artery fistula	12–24	No	88	91	Abnormal	–	–	Serious	TP
DORV, VSD	0–3	Yes	93	94	Normal	88	90	Serious	TP
TOF	0–3	Yes	93	94	Abnormal	–	–	Serious	TP
Hypoplastic AA, PDA	24+	No	97	92	Normal	99	97	Serious	FN
Anomalous LCA	12–24	No	100	100	Normal	–	–	Serious	FN

Lesion	Timing of test (hours)	Antenatal diagnosis	Pulse oximetry		Examination	Pulse oximetry 2		CHD category	Result
			Hand	Foot		Hand	Foot		
VSD	24+	No	98	97	Normal	–	–	Serious	FN
PDA	24+	No	99	100	Normal	–	–	Serious	FN
Aortopulmonary window	12–24	No	98	100	Normal	–	–	Serious	FN
PS	3–6	No	97	98	Normal	–	–	Serious	FN
VSD	24+	No	100	100	Normal	–	–	Serious	FN
PS	24+	No	97	96	Abnormal	–	–	Serious	FN
AVSD	24+	No	97	95	Normal	–	–	Serious	FN
CoA, VSD	24+	No	98	100	Normal	–	–	Serious	FN
VSD	3–6	No	99	97	Normal	–	–	Serious	FN
CoA, hypoplastic AA, VSD	24+	No	100	100	Normal	–	–	Serious	FN
VSD	6–12	No	96	97	Normal	–	–	Serious	FN
PS	6–12	No	97	98	Normal	–	–	Serious	FN
PS, VSD	12–24	No	100	100	Normal	–	–	Serious	FN
PS	12–24	No	99	98	Normal	–	–	Serious	FN
VSD, ASD	24+	No	97	98	Normal	–	–	Serious	FN
VSD	12–24	No	97	99	Normal	–	–	Serious	FN
TOF	12–24	No	98	99	Normal	–	–	Serious	FN
VSD	3–6	No	100	100	Normal	–	–	Serious	FN
TOF	24+	No	99	100	Normal	–	–	Serious	FN
Dextrocardia, DORV, AVSD	3–6	Yes	83	95	Normal	87	92	Significant	FP
VSD	6–12	No	99	94	Normal	98	92	Significant	FP
VSD, hypoplastic RV, ASD	6–12	No	91	90	Normal	90	92	Significant	FP
AVSD	12–24	Yes	92	92	Normal	98	94	Significant	FP
Ebstein's anomaly	0–3	No	93	96	Abnormal	–	–	Significant	FP
Persistent PDA	6–12	No	98	92	Normal	98	92	Significant	FP

AA, aortic atresia; ASD, atrial septal defect; DILV, double inlet left ventricle; DORV, double outlet right ventricle; LCA, left coronary artery; PA, pulmonary atresia; RV, right ventricle; TA, truncus arteriosus; TAPVD, total anomalous pulmonary venous drainage.

Appendix 3

Model inputs for cost-effectiveness evaluation

In this appendix, we illustrate, with the aid of flow diagrams, how the proportions used in Group 1 (see *Table 17*) were derived from the prevalence of CHDs (see *Table 16*) following varying modes of testing.
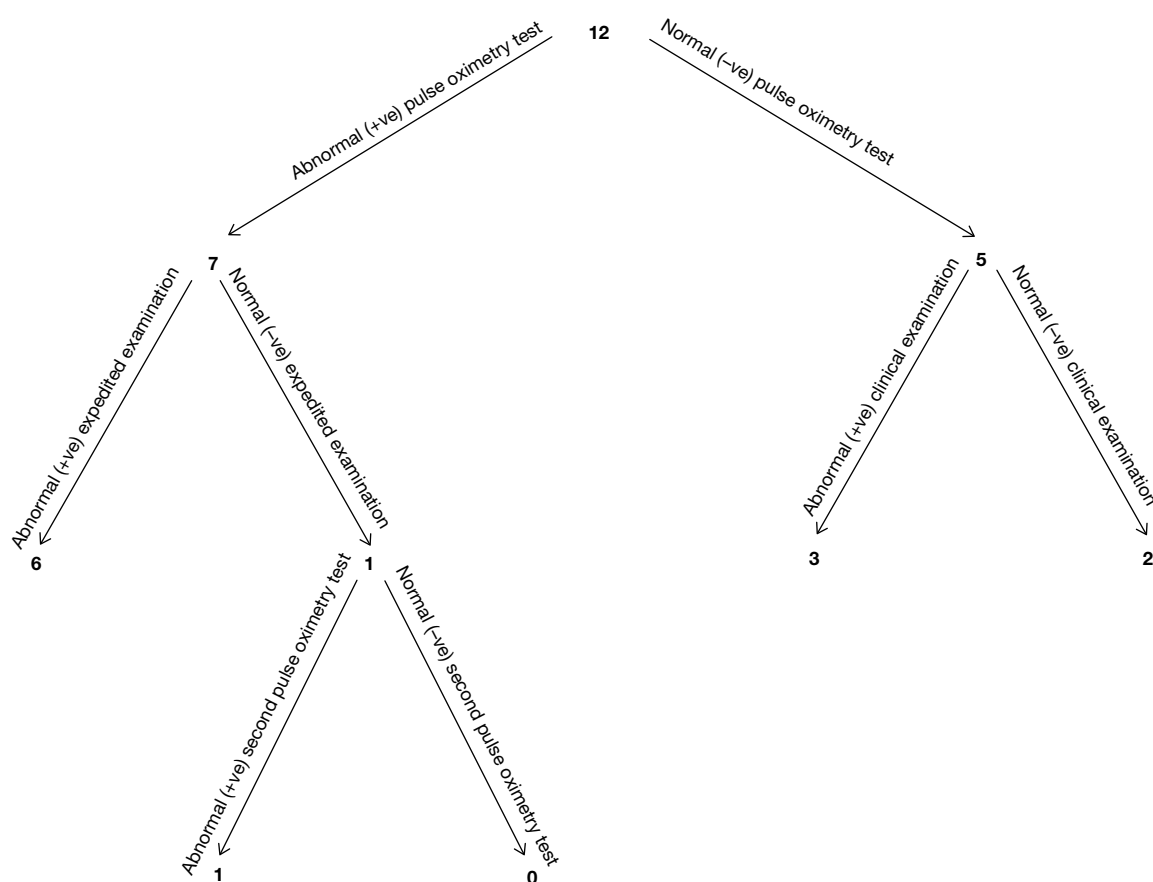

FIGURE 13a Flow chart of newborns with a critical CHD.

The prevalence of newborns with a critical CHD is 12 per 20,032; from these 12 newborns, seven were detected through first pulse oximetry screening test. Newborns here further received an ECE that targeted the lungs and heart. Six out of the seven newborns had abnormal results from the expedited examination. The remaining one newborn with a normal expedited examination result further received a second pulse oximetry test for which the baby had an abnormal result.

The remaining five cases that had initial normal pulse oximetry results received a routine clinical examination. Three babies out of the five had abnormal clinical examination results.
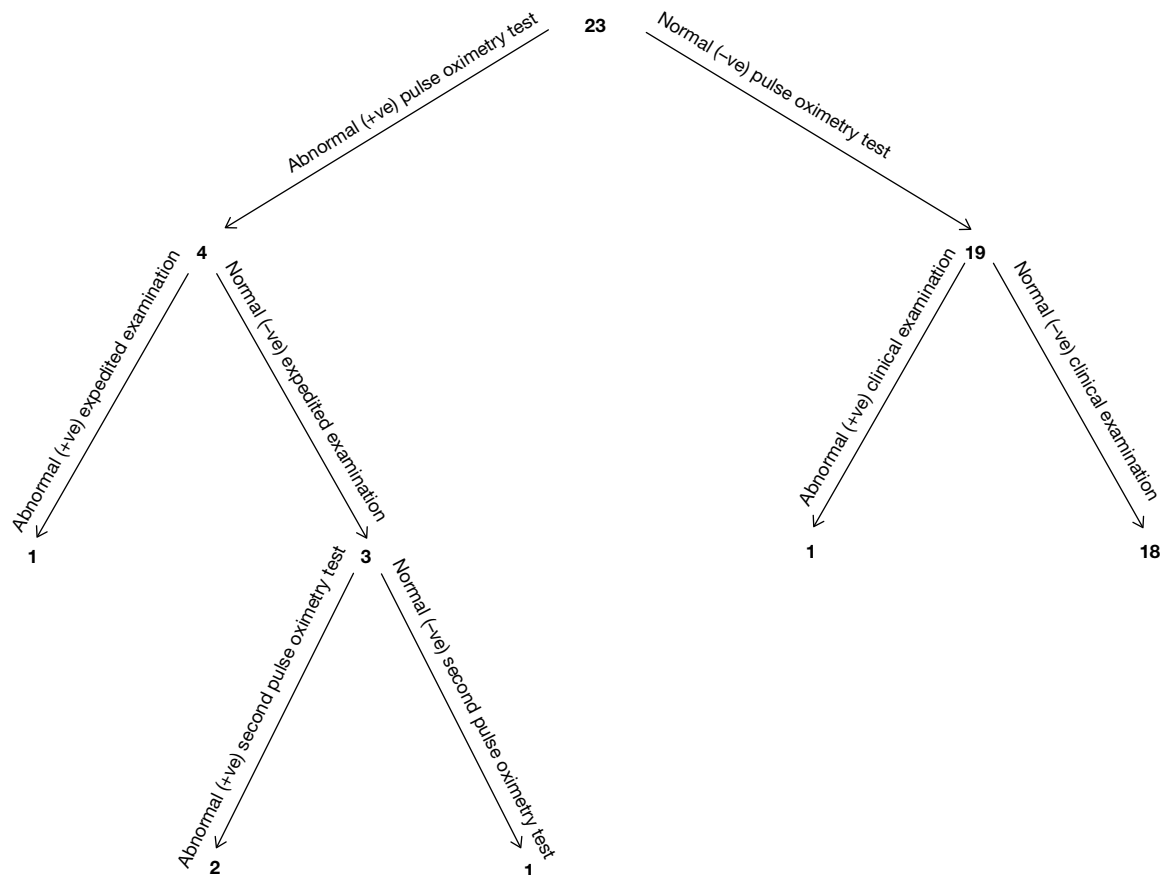

FIGURE 13b Flow chart of newborns with a serious CHD.

The prevalence of newborns with a serious CHD is 23 per 20,032. Of these 23 newborns, four were detected through abnormal first pulse oximetry test results. These four babies further received an ECE, for which one baby had an abnormal result and three babies had normal expedited examination results. The three babies with normal expedited examination results received a second pulse oximetry, for which two newborns had abnormal results.

From the remaining 19 cases that had an initial normal pulse oximetry test result, after routine clinical examination, one baby was found to have an abnormal (+ve) clinical examination and the remaining 18 had a normal (-ve) test.
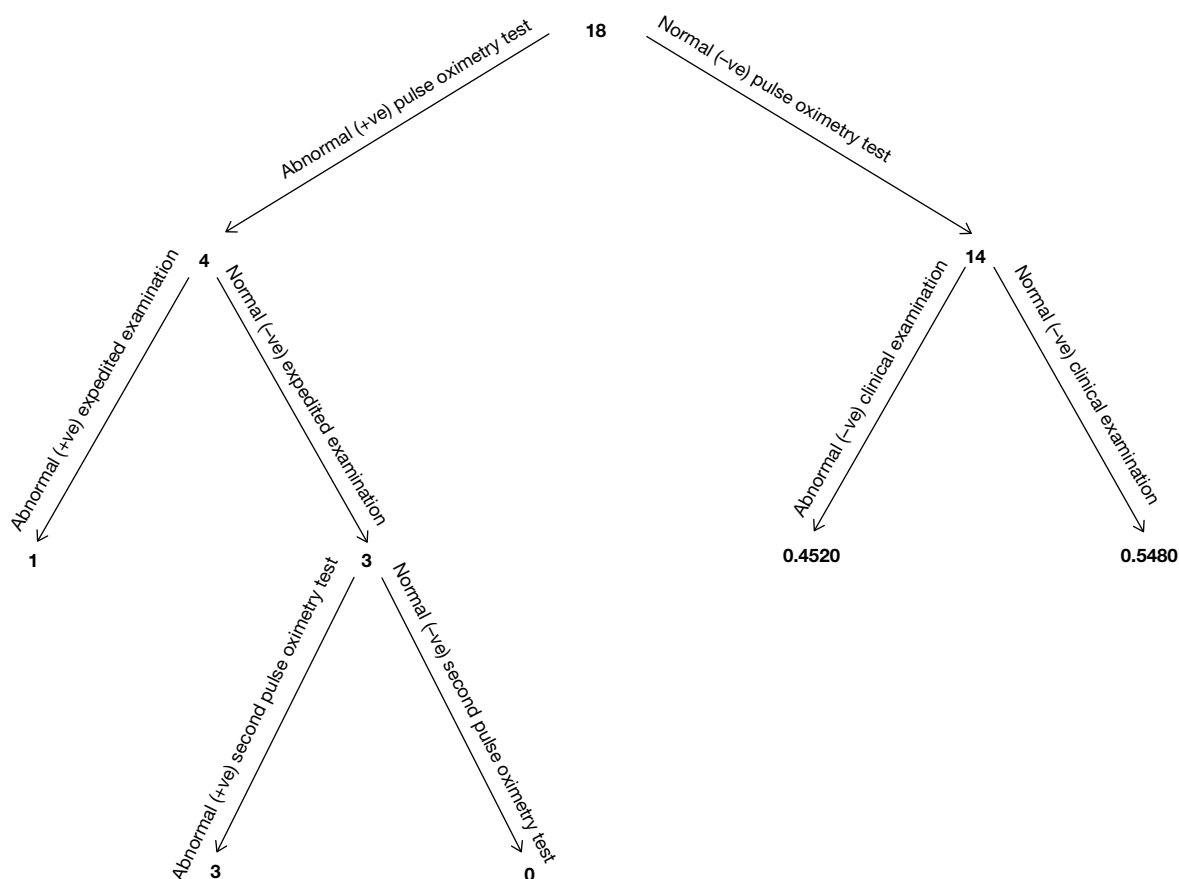
FIGURE 13c Flow chart of newborns with a significant CHD.

Of the 18 newborns with a significant CHD, four babies had an abnormal initial pulse oximetry test results. These babies further received an expedited clinical examination, from which one out of the four babies had an abnormal examination. The remaining babies (three) with normal ECE results received a second pulse oximetry test for which they had abnormal results.

The 14 babies with initial normal pulse oximetry test results further received routine clinical examination. We calculated an estimated probability of 0.4520 will have an abnormal clinical examination. This probability was calculated from the Wren *et al.* study,[26] based on the echocardiographic definitions used in the current study. These proportions were not collected in the current study.
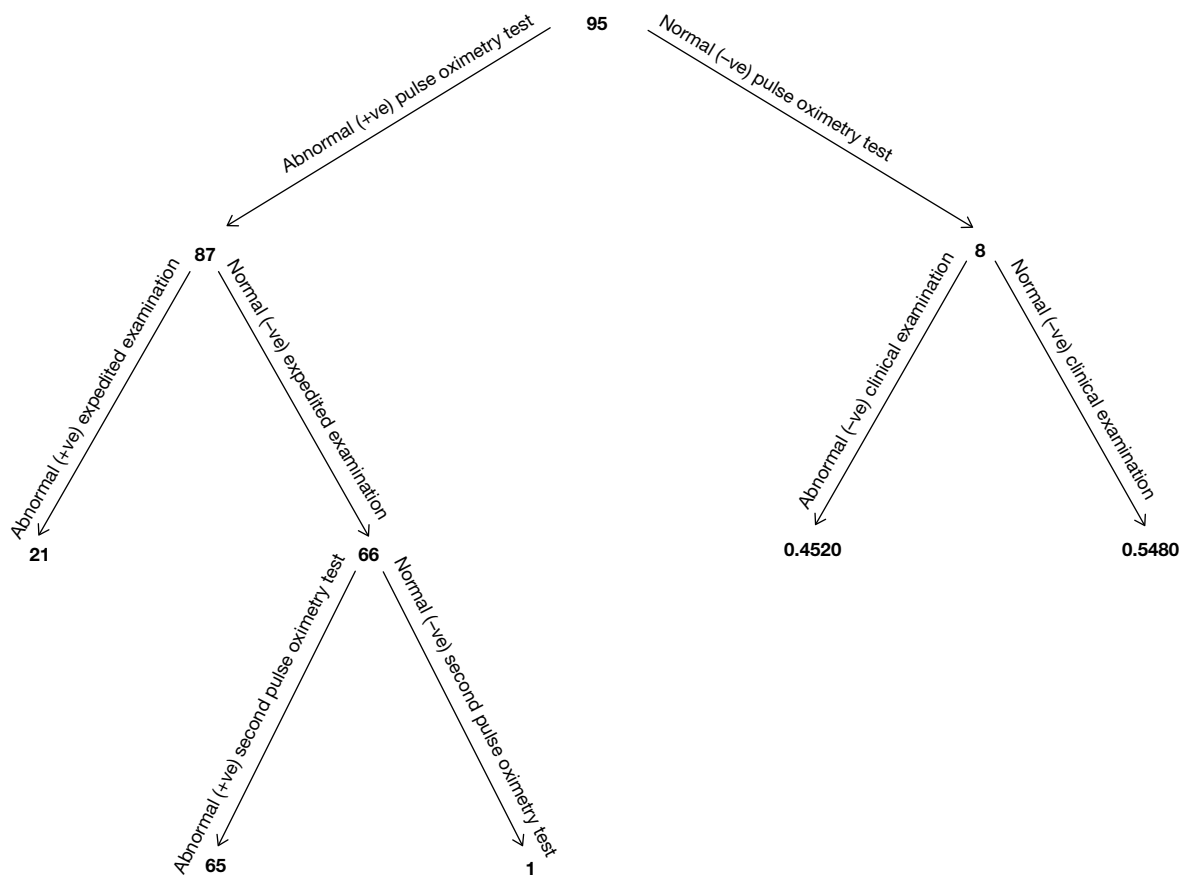
FIGURE 13d Flow chart of newborns with a non-significant CHD.

Of the 95 newborns that were considered to have a non-significant CHD, 87 had an abnormal pulse oximetry test result. These babies further received an expedited examination; 21 of these babies had an abnormal expedited examination. The remaining 66 babies, with normal expedited examination results, received a second pulse oximetry test, and 65 of these had abnormal second pulse oximetry test results.

The eight babies in whom the initial pulse oximetry test was normal underwent a routine clinical examination. We estimated from the Wren *et al.*[26] study that the probability of an abnormal clinical examination would be 0.4520. This probability was based on the echocardiography definitions (see *Table 1*) used in the study as these proportions were not measured in the study.
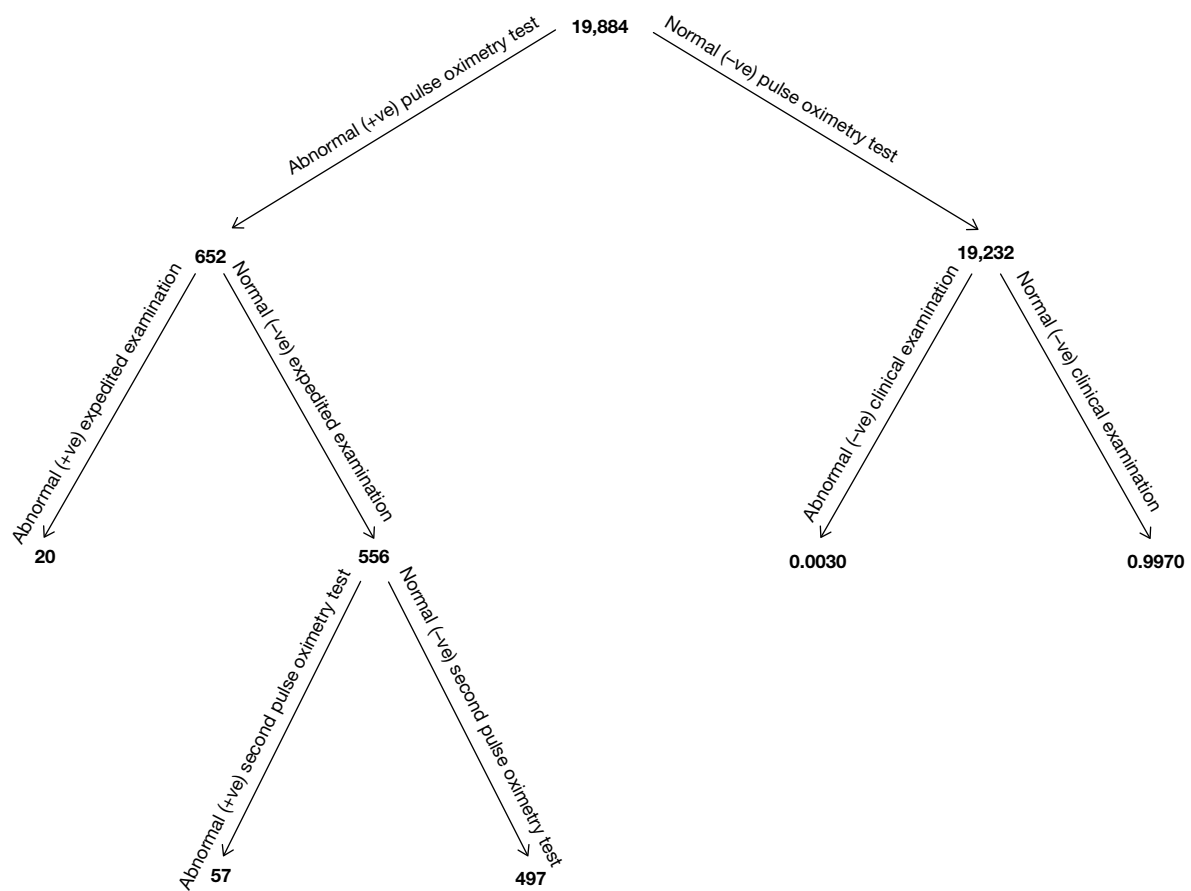
FIGURE 13e Flow chart of newborns without a CHD.

