Cost-effectiveness of newborn blood spot screening for Tyrosinaemia type 1 using tandem mass spectrometry - Final report

Review group: Peter Auguste
Julia Geppert
Sian Taylor-Phillips
Chris Stinton
Samantha Johnson
Aileen Clarke
Hema Mistry

Correspondence to: Mr Peter Auguste
Warwick Evidence
Division of Health Sciences
Warwick Medical School
University of Warwick
Coventry, CV4 7AL
Email: p.auguste@warwick.ac.uk

Date completed: 7th July 2021

The views expressed in this report are those of the authors and not necessarily those of the UK NSC. Any errors are the responsibility of the authors. The authors have no conflicts of interest.

Table of contents

List of tables

List of figures

List of abbreviations

CEAC	Cost-effectiveness acceptability curve
DBS	Dried blood spot
EVPI	Expected value of perfect information
HCHS	Hospital and Community Health Services
ICER	Incremental cost-effectiveness ratio
LYG	Life years gained
MS/MS	Tandem mass spectrometry
NBS	Newborn blood spot
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NSC	National Screening Committee
ONS	Office for National Statistics
Phe	Phenylalanine
PKU	Phenylketonuria
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALYs	Quality Adjusted Life Years
SUAC	Succinylacetone
Tyr	Tyrosine
TYR1	Tyrosinaemia type I
VoI	Value of information
WTP	Willingness-to-pay

Plain English Summary

Tyrosinaemia type 1 (TYR1) is a rare inherited disorder where a person has raised blood levels of a protein known as amino acid tyrosine. One person in 100,000 is affected with TYR1 globally, but it may be more common in some areas. In the UK, approximately seven babies are born each year with TYR1. Symptoms include diarrhoea and, jaundice may appear within the first few months of life. TYR1 can lead to liver and kidney failure, and there is also an increased risk of learning difficulties and liver cancer. There is no cure for TYR1. TYR1 is treated with a special diet and a drug called nitisinone, both of which continue for the rest of life. If untreated, death from liver failure or liver cancer occurs before the age of 10 years.

Currently in the UK, there is no universal screening of newborns for TYR1. Screening is in place for nine conditions (e.g. sickle cell disease and cystic fibrosis) through using the newborn bloodspot test. Newborns with siblings living with TYR1 are identified through genetic testing and others may be identified when undergoing screening for phenylketonuria (PKU). Incidental finding of newborns with TYR1 through PKU screening is not ideal, as three babies are likely to be missed and then perhaps diagnosed later in life with TYR1.

Evidence is available to support that there is clinical benefit in the early diagnosis of TYR1 in newborns. However, clinical benefit alone is not enough to justify funding healthcare technologies, such as a screening programme. To fund a technology, it must not only be clinically effective, but must also offer good value for money. Undertaking an economic analysis can show whether a technology offers good value for money and that can also save lives, and/or improve health. An economic analysis allows decision makers to distribute scarce healthcare resources with a limited budget.

Economic analyses often use economic models which are developed using clinical and economic evidence in a form that can be used to aid the decision-making process. These model-based economic analyses are part of the decision-making process for the approval of healthcare technologies in the UK such as by the National Institute for Health and Care Excellence (NICE). Economic models can be used to extend beyond the initial period of a clinical trial and are often conducted under uncertainty, which is common in a rare disease setting. Economic models are often developed with input from clinical experts.

A systematic review of the evidence showed that there were no economic analyses that assessed using tandem mass spectrometry (MS/MS) in newborn bloodspot screening using succinylacetone as a marker for identifying newborns with TYR1 in the UK. Hence, a new economic model was required to address the research question. Given the complexity of the model, we developed a preliminary

model structure to show the clinical pathway for newborns undergoing current screening and if universal screening for TYR1 was implemented. The pathway also included the diagnosis of TYR1 using a confirmatory testing and the treatment of sequelae that may be associated with TYR1 or treatment of TYR1 using nitisinone. Several iterations of the model structure were developed alongside discussions with the clinical experts before a final conceptual model was agreed.
There was a lack of robust published clinical and economic evidence, so we drew on clinical expert opinion. Group discussions were held with the clinical experts, and they provided opinion on patient pathways, helped to describe/define the outcomes by which screening should be measured, identified information sources, and discussed inputs for and meaning of the findings of the model.

