PULSE OXIMETRY AS A SCREENING TEST FOR CRITICAL CONGENITAL HEART DEFECTS AND OTHER SIGNIFICANT DIAGNOSES IN NEWBORN INFANTS

A COST-EFFECTIVENESS ANALYSIS

Report commissioned by Public Health England.

Date of Completion: 12th February 2019.

AUTHORS

Tracy Roberts PhD[1] ***Professor of Health Economics*** [*]
Karen Pickering MSc[1] ***Research Fellow in Health Economics***
Pelham Barton PhD[1] ***Reader in Mathematical Modelling***
Andrew Ewer MD[2] ***Professor of Neonatal Medicine***

- **Health Economics Unit, Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham**
- **Institute of Metabolism and Systems Research, College of Medical and Dental Sciences, University of Birmingham**

*Author for correspondence: Professor Tracy Roberts, University of Birmingham (email: T.E.Roberts@bham.ac.uk)

Competing Interests: None

ABSTRACT

Title

Pulse Oximetry as a screening test for critical congenital heart defects and other significant diagnoses in new-born infants – a cost-effectiveness analysis.

Background

Critical Congenital Heart Defects (CCHD) are potentially life-threatening and it is vital to for them to be detected early in order to improve outcomes for infants. Pulse Oximetry screening (POS) is a simple test that has previously been shown to be effective and likely cost-effective in detecting CCHDs in new-born infants. A recent Pilot study carried out by Public Health England found that many infants given a 'false-positive' result for CCHD did in fact have a potentially serious non-cardiac condition. The aim of this study was to conduct a post-hoc exploration of the relative cost-effectiveness of Pulse Oximetry screening as an adjunct to clinical examination versus clinical examination alone for the detection of *CCHD and other significant diagnoses* based on data collected in a Pilot study.

Methods

A cost-consequence analysis was attempted and decision-analytic model was constructed. The analysis was carried out from the UK NHS perspective. To complete the model-based analysis many assumptions were used, informed by data from other sources and expert opinion. Estimations were made for the number of unnecessary tests in the model based comparators. The cost effectiveness-analysis was based on an outcome of timely diagnosis. Results are presented as an incremental cost-effectiveness ratio, namely the additional cost timely diagnosis detected.

Results

The assumptions from two different data sets to complete two different pathways in the model led to two different implied prevalence rates for CCHD and other significant diagnosis. Consequently two separate analyses were attempted to estimate a range for the cost effectiveness ratio and different scenarios in terms of unnecessary tests experienced by infants as a result of Pulse Oximetry screening. Incremental cost-effectiveness ratios (ICERs) in terms of cost per case of timely diagnosis were estimated.

The results of the analyses are considered unsound because there is no nationally accepted prevalence for the target condition of *CCHD and other significant diagnosis* with which to calibrate

the many assumptions. Consequently the results do not predict the data we do know from the Pilot.

Discussion

Further model based analyses are unlikely to shed light on the issue of whether POS is an appropriate use of NHS resources unless new and more complete data become available. If policy recommendations cannot be made based on the previous analyses, then further research should take the form of a randomised controlled trial so that the true non screening comparator can be include with all infants followed up. Future studies should also include qualitative information from parents about their concerns regarding unnecessary invasive tests.

Funding

This study was funded by the National Screening Committee (NSC).

TABLE OF CONTENTS

1. INTRODUCTION

Congenital heart defects (CHDs) are the most common forms of birth defects, affecting up to 9 in every 1,000 babies born in the UK (1). Some CHDs are minor, requiring only observation or minor treatment. However, around 2 per 1,000 livebirths have a critical CHD (CCHD), which is potentially life-threatening and usually requires urgent treatment. In these cases it is advised that the CCHD is detected sufficiently early, as timely recognition of these conditions is known to improve outcomes (2-4).

Pulse Oximetry screening (POS) is a simple test that can be carried out on new-born infants to measure the amount of oxygen in their blood. The test is non-invasive, and involves placing a sensor on the infant's hand and foot (5). Previous studies have shown POS to be effective in detecting CCHD in neonates (3, 6-8) and meet the criteria for a screening test. POS is included as part of national guidelines in several countries including the USA, Ireland, Norway, Poland, Sri Lanka and Switzerland. Although the uptake of POS as routine practice is increasing in the UK (9), it is yet to be incorporated into national guidelines for the UK.

A recent Pilot study carried out by Public Health England (6), analysed the effect on clinical services when POS is undertaken as part of the Newborn and Infant Physical Examination (NIPE) Programme. The Pilot study showed POS is effective in detecting CCHD and suggested that many of the infants given a false-positive result for CCHD did in fact have a potentially serious non-cardiac related condition that might otherwise have gone undetected if not for POS. From the Pilot study these appeared to include: Congenital pneumonia, early-onset sepsis (culture positive and negative), persistent new-born pulmonary hypertension (PPHN), meconium aspiration, respiratory distress syndrome, pneumothorax, transient tachypnoea of the new-born (TTN) requiring oxygen and lung malformation. These findings in the Pilot raised the question of whether early identification of these non-cardiac conditions represented a potential secondary target for POS and could be an important additional advantage of the test. However, the inclusion of these 'secondary targets' would further introduce risks associated with overdiagnosis which might occur from unnecessary invasive test and/or treatment of such conditions once a baby has tested positive.

The planned objective for this commissioned study was to conduct a post-hoc exploration of the relative cost-effectiveness of POS to target *"CCHD and other significant diagnoses"* when performed on infants within 24 hours of birth versus no Pulse Oximetry screening (routine practice). POS for the detection of CCHD alone has already been suggested to be cost-effective when performed on babies within 24 hours of birth (4, 10). The motivation for this study was that if POS for CCHD could also identify 'other significant non-cardiac conditions', any additional benefits, harms and cost implications are identified and considered for the health care sector.

2. SERIOUS CHALLENGES AND LIMITATIONS TO STUDY OBJECTIVE AND RESULTS

In the event, the study objective has proved unattainable. Many significant challenges to the set objective were identified at the outset and were considered manageable through the use of accepted processes of modelling and use of expert opinion. The purpose of the report presented here is to present the methods, assumptions and processes that were used in the attempt to reach the initial objective and to explain the challenges that have led this piece of work to an unsatisfactory end in the hope that they will not be repeated by researchers in the future.

The study reported here was originally commissioned and funded as **a cost-effectiveness analysis comparing Pulse Oximetry screening (POS) to no screening on an outcome of cost per case of timely diagnosis: A model-based analysis.** A similar objective had been set before (4). The work here was intended to be an update of that analysis. The main difference this time is that the target condition was not just CCHD in neonates as before, but as a result of the Pilot study (6) and the implication that POS could detect more than just CCHD. Thus the new target condition in the updated analysis was proposed ***as 'CCHD plus other significant diagnosis'***.

Given our previous study (4) suggested that POS was likely to be considered cost-effective and should be implemented, it should be intuitively acceptable that if the same test can be revealed to have additional benefits with fewer false positives because of the benefit of identifying other significant conditions, then it can be intuitively anticipated that the results are likely to be even more favourable on cost-effectiveness grounds. However, it was considered appropriate to try and quantify the additional benefit.

A new model was constructed because this latest analysis would be based on a new target condition of 'CCHD plus other significant diagnosis'.
However there were clear limitations from the outset. The Pilot study (6) had only identified and followed cases that were positive to POS for CCHD. The Pilot study noted and recorded that there were some 'other significant diagnosis' but this was apparent only in hindsight and after the initial analysis of the Pilot results (6). Consequently there was no information from

the Pilot study for infants with other significant diagnosis who screened negative to POS because those infants were not followed because at the time this was not necessary for its objective.
Also the Pilot study was only focussing on feasibility of POS for CCHD and so there was no available comparison for infants who were not screened at all. Yet such a comparator group would be necessary to carry out an economic evaluation.

- *Serious Limitation 1: - No accepted prevalence for new target condition*

In the earlier study (4) which was also based on a cohort study of infants with no comparator group, these other unknown pathways were modelled and informed by expert opinion. These modelled pathways were considered plausible because there was appropriate data to which the assumptions could be calibrated, most fundamentally, there was a nationally accepted prevalence for CCHD (4).

In the new scenario reported here, there is no prevalence for **'CCHD and other significant diagnosis'** because this is a unique group. All that is known about these infants is that they may be identified by POS but otherwise are unlikely to be a definable homogenous population of infants with an easily defined condition.

The assumptions made to inform this model based analysis were required to inform two different pathways in the model and the balancing of these has been extremely challenging. Assumptions were required both for:

(i) The infants who tested negative in the POS intervention (who were not followed up the Pilot study). These were based on Birmingham Women's Hospital data and expert opinion.

And

(ii) The assumptions required to model a comparator of no screening arm. These were based on matched control from the Neonatal Data Analysis Unit (NDAU) and supported by expert opinion.

Both sets of assumptions defined above were necessarily created in the absence of a robust calibration point. This is because there is no nationally accepted prevalence for **'CCHD and other significant diagnosis',** consequently both these sets of assumptions imply their own prevalence for the new target condition which were actually revealed in the analysis to be significantly different from each other.

The assumptions were made by experts in good faith, though based on different data. In the absence of true prevalence data for the new target condition the use of these assumptions reveal inconsistent implications.

To try to address this problem, two analyses were undertaken. The first used the implied prevalence (from assumptions) present in (i) above and a second analysis used the other implied prevalence based on assumptions presented in point (ii) above. This has presented us with a range for a cost-effectiveness ratio – which on the basis of the outcome of timely diagnosis (for the new target condition) is very close to the estimates from the previous study of 2012 (4). This could be seen as reassuring and within expectations: POS is detecting fewer false positives because the false positives do have another condition which is helpful to identify.

- *Serious Limitation 2: - Limited interpretation for results of cost effectiveness analysis presented in terms of cost per timely diagnosis*

This study was commissioned as a cost- effectiveness analysis based on timely diagnosis like the initial study (4). To attempt to present the results in terms of cost per Quality Adjusted Life Year (QALY) for this population of new-born infants in the absence of robust quality of life data would require additional heroic assumptions and this is an area well recognised as being fraught with challenges and is likely to produce misleading results.