The study reported that 19,884 newborns had no CHD present and were considered normal. Of these 19,884, 652 babies had an abnormal pulse oximetry test result. These babies further received an ECE, for which 20 babies had an abnormal examination. Results for 76 of these babies were missed. It was assumed that these babies were normal (i.e. no CHD present), as they did not appear on the congenital anomalies register (CAR). Babies with a normal expedited examination further received a second pulse oximetry test. In total, 57 babies had an abnormal second pulse oximetry test result. Results for two babies were missed and it was assumed that these babies are normal as they did not appear on the CAR. The remaining 19,232 babies who had initial normal pulse oximetry results received a routine clinical examination. A probability of 0.0030 was calculated, based on the specificity of a routine clinical examination (0.9970) used in the Knowles *et al.* study.[6] These proportions were not measured in the study.

Appendix 4

Additional items supporting results presented in *Chapter 4* (economic evaluation)

Appendix 4.1 Model structure for Analysis 2

A decision tree structure was developed for Analysis 2 to compare three strategies: pulse oximetry as an adjunct to clinical examination, clinical examination alone and pulse oximetry.
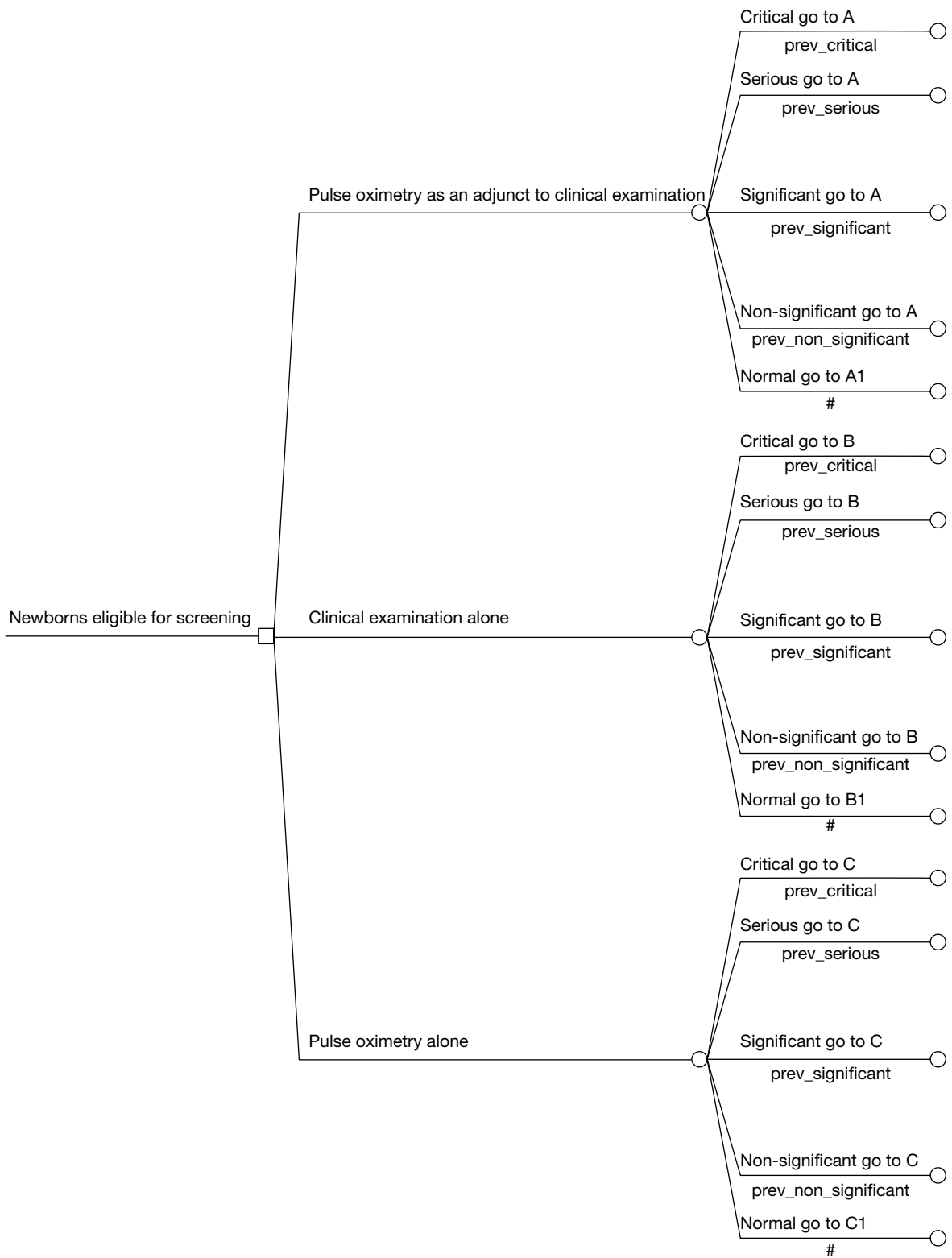
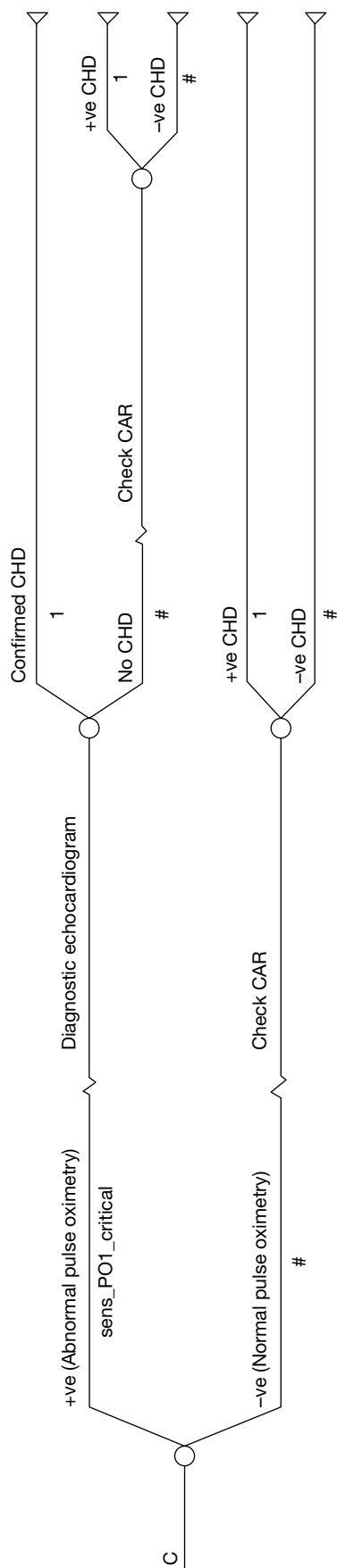**FIGURE 14a** Model structure of Analysis 2.

FIGURE 14b Pathway of the pulse oximetry-alone strategy for newborns with a CHD.
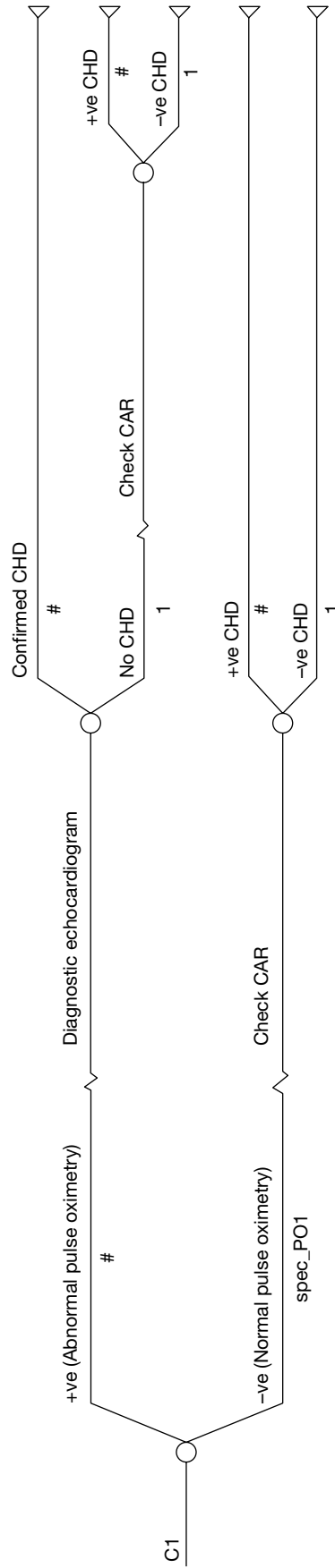
FIGURE 14c Pathway of the pulse oximetry-alone strategy for newborns without a CHD.

Appendix 4.2 Proportions and probabilities required for Analysis 2

TABLE 23 Proportions and probabilities required for sensitivity analysis by changing the threshold to <95% in either limb or a differential of >3%

Proportions and probabilities required	Critical	Serious	Significant	Non-significant	Normal
Pulse oximetry as an adjunct to clinical examination strategy					
Newborn having an abnormal (positive) pulse oximetry screen result (p_PO)	7/12	4/23	4/18	72/95	328/19,884
Newborn having a negative screen (specificity/detection rate) (1−p_PO)	N/A	N/A	N/A	N/A	19,556/19,884
Proportion of newborns having an abnormal expedited examination given an abnormal pulse oximetry result (p_EE)	6/7	1/4	1/4	17/72	18/324[a]
Proportion of newborns with an abnormal repeat pulse oximetry result having passed the expedited examination given an initial abnormal pulse oximetry result (p_2PO)	1/1	2/3	3/3	48/55	33/305[b]
Proportion of newborns having an abnormal (positive) clinical examination given normal (negative) pulse oximetry result (p_CE)	3/5	1/19	0.4520[c]	0.4520[c]	0.0030[d]
Clinical examination alone screening strategy					
Proportion of newborns having an abnormal clinical examination (sensitivity) (sens_CE)[e]	9/12	2/23	0.4070	0.2880	N/A
Proportion of newborns having a normal clinical examination (specificity) (spec_CE)[f]	N/A	N/A	N/A	N/A	0.9960

N/A, not applicable.
a   Four missing results here.
b   One missing result here.
c   Probability calculated from the Wren *et al.*[26] study based on the echocardiography definitions used in the study.
d   Probability calculated from the Knowles *et al.*[6] study based on the specificity of clinical examination alone (0.9970).
e   Proportion and probabilities calculated based on the weighted average of newborns with abnormal expedited and clinical examinations.
f   Probability calculated based on the weighted average of newborns with normal expedited and clinical examinations.

TABLE 24 Proportions and probabilities required for sensitivity analysis using threshold of <95% for foot saturations alone

Proportions and probabilities required	Critical	Serious	Significant	Non-significant	Normal
Pulse oximetry as an adjunct to clinical examination strategy					
Newborn having an abnormal (positive) pulse oximetry screen result (p_PO)	6/12	3/23	3/18	49/95	148/19,884
Newborn having a negative screen (specificity/detection rate) (1–p_PO)	N/A	N/A	N/A	N/A	19,736/19,884
Proportion of newborns having an abnormal expedited examination given an abnormal pulse oximetry result (p_EE)	5/6	1/3	0/3	10/49	10/146[a]
Proportion of newborns with an abnormal repeat pulse oximetry result having passed the expedited examination given an initial abnormal pulse oximetry result (p_2PO)	1/1	1/2	3/3	34/39	16/136
Proportion of newborns having an abnormal (positive) clinical examination given normal (negative) pulse oximetry result (p_CE)	3/6	1/20	0.4520[b]	0.4520[b]	0.0030[c]
Clinical examination alone screening strategy					
Proportion of newborns having an abnormal clinical examination (sensitivity) (sens_CE)[d]	9/12	2/23	0.3770	0.3240	N/A
Proportion of newborns having a normal clinical examination (specificity) (spec_CE)[e]	N/A	N/A	N/A	N/A	0.9970

N/A, not applicable.
a   Four missing results here.
b   Probability calculated from the Wren *et al.*[26] study based on the echocardiography definitions used in the study.
c   Probability calculated from the Knowles *et al.*[6] study based on the specificity of clinical examination alone (0.9970).
d   Proportion and probabilities calculated based on the weighted average of newborns with abnormal expedited and clinical examinations.
e   Probability calculated based on the weighted average of newborns with normal expedited and clinical examinations.

Appendix 4.3 Results of the base-case analysis using costs reported from Knowles *et al.*

TABLE 25 Deterministic results from Analysis 1 using the costs from Knowles *et al.*[6] (2005), 2000–1 prices

Strategy	Screening strategy	Total cost per strategy	Difference in costs	Effectiveness CHD detected	Incremental CHD detected	ICER (£)
A	Clinical examination alone	168,800	–	91	–	–
B	Pulse oximetry as an adjunct to clinical examination (Strategy B vs A)	401,000	232,200	121.4	29.9	7800

Appendix 5

Data collection forms

PULSE OX STUDY DATA COLLECTION FORM

Section 1: Eligibility

Is the baby well?   Yes ☐   No ☐

If the baby is **well**, they are **eligible** for the study. Please ensure consent has been taken and continue with **sections 2 – 5 below**.

If the baby is **not well**, they are **not eligible** for the study. Please complete the '**Ineligibility Form'**, put in the participant pack and put it in the Pulse Ox collection box. Treat the baby as normal.

Section 2: Identifying Details

Baby Hospital Number   ☐☐☐☐☐☐☐☐☐☐☐☐

Baby NHS number   ☐☐☐☐☐☐☐☐☐☐

Baby date of birth   _____ / _____ / _____

Time (24hr)   _____ : _____

Male ☐   Female ☐

Gestational age _____ /40 weeks

Mother's Forename _____

Mother's Surname _____

Hospital Number ☐☐☐☐☐☐☐☐☐☐

Mother's NHS Number ☐☐☐☐☐☐☐☐☐☐

Mother's date of birth _____ / _____ / _____

(or attach sticker)

Mother's gravida _____ parity _____

Section 3: Language

Was a translator and/or a translated version of the Information Sheet and Consent required?   Yes ☐   No ☐

If yes, specify language: _____

Section 4: Pulse Oximetry Saturations

First reading (This should be taken within 24 hours of age or before discharge)

Performed in: Delivery Suite ☐   Post Natal Unit ☐   Neonatal Unit ☐

Performed by:

    PRINT NAME: _____

Doctor ☐   Nurse ☐   Midwife ☐   Midwifery Assistant ☐   Care Assistant ☐   Neonatal nurse ☐

Date	Time	Foot sats (%)	Right hand sats (%)	% difference	Pass/Refer

If the result is normal (i.e. **Pass**), nothing further needs to be done. Please **ensure sections 1-5 are completed**. Please put this form in the participant pack and put it in the Pulse Ox collection box.

If less than 95% in either limb or ≥3% difference between limbs, this is a **Refer**, the baby needs to be examined/ reviewed by someone trained in neonatal examination. **Please continue with section 6 overleaf.**

Section 5: Antenatal Diagnosis

Is there a family history of congenital heart defects?   Yes ☐   No ☐

Was CHD suspected on antenatal ultrasound?   Yes ☐   No ☐

Data Collection Form   Version 3.0 06/11/2008

pulse♥x

The following sections only need to be completed if the first pulse oximetry reading was <95% in either limb or ≥3% difference between limbs i.e. a <u>Refer</u>.

Section 6: Clinical Examination

Undertaken by: Nurse/ Midwife ☐   SHO/ Registrar ☐   Consultant ☐   Other ☐

How long did the exam take?   _____ minutes

Did the baby show signs of CHD at examination?   Yes ☐   No ☐

     If yes, was this   murmur ☐   cyanosis ☐   abnormal pulses ☐

Overall exam finding   Normal ☐   Abnormal ☐   Details:_____

*If **exam is normal**, a second PO reading needs to be taken. This should be taken 1-2 hours after the first reading. **Please complete section 7 below.***

*If **exam is abnormal**, the baby needs to be referred straight to the Neonatal team for further investigation and management. A second PO reading does not need to be taken. **Please continue with sections 9-14 overleaf.***

Section 7: Second Pulse Oximetry Reading

Performed in: Delivery Suite ☐   Post Natal Unit ☐   Neonatal Unit ☐

Performed by:

     PRINT NAME: _____

Doctor ☐   Nurse ☐   Midwife ☐   Midwifery Assistant ☐   Care Assistant ☐   Neonatal nurse ☐

Date	Time	Foot sats (%)	Right hand sats (%)	% difference	Pass/Refer

*If less than 95% in either limb or ≥3% difference between limbs, this is a **Fail**, the baby needs to be referred to the Neonatal team for further investigation and management. **Please continue with sections 8-14 overleaf.***

*If the result is normal (i.e. **Pass**), nothing further needs to be done. Please ensure all above sections are completed. Please put this form in the participant pack and put it in the Pulse Ox collection box.*

Data Collection Form   Version 3.0 06/11/2008

pulseOx

The following sections only need to be completed if the clinical examination found an abnormality or if the second pulse oximetry reading was a Fail and the baby has been referred to the Neonatal team for further investigation and management.

Section 8: Hyperoxia Test (only to be completed if second PO reading was a fail)

Performed by: Nurse ☐   SHO ☐   Registrar ☐   Consultant ☐

Did saturation increase to 95% or above?   Yes ☐   No ☐

Section 9: Clinical Symptoms

If baby showed signs of CHD at birth, describe

Collapse ☐   Cyanosis ☐   Acidosis ☐   Respiratory distress ☐

Did this baby have a respiratory illness or any illness other than CHD?

Yes ☐   No ☐   If yes details: _____

Section 10: Echocardiography

Date of echocardiography ___ / ___ / ___   Time ___ : ___ (24 Hour)

Echocardiographer: Cardiology Research Fellow ☐   Other Echocardiographer ☐

Result: Normal ☐   Abnormal ☐   Uncertain ☐

Details: _____

Section 11: Review of Echocardiography

Reviewed by: Cardiology Research Fellow ☐   Consultant ☐   Other ☐   (Please tick more than one box if applicable)

Date of review ___ / ___ / ___

Result: Normal ☐   Abnormal ☐   Uncertain ☐

Details: _____

If there is a **disagreement between the two results** or the **result of the review is uncertain**, a second echocardiogram needs to be performed. This will be done by a Consultant Paediatric Cardiologist at Birmingham Children's Hospital

Date of second echocardiography ___ / ___ / ___

Result: Normal ☐   Abnormal ☐   Details: _____

Please continue with sections 12-14 overleaf.

pulseOx

Section 12: Admission Summary

Number of days on Postnatal unit? _____ days

Number of days on Neonatal unit? _____ days

Section 13: Discharge Details

Home ☐   Cardiac Unit ☐   Other ☐   Details: _____

If home, date of discharge _____ / _____ / _____   Time of discharge (24hr) _____ : _____

If cardiac unit, number of days on unit _____ days

Section 14: Death Details

Death ☐   Date of death _____ / _____ / _____   Time of death (24hr) _____ : _____

Cause of Death: Cardiac ☐   Non-cardiac ☐

　　　　If cardiac: Post-operative ☐   Other ☐   If other, please specify: _____

　　　　If non-cardiac, please specify: _____

Data Collection Form   Version 3.0 06/11/2008

pulseOx

PULSE OX STUDY
CONGENITAL HEART DEFECT CASE REPORT FORM

Date of Diagnosis   _____ /_____ /_____

How was the CHD detected and confirmed?

Pulse oximetry at birth then echocardiogram   ☐

Clinical examination at birth then echocardiogram   ☐

Post-mortem   ☐

Later echocardiogram by paed cardiologist   ☐

Congenital anomaly registry   ☐

Type of CHD?

Aortic (valve) stenosis	☐ Persistent (patent) ductus arteriosis	☐
Atrial septal defect	☐ Tetrology of Fallot	☐
Coartation of aorta	☐ Total anolomous pulmonary venous connection	☐
Complete artioventricular septic defect	☐ Transposition of great arteries	☐
Hypoplastic left heart syndrome	☐ Ventricular septal defect	☐
Interruption of aortic arch	☐	

Other (specify)   _____

Has the baby had cardiac surgery?   Yes ☐   No ☐   Planned ☐

If yes, when was the surgery _____ /_____ /_____

Type of surgery?   _____

Death ☐   Date of death _____/_____/_____   Time of death (24hr) _____ :_____

Cause of Death: Cardiac ☐   Non-cardiac ☐

If cardiac: Post-operative ☐   Other ☐

If non-cardiac, please specify: _____

pulseOx

Pulse Oximetry Study
Acceptability Questionnaire

Soon after your baby was born s/he was given a Pulse Oximetry test to see if there was a problem with the heart.

This involved putting a sensor on the baby's hand and foot to measure the oxygen levels (how 'pink' the baby was).

We would like you to answer some questions about the test and about yourself.

All your answers will be confidential and will **not** affect your treatment or your baby's treatment.

When are you filling in the questionnaire?

Date	Month	Year

For each question, please tick the box which best represents your views or fill in your answer.

Version 1.0 Date 10/12/2007

Before your baby was born, the midwife gave you a leaflet on the Pulse Oximetry test.

How satisfied are you with:

	Very satisfied	Satisfied	Neither	Dissatisfied	Very dissatisfied
1. The information in the leaflet on what the PO test involved?					
2. The information in the leaflet on heart disease in new born babies?					
3. The information in the leaflet on what would happen if the PO test showed there might be a problem?					

Before the test was done, how satisfied were you with:

4. The opportunities you had to discuss the PO test before you gave consent to have it?					
5. The opportunities you had to change your mind about having the test?					

Having the PO measure taken

Please say how much you agree with these statements

	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree
6. Overall I was happy with the way the test was done.					
7. My baby was very comfortable when the test was done					
8. I felt very stressed while the test was being done.					
9. I am very confident that the test was done properly					
10. I was sure results of the test were correct					

After the PO test was done, how satisfied were you with:

	Very satisfied	Satisfied	Neither	Dissatisfied	Very dissatisfied
11. The way you the doctor or nurse gave you the results?					
12. The information you were given on what would happen next?					
13. The opportunities you had to discuss the results?					

General feelings about the PO test

	Yes definitely	Probably	Don't know	Probably not	Definitely not
14. Do you think it was important for your baby to have the test?					
15. Do you think it is important for all babies to have the test?					
16. Do you think the doctors and nurses would have found any heart problem without the test					
17. Would you recommend the test to someone else?					