To address the research question, an economic model was developed. The model can be used to show the benefits (and any harms that may exist) of early diagnosis and treatment of TYR1 compared to current practice and, if it offers good value for money. The model's main outputs were the incremental benefit associated with correct diagnoses, life-years gained (LYG) and quality-adjusted life-years (QALY). These outputs were chosen because they show the short-term and long-term benefits of introducing population screening for TYR1, and more importantly the cost-effectiveness. Understanding the long-term benefits of an intervention is preferred in decision making, with results commonly reported in terms of cost per QALY gained. This allows decision-makers to choose how to allocate resources between groups competing for scarce resources by providing insight on the likely benefits from investing in new technologies such as screening. QALYs consider both the quantity (life expectancy) and quality of life of the remaining life years generated by health care interventions. For example, four years of living at an adjusted value of 0.5 is equal to two QALYs. This is also equivalent to somebody living for two years in full health. Some people may be indifferent between living four years with a condition valuing 0.5 and two years in full health.

The economic model predicted that if we continue with current practice (incidental finding of TYR1 by using PKU), we are likely to identify three babies with TYR1 later in their life when they present with symptoms. Most of these babies may present with liver disease and may require liver transplant later in life. Conversely, with universal screening for TYR1, we are likely to identify four babies through early detection, with these babies avoiding liver disease thus not requiring liver transplant. The model showed that more screened detected babies who are asymptomatic are likely to develop learning difficulties compared to babies who have presented symptomatically. There is no consensus if learning difficulties are due to early treatment with nitisinone or TYR1 itself. With regards to life expectancy, the model found that there to be modest gains between babies identified through PKU screening and those identified via universal screening. This result was expected as babies who had been missed would present early with symptoms and be diagnosed, and thus would not have been at an increased risk of death if they had presented with symptoms later in life. Hence, the small gains in

life years gained. The difference in QALYs gained was more noticeable but still modest gains, which is largely due to avoiding liver disease. Despite higher numbers of babies with learning difficulties, the utility value associated with learning difficulties is higher than with people living with liver disease.

The proposed universal screening of using tandem MS/MS in newborn bloodspot screening using succinylacetone at day five as a marker in addition to cascade testing should result in identifying babies with TYR1 earlier and less chance of babies receiving a false positive result. Early detection is likely to avoid babies presenting or developing liver disease and thus avoiding the need for liver transplant. If there are three symptomatic cases each year, of which 17% require transplants per year equates to approximately one liver transplant every two years. However, with the introduction of universal screening may lead to more cases of learning difficulty. These outcomes expected come at a cost of approximately an additional £1.4 million when screening 700,000 newborns. This cost comprises additional screening costs, treatment costs and costs associated with treatment of sequelae. This cost also includes the costs that would be avoided by treating liver disease and avoiding liver transplant. These results highlight the burden of the disease from which universal screening aims to prevent.

Due to the lack of evidence, the economic model was heavily reliant on clinical expert opinion, and the uncertainty in the model was addressed by undertaking additional analyses to see the impact on the results. One such analysis included bringing all this uncertainty together, then recalculating the results 10,000 times. These results showed that implementing universal screening compared to current practice is likely to offer good value for money, as it would be considered acceptable by NICE based on society's willingness-to-pay per QALY in their orphan drugs assessment. However, due to limited evidence, there is a need for further research to refine these results.

Scientific summary

Introduction

Tandem mass spectrometry (MS/MS) used in newborn blood spot (NBS) screening can be expanded using succinylacetone (SUAC) as a marker for identifying babies with tyrosinaemia type 1 (TYR1). We aimed to investigate whether extending the current screening programme for inborn errors of metabolism in the UK offers good value for money. This requires assessing the costs and benefits of the proposed programme compared with standard practice. We developed a decision analytical model that comprised two stages. The first stage uses a decision tree structure and predicts the number of TYR1 cases identified in the current (without universal screening for TYR1) and proposed (adding TYR1 to the current NBS screening programme) approaches by NBS screening, by cascade testing because of previously affected sibling(s) or by symptomatic presentation. The second stage of the model uses a Markov structure and considers the treatment and management of TYR1 and its long-term sequelae.