However, to shed as much light as possible on the results, an attempt was made, post-hoc, to investigate the extent of unnecessary tests imposed on infants as a result of screening through a cost-consequence analysis and to use the implications produced within the modelling framework to explore this. Had this not been attempted, (and it was not initially part of the

commissioned study design) the results would have been presented for the range of prevalence estimates as indicated by assumptions (i) and (ii). Furthermore, because of the relative concordance with the results of our earlier study (4) the new results may well have appeared to have presented a range of plausible results. But the detail of the results in terms of the estimation of infants who have necessary and unnecessary tests - both invasive and non-invasive based on the model - casts severe doubt on the results of the model since the results are at odds with the data that is known with certainty, which is how many unnecessary tests were imposed on infants in the Pilot study (6).

Consequently, the results presented in the enclosed report are seriously limited and must be viewed with caution. They could prevent POS from being put into policy when in fact it is something that could improve the lot of babies and their families. Alternatively, they could support POS wrongly and waste resources which could be used somewhere else better.

However, this small scale study was commissioned using public money, and the report here, although unsatisfactory, is presented to explain methods results and limitations, primarily to ensure it provides insight to future researchers who may attempt to address the same question.

For the reasons emphasised above the result of this analysis are likely to be unreliable and must be viewed and interpreted with great caution.

3. METHODS

The most appropriate approach to address this question was deemed to be a model-based economic evaluation since a modelling framework can most easily combine clinical and cost data from varying sources and time points and analyse beyond the observed data (11). A cost-consequence analysis was carried out in the first instance to compare the costs and outcomes of the strategies. A decision-analytic model was constructed, using as a starting point, the model developed by Roberts et al. (4). The analysis was carried out from the perspective of the UK NHS as this was the location of the study, and so only direct health care costs were included in the analysis. The analysis was based on an outcome of timely diagnosis, the same outcome measure used in Roberts et al. (4). Timely diagnosis refers to a correct diagnosis before the infant is discharged from hospital after birth and was chosen to reflect evidence that detecting CCHD early results in a better health outcome for the infant. This is an intermediate outcome and implicitly assumes a timely diagnosis can prevent a catastrophic event. This is considered a necessary and pragmatic outcome in the absence of: robust data from an appropriate randomised controlled trial evaluating the intervention; a cohort study with long follow up of all screened; or a suitable comparator with data reporting quality of life for both. It must be emphasised that the outcome measure of timely diagnosis used in Roberts et al. (4) cannot be directly compared with the outcome measure used in this study. In this study the aim is to detect "CC*HD and other significant diagn*oses". From here on this is defined as the "target condition." In Roberts et al. the aim was to detect CCHD ***only***. This study therefore developed a completely new model structure and new analysis. The results of the analysis are presented in terms of the incremental cost-effectiveness ratio (ICER), namely the additional cost per timely diagnosis.

In this study the definition of 'other significant diagnosis' refers to the infants who, in addition to those diagnosed with serious or significant CHD require admission to the neonatal unit, or receive either high dependency or intensive care, and are diagnosed with one of the following conditions unrelated to CHD: Persistent new-born pulmonary hypertension (PPHN), meconium aspiration, respiratory distress syndrome, pneumothorax, congenital pneumonia, sepsis (culture positive and negative), transient tachypnoea of the new-born (TTN) requiring oxygen or lung malformation.

3.1. Model Structure

A decision tree model was developed in TreeAge Pro 2017 (TreeAge Software, Inc., Williamstown, MA, USA) to represent two alternative strategies: POS as an adjunct to clinical examination versus clinical examination alone (routine practice). A decision tree was chosen as the most appropriate model due to the relatively short time of the intervention (12). See Figures 1,2 & 3.

The clinical pathways for the Intervention arm of the model represent, as far as possible, the clinical steps carried out in the Pilot study (6). Likewise, usual care, where POS is not used, is modelled to follow, as far as possible, current practice in those UK hospitals not performing such screening.

3.2. Screening Strategies

The model comprises two alternate pathways to reflect the two strategies being compared. These are:

1. Routine practice (clinical examination alone) (Figures 1&2, to include branch node B) [Comparator Arm].

   Routine practice consists of a clinical examination, namely the New-born and Infant Physical Examination Programme (NIPE) examination (12). The NIPE examination is a physical examination performed within 72 hours of birth and again between 6 to 8 weeks after birth. The test checks for conditions related to the heart, hips, eyes and testes (12).

2. Pulse Oximetry screening given as an adjunct to clinical examination (Figures1 &3, to include branch node C) [Intervention Arm].

   Pulse Oximetry (PO) readings are taken from the right hand and either foot (pre- and post-ductal).

   - A healthy infant would normally have blood oxygen saturation levels of 95-99%. As per the Pilot study, if both (pre- and post-ductal) readings are more than 94%, with a difference between the readings of less than 2%, this would be recorded as a negative screen in the study.

- If either reading is less than 90% the infant is deemed to have a positive screen and would be sent for an urgent medical assessment.
- If either (pre- or post-ductal) reading is 90-94% or a difference of more than 2% is recorded, a definitive positive screen cannot be assumed. The infant will be reviewed by a healthcare professional and a repeat screen may be performed after one to two hours.
    - If the infant is deemed unwell by the healthcare professional, the result of the screen is considered positive and the infant would be sent for further assessment and investigations as appropriate.
    - If the infant is not deemed unwell, they will be sent for a repeat screen. If either (pre- or post-ductal) reading in the repeat screen is less than 95% or there is a difference greater than 2% between the readings, this would be recorded as a positive screen, and the infant would be sent for further assessment and investigations as appropriate.

3.2.1. Routine Practice (clinical Examination alone) [Comparator Arm]

The clinical examination can be considered "abnormal" or "normal". If the clinical examination is considered "abnormal", this means the infant will be suspected of having the target condition. After a period of observation, the infant will either be diagnosed as negative for the target condition, or receive further assessment and investigations as appropriate.

Following further assessment and investigations, the infant will either be diagnosed as positive or negative for the target condition. If the infant is diagnosed as positive, they will be sent for treatment as appropriate. If the infant is diagnosed as negative, they will be reviewed by a senior clinician. Following the review by a senior clinician, the infant will either be assumed not to have the target condition, or the infant will be sent for a diagnostic echocardiogram (ECHO), if the diagnosis of the infant is still unclear.

If an ECHO is deemed appropriate, the infant will be diagnosed as either positive or negative for CCHD. At this stage, if negative for CCHD they will be assumed not to have the target condition and will be considered normal unless other indications arise.

3.2.2. Pulse Oximetry screening (POS) as an adjunct to clinical examination [Intervention arm]

Following a negative PO test, the infant will proceed to receive the clinical examination as in the routine practice strategy. Thus, the infant will follow the same pathways as described in "routine practice", as they would if they had not received the Pulse Oximetry screening test.

If PO is positive, the infant will undergo a medical assessment. If the medical assessment detects an abnormality, the infant will either receive further tests as appropriate (see above), or will have the target condition ruled out after a period of observation.

If the medical assessment is considered normal the infant will receive a repeat PO screen. If the repeat PO screen is positive, the infant will follow the same pathways as if they had an abnormal medical assessment.

If the repeat PO screen is normal the infant will be considered to have passed the screening test and will receive the clinical examination as in the routine practice strategy.

3.3. Clinical Data

The decision model was populated, as far as possible, using prevalence data from the recent Pilot study carried out by Public Health England (6). Fifteen trusts across England, ranging in number of deliveries, level of access to neonatal intensive care and paediatric cardiology and in geographical location, participated in the Pilot study. The Pilot study was carried out over a 6 month period from July 2015-December 2015, and 32,836 infants were screened. Infants eligible for POS were all asymptomatic newborns greater than 34 weeks gestation who were not on the neonatal unit.

Other exclusion criteria for POS were as follows:

- presence of a suspected cardiac lesion from the fetal anomaly scans;
- symptomatic new-born with symptoms relating to potential cardiorespiratory problem prior to POS;

The Pilot study only collected outcome data on the infants that screened positive to POS. This is a critical point. There was no comparator arm in the Pilot for the outcome of infants not exposed to screening and also no follow up for those who were screened but tested negative to POS.

For infants who were not exposed to any screening, another source of data was required to provide information to populate the 'no screening' pathway. These data were sourced from the Neonatal Data Analysis Unit (NDAU) and based on matched controls and a summary of the matching approach is presented in Appendix A. Fifteen centres not currently practising POS that matched the characteristics of the fifteen centres in the Pilot study were identified and the data extracted from NDAU were from these fifteen centres. Data were extracted from NDAU on all admissions (gestation >34 weeks) to the neonatal unit in these centres during the time of the Pilot study. Information extracted included the diagnosis of the infants and their length of stay on the NNU along with the level of care received.

3.3.1. Infants meeting the exclusion criteria for POS

The Pilot study excluded infants that were *preterm* (≤34 weeks gestation), *symptomatic* or *had had an antenatal diagnosis of CHD*. These account for 6.47% of all infants born in the screening units during the time of the Pilot study. The comparator data sourced from the NDAU also excludes infants that are preterm, but includes infants that are *symptomatic* or *had had an antenatal diagnosis*. For consistency across both strategies, a proportion of infants from the NDAU data were removed, to account for those that would have not received the clinical examination as normal, as result of being *symptomatic* or receiving *an antenatal diagnosis*.

To ensure the comparator data were consistent with the Pilot data it was necessary to assume that the same percentage of infants were ineligible for the clinical examination in the routine practice strategy and so assumptions were required to remove this proportion from the matched controls, whilst maintaining the ratio relative ratios for diagnosis of the target condition (as revealed in the Pilot but not known in NDAU data). It is acknowledged that this removal is random as opposed to the more targeted process of the Pilot and not does not benefit from the knowledge of the cases as occurred in the Pilot arm. This is another limitation.

The infants were excluded from each pathway in order to maintain the proportions in each pathway using the same ratios as the Pilot study.

i. Infants diagnosed as positive for the target condition (with a true diagnosis of positive for the target condition) before 24 hours (This accounted for 40% of infants in the Pilot data);
ii. Infants diagnosed as negative for the target condition after spending one day or less on the NNU in special care (with a true diagnosis as negative for the target condition); (This accounted for 4% in the Pilot data);
iii. Infants diagnosed as negative for the target condition after having investigations and possibly a diagnostic echocardiogram (with a true diagnosis as negative for the target condition). (This accounted for 46% of infants in the Pilot data);

In removing infants from the matched controls the ratios defined above in parenthesis were retained.