18. **What do you think the Pulse Oximetry test showed? Please tick one box only**

There is **no** problem with my baby's health	
My baby has a minor _heart_ condition	
My baby has a minor health condition, but not a _heart_ condition	
My baby has a serious _heart_ condition	
My baby has a serious health condition, but not a _heart_ condition	
Don't Know	
Other (please state)	

The following questions ask you about your views on Heart Disease in babies.

This is **not** a test of how much you know.
On a scale of 1 to 5, please circle the number that represents your views.

19. **How much would heart disease affect a baby's life?**

     No effect at all   1   2   3   4   5   Severe effect

20. **How long do you think heart disease would last if not treated?**

     A very short time   1   2   3   4   5   Forever

21. **How much do you think treatment helps a baby with heart disease?**

     Not at all   1   2   3   4   5   Extremely helpful

22. **How well do you feel you understand heart disease in babies?**

     Do not understand at all   1   2   3   4   5   Understand very clearly

The following questions are about how you feel about yourself.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
23. I hardly ever expect things to go my way					
24. I always look on the bright side					

Version 1.0 Date 10/12/2007

Below are a number of statements which people have used to describe themselves. Tick the box to indicate how you feel **right now,** that is **at this moment**

	Not at all	Somewhat	Moderately so	Very much so
25. I feel calm				
26. I am tense				
27. I feel upset				
28. I am relaxed				
29. I feel content				
30. I am worried				

Read each statement below and tick the answer that comes closest to how you have been feeling in **the past week**. Please don't take too long over your replies. Tick only one box for each statement

31. **I still enjoy the things I used to enjoy:**

Definitely as much ☐   Not quite so much ☐
Only a little ☐   Hardly at all ☐

32. **I can laugh and see the funny side of things:**

As much as I always could ☐   Not quite as much now ☐
Definitely not so much now ☐   Not at all ☐

33. **I feel cheerful:**

Not at all ☐   Not often ☐
Sometimes ☐   Most of the time ☐

34. **I feel as if I am slowed down:**

Nearly all of the time ☐   Very often ☐
Sometimes ☐   Not at all ☐

35. **I have lost interest in my appearance:**

Definitely ☐   I don't take as much care as I should ☐
I may not take quite as much care ☐   I take as much care as ever ☐

36. **I look forward with enjoyment to things:**

As much as I ever did ☐   Rather less than I used to ☐
Definitely less than I used to ☐   Hardly at all ☐

37. **I can enjoy a good book or radio or TV programme:**

Often ☐   Sometimes ☐
Not often ☐   Very seldom ☐

Thank you for answering these questions

Please use the space below to add any ideas you have on what we could do to improve the Pulse Oximetry test procedure for parents, or any other comments you would like to make

[blank comment box]

pulseOx

Pulse Oximetry Study
Follow-up Acceptability Questionnaire

Soon after your baby was born s/he was given a Pulse Oximetry test to see if there was a problem with the heart. This involved putting a sensor on the baby's hand and foot to measure the oxygen levels (how 'pink' the baby was).

We are contacting you again because we would be interested in what your views are now, a year after your baby was born.

All your answers will be confidential and will **not** affect your treatment or your baby's treatment.

For each question, please tick the box which best represents your views or fill in your answer.

When you have completed the questionnaire please return it in the envelope provided

When are you filling in the questionnaire?

Date	Month	Year

Follow-up Acceptability Questionnaire   Version 1.0 Date 14/04/2009

Thinking back to when the PO test was done, how satisfied are you now with:

	Very Satisfied	Satisfied	Neither	Dissatisfied	Very Dissatisfied
1. The way you the doctor or nurse gave you the results?					
2. The information you were given on what would happen next?					
3. The opportunities you had to discuss the results?					

General feelings about the PO test

	Yes Definitely	Probably	Don't know	Probably not	Definitely not
4. Do you think it was important for your baby to have the test?					
5. Do you think it is important for all babies to have the test?					
6. Do you think the doctors and nurses would have found any heart problem without the test					
7. Would you recommend the test to someone else?					

Below are a number of statements which people have used to describe themselves. Please tick the box to indicate how you feel <u>right now,</u> that is <u>at this moment</u>

	Not at all	Somewhat	Moderately so	Very much so
8. I feel calm				
9. I am tense				
10. I feel upset				
11. I am relaxed				
12. I feel content				
13. I am worried				

Finally, how you feel about your baby's health now <u>at the present time</u>

	Not at all	A little bit	Moderately	Quite a bit	A great deal
14. Do you think your baby is healthy?					
15. Do you ever think there is something seriously wrong with your baby?					
16. Do you worry a lot about your baby's health?					
17. Is it hard for you to believe a doctor or health visitor when they tell you there is nothing to worry about?					
18. If a disease is brought to your attention (e.g. on the TV, radio, newspapers or by someone you know) do you worry about your baby having it?					

Please use this space to add any ideas you have on what we could have done to improve the Pulse Oximetry test for you, or any other comments you would like to make

Thank you very much for answering these questions

Please return your questionnaire in the envelope provided

Appendix 6

National Screening Committee criteria for appraising the viability, effectiveness and appropriateness of a screening programme

Ideally all of the following criteria should be met before screening for a condition is initiated:

The condition

1. The condition should be an important health problem.
2. The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be a detectable risk factor, disease marker, latent period or early symptomatic stage.
3. All of the cost-effective primary prevention interventions should have been implemented as far as practicable.
4. If the carriers of a mutation are identified as a result of screening the natural history of people with this status should be understood, including the psychological implications.

The test

5. There should be a simple, safe, precise and validated screening test.
6. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.
7. The test should be acceptable to the population.
8. There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals.
9. If the test is for mutations the criteria used to select the subset of mutations to be covered by screening, if all possible mutations are not being tested, should be clearly set out.

The treatment

10. There should be an effective treatment or intervention for patients identified through early detection, with evidence of early treatment leading to better outcomes than late treatment.
11. There should be agreed evidence-based policies covering which individuals should be offered treatment and the appropriate treatment to be offered.
12. Clinical management of the condition and patient outcomes should be optimised in all health-care providers prior to participation in a screening programme.

The screening programme

13. There should be evidence from high-quality randomised controlled trials that the screening programme is effective in reducing mortality or morbidity. Where screening is aimed solely at providing information to allow the person being screened to make an 'informed choice' (e.g. Down syndrome, cystic fibrosis carrier screening), there must be evidence from high-quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual being screened.

14. There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/intervention) is clinically, socially and ethically acceptable to health professionals and the public.

15. The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment).

16. The opportunity cost of the screening programme (including testing, diagnosis and treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (i.e. value for money). Assessment against this criterion should have regard to evidence from cost–benefit and/or cost-effectiveness analyses and have regard to the effective use of available resource.

17. All other options for managing the condition should have been considered (e.g. improving treatment, providing other services), to ensure that no more cost-effective intervention could be introduced or current interventions increased within the resources available.

18. There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards.

19. Adequate staffing and facilities for testing, diagnosis, treatment and programme management should be available prior to the commencement of the screening programme.

20. Evidence-based information, explaining the consequences of testing, investigation and treatment, should be made available to potential participants to assist them in making an informed choice.

21. Public pressure for widening the eligibility criteria for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated. Decisions about these parameters should be scientifically justifiable to the public.

22. If screening is for a mutation the programme should be acceptable to people identified as carriers and to other family members.

References

1. Department of Health (DoH). *Screening of pregnant women for hepatitis B and immunisation of babies at risk*. London: DoH; 1998 (Health Service Circular: HSC 1998/127).

2. Wilson JMG, Jungner G. *Principles and practice of screening for disease*. Public Health Paper No. 34. Geneva: WHO; 1968.

3. Cochrane AL, Holland WW. Validation of screening procedures. *Br Med Bull* 1971;**27**:3.

4. Sackett DL. Controversy in the detection of disease. *Lancet* 1975;**2**:357–9.

5. Wald NJ (editor). *Antenatal and neonatal screening*. Oxford: Oxford University Press; 1984.

6. Holland WW, Stewart S. *Screening in healthcare*. London: The Nuffield Provincial Hospitals Trust; 1990.

7. Gray JAM. *Dimensions and definitions of screening*. Milton Keynes: NHS Executive Anglia and Oxford, Research and Development; 2007.

8. Raffle A, Gray M. *Screening evidence and practice*. Oxford: Oxford University Press; 2007.

Appendix 7

Protocol

pulseOx

Pulse oximetry as a screening test for congenital heart disease in newborn babies.
PROTOCOL

The Pulse Ox Study aims:

- To develop pulse oximetry (PO) as a feasible screening tool for congenital heart disease (CHD) in newborn babies.

- To determine the accuracy (sensitivity, specificity, predictive values, and likelihood ratios) of PO for diagnosing critical and serious CHD in newborn, using echocardiography / clinical follow up / congenital malformation registries' data as reference standards.

- To determine the psychosocial effect of PO as a screening method among parents and acceptability to health professionals.

- To compare the cost and cost effectiveness of PO and other screening tests (routine neonatal clinical examination and antenatal screening) for improving outcomes of CHD in the newborn using model based economic evaluation.
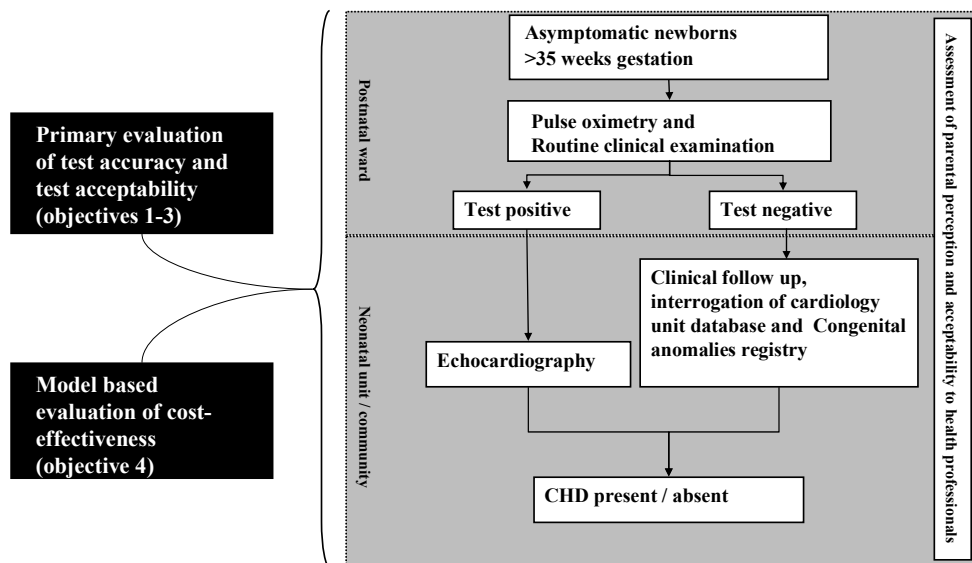
All women booking into six large obstetric units, in the West Midlands area, will be invited to participate. The study aims to recruit 20,000 women over 12 months for the assessment of test accuracy and acceptability. The main analysis for diagnostic accuracy will estimate sensitivity, specificity, predictive values, likelihood ratios and their confidence intervals. Using multivariable logistic regression analysis, predictive probabilities for various combinations of history, antenatal tests and oximetry results will be generated.

Pulse Ox Study Management Committee

Chief Investigator
Dr Andrew Ewer
Birmingham Women's Hospital

Clinical Co-investigators
Prof. Khalid Khan
Consultant in Obstetrics and Gynaecology
Birmingham Women's Hospital

Dr John Wright
Paediatric Cardiologist
Birmingham Children's Hospital

Statistics
Prof. Jon Deeks
University of Birmingham
Lee Middleton
University of Birmingham

Midwifery
Sarah Hooper
Birmingham Women's Hospital

Psychology
Dr Helen Pattison
University of Aston

Health Economics
Dr Tracey Roberts
University of Birmingham

Systematic Reviewer
Dr Shakila Thangaratinam
Birmingham Women's Hospital

Study Management
Ms Jane Daniels
Ms Alexandra Furmston
Ms Leanne Lynch
University of Birmingham

Study Steering and Data Monitoring and Ethics Committee

For interim analysis and response to specific concerns:

Dr Gerben ter Riet, Academisch Medisch Centrum

Suzie Hutchinson, Little Hearts Matter

Dr Carole Cummins, University of Birmingham

Dr Sam Richmond, Sunderland Royal Hospital

Dr Stavros Petrou, University of Oxford

Pulse Ox Study Office

The University of Birmingham Clinical Trials Unit,

Division of Medical Sciences, Robert Aitken Institute, University of Birmingham

Edgbaston, Birmingham, B15 2TT

Telephone: 0121 415 9112 (Voicemail outside office hours)

Fax: 0121 415 9136   E-mail: pulseox-study@contacts.bham.ac.uk

Co-ordination: Alexandra Furmston   Data management: Leanne Lynch
Statistics: Lee Middleton   Computing: Ed Tyler

Clinical queries should be directed during office hours to an appropriate member of the Management
Committee. Other queries should be directed to the Pulse Ox Study Office.

Study Principal Investigators

Principal Investigator (Birmingham Women's Hospital)
Dr Andrew Ewer

Principal Investigator (Heart of England)
Dr Richard Mupanemunda

Principal Investigator (City)
Dr Shanmugasundaram Sivakumar

Principal Investigator (New Cross)
Dr Babu Kumararatne

Principal Investigator (Royal Shrewsbury)
Dr Sanjeev Deshpande

Principal Investigator (Coventry and Warwickshire)
Dr Prakash Satodia

Version Number
Version 4.1 dated 18/03/2009

Protocol Versions
1.0 Submitted to MREC
      1.1 Incorporates MREC comments
      1.2 Amendment number 1 – Minor - Administrative changes
      1.3 Amendment number 2 – Minor - Administrative changes

2.0 Amendment Number 3 – Substantial – Change of eligibility criteria to include babies with extracardiac manifestations of CHD; submission of acceptability questionnaire; submission of Participant Information Booking Flyer, GP Poster and Participant Information Video; minor administrative changes.

3.0 Amendment Number 4 – Substantial – Clarification on data collection and administration of the acceptability questionnaires; clarification on consent procedure; definitions of critical and serious CHD

4.0 Amendment Number 5 – Substantial – Clarification and changes to the methods and analysis of acceptability to health professionals
     4.1 Incorporates MREC comments with inclusion of staff information sheets. In addition, version number correction of previously submitted staff acceptability questionnaire

Funding Body
NHS Research and Development Health Technology Assessment Programme

Sponsor

The University of Birmingham is the sponsor. Dr A K Ewer, Consultant Neonatologist, Birmingham Women's Hospital, is the Chief Investigator.

The University of Birmingham is responsible for obtaining necessary approvals and for safety monitoring, the Study Management Committee is jointly responsible for overseeing good clinical practice and the Investigators are responsible for obtaining informed consent and care of the participants.

CONTENTS

1. BACKGROUND

1.1. Clinical Context

CHD is the commonest group of congenital malformations and affects 7-8/1000 live born 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. A significant proportion of these children require surgery in the first year. Survival rates for infants with CHD have increased dramatically in recent years and over 80% of children born with CHD will survive to the age of 16 years; this is due largely to improvements in surgical techniques (3). Although the commonest types of CHD do not usually develop serious problems in the neonatal period, there are a number of important cardiac defects which, if undiagnosed, can present with potentially life-threatening cardiovascular collapse in the first few days of life. Although individually less common, when taken together, these lesions contribute significantly to death in infancy. As death or poorer outcome following surgery is much more likely if cardiovascular collapse occurs prior to diagnosis, timely recognition of these cardiac defects prior to collapse is vital in order to improve outcome.

Currently in the UK, all newborn babies undergo a routine screening examination, usually in the first 24 hours after birth, during which, among other things, a careful assessment of the cardiovascular system is undertaken. However, it is estimated that over 50% of babies with undiagnosed CHD failed to be picked up by routine neonatal examination. In addition, some hospitals are deferring this examination and it is then performed by the GP – sometimes up to 10 days after birth. This increases the risk of a baby with CHD presenting with acute deterioration before the screening exam has taken place. Antenatal screening for CHD in high risk pregnancies by fetal echocardiography has been developed and in experienced hands appears a useful technique (4), however a recent HTA assessment review of *routine* ultrasound screening in pregnancy concluded that detection rates for CHD were low (5). The need for an accurate, simple, non-invasive test for CHD in the early neonatal period has led a number of investigators to examine the use of PO and although results are encouraging, as we demonstrated in our systematic review (see below), there is a clear need for a larger, robust, well conducted study to confirm the value, acceptability and cost effectiveness of such a screening programme.

This large multi-centre study will determine the accuracy of PO screening for diagnosing critical and serious CHD in newborns. It will evaluate the acceptability of PO to both parents and health professionals. The study will also assess the costs and cost effectiveness of utilising such screening in combination with clinical examination in the early detection of potentially life-threatening CHD.

1.2. Literature review

1.2.1 Accuracy of PO as a screening test for CHD in newborns

The NHS Research and Development Health Technology Assessment Programme (HTA) commissioned a systematic review of screening for CHD in newborns which was published in 2005 (6). This report included 4 studies examining the role of PO as a screen for CHD in newborns. In 2007, we published a further systematic review of this screening test (7). An extensive search (from database inception to 2006) to retrieve primary and review articles was performed and from a total of 233 citations, 8 relevant studies - 4 more than those included in the previous HTA report (6) - were identified (Table 1). A careful and detailed analysis of the findings were summarised by meta-analysis using the bivariate model, which allows for some variation in threshold. The review can be summarised as follows:

- The quality of the studies was varied in both design and conduct: six studies recruited newborns consecutively (8-13); two were case control studies (14;15) - a

design that biases the results by overestimating the diagnostic odds ratio; two were not prospective (14;15); three studies raise concerns about spectrum bias due to exclusions of antenatally diagnosed cases (10;11;13); none calculated power *a priori*; and sample size was often inadequate to precisely estimate sensitivity.

- There were various thresholds for abnormality of test results (Table 1).
- Only one study explored the added value of PO above the accuracy achieved through clinical examination. The combination of PO 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%) (8).
- None of the studies evaluated acceptability of testing to parents and the psychosocial impact of false positive results or identification of non critical CHD.
- None of the primary studies provided information on cost and cost effectiveness.

Table 1 The accuracy of PO for diagnosing CHD in asymptomatic newborns

Test	No. of patients	Sensitivity % (95% CI)	Specificity % (95% CI)	False Positive Rate% (95% CI)	Timing of the test after birth
Most commonly used threshold					
saturation[*] ≤95% foot	11281	60 (14.7-94.7)	100 (100-100)	0	After 24 hrs or discharge
saturation[*] ≤95% foot or hand	2114	66.7(9.4-99.2)	99.9(99.7-100)	0.1(0-0.3)	As close to discharge as possible
saturation[*] ≤ 95% foot	3262	96.8(73.6-100)	99.7(99.5-99.9)	0.3(0.1-0.5)	Between 6 and 12 hours
saturation[§] <95% foot	5626	25(12.7-41.2)	99.6(99.4-99.7)	0.4(0.3-0.6)	Between 2 hrs and discharge
saturation[*] ≤ 95% foot	5292	66.7(9.4-99.2)	100(99.9-100)	0(0-0.1)	>24 hrs
saturation[§] <95% hand or foot	5211	30.8(9.1-61.4)	100(99.9-100)	0(0-0.1)	Prior to discharge
saturation[*] <95% foot	2733	75(57.8-87.9)	87.9(86.6-89.1)	12.1(10.9-13.4)	<6 hours of life, at 24hrs of life and/or at discharge
saturation[*] <95% foot	266	89.4(79.4-95.6)	99(96.4-99.9)	1(0.1-3.6)	12 hours (controls) or prior to surgery (cases)
Most accurate threshold					
saturation[*] <95% in both limbs or a differential of >3%	266	98.5(91.8-100)	96(92.3-98.2)	4(1.7- 7.7)	12 hours (controls) or prior to surgery (cases)
Summary estimate	**35785**	**63.4(38.7-82.5)**	**99.8(99-100)**	**0.2(0-1)**	summary using bivariate method of meta analysis

* functional oxygen saturation   § fractional oxygen saturation

The highest sensitivity seemed to be obtained with a cut off level of functional saturation <95% in both limbs or a differential of >3% in saturation between foot and right hand (15). The false positive rate could be influenced by the age of screening. Performance of the test after 24 hours of birth seems to have the highest specificity (100%) with the lowest false positive rates (10;11). These findings need further evaluation; there are potential concerns that some babies with CHD may present with clinical deterioration before 24 hours and therefore be missed by later screening. There is a clear need for an appropriately designed large test accuracy study evaluating PO as a screening tool for CHD in newborns.

1.2.2 Evidence on effectiveness of early intervention for congenital heart anomalies

In recent years, there has been a significant increase in survival of children with CHD. Developments in ultrasonography have resulted in more precise diagnosis without the

need for invasive procedures, and advances in surgical techniques mean that cardiac conditions which were considered lethal (e.g. hypoplastic left heart) are now routinely offered surgery. However, some cardiac lesions present with acute cardiovascular collapse or death prior to diagnosis. It is clear that if a baby presents with a clinical deterioration before surgery then this leads to worse outcomes (16). It is this group of babies in particular who would benefit from early detection via an appropriate and timely screening technique. This study will pay particular attention to timeliness of PO testing.

1.2.3 Evidence on cost effectiveness of PO as a screening test for CHD

None of the studies identified in the above systematic review had any cost or cost effectiveness data. The HTA systematic review of Knowles *et al* (6) reported the use of a decision analytic model to estimate the cost-effectiveness of clinical examination either alone, or with PO, or with screening echocardiography. The outcome of interest was timely diagnosis of life-threatening CHD, with clinically significant CHD as a secondary outcome. Data from the Northern Region study (1) were used to estimate prevalence of defects and probabilities of outcomes. Unit cost data was obtained by direct observation of a small number of screening tests, contact with manufacturers and published standard costs of management of CHD cases.

Using this model to estimate cost-effectiveness in a hypothetical population of 100,000 live born infants, it was estimated 82 cases of life-threatening CHD would be identified using a combination of PO and clinical examination, compared with 39 cases by clinical examination alone. The costs of the screening programmes would be £480,000 and £300,000 respectively, or an additional cost of £4,900 per additional timely diagnosis for the combined strategy. The cost of using screening echocardiography alongside clinical examination was prohibitive at £4.5m per additional timely diagnosis with a four-fold higher false positive rate. Although sensitivity analysis showed the findings are robust to many parameters, detection rates and screening tests costs influence cost-effectiveness greatly. The model assumes the screening tests are performed at 24 hours of age, whereas, if this test were to become routine, it would most likely be performed much earlier. The time of diagnosis is a major difference from this study: detection rates influence the model yet no data exists to demonstrate the impact of timing of diagnosis on test performance.