Methods

The model started with a hypothetical cohort of babies who received universal screening for TYR1 compared to a cohort receiving no universal TYR1 screening. The model then followed the pathway to reflect the mode of identification and treatment of TYR1 and its sequelae. The Markov component of the model had different cycle lengths to reflect monitoring of babies. The cycle lengths were 4-monthly for the first year of life, then 6-monthly. The relevant clinical information on sensitivity and specificity of each screening strategy, on birth prevalence, and on the probabilities of developing long-term complications were linked to resource use and cost information and to outcomes. Identified evidence on the magnitude of any potential benefits of earlier detection and treatment following screening on health outcomes was limited and at high risk of bias. The economic analyses were undertaken from the National Health Service (NHS) and Personal Social Service (PSS) perspective, with all future costs incurred and benefits accrued being discounted at 3.5% per annum. The cost-effectiveness was conducted over a lifetime horizon, with the primary outcome based on the quality-adjusted life years (QALYs). Other secondary outcome measures were also considered: cases of TYR1 correctly identified by screen detection and life-years gained. The results were presented in the form of an incremental cost-effectiveness ratio (ICER) expressed as cost per QALY gained.

Deterministic analysis was undertaken for the base-case results for the primary and secondary outcome measures. Additionally, probabilistic sensitivity analysis was undertaken for the outcome measure of cost per QALYs gained, to incorporate the joint uncertainty in the key model input parameters of prevalence, sensitivity and specificity, and utility values. Scenario and sensitivity analyses were also undertaken to determine the key drivers of cost-effectiveness.

Results

Deterministic base-case results estimated that the proposed approach with TYR1 added to the current NBS screening programme is both more costly and more effective than the current practice without universal TYR1 screening. Expanding the NBS screening programme for TYR1 was £214,094 more costly per 100,000 live births compared to current practice and was expected to yield an additional 0.50 life-years and 3.5 QALYs, respectively. This equates to an ICER of £423,000 per life-year gained and £61,800 per QALY gained, respectively. The probabilistic sensitivity analysis showed that expanding NBS screening using SUAC as a marker for identifying TYR1 had a near zero probability of being cost-effective at a willingness-to-pay threshold of £30,000 per QALY gained compared to the current approach. However, at a willingness-to-pay threshold of £100,000 per QALY gained it had a near one hundred percent probability of being cost-effective. The results from the scenario and sensitivity analyses showed that information about the costs of each strategy, discount rate and transition probabilities for liver disease were key drivers for the base-case incremental cost-effectiveness ratio. There were serious limitations in the clinical evidence that underpinned the economic model, so results should be interpreted with caution.

Assuming a population of 655,000 screened newborns, our model estimates that introducing universal screening for TYR1 will detect an extra three cases which would have otherwise been detected symptomatically around six-months of age, costing approximately an additional £1.4million per year. Overall, we estimate that this will give an extra 24 quality adjusted life years compared to the current practice of phenylketonuria testing. This was largely through improvements in quality of life through avoiding liver disease. The model estimates 89 fewer false positive results per year, but with uncertainty in test accuracy estimates underpinning them. We are very unsure of some of the values included in the economic model, so the results should be interpreted with caution. The model estimated more cases of learning disability, through the earlier administration of nitisinone, but the scientific community remains unsure whether the learning difficulties are as a result of nitisinone treatment or tyrosinaemia itself.

1 Introduction

Tandem mass spectrometry (MS/MS) used in newborn blood spot (NBS) screening can be expanded to identify babies with tyrosinaemia type 1 (TYR1). To evaluate if extending the NBS screening programme for inborn errors of metabolism in the UK offers good value for money requires assessing the costs and benefits of the proposed programme compared with standard practice. The objective of this economic analysis is to assess the cost-effectiveness of adding TYR1 screening using MS/MS measurement of succinylacetone (SUAC) to the current NBS screening programme compared with current practice (no universal TYR1 screening).