3.3.2. Infants that screen negative to POS

The number of infants with CCHD that were potentially missed by POS (screening negative initially) were identified. It was only after the Pilot that the additional potential benefits that POS may have in detecting 'other significant diagnosis' in those that tested positive was realized but it was too late to follow up the negative cases who did not have CCHD. Thus, assumptions were required to populate the pathway of the intervention arm of the model (Figure 2) for infants that screened negative to initial POS but had other significant conditions.

Data on live births and admissions to the NNU in Birmingham Women's hospital (BWH) during 2016 were obtained from the BadgerNet system in a personal communication (Professor Ewer) to underpin these required assumptions. In 2016, there were 8,370 live births in BWH, and 1,583 admissions to the NNU (gestation >34 weeks). Infants admitted to the NNU in BWH at over 6 hours of age were assumed (and likely) to have had POS. Based on assumption, 167 infants over 6 hours of age were assumed to have been admitted with a negative PO screen, and 6 of these infants were assumed to have had an "other significant diagnosis." More detail of the assumptions based on Birmingham Women's Hospital Data (BWH) to inform the POS pathway for those that screened negative is presented in Appendix B.

3.4. Resource use and Costs

All costs sourced are reported in 2016-2017 UK prices, having been appropriately inflated or deflated if necessary. Key costs are presented in Table 1.

3.4.1. Pulse Oximetry (PO) Test

The time taken to carry out the PO screen was 6.9 minutes, as recorded in a time-and-motion study carried out by Roberts et al.(4). The Pilot study recorded the staff that carried out the Pulse Oximetry tests, and the cost of staff time for carrying out the test was recorded as a weighted average, and calculated using salary data and costs presented in Curtis and Burns (13). The weighted average cost of staff time to carry out the PO test was £4.21. The main components of carrying out the test are the equipment itself and the staff time. Use of cables and reusable sensors was included. The cables and sensors cost £119 and £137 respectively, and three cables and three sensors per PO machine were provided and last approximately 6 months. Use of disposables, tape to attach the sensors to the infant and disposable wipes to clean the sensors, was also included. The equipment used for the test is a PO machine that costs £538 which includes a carry case for the machine at £60. The machine is assumed to have a life of 5 years. The annuitized cost of the equipment is estimated following the method of Drummond et al. using a discount rate of 3.5% (15). An annual maintenance cost of 10% is then added, which is the usual maintenance cost typically applied to all technical equipment. There were 15 centres in total and 157 pulse oximeters. 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 £4.26. We added £0.03 per infant to cover the cost of disposables.

The total cost of carrying out the PO test, including staff time, equipment and disposables based on a 5-year lifespan was approximately £8.50.

Following Roberts et al.(4), we assumed that a repeat PO screen will require the same average time for completion as the first PO screen. We used the Pilot study data to include the weighted average cost of staff time carrying out the repeat screen, which was £4.24. The cost of carrying out a repeat screen was £8.44. This cost is slightly lower than the cost of the initial Pulse Oximetry

screen due to differences in who carried out the repeat Pulse Oximetry test (i.e. the cost of staff time).

3.4.2. Assessments/Examinations

The cost of the NIPE examination was excluded from the model as the NIPE examination was carried out on every infant in the model.

The time taken to carry out a medical assessment, following a positive PO screen, was 8.57 minutes (4). A medical assessment can now be assumed to be carried out by a junior doctor and so the cost of a medical assessment is revised to amount to £4.29.

A senior clinician review, after an infant has tested negative to investigations, was assumed to be carried out by a Registrar and to take 30 minutes. The cost of carrying out a senior clinician review was £21.50.

3.4.3. Investigations

Investigations carried out include: Further clinical examination by a senior clinician, Chest X-ray, Blood Gases, Blood Cultures, C - reactive protein, Urea and Electrolytes, Full Blood Count, lumbar puncture, electrocardiography, and 4 limb Blood Pressure (BP).

Further clinical examination by a senior clinician, Chest X-ray, Blood Gases, Blood Cultures, C-Reactive Protein, Urea and Electrolytes and Full Blood Count were defined as a standard "package of investigations", as these investigations were often carried out together. A lumbar puncture is defined as an invasive investigation, due to the intrusive nature of the procedure.

The "Further examination" by a senior clinician was estimated to cost £21.50. This estimate is based on an assumption that it would be carried out by a Registrar and to take 30 minutes. Costs for blood culture, full blood count, chest X-ray, urea and electrolytes and lumbar puncture were sourced from the literature (15-18). Costs for taking blood gases and a C - reactive protein test were sourced from laboratory costs (19). The cost of an electrocardiography and a diagnostic echocardiogram were sourced via NHS Reference costs (20). A 4-limb Blood Pressure (BP) was

assumed to be carried out by a Band 4 nurse and to take 15.5 minutes (13, 22). Infants who required transport to a regional cardiac centre incurred an additional cost of £396.17 (23).

3.4.4. Length of Stay

Total length of stay on the neonatal unit (NNU) and the level of care received whilst on the NNU were recorded in both the Pilot study data and NDAU data. The cost to stay on the NNU for 24 hours was recorded via level of care (20). Infants delivered via home birth that needed to be transferred to hospital, incurred an additional cost of £236 to cover the cost of the ambulance call out and transfer to hospital (20). It was acknowledged that infants may also spend time on the maternity unit. These data were incomplete in the Pilot study and not available in the NDAU data. It was decided to remove the cost of stay on the maternity unit from the analysis as diagnosis was confirmed by admission to the NNU.

3.5. Assumptions

Several assumptions were required in order to develop a workable model. These are summarised and described below and divided into 3 categories: Model Pathways, Clinical Data and Costs and Resource Use. The Clinical Data category includes a sub-category on assumptions specifically relating to the data extracted from the NDAU.

3.5.1. Model Pathway Assumptions

- All infants adhered to the screening pathway.
- It is acknowledged that in practice, following the POS pathway, some infants will require an urgent medical assessment (reading <90%) and some will require an expedited assessment (reading 90-94%, or differential >2). However, in the Pilot study data there are no details about the expedited assessment or the urgent medical assessment. It is therefore assumed that 100% of infants requiring an expedited assessment will have this carried out by a junior doctor. It is also assumed that 100% infants requiring the urgent medical assessment will have this carried out by a doctor.

  The cost attached to each of these assessments therefore will be the identical, and given it is not possible to differentiate between these two in the Pilot data, an umbrella term for these assessments will be “medical assessment.”

- All infants that received a repeat PO screen are assumed to have had a normal medical assessment.
- It is acknowledged that some infants may have had periods of observation at certain points during their pathway; however, these were not explicitly included in the model, as it was acknowledged that the costs attached to observation would be included in the length of stay. To avoid double-counting, there would therefore be no cost or resource use attached to periods of observation, and so 'observation' was excluded from the model pathways.

3.5.2. Clinical Data Assumptions (Pilot Study)

- In the Pilot study 32,836 infants were screened in total, and there was an overall screen positive rate of 0.73% (n=239). Complete data were only available for 226 infants, so the overall number of screened infants was adjusted to maintain the screen positive rate of 0.73%. The total number of infants screened in the model therefore is 31,050.
- There are the same number of babies in the routine practice arm of the model as in the POS arm of the model.
- As indicated by the Pilot study, "other significant diagnosis" is defined as being diagnosed with one of the following: CHD (serious or significant), Persistent new-born pulmonary hypertension (PPHN), meconium aspiration, respiratory distress syndrome, pneumothorax, congenital pneumonia, sepsis (culture positive and negative), transient tachypnoea of the new-born (TTN) requiring oxygen or lung malformation. If the infant is diagnosed with a condition that is not CCHD and is not described as "other significant diagnosis", then the infant will be diagnosed as negative for the target condition. Infants are diagnosed as having a *significant* other diagnosis if the infant is admitted to the Neonatal unit (NNU) and receives high dependency or intensive care.
- As indicated by the Pilot study, further investigations include further assessment by a senior clinician (Registrar) , blood culture, blood gas, C-reactive protein test (CRP), full blood count (FBC), chest X-ray (CXR), urea and electrolytes (U+E), lumbar puncture (LP) and an electrocardiography (ECG). We are aware that not all babies will receive all of these tests. Investigations will therefore be included in the model as a weighted average, as per the Pilot study results. Any other investigations recorded in the Pilot study data (i.e. MRI scan, chromosomal studies) were excluded from the analysis as they deemed unrelated to the conditions that are being targeted in this study.

3.5.3. Clinical Data Assumptions (NDAU data)

- The percentage of infants with CCHD that are missed by routine practice (clinical examination) is 30% (23).
- All infants suspected of having the target condition are assumed to have been admitted to the NNU.
- All infants admitted to the NNU that are recorded as being diagnosed with a cardiac-related diagnosis, respiratory-related diagnosis or sepsis, are assumed to be suspected of having the target condition. All other infants admitted to the NNU are assumed to have had a 'normal' clinical examination for these above conditions and are on the NNU for non-cardiac, non-respiratory and non-sepsis reasons (e.g. jaundice, cleft palate).

  As the NDAU data only provides us with details of infants admitted to the NNU, by making this assumption we are assuming that we have not missed any infants from the model.
- All infants suspected of having a "cardiac or non-cardiac diagnosis" but only spend one day or less in special care on the NNU are assumed to have had no significant investigations. All infants suspected of having a "cardiac or non-cardiac diagnosis" that spend more than one day in special care, or any number of days in high dependency care or intensive care on the NNU, are assumed to have received significant investigations.

  We recognize that if an infant was to be admitted to the neonatal unit but have no investigations they would probably be on the NNU for less than one day (six-twelve hours is more realistic). However, length of stay in the NDAU data is recorded in "days", which is why we have assumed this to be 1 day.
- All infants receiving investigations were assumed to have received the following investigations:

  i. Further clinical examination by a senior pediatrician.

     This was assumed to be carried out by a registrar and to last 30 minutes.
  ii. Chest X-ray
  iii. Blood Gases
  iv. Blood Cultures
  v. C-Reactive Protein
  vi. Urea and Electrolytes
  vii. Full Blood Count

- A proportion of infants receiving investigations also received a lumbar puncture and ECG. It was assumed that the proportion of infants receiving a lumbar puncture or ECG was the same as the proportion of infants that received a lumbar puncture or ECG in the Pilot study data.
- The proportion of infants receiving a diagnostic echocardiogram was assumed to be the same proportion of infants that received echocardiograms in the Pilot study.