1.3. Objectives of the Pulse Ox Study

The objectives of the study have been framed to assess PO for screening for CHD in newborn babies in a hierarchical fashion, based on methodologically robust frameworks for evaluation of diagnostic tests outlined by Guyatt *et al* (17) and Fryback and Thornbury (18), as follows.

- To develop PO as a feasible screening tool for CHD in newborn babies.

- To determine the accuracy (sensitivity, specificity, predictive values, and likelihood ratios) of PO for diagnosing critical and serious CHD in newborns, using echocardiography / clinical follow up / congenital malformation registries' data as reference standards (see Appendix E for CHD definitions).

- To determine the psychosocial effect of PO as a screening method among parents and acceptability to health professionals.

- To compare the cost and cost effectiveness of PO and other screening tests (routine neonatal clinical examination and antenatal screening) for improving outcomes of CHD in the newborn using model based economic evaluation.

2. STUDY DESIGN

2.1. Brief summary

The current neonatal screening for CHD involves a clinical examination undertaken either by a trained health professional prior to discharge from hospital, or by a GP up to 10 days following delivery. The study will evaluate screening using PO undertaken by a midwife within 24 hours of birth. The primary analysis will be accuracy of PO in detecting life-threatening CHD. A comparison of the performance of PO either against, or in addition to, antenatal screening will be made by multivariable logistic regression analysis; another with postnatal clinical examination will be made by decision analytic modelling. The study will also assess parental perception and acceptability to healthcare staff. A cost-effectiveness analysis comparing different screening strategies will be undertaken, using direct study data where possible, in a model-based economic evaluation.

2.2. Overview

Antenatal booking visit
Provision of information with inclusion of study details in handheld notes and written consent sought for inclusion into study

Postnatal ward or after delivery in delivery suite
Reinforcement of information regarding study and verbal confirmation of consent previously obtained

Pulse oximetry (PO) before 24 hours or discharge (preferably between 3- 6 hrs) of age in postnatal ward measuring functional oxygen saturation in right upper and lower limbs
Routine clinical examination by trained health professional

Not eligible

Eligible but declines

SaO2 <95% in either limb or ≥3% difference

SaO2≥95% in both limbs

Clinical examination

Abnormal

Normal

Repeat PO in 1-2 hrs

SaO2 <95% in either limb or ≥3% difference

SaO2≥95% in both limbs

Hyperoxia test

Echocardiography

Clinical follow up, interrogation of cardiology unit database and Congenital anomalies registry

CHD present

CHD absent

CHD present

CHD absent

Figure 1: Flow chart of study organisation

This study will assess the accuracy, effectiveness and efficiency of PO in a delayed-type cross-sectional accuracy study. First, PO will be performed soon after delivery in all cases, which will have the advantage of identifying potentially serious low oxygen saturations in a timely fashion before symptoms are apparent. Integrating PO into the assessments performed routinely by the midwife when mother and baby are transferred from the delivery suite to the post-natal ward will make this design practicable and achieve maximum recruitment. Testing will occur early but over a range of times within the first 24 hours after birth (usually within 3-6 hours) or before discharge, allowing analysis of the optimum time for PO screening. After conducting PO, we will use different reference standards for test positive and test negative cases: echocardiography for the former and follow-up for the latter.

In this design, it would be unlikely that the subsequent clinical examination could be performed independently of the PO results, as midwives would want to alert the medical staff to a low saturation. This precludes a direct comparison of PO against clinical examination, as the latter would be prone to work-up bias. This comparison will be made through decision-analytic modelling using unblinded and historical data for input and appropriate sensitivity analyses conducted.

Clearly, in the presence of a feasible and accurate CHD screening strategy, there will be a need to consider the cost-effectiveness. The cost-effectiveness of PO in combination with postnatal clinical examination will be evaluated against examination alone and examination and antenatal screening, in a model-based analysis. Two systematic reviews of the performance of antenatal screening for CHD in first trimester (4) and second trimester (19) have been performed. Data on this group will be collected as part of our study. Using directly obtained information on outcomes and costs from a large number of babies in a real-life setting, the implications of all methods of identification of CHD will be modelled, thereby improving the generalisability of the results.

2.3. Test accuracy study design

A test accuracy study is different to an effectiveness study in that randomisation of subjects is not involved. An outline of the test accuracy study is shown in Figure 2. It is designed to generate a comparison of measurements obtained by index tests with those obtained by reference standards. In this way the accuracy of index tests can be estimated. A reference standard is a test that confirms or refutes the presence or absence of disease beyond reasonable doubt. Therefore it is sometimes also known as the gold standard. PO is the index test whereas the reference standard will be the identification of CHD cases up to one year through echocardiography, databases maintained by the West Midlands congenital anomaly register (CAR), regional cardiology referral unit, regional perinatal mortality survey, hospital information departments and, if necessary, READ codes of primary care trusts.
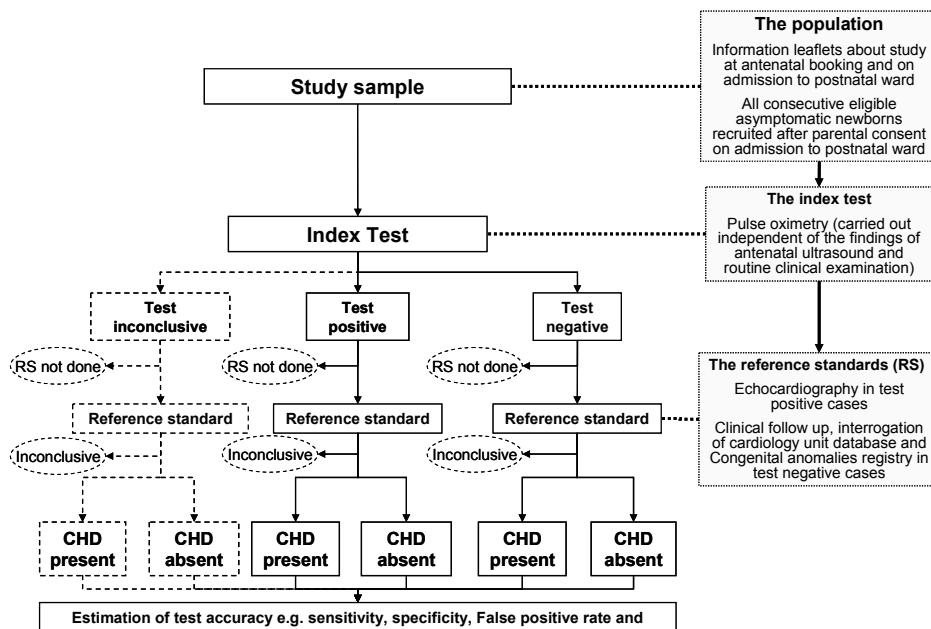

Figure 2 Flowchart of test accuracy study of PO as a screening tool for CHD in newborns

5

2.4. Large, simple study: minimal extra workload

In order to obtain the large number of patients necessary for the reliable evaluation of the PO test, the study will need the participation of more than one centre. To make these practicable, study procedures need to be kept simple, with the minimal extra workload placed on participating clinicians, beyond that required to manage their patients. This will be achieved by simple entry procedures, early consent of women, the use of standard local testing regimens, minimising documentation and streamlining data collected procedures. Regular newsletters will keep collaborators informed of study progress, and regular meetings will be held to report progress of the study and to address any problems encountered in the conduct of the study.

2.5. Setting

The setting will be 6 large maternity units in the West Midlands region. These units are based in the following hospitals: Birmingham Women's, Heart of England, City (Birmingham), New Cross (Wolverhampton), University Hospital of Coventry & Warwickshire (Coventry), and Royal Shrewsbury. These units serve a large, socio-economically and ethnically diverse population, which will aid generalisability of findings and have a large total number of deliveries (over 30,000 per annum). The units also represent the spectrum of obstetric settings, from a busy district general hospital to a specialised tertiary referral centre.

3. ELIGIBILITY

3.1. Eligibility criteria for the test accuracy study

The following inclusion / exclusion criteria will be used:

3.1.1 Inclusion criteria

Mother   Given written informed consent for screening and follow-up through databases

Baby   Gestational age ≥35 weeks

Asymptomatic at birth, regardless of antenatal risk factors or ultrasound findings

3.1.2 Exclusion criteria

Mother   Unable to give consent through incapacity, inability to speak English and lack of interpreter

Baby   Symptoms of cardiovascular abnormalities at birth

3.1.3 Inclusion of babies with suspected CHD from antenatal ultrasound

Babies suspected of having CHD, following antenatal ultrasound, will be included in the study. These pregnancies will be managed under the care of a consultant and will be transferred shortly after birth to a neonatal unit, where they will routinely have PO monitoring. The first reading, prior to any therapy or oxygen administration, will be recorded for the purposes of the study (provided the baby is asymptomatic at the time).

The evidence from the review about the accuracy of midtrimester fetal ultrasonography (19) showed that this antenatal test was not without error and it did not lend support to its routine use among unselected and low risk populations. In fact, practice concerning the intensity of such screening across the UK varies considerably, as does associated test performance. The West Midlands region, however, has an active, specialist-based fetal echocardiography screening programme for high-risk mothers, which detects a significant proportion of fetuses with CHD. Most babies born with a positive antenatal test result are

asymptomatic at birth, would comfortably meet all our inclusion criteria (see sections 3.1.1 and 3.1.2), and would have been eligible for mass screening with PO, regardless of antenatal screening, whether risk or ultrasound based. In babies with positive antenatal scan, PO is currently routinely carried out after birth on admission to the neonatal unit, so it would be unwise to disregard this information. There is no risk of introducing bias in PO estimation due to prior the suspicion of CHD as PO is an objective test. By including this subgroup, the overall CHD prevalence will be increased, thereby increasing the power and precision. It will allow the opportunity to explore variation in test accuracy in different population spectra and will help determine the added value of PO over and above what is achieved by antenatal screening. This information will be critical input into the decision-analytic model.

3.2. Recruitment of participants

Although PO will be introduced as a routine post-natal test, consent will still need to be obtained from each mother for the purposes of the study. Ideally consent is sought under unhurried circumstances, when entry criteria are fulfilled. These two issues may not be easily satisfied simultaneously in this study, as the setting involves obtaining PO readings in the first few hours of the baby's life. Consent will be sought in stages:

- A patient information leaflet will be given to all pregnant women at the time the mid-trimester ultrasound scan, or at any time after the scan but before birth, by the community or antenatal clinic midwives. The leaflet will also be made generally available and prominently displayed in various areas within the participating hospitals and their community antenatal clinics. This leaflet will be available in different local languages to reflect the ethnic communities at each centre.

- The leaflet will emphasise that PO is being introduced for routine use in the participating hospitals for the duration of the study, but that parents are entitled to choose whether or not they want their baby screened. Parents will be advised to discuss the study with the community or antenatal clinic midwife if they have any questions; and the phone number of the local coordinating midwife for the study will be provided.

- A coloured sticker confirming that the leaflet has been passed to the mother will be attached to the front of her notes. A copy of the information leaflet and consent form will be included in the hand-held notes. Any parent declining PO before delivery will have this recorded in the maternity notes.

- The consent form can be signed by the mother at any point before delivery and countersigned by the community or antenatal clinic midwife. At the time of testing, the consent will be reconfirmed verbally and any change recorded on the consent form.

Where necessary, appropriate interpreters will be asked to aid discussion relating to study participation. Some of the study centres employ link workers, who cover the role similar to community midwives for non-English speaking mothers. They are ideally suited to taking consent for non-English speaking mothers for this study.

It is anticipated that acceptability of the pulse oximetry test and willingness to participate in the study may potentially vary between ethnic groups. We will record baseline demographic and ethnicity information from all mothers invited to take part, including those who decline to take part, those who miss consent, or those who are later identified as ineligible. We will collect mother's hospital number, age, parity, ethnicity and language. These details will be stored anonymously. This will establish the take-up rate of the study.

By making PO a routine test in each hospital, it will become embedded in the care pathway for women post-delivery, removing the possibility that women are missed by

mistake, apathy or poor organisation. In this study, reasons for opting out will not be collected (unless it becomes apparent that this is an issue), nor will it be necessary to obtain all mothers' addresses. We anticipate a high take-up, making it unfeasible and unnecessary to collate excess information when it will be available from routine hospital data or can be collected postnatally.

All community midwives and link workers will receive training regarding the introduction of PO screening, information about the study and instruction on their roles from the local coordinating midwives. This will occur during community team meetings and the information provided will be reinforced periodically throughout the study by further meetings and newsletters from the Study Office.

3.3. Consent

The conduct of the study will be in accordance with the Medical Research Council (MRC) Guidelines for Good Clinical Practice 1998 and any subsequent amendments. The mother's written informed consent (according to usual local practice) to participate in the study must be obtained before testing. As the study is of short duration and has no consequence on the management of pregnancy or any post-natal care, the women's GP will not be notified of her participation in the study.

3.4. Organisation of Recruitment

Recruitment will be organised and supported by dedicated midwife leads, who will work with local community and antenatal clinics midwifery teams and obstetrics and neonatal leads. We believe that that the following strategy is likely to be successful in achieving maximum recruitment.

- Appointment of a dedicated coordinating midwife at each centre with responsibility for overseeing preliminary consent in the community and at the antenatal clinic of her centre, PO testing on post-natal ward and for data collection and problem resolution.
- Appointment of a lead research midwife at Birmingham Women's Hospital, who will liaise with all the coordinating midwives at each centre and coordinate the screening at this hospital, provide training and trouble-shoot recruitment and midwifery problems.
- Provision of simple written study information (similar to the "Newborn blood-spot screening" leaflet), supported by face to face discussion with midwifery staff in antenatal clinics and the community.
- Regular, close communication with midwifery staff in the community, antenatal clinic and post-natal wards.
- Training of midwifery staff on post-natal ward in obtaining PO readings, as the coordinating midwife will not be available to undertake this task on all babies.
- Provision of regular feedback on progress in study recruitment, including individual hospital teams' performance and progress against targets.
- Regular newsletters to all relevant staff involved in the study.

4. TESTS AND PROCEDURES

4.1. The index test

It is essential to explore the technical and practical aspects of conducting the PO test before the commencement of the accuracy study. A ward-based technical feasibility study will be carried out to establish the most practicable methods of testing and reporting results in the postnatal ward setting. This will also allow the development of programmes

for quality assurance and for training staff to perform PO. This will be one of the roles of the coordinating midwives.

The pulse oximeter used in the study is from the Radical® series from Masimo (Irvine CA, USA). These have been shown to outperform other oximeters in that the recording is free from motion artefact and is capable of achieving stable, accurate readings in an active subject and also when perfusion is low. Five oximeters will be available in each maternity unit in order to ensure readings can be taken at all times and to ensure readings are not missed because of faulty or misplaced machines. An oximeter will be available on both the postnatal ward and delivery suite.

The participating centres have agreed to make it standard practice to perform PO within each unit and the clinical protocols changed accordingly. PO will be performed before 24 hours or discharge, preferably at around 3-6 hours of age which corresponds to admission to the postnatal ward. PO is routinely performed on babies on the Neonatal Unit. It is painless, extremely well tolerated by babies and requires minimal clinical training. It is estimated that to perform saturations on one hand and one foot will take no more than five minutes. The results of the PO will be recorded as percentage functional saturation for each limb. The date and time of testing will also be recorded. A cut off of <95% in either limb or a difference of ≥3% between the limb readings will be considered to be abnormal. This threshold has been shown in previous studies to have the highest sensitivity. It also has the potential to detect coarctation of aorta, a treatable condition that has been missed in earlier studies with different thresholds (14,15) It is anticipated that the maximal sensitivity of the test is achieved by about 6-12 hours of age (9;14;15), by which time we anticipate the majority of babies in this study will have been screened. If PO is low and the clinical examination is unremarkable, the PO will be repeated 1-2 hours later for a definitive definition of abnormality. If the saturations remain low, oxygen will be administered (nitrogen washout or hyperoxia test) to identify potential respiratory causes for low saturations. An echocardiogram will be performed in all case of persisting low saturations (i.e. test positive cases)

4.2. Clinical examination

Physical examination of newborns is recommended in the NICE clinical guidelines "Postnatal Care of Women and their Babies section 1.4.11' and should occur within 72 hours of birth. Assessment of the cardiovascular system should include checking the position of the heart, heart rate, rhythm and sounds, murmurs and femoral pulse volume. Each hospital will have a protocol for the timing and content of the postnatal examination and this study will not attempt to alter clinical practice, although the NICE guidelines will be promoted. The assessment is usually performed before discharge from hospital by senior house officers or registrars but may also be performed by senior midwives or advanced nurse practitioners.

Some of the participating units, in certain circumstances, defer the clinical examination, to be performed by the GP up to 10 days after birth. All GPs in the catchment areas for the participating hospitals will be informed of the study through a flyer, together with posters for the waiting room and supplies of the information leaflet. Where GPs may be called upon to perform the clinical examination, further information regarding the study will be provided in the mother's postnatal notes, together with copies of the data collection form and envelopes to return the information to the hospital.

4.3. The reference standard

The reference standard will be a combination of the following approaches.

4.3.1 Echocardiography

Echocardiography is a resource intensive procedure, and will only be undertaken on those with symptoms or signs at clinical examination (as is current practice) or those with low PO readings. Given an assumed performance of 75% sensitivity and 99.5% specificity, and an overall prevalence of 5 per 1000, in our design, we would expect 9 per 1000 to undergo echocardiography based on low saturation by PO or symptoms apparent at clinical examination and between 40-50% of these will have clinically significant CHD. Thus we anticipate between 200 and 300 extra echocardiographic examinations during the study. To enable the participating hospitals to cope with this increased demand, a specially appointed clinical research fellow will be able to visit centres to perform echocardiograms when the hospital cannot accommodate them themselves.

4.3.2 Congenital anomalies registry

At periodic intervals after recruitment, the regional congenital anomalies register (CAR) and mortality register will be queried using the diagnostic codes for cardiac abnormalities. Information retrieved will include the baby date of birth, location, date and nature of the diagnosis, any known outcome, including death and the baby's NHS number We will also collect details on the diagnosis of Down's syndrome or any other congenital anomaly. Other identification details will be stripped to enable data transfer compliant with the Data Protection Acts. This will be sent to the Study Office to be cross-referenced with the babies already on the study database, matching for Baby and Mother NHS numbers and baby date of birth. For those that match, diagnostic information will be transferred to the study database for analysis. Unmatched babies will either have been missed at study hospitals, been born at non-participating hospitals within the region or have been born outside the region. These figures will help inform the study of completeness of verification. Additionally, registry holders in adjacent regions (e.g. Mersey and Trent) will be asked to be vigilant for notification of CHD cases from the West Midlands region.

A similar process will occur at the Birmingham Children's Hospital, which is the tertiary referral centre for the entire region. As the study progresses, if it appears that not all cases are being collected through either of these two methods, regional primary care trusts may be approached to search their databases using diagnostic READ codes. This will be substantially more labour intensive, so methods to maximise identification of CHD through the CAR and Children's Hospital sources will be explored first.

4.4. Compliance and follow-up issues

The issues of compliance and follow-up in diagnostic accuracy studies are of a different nature to those of randomised trials of interventions. In the systematic review of PO as a screening tool, none of the studies reported specific problems about parental non-compliance, i.e. declining participation after having given consent to take part. The key issue for us is the timing of the PO test to avoid loss of participation after consent has been obtained. The study has been designed to tackle this issue by performing the test in the first few hours after birth, on admission to the postnatal ward, which has the added advantage of early detection of life-threatening cardiac disease prior to the onset of symptoms, a situation that can worsen the prognosis. Although it not possible to anticipate how many babies presenting late with CHD will not be captured by the study, it is not anticipated that this will be above 5%, given the measures described in section 4.3. This figure would be estimated from the number of mis-matches arising from the CAR and presented at the interim analysis to provide an opportunity to re-calculate sample size, if required.

4.5. Quality Control

Quality assurance of testing will begin with a clearly documented staff training programme. A register of staff who have been trained, and their competence assessed will be maintained, and only staff whose names appear on this list will be permitted to undertake testing. Staff will also receive regular update training, and periodic reassessment of their competence.

4.6. Serious and unexpected adverse events

There are no foreseeable risks of mortality or significant morbidity associated with testing. Every effort will be made to minimise any risk through training. All **serious adverse events**[*] believed to be associated with the study should be reported by fax to the Study Office as soon as possible. This report should be followed within 2 weeks by a completed SAE form.

4.7. Other management at discretion of local doctors

All other aspects of patient management are entirely at the discretion of the local doctors. Mother and child are managed in whatever way appears best for them, with no special treatments and no extra follow-up visits.

5. OUTCOME MEASURES

5.1. Protection from bias

There are many possible sources of bias in accuracy studies (20) and these have recently been highlighted in the STARD statement (21). Selection bias may arise if the sample is not suitably representative of the population. This is likely to occur with use of non-consecutive or convenience sampling. The study will seek to recruit all consecutive eligible newborns. A related issue is that of spectrum bias whereby the accuracy of tests varies among study samples with differences in disease spectrum i.e. prevalence of CHD. It is possible that centres that perform fetal echocardiography may diagnose more newborns antenatally, thus affecting the spectrum composition. Information on PO testing and clinical examination of symptomatic newborns with suspected CHD will be collected and those newborns that were antenatally diagnosed to have CHD. Sensitivity analysis will explore the variation in test accuracy due to spectrum composition.

The PO testing procedure will be standardised in the first phase of the study and criteria for interpretation and thresholds have been determined *a priori* (7).