To address our objective, we developed a *de novo* economic model to structure evidence of clinical and economic outcomes in a form that can be used to inform decisions on clinical practices and allocation of resources in order to achieve maximum health benefits.[(1)] Prior to developing the economic model, the authors of this report undertook a rapid review of the published literature to identify existing cost-effectiveness models that compared screening for TYR1 using MS/MS measurement of SUAC from dried blood spots (DBS) in newborns versus no universal TYR1 screening. Additionally, searches in the form of rapid reviews were undertaken to identify studies which reported information about the birth prevalence and methods of detection of TYR1 in the UK, test accuracy (sensitivity and specificity) of SUAC measurement in DBS using MS/MS for TYR1 screening, long-term outcomes/complications in pre-symptomatically and symptomatically detected patients, resource use and costs of diagnosing and treating TYR1, utilities and harms/side-effects of treatments for TYR1.

This report will outline the methods which were used for underpinning the *de novo* economic model including the model structure and assumptions, transition probabilities, resource use and costs and utility values. The report will present the preliminary draft results including scenario analyses which were sent to the UK National Screening Committee in September 2019 and before feedback was received from the experts at the Tyrosinaemia workshop which was held on 22nd July 2020. Following on from this workshop, further additional analyses were undertaken, and a final workshop took place on 6th May 2021. These additional analyses are the final base-case results and scenario analyses.

2 Methods

2.1 Rapid review methodology

We used an enhanced rapid evidence assessment approach which is the core approach used by the UK National Screening Committee (NSC) for their evidence review process.[(2)] This uses systematic review methodologies as a reference point to ensure transparent reporting and rigorous conduct and

we followed the same methodology used in the 2016 UK NSC review on NBS screening for TYR1 in the UK.[3]

For the clinical effectiveness review, we updated our three searches (optimised to identify evidence on TYR1 incidence, screening test accuracy, and treatment, respectively) from the previous 2016 UK NSC review.[3] We searched MEDLINE (Ovid), MEDLINE In-Process & Other Non-Indexed Citations (Ovid), EMBASE (Ovid), Web of Science and the Cochrane Library from the search dates of the previous review (incidence search)[3] or its updated searches (screening test accuracy and treatment searches) for publications in peer-reviewed journals,[4; 5] respectively, up to 21 May 2018 (incidence and screening test accuracy searches) or 11 June 2018 (treatment search). These were supplemented with included studies from the 2016 UK NSC review by the University of Warwick and its related publications.[3; 4; 5] We also re-screened all records included at title/abstract stage in the 2016 UK NSC review and its updated searches for publications in peer-reviewed journal for studies providing information on resource use and costs of diagnosing and treating TYR1, utilities and harms/side-effects of treatments for TYR1.

For the cost-effectiveness review, we searched MEDLINE (Ovid), MEDLINE In-Process & Other Non-Indexed Citations (Ovid), EMBASE (Ovid), Web of Science, EconLit, and the Cochrane Library on 21 May 2018 without a date limit.

For all the searches, we also screened the reference lists of all included articles and relevant systematic reviews.

One reviewer screened the titles/abstracts and full texts of all records identified by the searches, performed data extraction and quality appraisal for all studies providing evidence on test accuracy or benefit of early versus late start of treatment, with a random 20% checked by a second reviewer. All included economic evaluations were extracted and quality appraised by one reviewer with 100% checked by a second reviewer. Included articles which provided information on TYR1 birth prevalence and methods of detection of TYR1 in the UK, resource use and costs of diagnosing and treating TYR1, utilities and harms/side-effects of treatments for TYR1 were not formally extracted and quality appraised.

2.2 Developing the model structure

An economic model was developed with clinical input. It was programmed in TreeAge pro (TreeAge Software Inc., Williamstown, MA, USA) to represent the clinical pathway that babies would take while being screened and treated for TYR1. An illustrative structure of the clinical pathway is shown in Figure 1. The decision-analytical model comprised two stages: 1) diagnosis of TYR1, and 2)

treatment and management of TYR1 and its sequelae. In the first stage, we used a decision tree structure to represent the pathway babies would take in the proposed approach with universal screening for TYR1 added to the NBS screening programme and compared it to the pathway in current practice (without universal NBS screening for TYR1) (see Figure 2). In the second stage, we used a Markov model to simulate the progression and treatment of long-term events related to TYR1 (see Figure 3).