3.5.4. Costs and Resource Use Assumptions

- It was assumed that no infants in the routine practice arm of the trial were transferred to hospital from home or a midwife-led unit.
- In the Pilot study 75% of babies with CCHD were transferred to a specialist cardiac unit. No other infants were transferred to a specialist cardiac unit from the Pilot study data. Similarly, in the routine practice arm of the model, 75% of infants with CCHD were assumed to have been transferred to a specialist cardiac unit. Transfer costs will be attached to these infants.
- It was acknowledged that infants will have varying lengths of stay on the maternity unit for reasons that may not be related to this study (e.g. issues with breastfeeding, issued relating to the mother), and so costs for length of stay on the maternity unit were dismissed from all arms of the model. The resource use and costs related to length of stay in the study therefore, only refer to length of stay on the NNU.

3.5.5. Assumptions Imposed by Approach to the Economic Evaluation.

- Cost-effectiveness analyses (CEA) are by definition based on one outcome that is usually clinically related (14). In the current analysis this outcome is timely diagnosis. In contrast a cost utility analysis is based on units in terms of quality adjusted life years (QALYs). In a CEA it is implicitly assumed that the chosen outcome 'cost per timely diagnosis' is the most important outcome. Carrying out a cost-effectiveness analysis as opposed to a cost-utility analysis was a necessity in the absence of robust data from randomised controlled trials, cohort studies with long follow up data and no available evidence on quality of life for infants or parents. These data are so difficult to ascertain it was not possible to inform assumptions on quality of life for either parents or infants.
- The design implicitly imposes the assumption that there are no lasting negative effects of screening on the parents in terms of anxiety caused by their babies undergoing the PO screen

first and/or any subsequent test to confirm diagnosis whether the infant receives either a true positive of a false positive diagnosis.

- Whilst the monetary costs associated with all test and interventions, whether unnecessary, and invasive or not are included in the analysis, the cost effectiveness approach imposes the assumption that there are no lasting detrimental effects on quality of life of mothers or infants as a result of unnecessary tests as this cannot be captured by design.

3.6. Analysis

Two separate analyses were carried out by necessity. The analysis is based on assumptions since the Pilot study did not follow up all infants that screened negative and no comparator arm was included in the Pilot study. Thus a true ***prevalence*** rate to represent infants with the target condition (CCHD and other significant diagnosis) is not currently known.

In the model described so far, in order to complete the pathways in the intervention arm of the model (since most infants who were negative to POS were not followed), assumptions were made about the number of infants truly positive to the target condition who were missed by POS and this was based on data from Birmingham Women's Hospital (as described in Appendix B). This process produces ***an implied prevalence*** for the target condition. **This first implied prevalence is used in Analysis 1.** Thus the prevalence for the target condition (CCHD plus other significant diagnosis) used in Analysis 1 is 0.2%.

Also there was no comparator arm in the Pilot study. In order to facilitate a comparator arm, the matched controls were collated from NDAU data as explained in Appendix A and above. To do this the NDAU data were examined closely by clinical experts to determine their likely pathway to their diagnosis (working backward from their likely final diagnosis provided in the NDAU data). From this process another implied prevalence rate for the target condition was presented. **This second implied prevalence is used in Analysis 2.** Thus the prevalence of the target condition (CCHD plus other significant diagnosis) used in Analysis 2 is 1.6%.

Consequently, the assumptions collated and approved independently by the clinical experts in the working group, caused the intervention arm and the comparator arm to provide two alternative prevalence rates for the target condition which were significantly different. The non-targeted removal of 6.47% of infants from the NDAU cohort may explain this to some extent.

Given both these proxy prevalence rates were based on the best available clinical assumptions for the infants with this condition there was no reason to prefer one proxy prevalence rate over another. Unlike the previous economic analysis (4) where the model was calibrated to the population prevalence of CCHD for new-born infants that was known to exist, there is no known population prevalence for the target condition of '*CCHD and other significant diagnosis*' that was the target condition in this study. The assumptions provide an estimated range for the overall prevalence for the target condition. Both are used in the model (Analysis 1 and Analysis 2 respectively) and these are perceived as the two extremes for the prevalence for this condition, in the absence of any better data.

Both Analysis 1 and Analysis 2 undertake a comparison of routine Practice with Pulse Oximetry screening:

- Routine practice (based primarily on clinical examination)
- Pulse Oximetry screening (POS) given as an adjunct to routine practice

The chosen outcome of cost per timely diagnosis was influenced by two previous UK studies evaluating POS; Roberts et al. (2012) (4) and Knowles et al. (2005) (10). The results are initially presented in terms of a cost-consequence analysis where the costs and outcomes are compared in a disaggregated manner. The results of the decision-analytic model are subsequently presented in terms of the incremental cost-effectiveness ratio (ICER), namely the additional cost per additional case of timely diagnosis of the target condition. Both analyses were carried out from an NHS perspective and based on the outcome of cost per timely diagnosis

- Analysis 1

This analysis is based on the prevalence that is implied when the assumptions (explained in appendix B) inform the pathway for the infants that screened negative to POS. In the model the only true data that exist are those produced in the Pilot study for infants that screened positive. So the assumptions used to inform the pathway for those that screen negative must reproduce the true data that was revealed in the Pilot study (screen positive branch). Consequently the screen negative branch of the intervention arm and the comparator arm in the model must absorb the residual data in order to adhere to the assumptions. The prevalence for the target condition (CCHD plus other significant diagnosis) used in Analysis 1 is 0.2%.

- Analysis 2

Analysis 2 is analogous to Analysis 1, but in this case the prevalence informing the model is that which was estimated from the separate assumptions required to inform the comparator arm (as described in Appendix A). Thus the prevalence of the target condition (CCHD plus other significant diagnosis) used in Analysis 2 is 1.6%. In using this prevalence, the model data are calibrated so that the infants in the intervention arm that tested positive to Pulse Oximetry continue to represent the true data that were reported in the Pilot study (for infants testing positive). The residual in the prevalence rate therefore impacts on the infants who follow the assumed pathway for those who tested negative initially for Pulse Oximetry, whose pathway are informed by assumption because these infants were not followed up in the Pilot study.

4. RESULTS

4.1.1 Infant outcomes:

Analysis 1 was based on the prevalence (0.2%) indicated by the assumptions used to inform the model pathway for infants that screened negative in intervention arm (Table 2a).
Analysis 2 was based on the prevalence of 1.6% indicated by the assumptions made when defining the comparator Arm (Table 2b).
In both analyses the results for those that screened positive to POS was fixed to the results of the Pilot study (6) which was the only known piece of data in the whole analysis.

The objective of presenting the results in the form of a cost-consequence analysis, presenting the costs and a full range of disaggregated outcomes was to try and use the model to ascertain the benefits and harms of screening with respect to the number of unnecessary investigation that are estimated by the model to have been undertaken. These results from the two model based Analyses 1 and 2 are presented in Tables 2a and 2b respectively but it is difficult to report these results in detail or give them too much attention because they are informed by assumptions that cannot be calibrated to a nationally accepted prevalence for the target condition of 'CCHD and other significant diagnosis' because such a prevalence is not known. The results from within the model do not correspond sufficiently to the real data that is known from the Pilot study, which suggests the assumptions made without any calibration point are not robust.

We present the results of Analysis 1 and 2 in Table 2a and 2b respectively, but urge the reader to contrast these results with the actual results of the Pilot presented in Appendix 3. Thus we feel that the model is not fit for the purpose of unpicking the implications regarding necessary and unnecessary tests caused by POS.

4.1.2 Costs

The breakdown of costs for the POS strategy for each Analyses 1 and 2 for are shown in Tables 3a and 3b respectively. Tables 4a & 4b present the total costs for the Routine Practice strategy for each Analysis 1 and 2 respectively.

Table 5a presents the incremental cost-effectiveness ratio (ICER) from Analysis 1 for POS compared with routine practice which is estimated to be £27,502 per timely diagnosis. This means that every additional timely diagnosis detected by POS as an adjunct to routine practice compared with routine practice alone, costs an extra £27,502.

The results of Analysis 2 (Table 5b) are based on the estimated prevalence of the target condition implied by the classification of the diagnosis, conducted by experts and based on NDAU data (1.6%).

The Analysis 2 incremental cost-effectiveness ratio (ICER) for POS compared with routine practice is estimated to be £9,996 per timely diagnosis. This means that every additional timely diagnosis detected by POS as an adjunct to routine practice compared with routine practice alone, costs an extra £9,996. This result for Analysis 2 is more favourable than the result of Analysis 1 because of the assumed higher prevalence for the target condition which was implied as a result on expert opinion when retrospectively diagnosing patients in the NDAU data. The higher prevalence but the fixed calibration to the Pilot study data for those testing positive mean that more cases are assumed to have missed a timely diagnosis in the Routine Practice arm.

These results, for the incremental cost effectiveness ratios have not been reported in detail because although superficially plausible estimates appear to have been produced, the misgivings that relate to the within model estimations regarding the number of necessary and unnecessary test imposed on infants as a result of screening are deemed implausible as they do not correspond with the data that is known from the Pilot (6). Thus it must follow that the estimated incremental cost- effectiveness ratios produces are subjects to the same serious concerns.

The assumption used to inform the model were made independently and before the impact on the results was known, as is good practice and to avoid bias. But only in seeing the results and comparing them to the actual data that is known from the Pilot (and presented in Appendix C), is it clear that in the absence of a consistent nationally accepted prevalence for the new target

condition of 'C*CHD and other significant diagnosis*', against which the assumptions could be calibrated, the results of the model based analysis reported here must be viewed with serious caution. The authors of the current report have tried in vain to refine model inputs in order to achieve more plausible results but we have limited known data, too much unknown data and too many possible variations and impacts caused by tweaked assumptions.

We view the objective of this study, given the limited available data, to be un-obtainable at the current time.

5. DISCUSSION

5.1. Principal Findings

The economic evaluation reported here intended to compare POS and Routine Practice in terms of their costs and effectiveness of achieving a timely diagnosis of the target condition. In this analysis the target condition was '***CCHD and other significant diagnosis***' in new born infants. There is no accepted prevalence for this target condition and this study has shown that in the absence of data from a randomized controlled trial comparing POS and Routine Practice for this population group, undertaking a model based analysis produces model results that cannot be considered robust because there are insufficient data, and there is no reference point to which the required assumptions can be calibrated.