The use of different reference standards for test positive and test negative cases is not ideal but this is the most practicable way to verify presence or absence of disease. Empirical studies have shown that studies with differential verification produce more biased estimates of accuracy than studies with complete verification by the preferred reference standard, particularly when differential verification is not pre-specified in the design or completely at random. The direction and magnitude of bias is likely to depend on whether differential verification will lead to different detection rates of CHD under different reference standards. If complete verification by the preferred reference is not possible and different reference standards have to be used, the best approach is to incorporate differential verification in the design (22), the approach used in this study. With this design, the estimates of positive and negative predictive values under the two reference standards, the key parameters for decision analytic modelling, will escape any bias or interpretive problems. The estimates of sensitivity and specificity will also escape any

---

[*] For the purposes of this study, "**serious**" adverse events are those occurring in either subjects or testers which are fatal, life-threatening, disabling or require some form of medical or surgical treatment.

difficulties in interpretation as the focus will be only on clinically significant CHD which is likely to be detected equally well by both reference standards.

One of the conclusions that arose from the recent systematic review is the need for a large well conducted study, aimed at improving the precision of the sensitivity estimate of the test. The sensitivity of PO as a screening test for CHD varied between 25% (95% CI, 13% to 41%) (12) and 98.5% (95% CI, 91.8% to 100%) (15) in previously conducted studies. The performance of the test, particularly estimates of sensitivity depends on the absolute number of patients with disease. Thus there is imprecision of the sensitivity estimates due to the low prevalence of CHD in relatively small studies. This study has been planned to be large enough to achieve good precision, as shown in section 6.1.

5.2. Primary outcome measures

These will be the accuracy of each index tests for detection of CHD, expressed as sensitivity, specificity, positive and negative predictive values and positive and negative likelihood ratios, together with 95% confidence intervals.

5.3. Secondary outcome measures

5.3.1 Clinical Examination

The presence or absence of irregularities in heart rate, rhythm and sounds or murmurs and suspicion of cyanosis observed through clinical examination will be noted. The date, time and grade of person performing the test will be recorded.

5.3.2 Assessment of patient acceptability

We define acceptability broadly in this study to incorporate the psychological impact of screening on parents. We will examine acceptability of pulse oximetry to parents using a structured questionnaire. The questionnaire will be designed specifically for the study incorporating standardised measures as appropriate. It will be administered as soon as is practicable after testing to limit recall bias whilst being sensitive to the distress of parents. As far as possible, the questionnaires will be administered before mother and baby are discharged. Otherwise, if appropriate, the questionnaire will be sent to the mother by post. Mothers will not be approached until the local clinician/midwife feels it is appropriate and face to face or telephone interviews will be offered as appropriate. Mother's address and telephone number will be collected as necessary.

Parents of all babies who screen positive, i.e. all true positive and false positive cases, will receive the questionnaire or a structured interview based on the questionnaire. All parents of children with CHD who screen negative, i.e. the false negative cases, will also be approached at a later stage. Finally a sample of the largest group will be approached: parents whose babies do not have CHD and who screen negative, i.e. the true negative cases. The inclusion of all four groups of parents will enable the differentiation of the impact of screening *per se* from that of the outcome of screening.

In designing the questionnaire it will be important to maximise the precision and validity of the instrument. As this is not a randomised-controlled trial, no comparison across testing procedures will be carried out where maximum differences in scores would be expected. Therefore the instrument must measure acceptability as accurately as possible so that any differences between sub-groups (e.g. socio-economic, ethnic) can be discerned. In order to maximise face and content validity, preliminary focus group discussions will be held with:

- parents of children with CHD diagnosed postnatally
- experienced midwives and paediatricians
- representatives of paediatric cardiology support groups (e.g. Little Hearts Matter)

From these, a set of items will be derived which will seem relevant to the participants and cover all the areas thought to be important by participants.

Pilot testing will be carried out to make certain the questionnaire is usable. It is anticipated that the questionnaire will measure acceptability and satisfaction in the following areas:

- the procedure(s) for testing
- the information provided when consent is obtained
- processes for giving test results.

Parents will also be asked whether they would be prepared to have PO testing after future births, or recommend it to others, since these give the most reliable measures of satisfaction. In addition, comparisons will be made using all data across social, ethnic, age and parity groups to see if testing is acceptable across all groups and, if it is not, the areas where acceptability is low and negative impact is high, so that procedures may be reviewed.

5.3.2.1. Measuring Distress

Psychological distress will be assessed across the populations using standardised instruments, so that levels of distress can be compared to population norms and to levels measured in other screening studies. Specifically we will use:

- State anxiety form of the Spielberger State-Trait Anxiety Inventory (STAI).   This instrument has been widely used and validated in a number of clinical studies on screening (23;24). It is a short, closed format questionnaire of current anxiety levels.
- Depression subscale of the Hospital Anxiety and Depression Scales (HADS). Again this instrument has been widely used and validated in clinical studies as a measure of mood, including previous screening research (23). Some of this literature has been reviewed in a systematic review (25).

In order to understand variations in acceptability and the impact of testing on distress levels, measures will also be taken of possible explanatory variables including:

- Illness perceptions of CHD which will be measured using a brief version of the Revised Illness Perceptions Questionnaire (26).
- General optimism which will be measured using the 2 item short form of the Life Orientation Test (27).

5.3.2.2. Assessment of parents who receive false positive results

The literature on antenatal and neonatal screening suggests that the effects of false positives may extend over considerable periods of time in some circumstances (28-31). In part this may be due to the information and support provided for such parents (32;33). However other work contradicts this and suggests the high levels of anxiety are dispelled by diagnostic tests showing the baby is healthy. Given the contradictions in the literature, parents receiving false positive results will be assessed after testing and at the one year follow up. This will give us an indication of both the degree and duration of distress caused by the outcome.   If there is significant change over this period, a second assessment will also give a better indication of the acceptability of PO screening to this group. Given that some of the literature on false positives results suggests an impact on the parent-child bond, a parental bonding questionnaire will be administered at both time points.

5.3.2.3. Assessment of parents who receive false negative results

Parents who have received false negative results will not be definitively identified until the 12 month follow-up stage of the project. So, although some may have been included in

the sample of parents receiving negative results, the main assessment of these parents will occur at that stage and no formal comparison with their immediate post-testing perceptions will be possible within the time-frame and scope of the project.

5.3.2.4. Assessment of acceptability to health professionals

The acceptability of PO as a routine procedure to health professionals will be addressed by holding focus groups with midwives, midwifery assistants and other staff who carry out the test (for example nursery nurses and health care assistants) at Birmingham Women's Hospital and one of the other study centres after they have had experience of this (see Appendix H for information sheet). Issues to be explored will include perceptions of the efficacy of testing, costs to staff in terms of time and effort involved, perceived benefits of testing, professional views on the impact on parents. The discussions of the focus groups will be recorded, transcribed and analysed using inductive thematic analysis.

To address the acceptability of PO to neonatologists, echocardiographers, registrars and senior house officers (SHOs), staff at all study centres will be invited to complete open-ended questionnaires administered by e-mail (see Appendix G for information sheet). Focus groups for these groups would not be feasible because of the rotation of posts for registrars and SHOs. Issues to be explored will be similar to those addressed in the focus groups (see Appendix F for questionnaire; precise wording/details may be amended following further discussion with professionals and piloting). In order to allow participants to respond to the thoughts of their peers, an anonymised summary of responses will be compiled and emailed to participants, with the invitation to add further thoughts. Participants will not be identifiable to each other; all participants will reply to the researcher and not to each other. Responses will be analysed using inductive thematic analysis.

5.4. Additional information sought

In addition to the PO and clinical examination information, a minimal demographic and clinical dataset will be collected (see Appendix D). Some additional data will be collected postnatally. This will include baby birth weight and mother ethnicity. If there was an antenatal diagnosis or suspicion of CHD, limited information on the ultrasound findings will be extracted from the maternity notes. The baby's NHS number will be used as the primary identifier. All babies are issued with unique NHS numbers and these are used to track individuals throughout the NHS, making it a safe identifier. Addresses will not be collected on all mothers for the purposes of the study, to reduce the amount of identifiable data transferred out of the hospital. Addresses and telephone numbers will be collected for only a small sample of mothers for the administration of the acceptability questionnaires (see section 5.3.2)

5.5. Health economic outcomes

5.5.1 Perspective and cost data collection

If PO screening is shown to be an effective adjunct to the standard practice of use of clinical examination to screen for CHD in newborn babies, then it is likely that important cost implications will be seen for the health care sector. For example, PO may detect additional cases of abnormality compared to standard clinical examination alone which increase the number of echocardiograms required and increase the number of cases of infants requiring cardiac surgery (i.e. babies that could have had surgery but died before this was possible). But the additional associated costs of early diagnosis and treatment may lead to a reduction in costs associated with undetected CHD and avoided complications. Given this, the economic evaluation will take the perspective of the NHS.

Resource use data will be collected to estimate the costs associated with the additional use of PO in screening newborn babies. Data on NHS resource use will be prospectively collect for a sub-sample of the study. The main resources to be monitored include:

1. the additional time for the procedure of PO screening and consultation/explanation, compared to current practice, during pregnancy (principally midwife and paediatrician time),
2. the equipment and resources associated with PO and knock-on costs associated with additional echocardiograms,
3. time and resources associated with clinical examination
4. neonatal cardiac surgery
5. admissions to neonatal intensive care.

Information on unit costs or prices will then be required to attach to each resource item in order that an overall cost per infant can be calculated. Cost data will be collected from two principal sources. First, the primary PO test accuracy study will provide the time (staff and equipment) and other resource use data to estimate cost incurred in administering the PO test and the knock on tests. Primary cost data for many of these resources will be collected from the participating hospital sites. Where possible other cost data, such as cost of midwife time etc to carry out the test will be collected from routine sources, including Netten *et al* (34) and hospital finance departments. Many cost data are already available in recently published sources. A study to investigate the costs of different levels of neonatal intensive care has already been carried out (35) and other cost studies with relevant costs and costs associated with pre term delivery are available to supplement these (36;37). Also recent systematic review and cost-effectiveness analysis on Newborn Screening For Congenital Heart Defects has been recently published (6). This study retrieved available literature costs and some primary costs which can be used as a comparison.

5.6. Data management and validation

5.6.1 Confidentiality of personal data

The study will collect personal data and sensitive information about the participating infants. Participants will be informed about the transfer of this information to the Pulse Ox study office at the University of Birmingham Clinical Trials Unit (BCTU) and will be asked to consent to this. Participant demographic data, test results and questionnaire answers will be stored on a secure server, inputted where possible via the internet using secure socket layer encryption technology. Remaining data will be returned by post to the BCTU. The use of the baby's NHS number will minimise the risk of disclosure of identifiable data. Data to be processed outside the BCTU, or cross-referenced with CARs, will be anonymised. Only registered study personnel will have access to the database.

All participant data will be processed and stored according to the MRC guidelines of use of personal data. All personal information obtained for the study will be held securely and treated as confidential. All staff, at the hospitals, in the community midwife teams or at the BCTU, share the same duty of care to prevent unauthorised disclosure of personal information. No data that could be used to identify an individual will be published.

5.6.2 Long-term storage of data

All participant data will be stored on computer for 20 years after recruitment in accordance with MRC guidelines on the archiving of personal medical data for research.

6. ACCRUAL AND ANALYSIS

6.1. Sample size

The approach used in the sample size calculations is to consider the power that a study of a certain size has for the lower limits of the confidence intervals for both sensitivity and specificity to exceed particular values (and hence prove that the test is statistically significantly superior to those values). Sample size computations have been undertaken assuming that the screening strategy will have a sensitivity of 75% and a specificity of 99.5%. Sample sizes for CHD prevalence between 2 and 5 per 1000 using a one-tailed significance level of 2.5% have been computed.

Simulations of the study design and analysis (using 10,000 repetitions) were performed to account for (1) sampling variability in the observed number of cases of CHD and (2) sampling variability in the observed sensitivity and specificity, and (3) to incorporate the use of binomial exact methods to estimate confidence intervals. Standard asymptotic sample size computations do not allow for all three of these issues to be considered simultaneously. Plots of power against the values of sensitivity and specificity that the study wishes to rule out were constructed for sample sizes of 10000, 15000, 20000, 25000, 30000 and 35000.

Based on these simulations it was observed that a sample size of 20,000 will have 80% power to demonstrate that sensitivity is above 61% at a prevalence of 5 per 1000 in the overall sample. For the subgroup of women who are not selectively screened antenatally, it is anticipated the prevalence to be around 2 per 1000, a level at which sensitivity of 52% can be detected with 80% power. If the sample size is increased to 25,000 the study will have 80% power to prove the sensitivity to be above 63% and 55% at 5 and 2 per 1000 CHD prevalences respectively; if it is increased to 30,000 the corresponding values are 64% and 57%. Variation of prevalence does not affect the power to detect differences in specificity. For sample sizes above 15,000 the study will have greater than 90% power to prove that the specificity is above 99.3%.

Further simulations using lower assumed values of the sensitivity and specificity for the screening strategy were undertaken. Statistical power to test absolute differences in sensitivity and specificity of the same magnitude was noted to be comparable to the above scenario for assumed values are within a reasonable range of those used above. Thus a sample size of 25,000 will have 80% power to prove that sensitivity is higher than a value 12% (prevalence 5 per 1000) or 20% (prevalence of 2 per 1000) below the assumed sensitivity of the test, and the specificity is above a value 0.2% below the assumed specificity.

The assumption about disease prevalence is quite conservative. A review of 62 prevalence studies since 1955 estimated a combined prevalence of 6 per 1000 live births of moderate to severe CHD (38). If prevalence in this study is higher than 0.5%, it will have more than the projected power to undertake subgroup analyses confidently. On this basis, considering the likely prevalence, the study aims to recruit a sample of 20,000 neonates.

6.2. Projected accrual and attrition rates

Accrual and attrition rates will be closely monitored against our target, and in the unlikely event that recruitment is insufficient, the Study Management Group have identified other maternity units likely to be able to participate.

6.3. Analysis for test accuracy study

The main analysis for diagnostic accuracy will include estimation of sensitivity, specificity, predictive values, likelihood ratios and their 95% confidence intervals. The baseline

characteristics of the patients enrolled in the study will be examined and planned subgroup analyses will be undertaken. Subgroup analyses are limited by statistical power and can produce spurious results particularly if many are undertaken. The recent literature review and consultation with obstetricians suggests that the accuracy of the index tests may vary according to risk factors (high risk history or antenatal suspicion on ultrasound) and timing of PO test. Therefore, secondary analyses will be limited to these subgroups only. The main and subgroup analyses are powered as outlined above based on conservative prevalence estimates. All estimates of accuracy for subgroups will be interpreted with appropriate caution.

As secondary analyses, the reference standard information obtained from echocardiography will be utilised to compare antenatal screening with PO. Such studies are commonplace in cancer research where invasive tests cannot be undertaken on those with low probability of disease. From such a design, it is possible to estimate the relative true (TPR) and false (FPR) positive rates of the two test strategies, but not the absolute sensitivity or specificity of either(39). It is also possible to estimate the trade-off between additional true positives and false positives related to the addition of PO to the standard screening programme (40). The statistical significance of the difference will be assessed using McNemar's test for paired data, and confidence intervals for the ratios computed using the methods of Cheng and Macalouso (41). The same approach can be utilised for comparing PO and clinical examination except that this analysis will be biased due to lack of blinding between the two tests. A "latent class" analysis, as described by Walter (42), will be used to estimate the sensitivity and specificity of the PO test and the antenatal screening from the screen positive study using only echocardiography as a reference standard. This approach involves an assumption of independence of test errors in the diseased and non-diseased groups, which will be tested as part of the investigation.

Using multivariable logistic regression analysis, predictive probabilities will be generated for various combinations of history, antenatal tests and oximetry results (43) Historical features to be included in the regression model will include gestational age (35-37, >37 weeks' gestation), antenatal scan findings, clinical examination (although this is likely to be affected by workup bias) and family history of CHD. In statistical terms, logistic modelling will aim to derive a diagnostic regression function, i.e. probability of CHD given test result. The analysis will be performed with presence or absence of CHD verified by reference standard as the binary dependent (outcome) variable. The models will allow a direct estimation of the post-test-combination disease probabilities that is needed for decision-making and for decision-analysis. Models of varying complexity may be compared through the familiar receiver operating characteristic (ROC) analyses. More importantly, the clinical situation where some information is already acquired, such as antenatal ultrasound or presentation with symptoms prior to undertaking PO, will be mirrored. In this way, for various index test results conditional disease probabilities will be generated directly taking into account any overlap of information that may exist between tests. This approach evaluates the extent to which the findings of the index tests add value to the babies' presentation. Its output is transparent, and is likely to enable production of simple clinical algorithms based on probabilities. The advantages of tackling diagnostic problems with logistic regression modelling are well known (44;45). The limitation associated with the regression approach lies mainly in its generalisability to other data sets or clinical practices. The recommended techniques, such as bootstrapping to enhance generalisability and estimate the amount of shrinkage will be applied for model validation (46;47). It is anticipated that the sample will comfortably meet the recommended events per variable rule to avoid overfitting the models (48-50) even if some data were missing. In a sensitivity analysis, missing data will be estimated by multiple imputation and maximum-

17

likelihood methods, as appropriate, to explore the potential bias and reduced statistical power associated with listwise deletion (51).

6.4. Handling missing data

Sensitivity analysis will be employed to explore the potential bias and reduced statistical power associated with listwise deletion of missing data, using multiple imputation and maximum-likelihood methods, as appropriate (52).

6.5. Economic analysis

There will be two components to the analysis: a within study analysis and a model-based analysis, the latter will seek to refine and develop the decision tree model used in the recently published HTA report by Knowles *et al*. (6) The model based analysis will allow projection of costs and benefits beyond the immediate screening study data. Data from the follow up assessment carried out at one year on those who screened positive will be available from the study. Data will be sought from the anomaly register for those who initially screened negative and were not further followed up by the study. The accuracy data on screening based on the PO, and the clinical examination will be collected directly from the current study. There is a potential source of bias arising from the fact that the discrete clinical examination results may not be blind to the results of the PO. However, effort will be made to interpret these data appropriately and record any known potential bias. The data collected in this study will refine the detection rate and other aspects of PO which can be used in an already existing model.

The model will consider treatment over total duration for an infant screened positive by PO and will include consideration of medical and/or surgical treatments provided in the longer term. The model-based analysis will adopt a short term outcome of 'cost per timely diagnosis of life-threatening CHD' and an outcome of cost per death avoided at one year to coincide with the final follow up. Depending on the data availability from published sources, the model outcome may be extended beyond the study outcome of one year. However, given that at present there is no consensus regarding the methodology for developing QALYs in children, a cost utility analysis will not be attempted (6).

6.5.1 Within study analysis

This will use only data collected within the accuracy study and so, for example, will draw upon the test performance data. Estimates of costs and benefits will therefore relate only to the period of follow-up, and no predictions for costs and benefits beyond the study will be made. The data available for this analysis will be patient-specific resource use and costs. Given the skewness inherent in most cost data and the concern of economic analyses with mean costs, we shall use a bootstrapping approach in order to calculate confidence intervals around the difference in mean costs (53;54). An incremental economic analysis will be conducted. The base-case analysis will be framed in terms of cost-consequences, reporting data in a disaggregated manner on the incremental cost and the important consequences, including data on the number of true positive cases of CHD detected, etc.

Three main strategies will be compared:

- Routine clinical examination alone

- PO screening as an adjunct to clinical examination

- PO alone

18

6.5.2 Discounting

If the outcome of the model coincides with that of the study, i.e. at one year, then discounting is not required. But if the model extends beyond the outcome of the study and given the potentially relatively long time horizons being considered in these analyses, many of the costs (and benefits) will be incurred (and experienced) in future years. Using discounting, adjustments will be made to reflect this differential timing. The base-case analysis will follow Treasury recommendations for public sector projects.

6.5.3 Presentation of results and sensitivity analysis

The results of these economic analyses will be presented using cost-effectiveness acceptability curves to reflect sampling variation and uncertainties in the appropriate threshold cost-effectiveness value. Both simple and probabilistic sensitivity analyses will be used to explore the robustness of these results to plausible variations in key assumptions and variations in the analytical methods used, and to consider the broader issue of the generalisability of the results.

7. DATA ACCESS AND QUALITY ASSURANCE

7.1. In-house Data Quality Assurance

7.1.1 Monitoring and Audit

Midwifery staff performing the PO tests will be trained by the lead research midwife, who will review their skills regularly. A sample of test results inputted in the maternity units will be cross-checked at the BCTU with paper records.

7.1.2 Statistical monitoring throughout the study

Real-time reports will be available to postnatal staff indicating missing test and questionnaire data for all participants at that centre. This will be supplemented by regular reminders from the Study Office for incomplete data. The study statistician will report on recruitment, compliance and completeness of verification to the Steering Committee quarterly.

7.2. Independent Supervision of the Study

The Study Steering Committee provides independent supervision for the study, providing advice to the investigators and the Sponsor on all aspects of the study and affording protection for patients by ensuring the study is conducted according to the MRC Guidelines for Good Clinical Practice in Clinical Trials.

If the clinical co-ordinators are unable to resolve any concern satisfactorily, collaborators, and all others associated with the study, may write through the study office to the chair of the SSC, drawing attention to any concerns they may have about the possibility of distortion of clinical practice, or of particular categories of patient requiring special study, or about any other matters thought relevant.

The study shall follow and comply with the MRC Guidelines on Good Clinical Practice, although its advice in relation to test accuracy studies is limited. The Study Team has made provisional recommendations regarding the independent supervision and data monitoring of test accuracy studies as a consequence of experiences in previous studies (55). One such recommendation is that, if desirable, the independent Data Monitoring and Ethics Committee (DMEC) should be formed as a sub-committee of the Study Steering Committee (SSC). For the purposes of this study, the SSC shall convene and nominate a three member independent DMEC from within its membership, that shall not include study researchers.

7.3. Data Monitoring and Ethics Committee: determining when clear answers have emerged

If the PO test has acceptable sensitivity and specificity, using echocardiography and clinical examination as a reference, then this may become apparent before the target recruitment has been reached. The assumed prevalence of CHD may prove to be inaccurate and require a recalculation of the sample size. Alternatively, the PO test may be found to be unworkable, new evidence of the effectiveness of the test might emerge from other sources or new technologies may be introduced to the market.