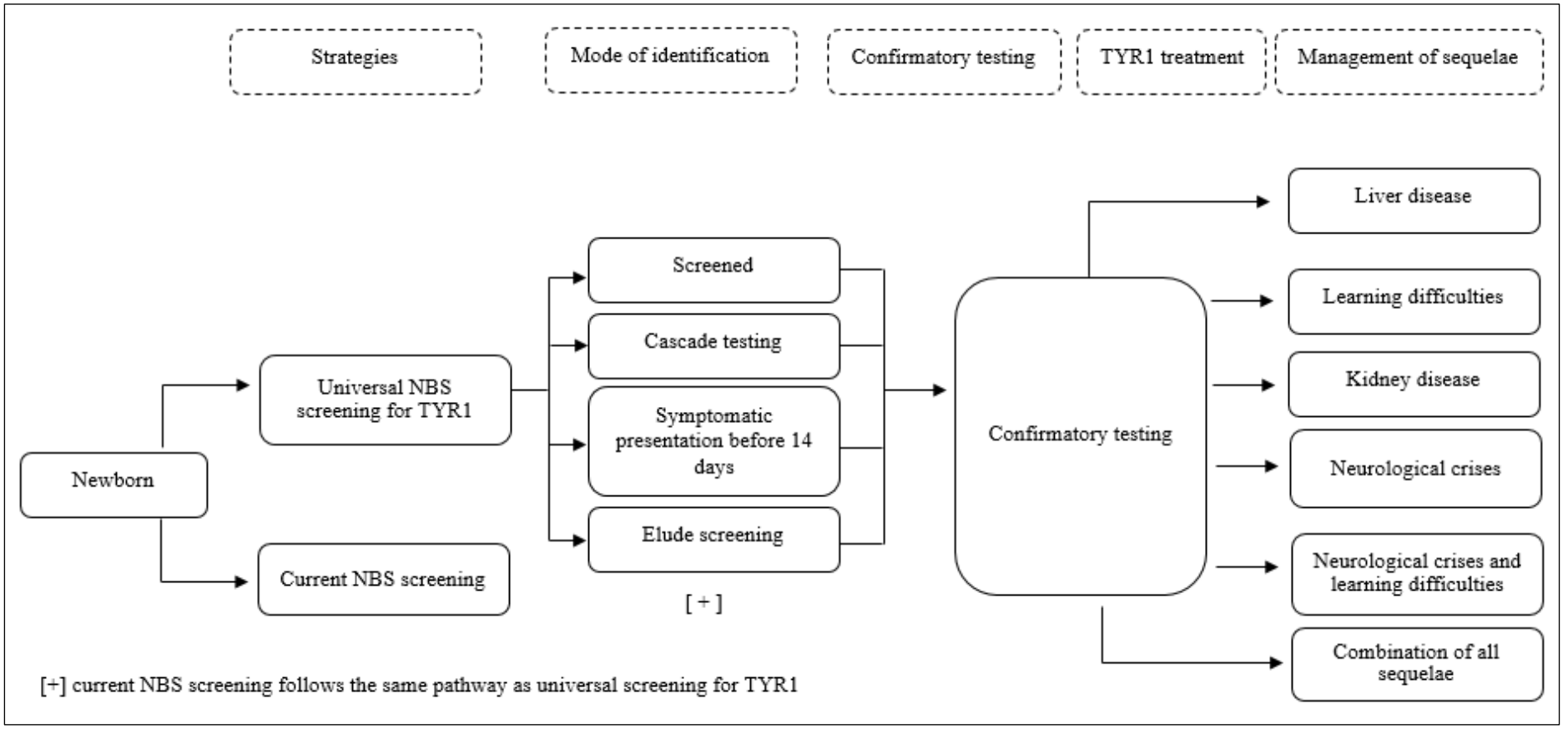

Figure 1: Illustrative structure of the clinical pathways

2.3 Pathway of babies in the two compared screening strategies

The model compares a TYR1 expanded NBS screening programme versus current practice (no universal TYR1 screening), with details of each strategy discussed below.