We conducted a micro level cost-consequence analysis and a model based analysis. The inputs for the intervention strategy of POS were based on a Pilot study recently carried out and reported in detail elsewhere (6). But the Pilot study had explored POS only, and had only retained and followed up infants who tested positive on initial screen for CCHD. Information on infants who screened negative for other significant diagnosis after POS was not recorded in the Pilot and was by necessity derived from other data based on assumptions informed by expert opinion. In addition, the data to populate the comparator strategy of Routine Practice in the analyses were derived from NDAU data and based on assumptions also informed by expert opinion. All assumptions were made a priori and before any impact of the assumptions on the results could be predicted, in order to maintain the integrity of the results.

The assumptions used to inform the first original planned analysis inadvertently caused a discrepancy in the prevalence for the intervention arm and the comparator arm which led to the model reporting an imbalance in the number of infants assumed to have the target condition between both arms.

As a result of the revealed discrepancy the results of the CEA for the planned analysis were deemed uninformative and misleading. Subsequently, two separate analyses were instead carried out.

Analysis 1 is based on the prevalence estimated from the intervention arm and concurs with the true data reported in the Pilot study for infants that ***tested positive*** as a result of POS.

Analysis 2 is analogous to Analysis 1, but in this case the prevalence is that which was estimated from the separate assumptions applied to the NDAU data in the comparator arm.

Thus the intention was to present two separate analyses to represent a range for the costs and outcomes associated with POS for the two implied prevalence rates. But the results are not considered robust because in un-picking the model to estimate the necessary and unnecessary tests likely to have been experienced by infants, the model results do not concur with data that we do know.

Thus it must be concluded that these results cannot be considered robust.

5.2. Strengths of the Study

The main strength of this analysis is that it is the first study to attempt to consider the cost-effectiveness of POS for the detection of other significant diagnoses, in addition to the detection of CCHD. The Pilot study data consisted of a large sample of infants, who received POS and were considered generalizable for the UK. A thorough attempt to apply appropriate costs at all points in the model based analysis was undertaken. All assumptions required to complete the modeling exercise were made by consensus from a wide range of clinicians involved in the study team (listed in the acknowledgements) – through an iterative process and were closely scrutinized. Furthermore, all assumptions were made, agreed and signed off before any analysis took place so that the impact of the assumptions on the analysis was unknown before the analysis took place. The only exception existed when the data revealed discrepancy in the implied prevalence between the proportions of infants with the target condition in each of the strategies. Thus two separate analyses were carried out to explore the range of possible impacts caused by the assumptions, which were made in good faith, without knowledge of the impact, based on expert opinion.
The study has made explicit the severe limitations of undertaking a model based analysis with insufficient data – and specifically without the prevalence rate for the target condition to which all assumptions could have been appropriately calibrated. We hope that we have provided a clear enough account to ensure future researchers avoid repeating this exercise only to face the same serious challenges.

5.3. Limitations of the Study

There are many limitations of this study. First the Pilot study which initiated the call for an economic analysis focused only on POS and did not directly examine pathways for babies not receiving screening and so no direct comparator data existed. In addition, the Pilot study only recorded data on infants that screened positive to POS, and those with CCHD and CHDs that were misdiagnosed as 'healthy'. Information on all other infants was not recorded and so, in the current analysis, had to be supplemented by data from other sources and on a number of potentially heroic assumptions. Furthermore, the supplementary data (from Birmingham Women's Hospital and NDAU) to complete the intervention arm of the model and to inform a comparator arm, were collected separately and so many assumptions had to be included to attempt a workable model.

An illustration of the impact of the assumptions on the model results are most clearly illustrated by Analysis 2. The prevalence and implications represented by Analysis 2 (based on an implied prevalence from NDAU data) are shown to be un-realistic. Not least because the true primary data in the Pilot study identified 74 and missed 22. The corresponding figures for Analysis 2 suggest that 488 infants would have the target condition and in the absence of POS, 436 would be subsequently picked up via the clinical examination – this is a result that is not plausible given the known accuracy of the test for CCHD and the prevalence of other significant diagnosis in the Pilot study.

The summary results for cost-effectiveness in terms of the incremental cost-effectiveness ratio (ICER) for each of Analysis 1 and Analysis 2 present results that might be deemed superficially plausible but these results should also be viewed with serious caution because of the closer scrutiny of the model implications for unnecessary tests. An always anticipated limitation of the study is that timely diagnosis is an intermediate outcome and does not include implications of the pathway after diagnosis, nor does it capture the impact of other outcomes on the screened and unscreened population. The monetary costs of any unnecessary tests would be captured by a cost-effectiveness analysis (CEA), but the anxiety or distress caused to the parents of the healthy infants directly as a result of unnecessary investigations (although not realized to be unnecessary) as a result of screening cannot be quantified in this analysis and is ignored in a cost -effectiveness

analysis by design, since CEA focuses on one principal outcome typically in clinical units (natural units).

Because timely diagnosis is an intermediate outcome, in the absence of robust data from a randomized controlled trial and without lengthy follow-up data we would not be able to show whether Pulse Oximetry was cost-effective in the long term compared with routine practice. Whilst there is much evidence already to show that early detection of CCHDs improves outcomes for infants (2-4) and if we assume that as POS is a cost-effective strategy in the short term for 'timely diagnosis', then based on the literature, it would result in better long-term outcomes for the infants.

However, the benefits of a timely diagnosis as a result of screening which will result in a small number of babies avoiding longer term catastrophe must be weighed against the anxiety and distress, if any, caused to parents as a result of ultimately unnecessary tests, on otherwise healthy infants, some of which are invasive and have risks of their own. It will be important to explore whether parents experience long lasting tangible distress from tests on their infants as a result of screening or whether the unfounded concern is mitigated by the ultimate reassurance that the baby is healthy.

Quality of life information was not available and so presenting results in terms of cost per QALY was never anticipated to be possible. Furthermore, a model-based economic evaluation with a lifetime time horizon would require strong assumptions regarding the lifetime quality of life and development of the infant. If it was possible to follow up infants beyond diagnosis and explore the long term impact of having a timely diagnosis via POS, as part of a randomized controlled trial a result in terms of QALYs may be possible but this is likely to be beyond the scope of many studies.

5.4. Comparison with Similar Studies

Two previous studies have been carried out in a UK setting to determine the cost-effectiveness of POS (4, 10). Neither focused on the harms of the screening test.

5.4.1. Knowles et al. (2005)

Knowles et al. (10) carried out a cost-effectiveness analysis comparing three strategies:

1. Clinical Examination
2. Pulse Oximetry with Clinical Examination
3. Screening echocardiography with Clinical Examination

In the base-case analysis screening takes place at around 24 hours of age which is longer than the suggest target time for screening used in this study (4-8 hours). Knowles et al. use secondary data to populate the decision tree model and the outcome used is 'timely diagnosis, like in this study. However, in Knowles et al., 'timely diagnosis' refers to a diagnosis made pre-operatively before the infants collapses or dies. Costs are included for each screening strategy, the cost of a diagnostic assessment following a positive screen, and costs for the management of any collapsed infants. No costs are included for admission to the neonatal unit following a positive screen.

The results show 'POS with clinical examination' and 'screening echocardiography with clinical examination' to have roughly the same effectiveness. Both of these strategies are around twice as effective as 'clinical examination.' The total cost of the 'POS with clinical examination' is found to be much less costly than the total cost of 'screening echocardiography with clinical examination' (£476,016 versus £3,457,233 per 100,000 live births). Knowles et al. conclude that there is a high probability that 'Pulse Oximetry with clinical examination' is a cost-effective screening strategy for congenital heart defect.

5.4.2. Roberts et al. (2012)

Roberts et al. (4) expand on the findings from Knowles et al. (10) to determine the cost-effectiveness of POS as an adjunct to clinical examination versus clinical examination alone. They assume that all infants with a positive Pulse Oximetry screen followed by an "abnormal" clinical examination or with a "normal" clinical examination and an "abnormal" 2nd PO test will have a diagnostic echocardiogram.

Roberts et al. also use the outcome of 'timely diagnosis' but use primary accuracy data and primary cost and resource use data. They use a time-and-motion study to measure the time taken

to carry out a PO test (6.9 minutes) and a clinical examination (8.57 minutes) which is longer than the 2 minutes assumed in Knowles et al. (10).

The total cost of POS in Roberts et al. (£1,358,800 per 100,000 live births) is reported as being much higher than in Knowles et al. The estimated ICER for 'POS as an adjunct to clinical examination' compared with 'clinical examination alone' is £24,900 per timely diagnosis.

Although the studies by Knowles et al. and Roberts et al. referred to above were model based economic evaluations, and have their limitations, they are not subject to the same serious concerns as the current study. This is because the target condition for both those studies was CCHD alone for which there is a nationally accepted prevalence rate and this was used to calibrate model inputs ensuring the model results were plausible.

5.5. Interpretation

The results of this study cannot be deemed robust. But intuitively if POS is deemed cost effective for detecting CCHD as indicated by the previous studies if it can detect other significant diagnosis in addition to CCHD that it is likely to still be considered as cost-effective, if one accepts the limitations of the previous studies.

5.6. Policy Recommendations

Policy recommendations should not be influenced by the study reported here.

5.7. Future Research

To date, analyses exploring the relative cost effectiveness of POS are model based analysis based on secondary data supplemented by assumption (4,10). Further model based analyses are unlikely to shed any more light on the issue of whether POS is an appropriate use of NHS resources unless new and more complete data become available. If policy recommendations cannot be made based on evidence from the existing previous analyses, then further research should take the form of a randomised controlled trial so that a true 'no screening' comparator can be include in the evaluation with all infants followed up. Future studies should also include qualitative information from parents about their concerns about unnecessary invasive tests.

6. ACKNOWLEDGMENTS

The authors would like to thank the following people,

Matt Cawsey, David Elliman, Claire Evans, Chris Gale, Rachel Knowles, Deepa Panjwani, Jill Walker.