To protect against this, at 3-4 months into recruitment to the study, interim analyses of major endpoints will be supplied to the DMEC along with updates on results of other related studies, and any other analyses that the DMEC may request. The DMEC will determine whether the assumptions underpinning the sample size are correct at 3-4 months after commencement of recruitment. In particular they will be asked to examine the proportion of babies showing reduced oxygen saturation by PO and those who were diagnosed by antenatal scanning. The interim analysis will also determine if the principal question on index test accuracy has been answered and will monitor adverse events. The combined SSC/ DMEC will decide if the accuracy of the tests shows both (a) "proof beyond reasonable doubt"* that one particular test is definitely superior or definitely inferior in terms of a net difference in the major endpoints, and (b) evidence that might reasonably be expected to influence the patient management of many clinicians. The SSC/DMEC can then decide whether to close or modify any part of the study. Unless this happens, however, the SSC, the collaborators and all of the central administrative staff (except the statisticians who supply the confidential analyses) will remain unaware of the interim results.

8. ORGANISATION AND RESPONSIBILITIES

The Chief Investigator is responsible for the management, central co-ordination of clinical and administrative aspects of the study, compliance with the Research Governance Framework and management of study budget. Relevant ethics committee and Trust research governance approval will be coordinated centrally for efficiency and speed.

All investigators are responsible for ensuring that the research they undertake follows the agreed protocol, for helping care professionals to ensure that participants receive appropriate care while involved in research, for protecting the integrity and confidentiality of clinical and other records and data generated by the research, and for reporting any failures in these respects, adverse reactions and other events or suspected misconduct through the appropriate systems.

8.1. Centre eligibility

Initially, six hospitals in six NHS Trusts will recruit women into the study. These centres will receive equipment for the PO test from the study and will be funded to employ a part-time research midwife to conduct the study. Other centres wishing to participate can do so provided their Trust will supply the above resources.

8.2. Local Co-ordinator at each centre

Each Trust has a designated Consultant Neonatologist to act as Principal Investigator and bear responsibility for the conduct of research at their centre. Close collaboration between

---

* Appropriate criteria of proof beyond reasonable doubt cannot be specified precisely, but a difference of at least three standard deviations in an interim analysis of a major endpoint may be needed to justify halting, or modifying, the study prematurely. If this criterion were to be adopted, it would have the practical advantage that the exact number of interim analyses would be of little importance, so no fixed schedule is proposed.

midwifery and clinical teams is particularly important in order to ensure that recruitment of babies is maximised. The responsibilities of the Principal Investigators will be to ensure that all medical and post-natal ward staff involved are well informed about the study. This will involve distributing protocols and patient information sheets to all relevant staff, displaying publicity material where it is likely to be read, and contributing to the regular newsletters. The Principal Investigators should liaise with the study administrator on logistic, data collection and administrative matters connected with the study.

8.3. Midwifery Co-ordinator at each centre

Each participating centre will have a designated research midwife who will act as Local Midwifery Coordinator. This person would be responsible for ensuring that all eligible babies are considered for the study, that patients are provided with study information sheets, and have an opportunity to discuss the study if required. The midwife will be responsible for the organisation of data collection and will be the first point of contact for data queries.

8.4. The Study Office

The Study Office at the University of Birmingham Clinical Trials Unit (BCTU) is responsible for providing all study materials, including the coded stickers and questionnaires. Additional supplies of any printed material can be obtained on request. The Study Office is also responsible for collection and checking of data (including reports of serious adverse events) and for analyses. The Study Office will help resolve any local problems that may be encountered in study participation.

8.5. Research Governance

The conduct of the study will be according to the principles of MRC Guidelines for Good Clinical Practice in Clinical Trials (1998) and the appropriate NHS Research Governance Frameworks.

All centres will be required to sign an Investigator's Agreement, detailing their commitment to accrual, compliance, Good Clinical Practice, confidentiality and publication. Deviations from the agreement will be monitored and the SSC will decide whether any action needs to be taken, e.g. withdrawal of funding, suspension of centre.

The Study Office will ensure researchers not employed by an NHS organisation who interact with individuals in a way that has direct bearing on the quality of their care hold an NHS honorary contract for that organisation.

8.6. Regulatory and Ethical Approval

Site specific approval from local research ethics committees (LREC) will be gained for each site following Multi-centre Research Ethics Committee (MREC) approval. The LREC and Trust Research and Development Office will assess each site for "locality issues" relating to their population, the investigators, the facilities and resources.

8.7. Funding and Cost implications

The research costs of the study are funded by a grant from the NHS R&D Health Technology Assessment Unit awarded to the University of Birmingham.

The study has been designed to minimise extra 'service support' costs for participating hospitals, with no extra visits to hospital and no extra tests. Additional costs associated with the study should be minimal. These costs should be met by accessing the Trust's budget.

8.8. Indemnity

There are no special arrangements for compensation for non-negligent harm suffered by patients as a result of participating in the study. The study is not an industry-sponsored

study and so ABPI/ABHI guidelines on indemnity do not apply. The normal NHS indemnity liability arrangements for research detailed in HSG96(48) will operate in this case.

However, it should be stressed that in terms of negligent liability, NHS Trust hospitals have a duty of care to a patient being treated within their hospital, whether or not that patient is participating in a clinical study. Apart from defective products, legal liability does not arise where there is non-negligent harm. NHS Trusts may not offer advance indemnities or take out commercial insurance for non-negligent harm.

8.9. Publication

A meeting will be held after the end of the study to allow discussion of the main results among the collaborators prior to publication. The success of the study depends entirely on the wholehearted collaboration of a large number of doctors, midwives and others. For this reason, chief credit for the main results will be given not to the committees or central organisers but to all those who have collaborated in the study. Collaborators will be permitted to publish data obtained from participants in the Pulse Ox Study that use study outcome measures but do not relate to the study objectives.

9. REFERENCE LIST

1. Wren C, Richmond S, Donaldson L 2000 Temporal variability in birth prevalence of cardiovascular malformations. Heart 83:414-419

2. Death Registrations in England and Wales, 2002: causes 2003 Health Stat Q 18:57-64

3. Wren C, O'Sullivan JJ 2001 Survival with congenital heart disease. Heart 85:438-443

4. Rasiah SV, Publicover M, Ewer AK, Khan KS, Kilby MD, Zamora J 2006 A systematic review of the accuracy of first-trimester ultrasound examination for detecting major congenital heart disease. Ultrasound Obstet Gynecol 28:110-116

5. Bricker L, Garda J, Henderson J, Mugford M, Neilson J, Roberts T, Martin M-A 2000 Ultrasound screening pregnancy: a systematic review of the clincial effectiveness, cost effectiveness and women's views. Health Technol Assess 4:

6. Knowles R, Griebsch I, Dezateux C, Brown J, Bull C, Wren C 2005 Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis. Health Technol Assess 9:1-168

7. Thangaratinam S, Daniels JP, Ewer AK, Zamora J, Khan KS 2006 The accuracy of Pulse Oximetry in Screening for Congenital Heart Disease in Asymptomatic Newborns: A Systematic Review. Archives of disease in childhood In press:

8. Bakr AF, Habib HS 2005 Combining pulse oximetry and clinical examination in screening for congenital heart disease. Pediatric Cardiol 26:832-835

9. Arlettaz R, Bauschatz A, Mankhoff M, Essers B, Bauersfeld U 2006 The contribution of pulse oximetry to the early detection of congenital heart disease in newborns. Eur J Pediatrics 165:94-98

10. Rosati E, Chitano G, Dipaola L, De F, Latini G 2005 Indications and limitations for a neonatal pulse oximetry screening of critical congenital heart disease. Journal of perinatal medicine 33:455-457

11. Koppel RI, Druschel C, Carter T, Goldberg B, Mehta P, Talwar R, Bierman F 2003 Effectiveness of pulse oximetry screening for congenital heart disease in asymptomatic newborns. Pediatrics 111:451-455

12. Richmond S, Reay G, Abu H 2002 Routine pulse oximetry in the asymptomatic newborn. Archives of disease in childhood Fetal and neonatal edition 87:F83-F88

13. Reich J, Miller S, Brogdon B, Casatelli J, Gompf T, Huhta J, Sullivan K 2003 The use of pulse oximetry to detect congenital heart disease. J Pediatrics 142:268-272

14. Hoke TR, Donohue PK, Bawa PK, Mitchell RD, Pathak A, Rowe PC, Byrne BJ 2002 Oxygen saturation as a screening test for critical congenital heart disease: a preliminary study. Pediatric Cardiol 23:403-409

15. de Wahl Granelli A, Mellander M, Sunnegardh J, Sandberg K, Ostman-Smith I 2005 Screening for duct-dependant congenital heart disease with pulse oximetry: a critical evaluation of strategies to maximize sensitivity. Acta Paediatrica 94:1590-1596

16. Bove EL, Bull C, Stark J, de Leval M, MaCartney FJ, Taylor JF 1983 Congenital heart disease in the neonate: results of surgical treatment. Arch Dis Child 58:137-141

17. Guyatt GH, Tugwell PX, Feeny DH, Haynes RB, Drummond M 1986 A framework for clinical evaluation of diagnostic technologies. CMAJ 134:587-594

18. Fryback DG, Thornbury JR 1991 The efficacy of diagnostic imaging. Med Decis Making 11:88-94

19. Randall P, Brealey S, Hahn S, Khan KS, Parsons JM 2005 Accuracy of fetal echocardiography in the routine detection of congenital heart disease among unselected and low risk populations: a systematic review. BJOG 112:24-30

20. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, Bossuyt PM 1999 Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 282:1061-1066

21. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG 2003 The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem 49:7-18

22. Rutjes AWS, Reitsma JB, Coomarasamy A, Khan KS, Bossuyt PM 2006 Methodological research in diagnostic tests: evaluation of diagnostic tests when there is no gold standard. Health Technol Assess in press:

23. Michie S, Bobrow M, Marteau TM 2001 Predictive genetic testing in children and adults: a study of emotional impact. J Med Genet 38:519-526

24. Watkin PM, Baldwin M, Dixon R, Beckman A 1998 Maternal anxiey and attitudes to universal neonatal hearing screening. Br J Audiol 32:27-37

25. Herrmann C 1997 International experiences with the HADS - review of validation data and clinical results. J Psychosomatic Research 42:17-41

26. Broadbent E, Petrie KJ, Main J, Weinman J 2006 The brief illness perception questionnaire. J Psychosomatic Research 60:631-637

27. Scheier MF, Carver CS 1985 Optimism, coping and health: assessment and implications of generalized outcome expectancies. Health Psychology 4: 219-247

28. Green J, Statham H 1993 Testing for fetal abnormality in routine antenatal care. Midwifery 9:124-135

29. Pollitt RJ, Green A, McCabe CJ, Booth A, Cooper NJ, Leonard JV 1997 Neonatal screening for inborn errors of metabolism: cost, yield and outcome. Health Technol Assess 1:

30. McCrindle BW, Shaffer KM, Kan JS, Zahka KG, Rowe SA, Kidd L 1995 An evaluation of parental concerns and misperceptions about heart murmurs. Clin Pediatr (Phila) 34:25-31

31. Fyro K, Bodegard G 1987 Four-year follow-up of psychological reactions to false positive screening tests for congenital hypothyroidism. Acta Paediatr Scand 76:107-114

32. Baillie C, Smith J, Hewison J, Mason G 2000 Ultrasound screening for chromosomal abnormality: women's reactions to false positive results. Br J Health Psychol 5:377-394

33. Farrell M, Certain L, Farrell P 2001 Genetic counseling and risk communication services of newborn screening programs. Arch Pediatr Adolesc Med 155:120-126

34. Netten A, Knight J 1999 Annuitizing the human capital investment costs of health service professionals. Health Economics 8:

35. O'Neill C, Malek M, Mugford M, Normand C, Tarnow-Mordi WO, Hey E, Halliday HL 2000 A cost analysis of neonatal care in the UK: Results from a multicentre study. Journal of Public Health Medicine 22:

36. Petrou S, Mehta Z, Hockley C, Cook-Mozaffari P, Henderson J, Goldacre M, Petrou S, Mehta Z, Hockley C, Cook-Mozaffari P, Henderson J, Goldacre M 2003 The impact of preterm birth on hospital inpatient admissions and costs during the first 5 years of life. Pediatrics 112:1290-1297

37. Petrou S, Petrou S 2003 Economic consequences of preterm birth and low birthweight. BJOG: An International Journal of Obstetrics & Gynaecology 110 Suppl 20:17-23

38. Hoffman JIE, Kaplan S 2002 The incidence of congenital heart disease. J Am Coll Cardiol 39:1890-1900

39. Schatzkin A, Connor RJ, Taylor PR, Bunnag B 1987 Comparing new and old screening tests when a reference procedure cannot be performed on all screenees. Am J Epidemiol 125:672-678

40. Chock C, Irwig L, Berry G, Glasziou P 1997 Comparing dichotomous screening tests when individuals negative on both tests are not verified. J Clin Epidemiol 50:1211-1217

41. Cheng H, Macalouso M 1997 Comparison of the accuracy of two tests with a confirmatory procedure limited to positive results. Epidemiology 8:104-106

42. Walter SD 1999 Estimation of test sensitivity and specificity when disease confirmation is limited to positive results. Epidemiology 10:67-72

43. Bachmann LM, ter Riet G, Clark TJ, Gupta JK, Khan KS 2003 Probability analysis for diagnosis of endometrial hyperplasia and cancer in postmenopausal bleeding: an approach for a rational diagnostic workup. Acta Obstet Gynecol Scand 82:564-569

44. Dawid AP 1976 Properties of diagnostic data distributions. Biometrics 32:647-658

45. Miettinen OS, Henschke CI, Yankelevitz DF 1998 Evaluation of diagnostic imaging tests: diagnostic probability estimation. J Clin Epidemiol 51:1293-1298

46. Harrell FE, Jr., Lee KL, Pollock BG 1988 Regression models in clinical studies: determining relationships between predictors and response. J Natl Cancer Inst 80:1198-1202

47. Steyerberg EW, Eijkemans MJ, Van Houwelingen JC, Lee KL, Habbema JD 2000 Prognostic models based on literature and individual patient data in logistic regression analysis. Stat Med 19:141-160

48. Concato J, Feinstein AR, Holford TR 1993 The risk of determining risk with multivariable models. Ann Intern Med 118:201-210

49. Khan KS, Chien PF, Dwarakanath LS 1999 Logistic regression models in obstetrics and gynecology literature. Obstet Gynecol 93:1014-1020

50. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR 1996 A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49:1373-1379

51. Sinharay S, Stern HS, Russell D 2001 The use of multiple imputation for the analysis of missing data. Psychol Methods 6:317-329

52. Sinharay S, Stern HS, Russell D 2001 The use of multiple imputation for the analysis of missing data. Psychol Methods 6:317-329

53. Briggs A, Gray A 1998 The distribution of health care costs and their statistical analysis for economic evaluation. Journal of Health Services & Research Policy 3:

54. Thompson SG, Barber JA 2000 How should cost data in pragmatic randomised trials be analysed? British Medical Journal 320:29

55. Thangaratinam S, Daniels JP, Gilbert R, Cummins C, Kleijnen J, Khan KS 2007 Independent monitoring of test accuracy studies. In preparation

APPENDIX A PATIENT INFORMATION SHEET:

Pulse Ox STUDY   PULSE OXIMETRY AS A   SCREENING TEST

FOR CONGENITAL HEART DISEASE IN NEWBORN BABIES

We would like to invite you to take part in a research study that may help newly born babies. The decision to take part will be yours, you do not have to join in or give us the reason why you choose not to. It is important for you to understand why the research is being done and what it will involve. Discover more by reading the following information. Feel free to spend time discussing it with anyone - it may help to talk to your midwife, or a member of the research team. You will have plenty of time to decide whether or not to take part.

What is the purpose of the study?

• Congenital Heart Disease (CHD), where the baby's heart is abnormal at birth, is the most common group of abnormalities in newborn babies and affects 7-8 per 1000 births.

• If some heart problems are not diagnosed early, they can very quickly cause the baby to become unwell

• Timely recognition of these heart defects is vital in order to improve outcome.

• Currently in the UK, all newborn babies undergo a routine examination, usually in the first 24 hours after birth, during which, among other things, a careful assessment of the heart is undertaken. ***However, it is estimated that over half of all babies with CHD will not be picked up by this examination.***

• Pulse Oximetry is a simple test which has been used routinely used on babies in the Neonatal Unit for many years.

• Pulse Oximetry measures the amount of oxygen being carried around in the blood by shining a special light through the skin of the babies' hand and foot. ***It is completely harmless and painless and takes only a couple of minutes to perform.***

• This test will pick up babies who do not have as much oxygen in their blood as they should and this is a common finding in the early stages of Congenital Heart Disease

• This hospital is now hoping to screen every newborn baby with pulse oximetry as a way of trying to detect those with heart problems before they become unwell.

• We do not know for certain how accurate this method will be at detecting heart problems - that is why we are undertaking this research study.

• **We do know that it is important to identify babies who have low oxygen levels and find out what is causing this.**

Commonly asked questions

Why have I been chosen?

All women being booked in for antenatal care at this hospital are being invited to take part. It is hoped 20,000 women from seven hospitals will take part in the study.

Do I have to take part?

It is up to you to decide whether or not to take part. You should keep this information sheet and consent form in your hand-held notes as you will be asked at your antenatal clinic

whether you are willing to take part. If you agree, you will be asked to sign the consent form. At the time of testing, you will be asked again if you still agree to participate. You are free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to be tested, will not affect the standard of care you receive.

What will happen to me if I take part?

If you agree to take part, we will perform the pulse oximetry test on your baby. We will do this before s/he is 24 hours old, preferably at around 3-6 hours of age. Two readings will be taken: one from the hand and one from the foot. **The test will take no more than five minutes to perform.**

We know from our experiences with babies on the neonatal unit what level of oxygen in the blood is acceptable.

If your baby has acceptable oxygen levels no further action will be taken and your baby will be managed as usual.

If your baby's oxygen measurement is below this level it may be for one of the following reasons.

1. The baby's circulation is still readjusting to the process of birth - this is quite common. The baby does not have a problem with the heart or lungs.
2. Your baby has a problem with the lungs which means not enough oxygen is getting into the body.
3. Your baby has a problem with the heart which means that not enough oxygen is getting around the body.

If your baby's oxygen measurement remains low then s/he will be carefully examined for signs of a problem with the heart or the lungs.

If we find a problem or if the oxygen measurement does not improve, then the baby will need further tests to find the cause – this will include a scan of the heart (echocardiogram) to look for congenital heart disease. Your doctor will explain exactly what this means for your baby and what will happen. We will use data from the echocardiogram and clinical follow up, and up to one year after birth, we will search regional databases for all babies with heart disease. We will compare this information with data from the oxygen measurement.

A low oxygen measurement does not necessarily mean your baby has a problem. It just means that they are more likely to – further tests will help us identify which babies have problems and which do not. If your baby has a low oxygen measurement and then the follow up tests are normal, this means your baby's heart is normal.

However, an acceptable oxygen measurement does not *completely exclude* other problems including heart disease.

What else do I have to do?

We (the researchers) would also like a sample of mothers to answer some questions of acceptability. You may be given an anonymous questionnaire to complete before you leave hospital. We want to find out the how you found the tests and the research study, the information given to you before the test, and the way you were given your test results.

What are the possible disadvantages and risks of taking part?

We have lots of experience with this test in babies on the Neonatal Unit and we do not expect there to be any problems or risks to the babies who take part.

What are the possible benefits of taking part?

There may be no benefit from taking part, however, we hope that a heart or lung problem would be detected early which would ensure that treatments for the disease could be started earlier.

Also, of course, the information we get from this study may in the future help us better treat newborn babies.

What if something goes wrong?

We do not believe that there is a risk of anything going wrong. However if your baby is harmed by taking part in this study, there are no special compensation arrangements. If your baby is harmed due to someone's negligence, then you may have grounds for a legal action but you may have to pay for it. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms should be available to you.

Will my taking part in this study be kept confidential?

Yes, all information collected in the study will remain strictly confidential in the same way as your other medical records. If you agree to take part, your midwife or doctor will send basic information about you and your baby to the study's central organisers. This information will be put into a computer and analysed. The questionnaires will be identified only by a code number and will not be seen by your doctor or midwife. All information will be held securely and in strict confidence. No named information about you will be published in the study report. Occasionally, inspections of clinical study data are undertaken to ensure that, for example, all participants have given consent to take part. But, apart from this, only the study organisers will have access to the data.

What will happen to the results of the research study?

The study will last for around two and a half years, after which we expect to publish the results in scientific journals.

Who is organising and funding the research?

The Pulse Ox study is funded by a grant from the National Health Service's Health Technology Assessment programme. The central study organisers are based at the University of Birmingham. The Clinical Trials Unit at the University of Birmingham will collect and analyse the data. The researchers, doctors and midwives involved are not being paid for recruiting women into the study. We cannot pay women to take part either, but we will be very grateful for their help in finding out more about the accuracy of this new test.

Do you have any other questions?

Having read this leaflet, we hope that you will choose to take part in the Pulse Ox Study. If you have any questions about the study now or later, feel free to ask your midwife. Their name and telephone numbers are given on the front of this leaflet.

What if I have any concerns?

If you have any concern or other questions about this study or the way it has been carried out, you should contact the investigator [name], or you may contact the [name] hospital complaints department.

APPENDIX B: PATIENT CONSENT FORM:

Pulse Ox STUDY

SCREENING TEST FOR CONGENITAL HEART DISEASE IN NEWBORN BABIES

I confirm that I have read and understand the information sheet (version 1.1, dated 22/06/2007) for the above study and have had the opportunity to ask questions.

☐ (Please initial)

I understand what is involved in the Pulse Ox study and agree to participate. I intend to participate in the study, but I understand that I am free to change my mind when I go into hospital without necessarily giving a reason. If I do withdraw, I can continue to expect the highest standard of care from my doctor or midwife.

☐ (Please initial)

I understand that the information will be used for medical research only and that I will not be identified in any way in the analysis and reporting of the results. I understand that sections of any of my medical notes may be looked at by responsible individuals from the University of Birmingham or from regulatory authorities where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.

☐ (Please initial)

DURING PREGNANCY

I consent to participate in the Pulse Ox Study

☐ (Please initial)

AT TIME OF PO TESTING

I still agree to participate in the Pulse Ox Study (midwife to indicate here)

☐

Name of Participant	Date	Signature

Name of Person taking consent	Date	Signature

Copies should be kept in hand-held notes until after delivery, then top copy should be kept in the mother's notes, pink copy put in the study collection box and the yellow copy given to the mother to keep.