Proposed expanded NBS screening for TYR1 using MS/MS

When screened at day 5, babies may have SUAC levels above (positive) or below (negative) the pre-specified cut-off. Babies with a positive result receive a diagnostic protocol to confirm if they have TYR1. We assumed that babies who have a confirmed diagnosis, commence treatment with diet and nitisinone (1 mg per kg body weight per day in a liquid form). We assumed that the diagnostic protocol is 100% accurate. Following TYR1 confirmation, we assumed that babies do not have any signs or symptoms suggestive of TYR1, hence are asymptomatic or do not yet have sequelae associated with TYR1. As babies get older, we assumed that they may develop sequelae associated with TYR1 or its treatment. In the model babies may develop liver disease, liver disease/cancer requiring transplant, kidney disease, learning difficulties, neurological crises or a combination of learning difficulties and neurological crises, or a combination of all sequelae. We assumed that babies with these sequelae are treated, and that they remain in these health states.

Current NBS screening programme

In the current NBS screening programme, there is no universal NBS screening for TYR1, but there may be some incidental detection of TYR1 as a result of NBS screening for phenylketonuria (PKU).[6] The clinical pathway for this strategy is similar to the proposed expanded screening pathway. When screened on day 5, babies with initial elevated phenylalanine (Phe) level who have elevated tyrosine (Tyr) levels on repeat duplicate testing (irrespective of the mean and repeat test Phe levels) are not suspected to have PKU but receive diagnostic testing for other diseases including TYR1. We assumed that these babies would immediately commence treatment with diet and nitisinone and that over time they may develop sequelae related to TYR1 or its treatment.

Symptomatic detection before 14 days of life

In this pathway, babies younger than 14 days (so before NBS screening results would be available) with symptoms suggestive of TYR1 receive a diagnostic protocol to confirm TYR1. Following confirmation, babies are treated with diet and nitisinone, and as they get older, they may develop sequelae related to TYR1 or its treatment.

Cascade testing

We assumed that newborns whose older sibling(s) have TYR1, receive the diagnostic protocol at 48-72 hours of age to test if they have TYR1. Following confirmation of TYR1, babies receive treatment

with diet and nitisinone. It was assumed that these babies would be in an asymptomatic health state. Over time, we assumed that these babies might develop sequelae related to TYR1 or its treatment.

Elude NBS screening and symptomatic detection after 14 days

It was assumed that newborns who were not screened at day 5 remained in an asymptomatic health state until they developed symptoms suggestive of TYR1 after 14 days of life. Following confirmation of TYR1, babies/children are treated with diet and nitisinone, and as they get older they may develop sequelae related to TYR1 or its treatment.

Long-term complications

It was assumed that following confirmatory diagnosis, all babies are free of long-term complications. However, over time babies can develop long-term complications related to TYR1 or its treatment. These complications include liver disease, learning difficulties, kidney disease, neurological crises, a combination of learning difficulties and neurological crises, and a combination of any sequelae. Figure 3 depicts the illustrative structure for the Markov component of the model. For simplicity and due to the paucity of information, it was assumed that babies/children developed these complications at a constant rate (see Section 0). Also, with the exception of the liver disease health state, it was assumed that if babies developed one long-term complication they would not develop another long-term complication. Babies/children who developed liver disease are at risk of liver cancer and thus the possibility of liver transplant is included in this pathway. It was assumed that babies would remain in the same health state until they die.

2.4 Model assumptions

A number of assumptions were made to allow us to develop an executable model in order to enable the analyses to be undertaken:-

- Babies are screened on day 5.
- The national protocol used to confirm/diagnose TYR1 is 100% accurate.
- Diagnostic results to confirm TYR1 are available soon after testing.
- Babies commence treatment for TYR1 (Nitisinone and a diet low in phenylalanine and tyrosine) as soon as they are diagnosed.
- All screen-detected newborns with a confirmed diagnosis do not show any signs or symptoms related to TYR1, but may develop long-term complications.
- Average dose of nitisinone treatment is 1 mg per kg body weight per day.
- People are 100% compliant with diet and nitisinone treatment.