7. REFERENCES

1. NHS. Congenital heart disease 2015 [Available from: https://www.nhs.uk/conditions/congenital-heart-disease/.
2. Morrison JA, Friedman LA, Gray-McGuire C. Metabolic syndrome in childhood predicts adult cardiovascular disease 25 years later: the Princeton Lipid Research Clinics Follow-up Study. Pediatrics. 2007;120(2):340-5.
3. Ewer AK, Middleton LJ, Furmston AT, Bhoyar A, Daniels JP, Thangaratinam S, et al. Pulse Oximetry screening for congenital heart defects in newborn infants (PulseOx): a test accuracy study. The Lancet. 2011;378(9793):785-94.
4. Roberts TE, Barton PM, Auguste PE, Middleton LJ, Furmston AT, Ewer AK. Pulse Oximetry as a screening test for congenital heart defects in newborn infants: a cost-effectiveness analysis. Arch Dis Child. 2012;97(3):221-6.
5. England PH. UK National Screening Committee recommends new test for newborn babies with heart disease 2014 [Available from: https://www.gov.uk/government/news/uk-national-screening-committee-recommends-new-test-for-newborn-babies-with-heart-disease.
6. Newborn Pulse Oximetry Screening Pilot End Project Report. Public.screeninghelpdesk@nhs.net; 2016.
7. de-Wahl Granelli A, Wennergren M, Sandberg K, Mellander M, Bejlum C, Inganas L, et al. Impact of Pulse Oximetry screening on the detection of duct dependent congenital heart disease: a Swedish prospective screening study in 39,821 newborns. BMJ. 2009;338:a3037.
8. Thangaratinam S, Brown K, Zamora J, Khan KS, Ewer AK. Pulse Oximetry screening for critical congenital heart defects in asymptomatic newborn babies: a systematic review and meta-analysis. The Lancet. 2012;379(9835):2459-64.
9. Mikrou P, Singh A, Ewer AK. Pulse Oximetry screening for critical congenital heart defects: a repeat UK national survey. Arch Dis Child Fetal Neonatal Ed. 2017;102(6):F558.
10. Knowles R, Griebsch I, Dezateux C, Brown J, Bull C, Wren C. Health Technology Assessment. 2005;9(44).
11. Barton P, Bryan S, Robinson S. Modelling in the Economic Evaluation of Health Care: Selecting the Appropriate Approach. Journal of Health Services Research & Policy. 2004;9(2):110-8.
12. England PH. Newborn and infant physical examination screening: programme overview 2013 [Available from: https://www.gov.uk/guidance/newborn-and-infant-physical-examination-screening-programme-overview.
13. Curtis L, Burns A. Unit Costs of Health and Social Care 2017. 2017.
14. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes. Third Edition ed. Oxford: Oxford University Press; 2005.
15. Pandor A, Eastham J, Beverley C, Chilcott J, Paisley S. Clinical effectiveness and costeffectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review. Health Technology Assessment. 2004;8(12).
16. Czoski-Murray C, Lloyd Jones M, McCabe C, Claxton K, Oluboyede Y, Roberts J, Nicholl JP, Rees A, Reilly CS, Young D, Fleming T. What is the value of routinely testing full blood count, electrolytes and urea, and pulmonary function tests before elective surgery in patients with no apparent clinical indication and in subgroups of patients with common comorbidities: a systematic review of the clinical and cost-effective literature. Health Technology Assessment. 2012;16(50).
17. Idec B. Natalizumab (Tysabri®) for the Treatment of Adults with Highly Active Relapsing Remitting Multiple Sclerosis.

18. 2014 [Available from: https://www.whatdotheyknow.com/request/199758/response/498958/attach/3/FOI23023%20ResponseAttachment%20Deakin.pdf.
19 Labratory Costs.
20. NHS Reference Costs 2016-2017 [Internet]. 2017.
21. Curtis L, Burns A. Unit Costs of Health & Social Care 2015. 2015.
22. Roberts TE et al. Economic evaluation and randomised controlled trial of extracorporeal membrane oxygenation: UK collaborative trial. BMJ. 1998;317:911-6.
23. Wren C, Reinhardt Z, Khawaja K. Twenty-year trends in diagnosis of life-threatening neonatal cardiovascular malformations. Arch Dis Child Fetal Neonatal Ed. 2008;93(1):F33-5.
24. Fenwick E, Marshall DA, Levy AR, Nichol G. Using and interpreting cost-effectiveness acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. BMC Health Serv Res. 2006;6:52.

8. FIGURES

Figure 1: The alternative screening pathways

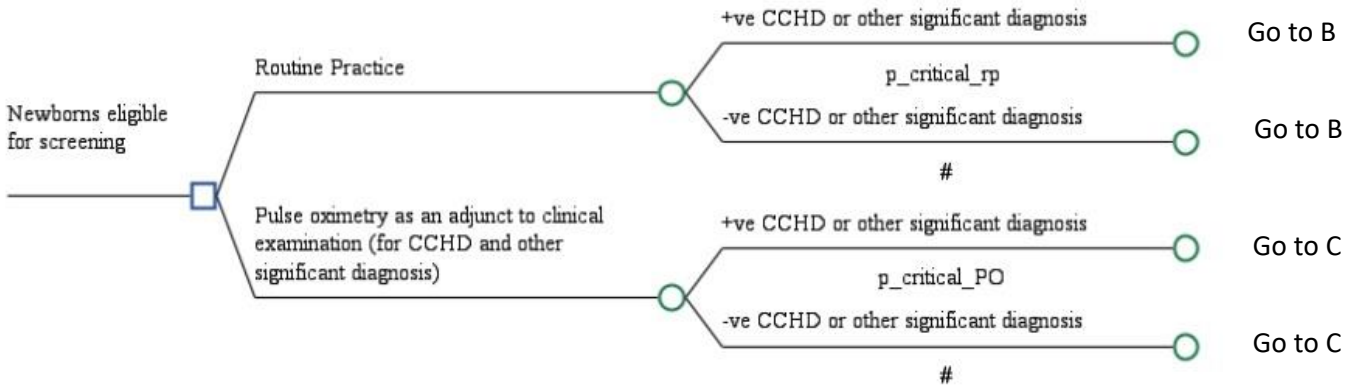


Figure 2: **Branch node B -** the pathways followed by those in Routine Practice.