APPENDIX C: SERIOUS ADVERSE EVENT FORM

Please report any **serious**, **unexpected** adverse events[2] believed to be due to the treatments given as part of the Pulse Ox study by sending or faxing the following details to the Pulse Ox Study Office (fax: 0121-415-9136) within 2 weeks of the event:

Patient's Full Name: ....

Date of Birth: .... Hospital Number: ....

Responsible doctor: ....

Date Treatment Started:(if known) ....

Date Event Started: .... Date Event Ceased: ....

Outcome (e.g. fatal, recovered, continuing): ....

Details of Adverse Event (please attach copies of relevant reports) ....

....

....

....

Did the event require or prolong hospitalisation? ....

Please give reasons why you consider the event to be treatment-related: ....

....

....

Name of Person Reporting (please print) ....

Telephone Number: .... Today's Date: ....

---

[2] For the purposes of this study, "**serious**" adverse events are those which are fatal, life-threatening, disabling or require hospitalisation. "**Unexpected**" adverse experiences are defined as those that would not be expected as a result of PO testing. It is not required to report in this way side-effects or events that might reasonably be expected.

APPENDIX D: PULSE OX STUDY
DATA COLLECTION FORM

pulse♥x

Section 1: Eligibility

Is the baby well?   Yes ☐   No ☐

*If the baby is **well**, they are **eligible** for the study. Please ensure consent has been taken and continue with **sections 2 – 5 below.***

*If the baby is **not well**, they are **not eligible** for the study. Please complete the '**Ineligibility Form'**, put in the participant pack and put it in the Pulse Ox collection box. Treat the baby as normal.*

Section 2: Identifying Details

Baby Hospital Number ☐☐☐☐☐☐☐☐☐☐

Baby **NHS** number ☐☐☐☐☐☐☐☐☐☐

Baby date of birth   ____ / ____ / ____

        Time (24hr) ____ : ____

Male ☐   Female ☐

Gestational age ____ /40 weeks

Mother's Forename ————————————

Mother's Surname ————————————

Hospital Number ☐☐☐☐☐☐☐☐☐☐

Mother's NHS Number ☐☐☐☐☐☐☐☐☐☐

Mother's date of birth ____ / ____ / ____

(or attach sticker)

Mother's gravida _____ parity _____

Section 3: Language

Was a translator and/or a translated version of the Information Sheet and Consent required?   Yes ☐   No ☐

If yes, specify language: ————————————————

Section 4: Pulse Oximetry Saturations

First reading (This should be taken within 24 hours of age or before discharge)

Performed in: Delivery Suite ☐   Post Natal Unit ☐   Neonatal Unit ☐

Performed by:

      PRINT NAME: ————————————

Doctor ☐   Nurse ☐   Midwife ☐   Midwifery Assistant ☐   Care Assistant ☐   Neonatal nurse ☐

Date	Time	Foot sats (%)	Right hand sats (%)	% difference	Pass/Refer

*If the result is normal (i.e. **Pass**), nothing further needs to be done. Please **ensure sections 1-5 are completed**. Please put this form in the participant pack and put it in the Pulse Ox collection box.*

*If less than 95% in either limb or ≥3% difference between limbs, this is a **Refer**, the baby needs to be examined/ reviewed by someone trained in neonatal examination. **Please continue with section 6 overleaf.***

Section 5: Antenatal Diagnosis

Is there a family history of congenital heart defects?   Yes ☐   No ☐

Was CHD suspected on antenatal ultrasound?   Yes ☐   No ☐

pulseOx

The following sections only need to be completed if the first pulse oximetry reading was <95% in either limb or ≥3% difference between limbs i.e. a <u>Refer</u>.

Section 6: Clinical Examination

Undertaken by: Nurse/ Midwife ☐   SHO/ Registrar ☐   Consultant ☐   Other ☐

How long did the exam take?   _____ minutes

Did the baby show signs of CHD at examination?   Yes ☐   No ☐

       If yes, was this   murmur ☐   cyanosis ☐   abnormal pulses ☐

Overall exam finding   Normal ☐   Abnormal ☐   Details:_____

*If **exam is normal**, a second PO reading needs to be taken. This should be taken 1-2 hours after the first reading. **Please complete section 7 below.***

*If **exam is abnormal,** the baby needs to be referred straight to the Neonatal team for further investigation and management. A second PO reading does not need to be taken. **Please continue with sections 9-14 overleaf.***

Section 7: Second Pulse Oximetry Reading

Performed in: Delivery Suite ☐   Post Natal Unit ☐   Neonatal Unit ☐

Performed by:
      PRINT NAME: _____

Doctor ☐   Nurse ☐   Midwife ☐   Midwifery Assistant ☐   Care Assistant ☐   Neonatal nurse ☐

Date	Time	Foot sats (%)	Right hand sats (%)	% difference	Pass/Refer

*If less than 95% in either limb or ≥3% difference between limbs, this is a <u>**Fail**</u>, the baby needs to be referred to the Neonatal team for further investigation and management. **Please continue with sections 8-14 overleaf.***

*If the result is normal (i.e. <u>**Pass**</u>), nothing further needs to be done. Please ensure all above sections are completed. Please put this form in the participant pack and put it in the Pulse Ox collection box.*

pulse⊙x

The following sections only need to be completed if the clinical examination found an abnormality or if the second pulse oximetry reading was a __Fail__ and the baby has been referred to the Neonatal team for further investigation and management.

Section 8: Hyperoxia Test *(only to be completed if second PO reading was a fail)*

Performed by: Nurse ☐   SHO ☐   Registrar ☐   Consultant ☐

Did saturation increase to 95% or above?   Yes ☐   No ☐

Section 9: Clinical Symptoms

If baby showed signs of CHD at birth, describe

Collapse ☐   Cyanosis ☐   Acidosis ☐   Respiratory distress ☐

Did this baby have a respiratory illness or any illness other than CHD?

Yes ☐   No ☐   If yes details: _____

Section 10: Echocardiography

Date of echocardiography   ____ / ____ / ____   Time ____ : ____   (24 Hour)

Echocardiographer: Cardiology Research Fellow ☐   Other Echocardiographer ☐

Result: Normal ☐   Abnormal ☐   Uncertain ☐

    Details: _____

Section 11: Review of Echocardiography

Reviewed by: Cardiology Research Fellow ☐   Consultant ☐   Other ☐   (Please tick more than one box if applicable)

Date of review   ____ / ____ / ____

Result: Normal ☐   Abnormal ☐   Uncertain ☐

    Details: _____

*If there is a **disagreement between the two results** __or__ the **result of the review is uncertain**, a second echocardiogram needs to be performed. This will be done by a Consultant Paediatric Cardiologist at Birmingham Children's Hospital*

Date of second echocardiography   ____ / ____ / ____

Result: Normal ☐   Abnormal ☐   Details: _____

Please continue with sections 12-14 overleaf.

Section 12: Admission Summary

Number of days on Postnatal unit? ———————— days

Number of days on Neonatal unit? ———————— days

Section 13: Discharge Details

Home ☐   Cardiac Unit ☐   Other ☐   Details: _____

If home, date of discharge _____ / _____ / _____   Time of discharge (24hr) _____ : _____

If cardiac unit, number of days on unit _____ days

Section 14: Death Details

Death ☐   Date of death _____ / _____ / _____   Time of death (24hr) _____ : _____

Cause of Death: Cardiac ☐   Non-cardiac ☐

If cardiac: Post-operative ☐   Other ☐   If other, please specify: _____

If non-cardiac, please specify: _____

pulseOx

PULSE OX STUDY

CONGENITAL HEART DEFECT CASE REPORT FORM

Date of Diagnosis _____ /_____ /_____

How was the CHD detected and confirmed?

Pulse oximetry at birth then echocardiogram ☐

Clinical examination at birth then echocardiogram ☐

Post-mortem ☐

Later echocardiogram by paed cardiologist ☐

Congenital anomaly registry ☐

Type of CHD?

Aortic (valve) stenosis	☐	Persistent (patent) ductus arteriosis	☐
Atrial septal defect	☐	Tetrology of Fallot	☐
Coartation of aorta	☐	Total anolomous pulmonary venous connection	☐
Complete artioventricular septal defect	☐	Transposition of great arteries	☐
Hypoplastic left heart syndrome	☐	Ventricular septal defect	☐
Interruption of aortic arch	☐		

Other (specify) _____

Has the baby had cardiac surgery? Yes ☐   No ☐   Planned ☐

If yes, when was the surgery _____ / _____ / _____

Type of surgery? _____

Death ☐   Date of death _____/_____/_____   Time of death (24hr) _____ : _____

Cause of Death: Cardiac ☐   Non-cardiac ☐

If cardiac: Post-operative ☐   Other ☐

If non-cardiac, please specify: _____

APPENDIX E: DEFINITIONS OF ECHOCARDIOGRAPHIC FINDINGS

Normal   No echocardiographic abnormalities

Non- Significant   No clinical signs (e.g. murmur, thrill, pulse abnormalities, hepatic enlargement)

No symptoms

Small patent ductus arteriosis (PDA) or small inter-atrial communication (patent foramen ovale (PFO), atrial septal defect (ASD) or muscular ventricular septal defect (VSD))

Mildly abnormal turbulence at branch Pulmonary Artery

Echocardiographic findings no longer detected at 6 months

Significant   Small patent ductus arteriosis (PDA) or patent foramen ovale (PFO) or muscular ventricular septal defect (VSD)

Mildly abnormal turbulence at branch Pulmonary Artery

Above where echocardiographic findings persist for longer than 6 months of age

Any cardiac lesion which requires regular monitoring beyond 6 months or drug treatment but does not fall into serious or critical category.

Serious   Any cardiac lesion not defined as critical which requires intervention (cardiac catheterisation or surgery) within 1 year of age

Critical   All infants with hypoplastic left heart, pulmonary atresia with intact ventricular septum, simple transposition of the great arteries or interruption of the aortic arch and all infants dying or requiring surgery within the first 28 days of life with the following conditions: coarctation of the aorta, aortic valve stenosis, pulmonary valve stenosis, tetralogy of Fallot, pulmonary atresia with ventricular septal defect, or total anomalous pulmonary venous connection.

APPENDIX F: ACCEPTABILTY QUESTIONNAIRE: NEONATOLOGY GROUP

- What is your role? (if your job has changed, please give your role at the time of the study?) (tick box(es)).

  Neonatologist (consultant)   □

  Echocardiographer   □

  Senior House Officer   □

   Registrar   □

- To what extent do you think that pulse oximetry has been worthwhile?
- How effective do you think pulse oximetry testing is?
- How did the pulse oximetry assessment affect your role?
- How do you think the women felt about having taken part in the study?
  - o To what extent do you think they felt it was worthwhile?
- Did you encounter any incidents where there were problems in connection with having had the test carried out? (yes/no)
  - o If so, please give details below.
- Did you encounter any incidents where having the test proved beneficial? (yes/no)
  - o If so, please give details below.
- How would you feel if the study showed pulse oximetry to be beneficial and working with test results became a routine part of your job?
- What do you think the benefits of pulse oximetry have been?
- What do you think the costs of pulse oximetry have been?
- How your thoughts about pulse oximetry have changed over time:
  - o How did you feel about the test at the start of the study? (or the start of your placement)
  - o How do you feel about the test now? What is it that makes you feel that way?
- In general, how do you feel about the use of monitoring equipment for new-born babies?

APPENDIX G: STAFF EMAIL SURVEY
INFORMATION SHEET

What is the purpose of the study?

In the Pulse Ox study, more than 20,000 new born babies were screened for congenital heart disease using pulse oximetry. Pulse Ox aims to determine the accuracy and cost effectiveness of the test, and to assess the psychosocial effect of the screening for parents. It is important to understand the experience of health care professionals, such as those receiving tested babies in neonatology, whose work has been affected by the study and to determine whether the test is acceptable to them also.

Why have I been chosen?

You have been invited to take part in this email survey because, in your role, you are likely to have encountered babies who had the pulse oximetry test as part of the Pulse Ox study. We are inviting staff in the neonatology units of all six hospitals taking part in the Pulse Ox study to complete this survey.

What will happen to me if I take part?

You will be asked to complete a short survey by email. The survey will include items such as how pulse oximetry affected your role and whether or not you consider the test to be beneficial. The researcher (Rachael Powell) will collate a summary of the comments raised. This summary will be emailed back to you and you will be invited to comment on its contents. For example, if someone has raised an issue that you had not considered, or that you disagree with, you will be able to respond.

Summaries will be created for groups of staff with similar roles. So, for example, if you are a Senior House Officer (SHO), your responses will be summarised with other SHOs' comments and will be emailed only to the SHO group for further comment.

Only the researcher (Rachael Powell) will know who made which comment; no comment will be directly emailed to the group.

Are there any potential risks in talking part in the study?

It is not expected that there are any risks to you in taking part in the study.

Do I have to take part?

Whether or not you take part is completely up to you. If you do choose to take part, you are free to withdraw at any time, without giving a reason.

Will my taking part in this study be kept confidential?

Taking part in this study is completely confidential. None of your colleagues will know who made which comment. Only members of the research team will have access to returned surveys. When we write up the results it will not be possible to identify anything that you tell us as being about you. All data will be kept in a locked filing cabinet or in password-protected computer files. Data will be stored at Aston University for 10 years and then destroyed.

What will happen to the results of the research study?

When the study is complete, the findings will be written up as part of the Pulse Ox study report to the Health Technology Assessment programme. They will also be written up and submitted for publication in scientific journals. **You will not be identifiable in any of these documents.**

Who is organising and funding the research?

The Pulse Ox study is funded by a grant from the National Institute of Health Research Health Technology Assessment programme to researchers at the University of Birmingham and Aston University. This email survey is being organised by Dr Rachael Powell and Dr Helen Pattison from Aston University, as part of the Pulse Ox study.

Who do I Contact if I need Further Information?

For further information about the email survey contact:

Dr Rachael Powell
Psychology
School of Life and Health Sciences
Aston University
Aston Triangle
Birmingham, B4 7ET.

Email: r.k.powell@aston.ac.uk
Phone: 0121 204 4188

APPENDIX H: FOCUS GROUP STAFF INFORMATION SHEET

What is the purpose of the study?

In the Pulse Ox study, more than 20,000 new born babies were screened for congenital heart disease using pulse oximetry. Pulse Ox aims to determine the accuracy and cost effectiveness of the test, and to assess the psychosocial effect of the screening for parents. It would not have been possible to screen these babies without the support of the staff who carried out the tests. It is important to understand the experience of health care professionals whose work has been affected by the study and to determine whether the test is acceptable to them also.

Why have I been chosen?

You have been invited to take part in this group discussion because you have been involved in carrying out the test on new born babies.

What will happen to me if I take part?

You will be asked to take part in one group discussion. The other people in the group will be colleagues who have also carried out the test. The researchers (Dr Rachael Powell and Dr Helen Pattison) will ask you to discuss your experiences of using the test, how it has affected your role at work and your views about testing neonates more generally. The discussion will last up to one hour and will be recorded on a digital voice recorder.

Are there any potential risks in talking part in the study?

It is not expected that there are any risks to you in taking part in the study.

Do I have to take part?

Whether or not you take part is completely up to you. If you do choose to take part, you are free to withdraw at any time, without giving a reason.

Will my taking part in this study be kept confidential?

Taking part in this study is completely confidential. Only members of the research team will listen to the recording or see the discussion transcripts. When we write up the results it will not be possible to identify anything that you tell us as being about you. All recordings and transcripts will be kept in a locked filing cabinet or in password-protected computer files. The voice recordings will be transcribed anonymously. As soon as we have written up the results, the voice recording will be destroyed and only the anonymous transcripts will be kept. These will be stored at Aston University for 10 years and then destroyed also.

What will happen to the results of the research study?

When the study is complete, the findings will be written up as part of the Pulse Ox study report to the Health Technology Assessment programme. They will also be written up and submitted for publication in scientific journals. **You will not be identifiable in any of these documents.**

Who is organising and funding the research?

The Pulse Ox study is funded by a grant from the National Institute of Health Research Health Technology Assessment programme to researchers at the University of Birmingham and Aston University. These group discussions are being organised by Dr Rachael Powell and Dr Helen Pattison from Aston University, as part of the Pulse Ox study.

Who do I Contact if I need Further Information?

For further information about the discussion groups, contact:

Dr Rachael Powell
Psychology
School of Life and Health Sciences
Aston University
Aston Triangle
Birmingham, B4 7ET.

Email: r.k.powell@aston.ac.uk
Phone: 0121 204 4188

Health Technology Assessment programme

Director,
Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology,
University of Liverpool

Deputy Director,
Professor Hywel Williams,
Professor of Dermato-Epidemiology,
Centre of Evidence-Based Dermatology,
University of Nottingham

Prioritisation Group

Members

Chair,
Professor Tom Walley, CBE,
Director, NIHR HTA
programme, Professor of Clinical
Pharmacology, University of
Liverpool

Professor Imti Choonara,
Professor in Child Health,
Academic Division of Child
Health, University of Nottingham
Chair – Pharmaceuticals Panel

Dr Bob Coates,
Consultant Advisor – Disease
Prevention Panel

Dr Andrew Cook,
Consultant Advisor – Intervention
Procedures Panel

Dr Peter Davidson,
Director of NETSCC, Health
Technology Assessment

Dr Nick Hicks,
Consultant Adviser – Diagnostic
Technologies and Screening Panel,
Consultant Advisor–Psychological
and Community Therapies Panel

Ms Susan Hird,
Consultant Advisor, External
Devices and Physical Therapies
Panel

Professor Sallie Lamb,
Director, Warwick Clinical Trials
Unit, Warwick Medical School,
University of Warwick
Chair – HTA Clinical Evaluation
and Trials Board

Professor Jonathan Michaels,
Professor of Vascular Surgery,
Sheffield Vascular Institute,
University of Sheffield
Chair – Interventional Procedures
Panel

Professor Ruairidh Milne,
Director – External Relations

Dr John Pounsford,
Consultant Physician, Directorate
of Medical Services, North Bristol
NHS Trust
Chair – External Devices and
Physical Therapies Panel

Dr Vaughan Thomas,
Consultant Advisor –
Pharmaceuticals Panel, Clinical
Lead – Clinical Evaluation Trials
Prioritisation Group

Professor Margaret Thorogood,
Professor of Epidemiology, Health
Sciences Research Institute,
University of Warwick
Chair – Disease Prevention Panel

Professor Lindsay Turnbull,
Professor of Radiology, Centre for
the MR Investigations, University
of Hull
Chair – Diagnostic Technologies
and Screening Panel

Professor Scott Weich,
Professor of Psychiatry, Health
Sciences Research Institute,
University of Warwick
Chair – Psychological and
Community Therapies Panel

Professor Hywel Williams,
Director of Nottingham Clinical
Trials Unit, Centre of Evidence-
Based Dermatology, University of
Nottingham
Chair – HTA Commissioning
Board
Deputy HTA Programme Director

HTA Commissioning Board

Chair,
Professor Hywel Williams,
Professor of Dermato-Epidemiology, Centre
of Evidence-Based Dermatology, University of
Nottingham

Deputy Chair,
Professor Jon Deeks,
Department of Public Health and
Epidemiology, University of Birmingham

Professor Tom Walley, CBE,
Professor of Clinical Pharmacology, Director,
NIHR HTA programme, University of
Liverpool

Members

Professor Ann Ashburn,
Professor of Rehabilitation and
Head of Research, Southampton
General Hospital

Professor Peter Brocklehurst,
Professor of Women's Health,
Institute for Women's Health,
University College London

Professor Jenny Donovan,
Professor of Social Medicine,
University of Bristol

Professor Jonathan Green,
Professor and Acting Head of
Department, Child and Adolescent
Psychiatry, University of
Manchester Medical School

Professor John W Gregory,
Professor in Paediatric
Endocrinology, Department of
Child Health, Wales School of
Medicine, Cardiff University

Professor Steve Halligan,
Professor of Gastrointestinal
Radiology, University College
Hospital, London

Professor Freddie Hamdy,
Professor of Urology, Head of
Nuffield Department of Surgery,
University of Oxford

Professor Allan House,
Professor of Liaison Psychiatry,
University of Leeds

Dr Martin J Landray,
Reader in Epidemiology, Honorary
Consultant Physician, Clinical
Trial Service Unit, University of
Oxford

Professor Stephen Morris,
Professor of Health Economics,
University College London,
Research Department of
Epidemiology and Public Health,
University College London

Professor Irwin Nazareth,
Professor of Primary Care and
Head of Department, Department
of Primary Care and Population
Sciences, University College
London

Professor E Andrea Nelson,
Professor of Wound Healing and
Director of Research, School of
Healthcare, University of Leeds

Professor John David Norrie,
Chair in Clinical Trials and
Biostatistics, Robertson Centre for
Biostatistics, University of Glasgow

Dr Rafael Perera,
Lecturer in Medical Statisitics,
Department of Primary Health
Care, University of Oxford

HTA Commissioning Board *(continued)*

Professor Barney Reeves,
Professorial Research Fellow
in Health Services Research,
Department of Clinical Science,
University of Bristol

Professor Martin Underwood,
Professor of Primary Care
Research, Warwick Medical
School, University of Warwick

Professor Marion Walker,
Professor in Stroke Rehabilitation,
Associate Director UK Stroke
Research Network, University of
Nottingham

Dr Duncan Young,
Senior Clinical Lecturer and
Consultant, Nuffield Department
of Anaesthetics, University of
Oxford

Observers

Dr Tom Foulks,
Medical Research Council

Dr Kay Pattison,
Senior NIHR Programme
Manager, Department of Health

HTA Clinical Evaluation and Trials Board

Chair,
Professor Sallie Lamb,
Director,
Warwick Clinical Trials Unit,
Warwick Medical School,
University of Warwick and Professor of
Rehabilitation,
Nuffield Department of Orthopaedic,
Rheumatology and Musculoskeletal Sciences,
University of Oxford

Deputy Chair,
Professor Jenny Hewison,
Professor of the Psychology of Health Care,
Leeds Institute of Health Sciences,
University of Leeds

Programme Director,
Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of
Clinical Pharmacology, University of Liverpool