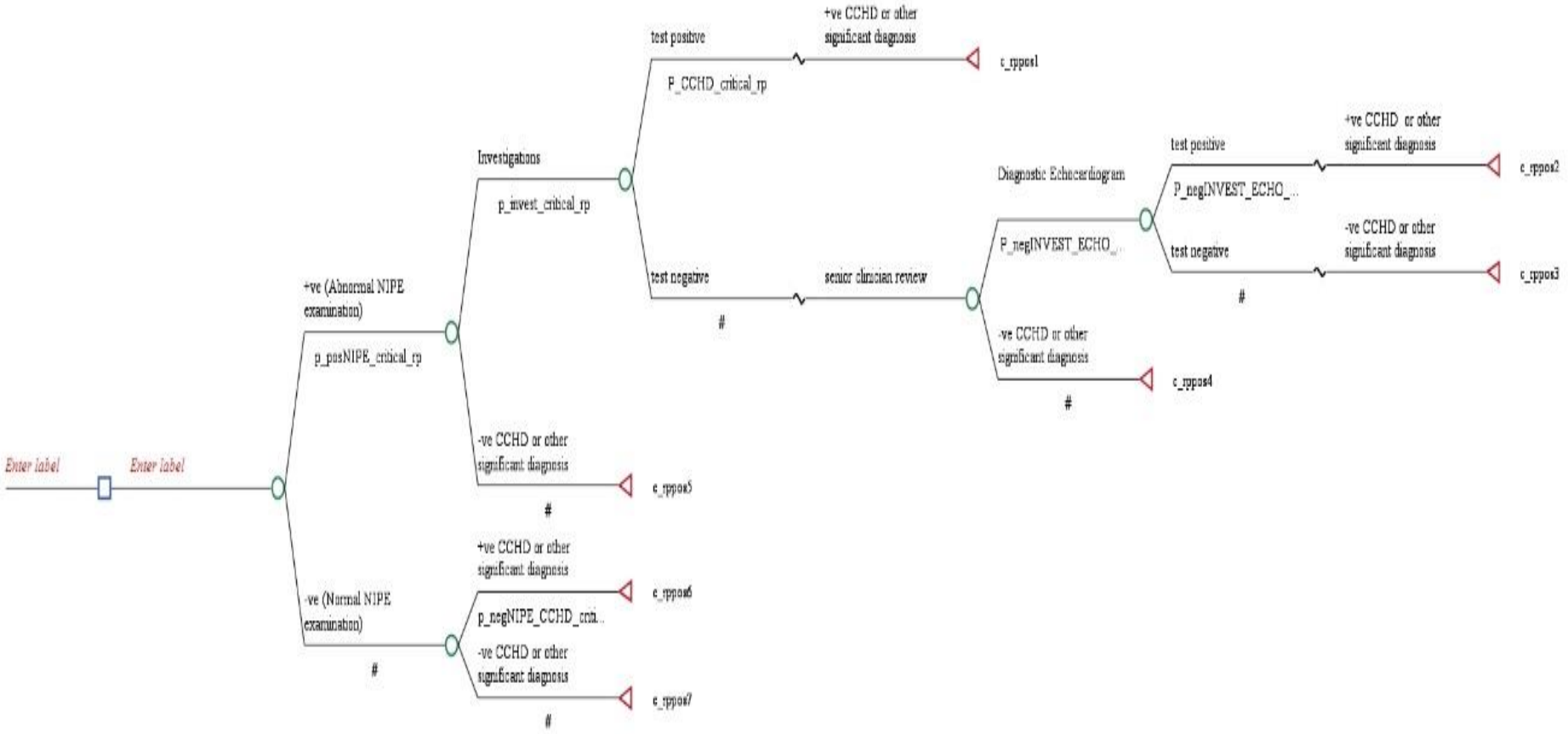


Figure 3: Branch node C: the pathways for those undergoing screening

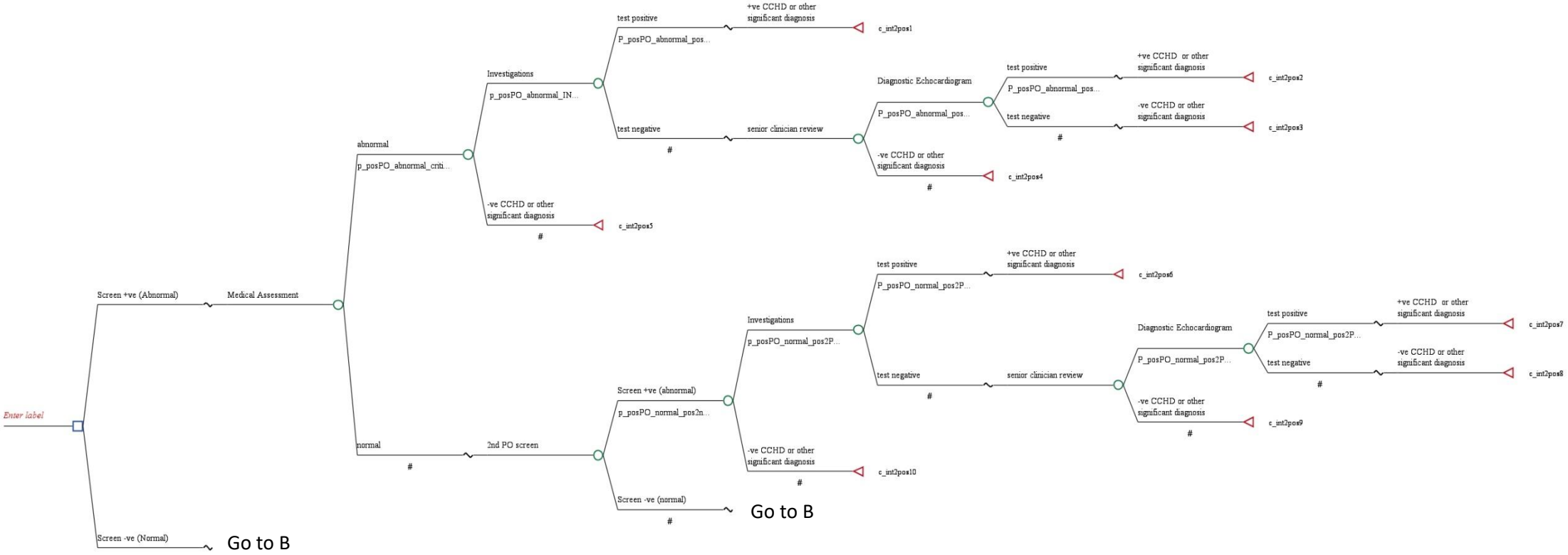


Table 1: Table of Costs

Item	Cost in 2016/2017 prices	Sources
Tests		
Pulse Oximetry test (first)	8.50	[1-3]
Pulse Oximetry test (second)	8.44	[1, 3, 2]
Assessments/Examinations		
Medical Assessment (urgent or expedited)	4.29	[1, 3, 2]
Senior clinician review	21.50	[1]
Investigations		
Further examination by a senior clinician	21.50	[1]
Blood Gas	4.45	[4]
Blood Culture	1.58	[5]
C-reactive protein test	3.26	[4]
Full blood count	7.62	[6]
Chest X-Ray	26.20	[7]
Urea and Electrolytes	5.08	[6]
Lumbar Puncture	125.15	[8]
Electrocardiography	81.02	[9]
4 limb BP	7.23	[10, 1]
Diagnostic Echocardiogram	139.61	[9]
Transport to cardiac centre	396.17	[11]
Home birth/MLU – transport to hospital	247	[9]
Neonatal unit costs (cost per 24hours)		
Intensive care	1,295	[9]
High Dependency	897	[9]
Special Care	423	[9]

1. Curtis LAaB, Amanda Unit Costs of Health and Social Care 20172017.
2. Roberts TE, Barton PM, Auguste PE, Middleton LJ, Furmston AT, Ewer AK. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a cost-effectiveness analysis. Arch Dis Child. 2012;97(3):221-6. doi:10.1136/archdischild-2011-300564.
3. Newborn Pulse Oximetry Screening Pilot End Project Report. Public.screeninghelpdesk@nhs.net; 2016.
4. Labratory Costs.
5. A Pandor JE, C Beverley, J Chilcott and S Paisley. Clinical effectiveness and costeffectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review. Health Technology Assessment. 2004;8(12).
6. C Czoski-Murray MLJ, C McCabe, K Claxton, Y Oluboyede, J Roberts, JP Nicholl, A Rees, CS Reilly, D Young and T Fleming. What is the value of routinely testing full blood count, electrolytes and urea, and pulmonary function tests before elective surgery in patients with no apparent clinical indication and in subgroups of patients with common comorbidities: a systematic review of the clinical and cost-effective literature. Health Technology Assessment. 2012;16(50).
7. . 2014. https://www.whatdotheyknow.com/request/199758/response/498958/attach/3/FOI23023%20ResponseAttachment%20Deakin.pdf.
8. Idec B. Natalizumab (Tysabri®) for the Treatment of Adults with Highly Active Relapsing Remitting Multiple Sclerosis.
9. NHS Reference Costs 2016-2017 [database on the Internet]2017. Accessed:
10. Burns LCaA. Unit Costs of Health & Social Care 20152015.
11. Group TERatEMOEWGobotEMOTS. Economic evaluation and randomised controlled trial of extracorporeal membrane oxygenation: UK collaborative trial. BMJ. 1998;317:911-6.

ANALYSIS 1

Table 2a: Infant outcomes for POS and Routine Practice (RP Prevalence calibrated to PO prevalence) (Analysis 1)

	POS Number of infants*	RP Number of infants*
Total Number of Infants	**31,050**	**31,050**
Infants with Target Condition	**74**	**74**
Infants falsely diagnosed as healthy	3	22
Infants with Target Condition detected via pulse oximetry screening	49	-
Infants with Target Condition detected via clinical examination	22	52
Infants with Target Condition having investigations	71	52
Infants with Target Condition having invasive investigations	25	18
Infants with Target Condition having an ECHO	29	21
Infants with Target Condition admitted to the NNU	69	52
Infants without Target Condition	**30,976**	**30,976**
Infants without Target Condition with a positive pulse oximetry screen	176	-
Infants without Target Condition with a positive clinical examination	-	528
Infants without Target Condition having investigations	620	493
Infants without Target Condition having invasive investigations	254	202
Infants without Target Condition having an ECHO	84	66
Infants without Target Condition admitted to the NNU	658	528

*Numbers rounded to nearest whole number

Table 3a: Costs associated with infants in the Pulse Oximetry Screening Strategy (RP Prevalence calibrated to PO prevalence) (Analysis 1)

Item	Cost item	Unit cost (£)	*n* (screen positive to PO screening)	*n* (screen negative to PO screening)	*n* (total)	Total cost (£)
Tests						
Pulse Oximetry Test (first)	Per test	8.50	226	30,824	31,050	263,925.00
Pulse Oximetry test (second)	Per test	8.44	195	0	195	1,645.80
					0	0.00
Assessments/Examinations					0	0.00
Medical Assessment (urgent or expedited)	Per examination	4.29	226	0	226	969.54
Senior clinician review	Per Review	21.50	82	568	650	13,975.00
					0	0.00
Investigations					0	0.00
Further examination by a senior clinician	Per examination	21.50	178	581	759	16,318.50
Blood Gas	Per test	4.45	92	581	673	2,994.85
Blood Culture	Per test	1.58	98	581	679	1,072.82
C-reactive protein test	Per test	3.26	98	581	679	2,213.54
Full blood count	Per test	7.62	100	581	681	5,189.22
Chest X-Ray	Per test	26.20	86	581	667	17,475.40
Urea and Electrolytes	Per test	5.08	75	581	656	3,332.48
Lumbar Puncture	Per test	125.15	42	237	279	34,916.85
Electrocardiography	Per test	81.02	24	178	202	16,366.04
4 limb BP	Per test	7.23	18	121	139	1,004.97
Diagnostic Echocardiogram	Per test	139.61	31	83	114	15,915.54
					0	0.00
Transport to cardiac centre	Per infant	396.17	6	0	6	2,377.02
Home birth/MLU – transport to hospital	Per infant	247	6	0	6	1,482.00
					0	0.00
Neonatal unit costs					0	0.00
Intensive care	Per 24 hours	1,295	114	43	157	203,315.00
High Dependency	Per 24 hours	897	97	44	141	126,477.00
Special Care	Per 24 hours	423	146	2,985	3,131	1,324,413.00
Total Costs for POS						2,055,379.57
Average cost per infant			31,050			66.20

Table 4a: Costs associated with infants in the Routine Practice Strategy (RP Prevalence calibrated to PO prevalence) (Analysis 1)

Item	Cost item	Unit cost (£)	*n*	Total cost (£)
Assessments/Examinations				
Senior clinician review	Per Review	21.50	514	11,051.00
Investigations				
Further examination by a senior clinician	Per examination	21.50	546	11,739.00
Blood Gas	Per test	4.45	546	2,429.70
Blood Culture	Per test	1.58	546	862.68
C-reactive protein test	Per test	3.26	546	1,779.96
Full blood count	Per test	7.62	546	4,160.52
Chest X-Ray	Per test	26.20	546	14,305.20
Urea and Electrolytes	Per test	5.08	546	2,773.68
Lumbar Puncture	Per test	125.15	220	27,533.00
Electrocardiography	Per test	81.02	124	10,046.48
4 limb BP	Per test	7.23	110	795.30
Diagnostic Echocardiogram	Per test	139.61	87	12,146.07
Neonatal unit costs				
Intensive care	Per 24 hours	1,295	102	132,090.00
High Dependency	Per 24 hours	897	104	93,288.00
Special Care	Per 24 hours	423	2,819	1,192,437.00
Total Costs for Routine Practice				1,517,437.59
Average cost per infant			31,050	48.87

Table 5a: Summary of Analysis 1 (RP prevalence calibrated to PO prevalence). Based on a cohort of 31,050 infants with an outcome of cost per timely diagnosis to detect CCHD or other significant diagnoses (2016-2017 prices) **(Analysis 1)**

Strategy	Expected total cost per screening strategy (£)	Incremental Cost (£)	Number of Infants receiving timely diagnosis (% of total infants)	Incremental timely diagnosis detected	ICER (£)
Pulse Oximetry Screening as an adjunct to Clinical Examination	2,054,889	521950	31,047 (99.99)	19	27,502*
Routine Practice: Clinical Examination	1,532,939	-	31,028 (99.93)	-	-

*slight discrepancy in calculations due to rounding

ANALYSIS 2

Table 2a: Infant outcomes for POS and Routine Practice (PO Prevalence calibrated to RP prevalence) (Analysis 2)