Members

Professor Keith Abrams,
Professor of Medical Statistics,
Department of Health Sciences,
University of Leicester

Professor Martin Bland,
Professor of Health Statistics,
Department of Health Sciences,
University of York

Professor Jane Blazeby,
Professor of Surgery and
Consultant Upper GI Surgeon,
Department of Social Medicine,
University of Bristol

Professor Julia M Brown,
Director, Clinical Trials Research
Unit, University of Leeds

Professor Alistair Burns,
Professor of Old Age Psychiatry,
Psychiatry Research Group, School
of Community-Based Medicine,
The University of Manchester &
National Clinical Director for
Dementia, Department of Health

Dr Jennifer Burr,
Director, Centre for Healthcare
Randomised trials (CHART),
University of Aberdeen

Professor Linda Davies,
Professor of Health Economics,
Health Sciences Research Group,
University of Manchester

Professor Simon Gilbody,
Prof of Psych Medicine and Health
Services Research, Department of
Health Sciences, University of York

Professor Steven Goodacre,
Professor and Consultant in
Emergency Medicine, School of
Health and Related Research,
University of Sheffield

Professor Dyfrig Hughes,
Professor of Pharmacoeconomics,
Centre for Economics and Policy
in Health, Institute of Medical
and Social Care Research, Bangor
University

Professor Paul Jones,
Professor of Respiratory Medicine,
Department of Cardiac and
Vascular Science, St George's
Hospital Medical School,
University of London

Professor Khalid Khan,
Professor of Women's Health and
Clinical Epidemiology, Barts and
the London School of Medicine,
Queen Mary, University of London

Professor Richard J McManus,
Professor of Primary Care
Cardiovascular Research, Primary
Care Clinical Sciences Building,
University of Birmingham

Professor Helen Rodgers,
Professor of Stroke Care, Institute
for Ageing and Health, Newcastle
University

Professor Ken Stein,
Professor of Public Health,
Peninsula Technology Assessment
Group, Peninsula College
of Medicine and Dentistry,
Universities of Exeter and
Plymouth

Professor Jonathan Sterne,
Professor of Medical Statistics
and Epidemiology, Department
of Social Medicine, University of
Bristol

Mr Andy Vail,
Senior Lecturer, Health Sciences
Research Group, University of
Manchester

Professor Clare Wilkinson,
Professor of General Practice and
Director of Research North Wales
Clinical School, Department of
Primary Care and Public Health,
Cardiff University

Dr Ian B Wilkinson,
Senior Lecturer and Honorary
Consultant, Clinical Pharmacology
Unit, Department of Medicine,
University of Cambridge

Observers

Ms Kate Law,
Director of Clinical Trials,
Cancer Research UK

Dr Morven Roberts,
Clinical Trials Manager, Health
Services and Public Health
Services Board, Medical Research
Council

Current and past membership details of all HTA programme 'committees' are available from the HTA website (www.hta.ac.uk)

Diagnostic Technologies and Screening Panel

Members

Chair,
Professor Lindsay Wilson Turnbull,
Scientific Director of the Centre for Magnetic Resonance Investigations and YCR Professor of Radiology, Hull Royal Infirmary

Professor Judith E Adams, Consultant Radiologist, Manchester Royal Infirmary, Central Manchester & Manchester Children's University Hospitals NHS Trust, and Professor of Diagnostic Radiology, University of Manchester

Mr Angus S Arunkalaivanan, Honorary Senior Lecturer, University of Birmingham and Consultant Urogynaecologist and Obstetrician, City Hospital, Birmingham

Dr Diana Baralle, Consultant and Senior Lecturer in Clinical Genetics, University of Southampton

Dr Stephanie Dancer, Consultant Microbiologist, Hairmyres Hospital, East Kilbride

Dr Diane Eccles, Professor of Cancer Genetics, Wessex Clinical Genetics Service, Princess Anne Hospital

Dr Trevor Friedman, Consultant Liason Psychiatrist, Brandon Unit, Leicester General Hospital

Dr Ron Gray, Consultant, National Perinatal Epidemiology Unit, Institute of Health Sciences, University of Oxford

Professor Paul D Griffiths, Professor of Radiology, Academic Unit of Radiology, University of Sheffield

Mr Martin Hooper, Public contributor

Professor Anthony Robert Kendrick, Associate Dean for Clinical Research and Professor of Primary Medical Care, University of Southampton

Dr Nicola Lennard, Senior Medical Officer, MHRA

Dr Anne Mackie, Director of Programmes, UK National Screening Committee, London

Mr David Mathew, Public contributor

Dr Michael Millar, Consultant Senior Lecturer in Microbiology, Department of Pathology & Microbiology, Barts and The London NHS Trust, Royal London Hospital

Mrs Una Rennard, Public contributor

Dr Stuart Smellie, Consultant in Clinical Pathology, Bishop Auckland General Hospital

Ms Jane Smith, Consultant Ultrasound Practitioner, Leeds Teaching Hospital NHS Trust, Leeds

Dr Allison Streetly, Programme Director, NHS Sickle Cell and Thalassaemia Screening Programme, King's College School of Medicine

Dr Matthew Thompson, Senior Clinical Scientist and GP, Department of Primary Health Care, University of Oxford

Dr Alan J Williams, Consultant Physician, General and Respiratory Medicine, The Royal Bournemouth Hospital

Observers

Dr Tim Elliott, Team Leader, Cancer Screening, Department of Health

Dr Joanna Jenkinson, Board Secretary, Neurosciences and Mental Health Board (NMHB), Medical Research Council

Professor Julietta Patnick, Director, NHS Cancer Screening Programme, Sheffield

Dr Kay Pattison, Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE, Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells, Principal Research Officer, Policy Research Programme, Department of Health

Disease Prevention Panel

Members

Chair,
Professor Margaret Thorogood,
Professor of Epidemiology, University of Warwick Medical School, Coventry

Dr Robert Cook, Clinical Programmes Director, Bazian Ltd, London

Dr Colin Greaves, Senior Research Fellow, Peninsula Medical School (Primary Care)

Mr Michael Head, Public contributor

Professor Cathy Jackson, Professor of Primary Care Medicine, Bute Medical School, University of St Andrews

Dr Russell Jago, Senior Lecturer in Exercise, Nutrition and Health, Centre for Sport, Exercise and Health, University of Bristol

Dr Julie Mytton, Consultant in Child Public Health, NHS Bristol

Professor Irwin Nazareth, Professor of Primary Care and Director, Department of Primary Care and Population Sciences, University College London

Dr Richard Richards, Assistant Director of Public Health, Derbyshire County Primary Care Trust

Professor Ian Roberts, Professor of Epidemiology and Public Health, London School of Hygiene & Tropical Medicine

Dr Kenneth Robertson, Consultant Paediatrician, Royal Hospital for Sick Children, Glasgow

Dr Catherine Swann, Associate Director, Centre for Public Health Excellence, NICE

Mrs Jean Thurston, Public contributor

Professor David Weller, Head, School of Clinical Science and Community Health, University of Edinburgh

Observers

Ms Christine McGuire, Research & Development, Department of Health

Dr Kay Pattison, Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE, Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

External Devices and Physical Therapies Panel

Members

Chair,
Dr John Pounsford,
Consultant Physician North Bristol NHS Trust

Deputy Chair,
Professor E Andrea Nelson,
Reader in Wound Healing and Director of Research, University of Leeds

Professor Bipin Bhakta,
Charterhouse Professor in Rehabilitation Medicine, University of Leeds

Mrs Penny Calder,
Public contributor

Dr Dawn Carnes,
Senior Research Fellow, Barts and the London School of Medicine and Dentistry

Dr Emma Clark,
Clinician Scientist Fellow & Cons. Rheumatologist, University of Bristol

Mrs Anthea De Barton-Watson,
Public contributor

Professor Nadine Foster,
Professor of Musculoskeletal Health in Primary Care Arthritis Research, Keele University

Dr Shaheen Hamdy,
Clinical Senior Lecturer and Consultant Physician, University of Manchester

Professor Christine Norton,
Professor of Clinical Nursing Innovation, Bucks New University and Imperial College Healthcare NHS Trust

Dr Lorraine Pinnington,
Associate Professor in Rehabilitation, University of Nottingham

Dr Kate Radford,
Senior Lecturer (Research), University of Central Lancashire

Mr Jim Reece,
Public contributor

Professor Maria Stokes,
Professor of Neuromusculoskeletal Rehabilitation, University of Southampton

Dr Pippa Tyrrell,
Senior Lecturer/Consultant, Salford Royal Foundation Hospitals' Trust and University of Manchester

Dr Nefyn Williams,
Clinical Senior Lecturer, Cardiff University

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Interventional Procedures Panel

Members

Chair,
Professor Jonathan Michaels,
Professor of Vascular Surgery, University of Sheffield

Deputy Chair,
Mr Michael Thomas,
Consultant Colorectal Surgeon, Bristol Royal Infirmary

Mrs Isabel Boyer,
Public contributor

Mr Sankaran Chandra Sekharan,
Consultant Surgeon, Breast Surgery, Colchester Hospital University NHS Foundation Trust

Professor Nicholas Clarke,
Consultant Orthopaedic Surgeon, Southampton University Hospitals NHS Trust

Ms Leonie Cooke,
Public contributor

Mr Seumas Eckford,
Consultant in Obstetrics & Gynaecology, North Devon District Hospital

Professor Sam Eljamel,
Consultant Neurosurgeon, Ninewells Hospital and Medical School, Dundee

Dr Adele Fielding,
Senior Lecturer and Honorary Consultant in Haematology, University College London Medical School

Dr Matthew Hatton,
Consultant in Clinical Oncology, Sheffield Teaching Hospital Foundation Trust

Dr John Holden,
General Practitioner, Garswood Surgery, Wigan

Dr Fiona Lecky,
Senior Lecturer/Honorary Consultant in Emergency Medicine, University of Manchester/Salford Royal Hospitals NHS Foundation Trust

Dr Nadim Malik,
Consultant Cardiologist/Honorary Lecturer, University of Manchester

Mr Hisham Mehanna,
Consultant & Honorary Associate Professor, University Hospitals Coventry & Warwickshire NHS Trust

Dr Jane Montgomery,
Consultant in Anaesthetics and Critical Care, South Devon Healthcare NHS Foundation Trust

Professor Jon Moss,
Consultant Interventional Radiologist, North Glasgow Hospitals University NHS Trust

Dr Simon Padley,
Consultant Radiologist, Chelsea & Westminster Hospital

Dr Ashish Paul,
Medical Director, Bedfordshire PCT

Dr Sarah Purdy,
Consultant Senior Lecturer, University of Bristol

Dr Matthew Wilson,
Consultant Anaesthetist, Sheffield Teaching Hospitals NHS Foundation Trust

Professor Yit Chiun Yang,
Consultant Ophthalmologist, Royal Wolverhampton Hospitals NHS Trust

Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

Current and past membership details of all HTA programme 'committees' are available from the HTA website (www.hta.ac.uk)

Pharmaceuticals Panel

Members

Chair,
Professor Imti Choonara,
Professor in Child Health,
University of Nottingham

Deputy Chair,
Dr Yoon K Loke,
Senior Lecturer in Clinical
Pharmacology, University of East
Anglia

Dr Martin Ashton-Key,
Medical Advisor, National
Commissioning Group, NHS
London

Dr Peter Elton,
Director of Public Health, Bury
Primary Care Trust

Dr Ben Goldacre,
Research Fellow, Epidemiology
London School of Hygiene and
Tropical Medicine

Dr James Gray,
Consultant Microbiologist,
Department of Microbiology,
Birmingham Children's Hospital
NHS Foundation Trust

Dr Jurjees Hasan,
Consultant in Medical Oncology,
The Christie, Manchester

Dr Carl Heneghan,
Deputy Director Centre for
Evidence-Based Medicine and
Clinical Lecturer, Department of
Primary Health Care, University
of Oxford

Dr Dyfrig Hughes,
Reader in Pharmacoeconomics
and Deputy Director, Centre for
Economics and Policy in Health,
IMSCaR, Bangor University

Dr Maria Kouimtzi,
Pharmacy and Informatics
Director, Global Clinical Solutions,
Wiley-Blackwell

Professor Femi Oyebode,
Consultant Psychiatrist and Head
of Department, University of
Birmingham

Dr Andrew Prentice,
Senior Lecturer and Consultant
Obstetrician and Gynaecologist,
The Rosie Hospital, University of
Cambridge

Ms Amanda Roberts,
Public contributor

Dr Gillian Shepherd,
Director, Health and Clinical
Excellence, Merck Serono Ltd

Mrs Katrina Simister,
Assistant Director New Medicines,
National Prescribing Centre,
Liverpool

Professor Donald Singer,
Professor of Clinical
Pharmacology and Therapeutics,
Clinical Sciences Research
Institute, CSB, University of
Warwick Medical School

Mr David Symes,
Public contributor

Dr Arnold Zermansky,
General Practitioner, Senior
Research Fellow, Pharmacy
Practice and Medicines
Management Group, Leeds
University

Observers

Dr Kay Pattison,
Senior NIHR Programme
Manager, Department of Health

Mr Simon Reeve,
Head of Clinical and Cost-
Effectiveness, Medicines,
Pharmacy and Industry Group,
Department of Health

Dr Heike Weber,
Programme Manager, Medical
Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA
programme, Professor of Clinical
Pharmacology, University of
Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy
Research Programme, Department
of Health

Psychological and Community Therapies Panel

Members

Chair,
Professor Scott Weich,
Professor of Psychiatry, University
of Warwick, Coventry

Deputy Chair,
Dr Howard Ring,
Consultant & University Lecturer
in Psychiatry, University of
Cambridge

Professor Jane Barlow,
Professor of Public Health in
the Early Years, Health Sciences
Research Institute, Warwick
Medical School

Dr Sabyasachi Bhaumik,
Consultant Psychiatrist,
Leicestershire Partnership NHS
Trust

Mrs Val Carlill,
Public contributor

Dr Steve Cunningham,
Consultant Respiratory
Paediatrician, Lothian Health
Board

Dr Anne Hesketh,
Senior Clinical Lecturer in Speech
and Language Therapy, University
of Manchester

Dr Peter Langdon,
Senior Clinical Lecturer, School
of Medicine, Health Policy and
Practice, University of East Anglia

Dr Yann Lefeuvre,
GP Partner, Burrage Road Surgery,
London

Dr Jeremy J Murphy,
Consultant Physician and
Cardiologist, County Durham and
Darlington Foundation Trust

Dr Richard Neal,
Clinical Senior Lecturer in General
Practice, Cardiff University

Mr John Needham,
Public contributor

Ms Mary Nettle,
Mental Health User Consultant

Professor John Potter,
Professor of Ageing and Stroke
Medicine, University of East
Anglia

Dr Greta Rait,
Senior Clinical Lecturer and
General Practitioner, University
College London

Dr Paul Ramchandani,
Senior Research Fellow/Cons.
Child Psychiatrist, University of
Oxford

Dr Karen Roberts,
Nurse/Consultant, Dunston Hill
Hospital, Tyne and Wear

Dr Karim Saad,
Consultant in Old Age Psychiatry,
Coventry and Warwickshire
Partnership Trust

Dr Lesley Stockton,
Lecturer, School of Health
Sciences, University of Liverpool

Dr Simon Wright,
GP Partner, Walkden Medical
Centre, Manchester

Observers

Dr Kay Pattison,
Senior NIHR Programme
Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health
Services and Public Health
Services Board, Medical Research
Council

Professor Tom Walley, CBE,
Director, NIHR HTA
programme, Professor of Clinical
Pharmacology, University of
Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy
Research Programme, Department
of Health

Expert Advisory Network

Members

Professor Douglas Altman,
Professor of Statistics in Medicine,
Centre for Statistics in Medicine,
University of Oxford

Professor John Bond,
Professor of Social Gerontology
& Health Services Research,
University of Newcastle upon Tyne

Professor Andrew Bradbury,
Professor of Vascular Surgery,
Solihull Hospital, Birmingham

Mr Shaun Brogan,
Chief Executive, Ridgeway
Primary Care Group, Aylesbury

Mrs Stella Burnside OBE,
Chief Executive, Regulation and
Improvement Authority, Belfast

Ms Tracy Bury,
Project Manager, World
Confederation of Physical Therapy,
London

Professor Iain T Cameron,
Professor of Obstetrics and
Gynaecology and Head of the
School of Medicine, University of
Southampton

Professor Bruce Campbell,
Consultant Vascular & General
Surgeon, Royal Devon & Exeter
Hospital, Wonford

Dr Christine Clark,
Medical Writer and Consultant
Pharmacist, Rossendale

Professor Collette Clifford,
Professor of Nursing and Head
of Research, The Medical School,
University of Birmingham

Professor Barry Cookson,
Director, Laboratory of Hospital
Infection, Public Health
Laboratory Service, London

Dr Carl Counsell,
Clinical Senior Lecturer in
Neurology, University of Aberdeen

Professor Howard Cuckle,
Professor of Reproductive
Epidemiology, Department
of Paediatrics, Obstetrics &
Gynaecology, University of Leeds

Professor Carol Dezateux,
Professor of Paediatric
Epidemiology, Institute of Child
Health, London

Mr John Dunning,
Consultant Cardiothoracic
Surgeon, Papworth Hospital NHS
Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon,
Gloucestershire Royal Hospital,
Gloucester

Professor Martin Eccles,
Professor of Clinical Effectiveness,
Centre for Health Services
Research, University of Newcastle
upon Tyne

Professor Pam Enderby,
Dean of Faculty of Medicine,
Institute of General Practice
and Primary Care, University of
Sheffield

Professor Gene Feder,
Professor of Primary Care
Research & Development, Centre
for Health Sciences, Barts and The
London School of Medicine and
Dentistry

Mr Leonard R Fenwick,
Chief Executive, Freeman
Hospital, Newcastle upon Tyne

Mrs Gillian Fletcher,
Antenatal Teacher and Tutor and
President, National Childbirth
Trust, Henfield

Professor Jayne Franklyn,
Professor of Medicine, University
of Birmingham

Mr Tam Fry,
Honorary Chairman, Child
Growth Foundation, London

Professor Fiona Gilbert,
Consultant Radiologist and NCRN
Member, University of Aberdeen

Professor Paul Gregg,
Professor of Orthopaedic Surgical
Science, South Tees Hospital NHS
Trust

Bec Hanley,
Co-director, TwoCan Associates,
West Sussex

Dr Maryann L Hardy,
Senior Lecturer, University of
Bradford

Mrs Sharon Hart,
Healthcare Management
Consultant, Reading

Professor Robert E Hawkins,
CRC Professor and Director of
Medical Oncology, Christie CRC
Research Centre, Christie Hospital
NHS Trust, Manchester

Professor Richard Hobbs,
Head of Department of Primary
Care & General Practice,
University of Birmingham

Professor Alan Horwich,
Dean and Section Chairman,
The Institute of Cancer Research,
London

Professor Allen Hutchinson,
Director of Public Health and
Deputy Dean of ScHARR,
University of Sheffield

Professor Peter Jones,
Professor of Psychiatry, University
of Cambridge, Cambridge

Professor Stan Kaye,
Cancer Research UK Professor of
Medical Oncology, Royal Marsden
Hospital and Institute of Cancer
Research, Surrey

Dr Duncan Keeley,
General Practitioner (Dr Burch &
Ptnrs), The Health Centre, Thame

Dr Donna Lamping,
Research Degrees Programme
Director and Reader in
Psychology, Health Services
Research Unit, London School of
Hygiene and Tropical Medicine,
London

Professor James Lindesay,
Professor of Psychiatry for the
Elderly, University of Leicester

Professor Julian Little,
Professor of Human Genome
Epidemiology, University of
Ottawa

Professor Alistaire McGuire,
Professor of Health Economics,
London School of Economics

Professor Neill McIntosh,
Edward Clark Professor of Child
Life and Health, University of
Edinburgh

Professor Rajan Madhok,
Consultant in Public Health, South
Manchester Primary Care Trust

Professor Sir Alexander Markham,
Director, Molecular Medicine
Unit, St James's University
Hospital, Leeds

Dr Peter Moore,
Freelance Science Writer, Ashtead

Dr Andrew Mortimore,
Public Health Director,
Southampton City Primary Care
Trust

Dr Sue Moss,
Associate Director, Cancer
Screening Evaluation Unit,
Institute of Cancer Research,
Sutton

Professor Miranda Mugford,
Professor of Health Economics
and Group Co-ordinator,
University of East Anglia

Professor Jim Neilson,
Head of School of Reproductive
& Developmental Medicine
and Professor of Obstetrics
and Gynaecology, University of
Liverpool

Mrs Julietta Patnick,
Director, NHS Cancer Screening
Programmes, Sheffield

Professor Robert Peveler,
Professor of Liaison Psychiatry,
Royal South Hants Hospital,
Southampton

Professor Chris Price,
Director of Clinical Research,
Bayer Diagnostics Europe, Stoke
Poges

Professor William Rosenberg,
Professor of Hepatology and
Consultant Physician, University
of Southampton

Professor Peter Sandercock,
Professor of Medical Neurology,
Department of Clinical
Neurosciences, University of
Edinburgh

Dr Philip Shackley,
Senior Lecturer in Health
Economics, Sheffield Vascular
Institute, University of Sheffield

Dr Eamonn Sheridan,
Consultant in Clinical Genetics, St
James's University Hospital, Leeds

Dr Margaret Somerville,
Director of Public Health
Learning, Peninsula Medical
School, University of Plymouth

Professor Sarah Stewart-Brown,
Professor of Public Health,
Division of Health in the
Community, University of
Warwick, Coventry

Dr Nick Summerton,
GP Appraiser and Codirector,
Research Network, Yorkshire
Clinical Consultant, Primary Care
and Public Health, University of
Oxford

Professor Ala Szczepura,
Professor of Health Service
Research, Centre for Health
Services Studies, University of
Warwick, Coventry

Dr Ross Taylor,
Senior Lecturer, University of
Aberdeen

Dr Richard Tiner,
Medical Director, Medical
Department, Association of the
British Pharmaceutical Industry

Mrs Joan Webster,
Consumer Member, Southern
Derbyshire Community Health
Council

Professor Martin Whittle,
Clinical Co-director, National
Co-ordinating Centre for Women's
and Children's Health, Lymington