	POS Number of infants*	RP Number of infants*
Total Number of Infants	**31,050**	**31,050**
Infants with Target Condition	**488**	**488**
Infants falsely diagnosed as healthy	3	144
Infants with Target Condition detected via pulse oximetry screening	49	-
Infants with Target Condition detected via clinical examination	436	344
Infants with Target Condition having investigations	485	344
Infants with Target Condition having invasive investigations	168	119
Infants with Target Condition having an ECHO	195	137
Infants with Target Condition admitted to the NNU	483	344
Infants without Target Condition	**30,562**	**30,562**
Infants without Target Condition with a positive pulse oximetry screen	176	-
Infants without Target Condition with a positive clinical examination	-	521
Infants without Target Condition having investigations	620	487
Infants without Target Condition having invasive investigations	254	200
Infants without Target Condition having an ECHO	84	65
Infants without Target Condition admitted to the NNU	658	521

*Numbers rounded to nearest whole number

Table 3a: Costs associated with infants in the Pulse Oximetry Screening Strategy (PO Prevalence calibrated to RP prevalence) (Analysis 2)

Item	Cost item	Unit cost (£)	*n* (screen positive to PO screening)	*n* (screen negative to PO screening)	*n* (total)	Total cost (£)
Tests						
Pulse Oximetry Test (first)	Per test	8.50	226	30,824	31,050	263,925.00
Pulse Oximetry test (second)	Per test	8.44	195	0	195	1,645.80
					0	0.00
Assessments/Examinations					0	0.00
Medical Assessment (urgent or expedited)	Per examination	4.29	226	0	226	969.54
Senior clinician review	Per Review	21.50	82	733	815	17,522.50
					0	0.00
Investigations					0	0.00
Further examination by a senior clinician	Per examination	21.50	178	994	1,172	25,198.00
Blood Gas	Per test	4.45	92	994	1,086	4,832.70
Blood Culture	Per test	1.58	98	994	1,092	1,725.36
C-reactive protein test	Per test	3.26	98	994	1,092	3,559.92
Full blood count	Per test	7.62	100	994	1,094	8,336.28
Chest X-Ray	Per test	26.20	86	994	1,080	28,296.00
Urea and Electrolytes	Per test	5.08	75	994	1,069	5,430.52
Lumbar Puncture	Per test	125.15	42	380	422	52,813.30
Electrocardiography	Per test	81.02	24	218	242	19,606.84
4 limb BP	Per test	7.23	18	164	182	1,315.86
Diagnostic Echocardiogram	Per test	139.61	31	248	279	38,951.19
					0	0.00
Transport to cardiac centre	Per infant	396.17	6	0	6	2,377.02
Home birth/MLU – transport to hospital	Per infant	247	6	0	6	1,482.00
					0	0.00
Neonatal unit costs					0	0.00
Intensive care	Per 24 hours	1,295	114	848	962	1,245,790.00
High Dependency	Per 24 hours	897	97	865	962	862,914.00
Special Care	Per 24 hours	423	146	5,307	5,453	2,306,619.00
Total Costs for POS						4,893,310.83
Average cost per infant			31,050			157.59

Table 4a: Costs associated with infants in the Routine Practice Strategy (PO Prevalence calibrated to RP prevalence) (Analysis 2)

Item	Cost item	Unit cost (£)	*n*	Total cost (£)
Assessments/Examinations				
Senior clinician review	Per Review	21.50	624	13,416.00
Investigations				
Further examination by a senior clinician	Per examination	21.50	830	17,845.00
Blood Gas	Per test	4.45	830	3,693.50
Blood Culture	Per test	1.58	830	1,311.40
C-reactive protein test	Per test	3.26	830	2,705.80
Full blood count	Per test	7.62	830	6,324.60
Chest X-Ray	Per test	26.20	830	21,746.00
Urea and Electrolytes	Per test	5.08	830	4,216.40
Lumbar Puncture	Per test	125.15	319	39,922.85
Electrocardiography	Per test	81.02	186	15,069.72
4 limb BP	Per test	7.23	139	1,004.97
Diagnostic Echocardiogram	Per test	139.61	202	28,201.22
Neonatal unit costs				
Intensive care	Per 24 hours	1,295	668	865,060.00
High Dependency	Per 24 hours	897	681	610,857.00
Special Care	Per 24 hours	423	4421	1,870,083.00
Total Costs for Routine Practice				3,501,457.46
Average cost per infant			31,050	112.77

Table 5a: Summary of Analysis 1 (PO prevalence calibrated to RP prevalence). Based on a cohort of 31,050 infants with an outcome of cost per timely diagnosis to detect CCHD or other significant diagnoses (2016-2017 prices) (Analysis 2)

Strategy	Expected total cost per screening strategy (£)	Incremental Cost (£)	Number of Infants receiving timely diagnosis (% of total infants)	Incremental timely diagnosis detected	ICER (£)
Pulse Oximetry Screening as an adjunct to Clinical Examination	4,892,859	1,413,396	31,047 (99.99)	141	9,996*
Routine Practice: Clinical Examination	3,479,463	-	30,906 (99.54)	-	-

*slight discrepancy in calculations due to rounding

10. APPENDICES

Appendix A: Matching of Trusts

The Trusts were independently matched by Claire Evans Clinical Lead for the National Screening Committee (NSC) PulseOx pilot using the following methods:

- Neonatal Units were review of on their respective websites to establish information about the size and level of intensive care provided (Level of Neonatal Unit (NNU) e.g. Level 1 (Special Care), 2 or Level 3 NICU), birth rate in each Trust and number of cots in each Neonatal Unit
- Geographic location of the Trust
- Trusts that were already performing PO screening were established from data from the recent UK National survey and Neonatal Units were contacted by phone and Remaining potential matching Trusts were asked if they performed PO screening and confirmed level of care provided and number of cots
- Non-PO screening Trusts were then matched the to the Trusts participating in the PO Pilot by birth rate, level and size of Neonatal Unit and geographic location

Appendix B: Collection of Birmingham Women's Hospital (BWH) data

In order to attempt to ascertain how many babies who were test negative might be admitted to NNU with a suspected significant other condition the BadgerNet data from Birmingham Women's Hospital Neonatal Unit was interrogated as follows:

Data on one year's total admissions were collected

Those outside the gestational range 35-42 weeks were excluded, of the remaining babies those who did not receive antibiotics (as a proxy for suspected significant other condition) were also excluded.

Admissions from delivery suite or within 6 hours (I.e. unlikely to have had PO screening) were then excluded.

Clinical data from remaining admissions from post natal ward (PNW) (assumed PO screened) was then reviewed and the number admitted because of a positive PO screen, were excluded. The number of the remaining admissions with other conditions, were identified and maximum level of care identified to assess significant illness (HDU or ITU care)

26 babies were admitted with a positive PO result and 167 admitted who were PO screen negative.

54% (14/26) of screen positive babies and 6% (6/103) of screen negative had significant other conditions and

For BWH admissions 1/1/16-31/12/16

Admission 35-42 week	= 1583
Admission who received antibiotics (as a proxy for suspected illness)	= 1033
Admissions from delivery suite or within 6 hours (I.e. unlikely to have been screened	= 832
Admissions form PNW or home	= 201
Admissions from home (therefore assumed to have passed both PulseOx and NIPE	= 8

Admissions from PNW (assumed PO screened) = 193

PO screen positives admitted = 26

Assumed PO screen negatives admitted = 167

No of screen negatives with condition of interest (i.e. respiratory/septic illness) = 103

No of PO screen negatives with conditions of interest (i.e. respiratory/septic illness) who have significant illness = 6/103 (6%)

No of PO screen positives with significant illness (i.e. respiratory/septic illness) = 14/26 (54%)

No of screen positives with NS illness = 12/26

No of screen negatives with NS illness = 97/103 (94%)

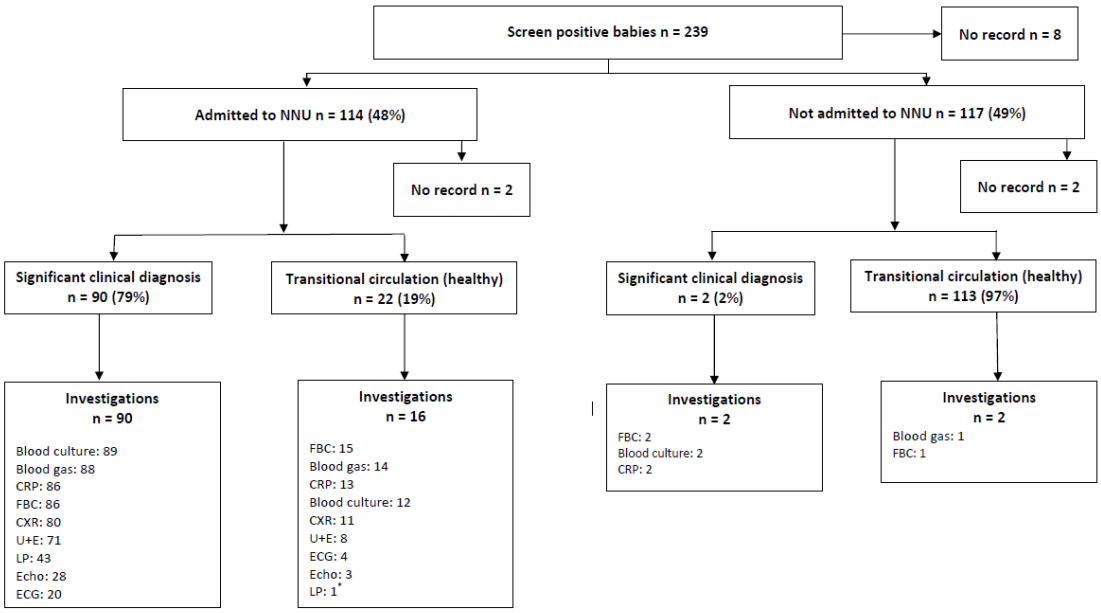
Appendix C: Investigations in PO positive babies in the Pilot study

Of the 231 babies who tested positive in the pilot (where full information was available), 110 had one or more investigations all but 4 were admitted to NNU also (121 had no investigations).

Of the 110 babies who had investigations the majority had blood tests and chest x-ray only.

43 babies had a lumbar puncture (invasive investigation) and 31 had an echocardiogram. Almost all babies who underwent these investigations has a significant clinical diagnosis (3 Echos only on babies who did not).

Figure 3: PO Screen positive babies – management and investigations



*LP was only investigation which suggests this was recorded in erro