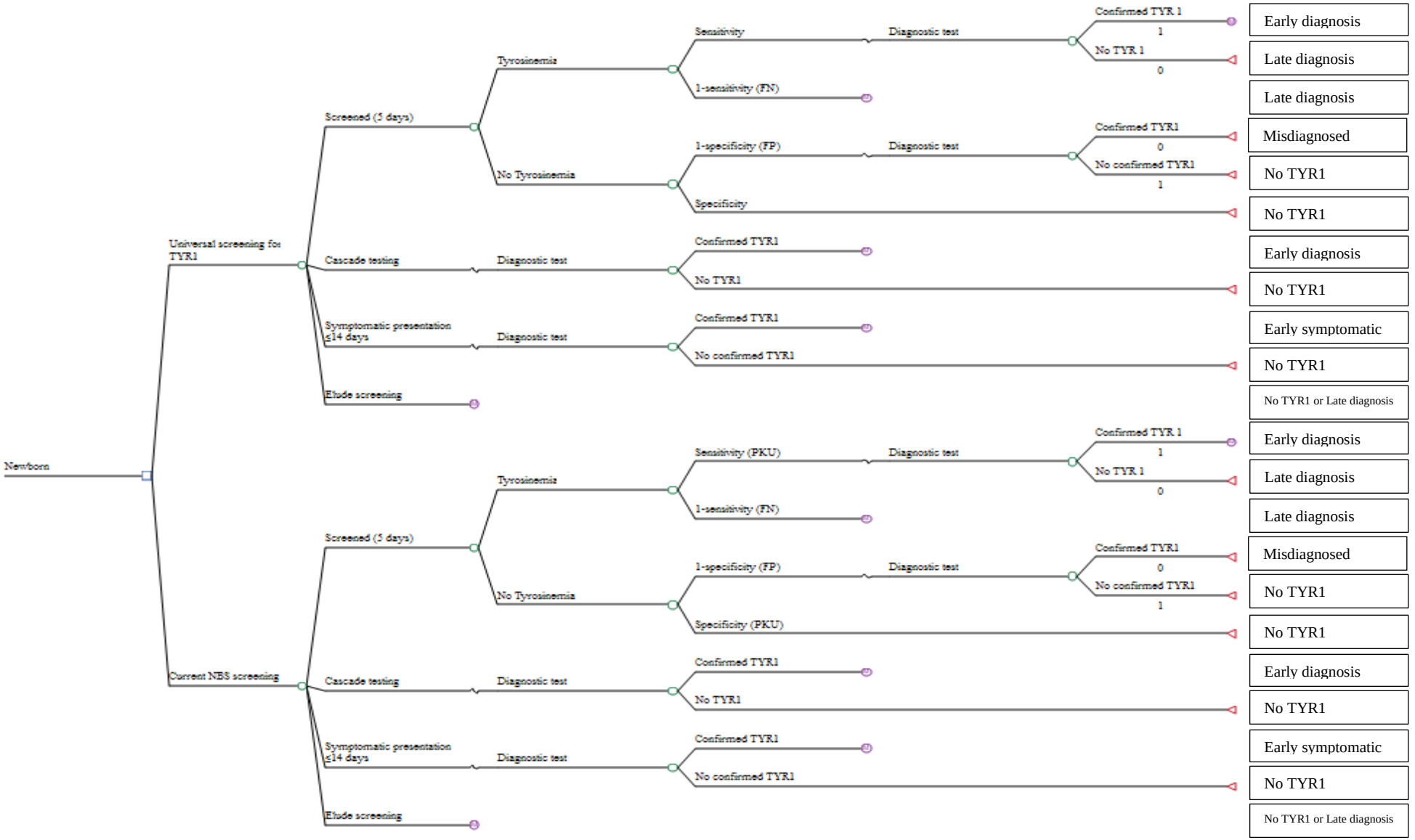

Figure 2: Illustrative decision tree structure for stage one

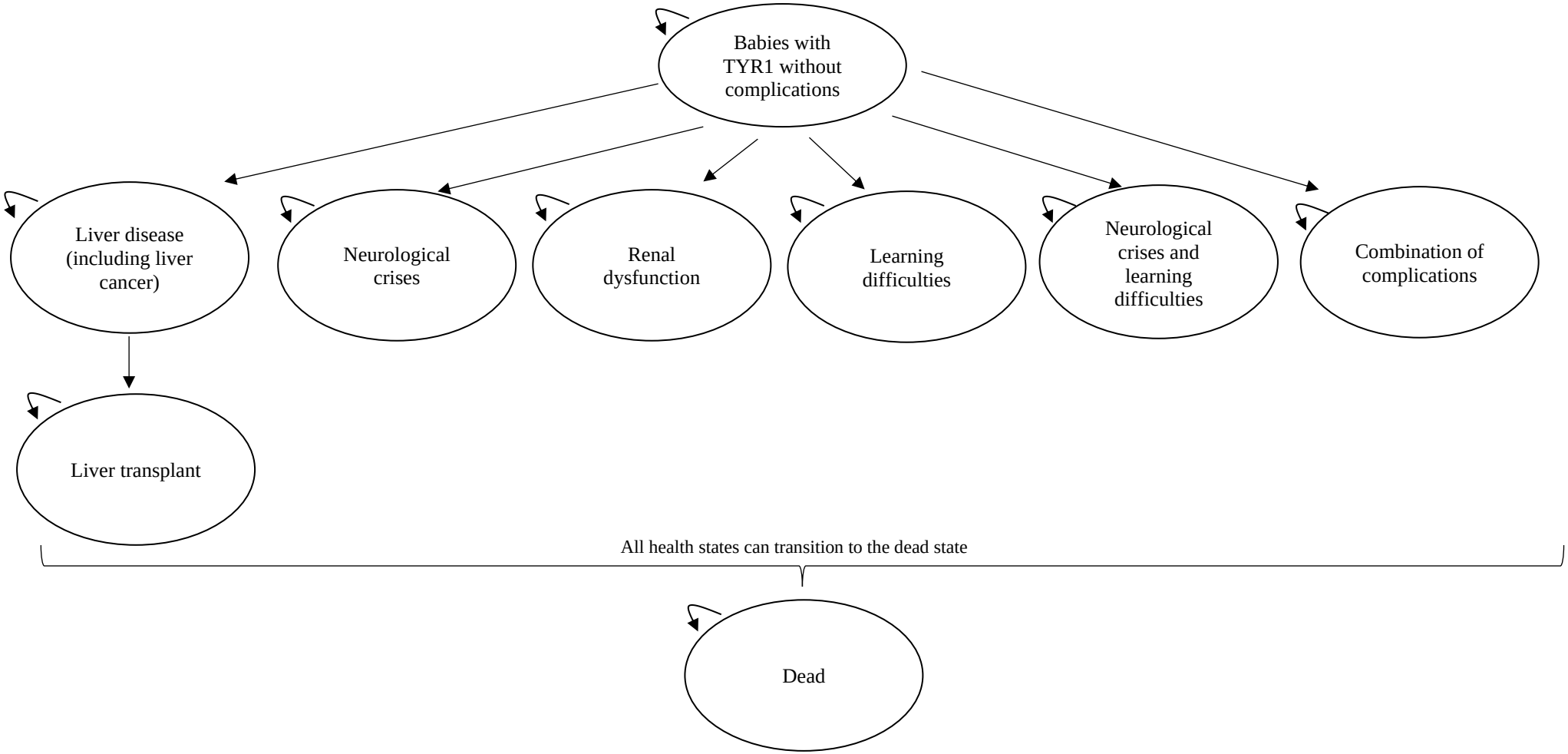

Figure 3: Illustrative Markov model structure for stage two

2.5 Data required for the model

The model was populated with information obtained from evidence identified by the previous 2016 NSC review[3] and its related publications,[4; 5] and by our clinical and cost-effectiveness reviews, and supplemented with information from secondary sources as well as expert clinical opinion. Information was required for the proportion of babies screened, the test performance of screening and confirmatory testing (sensitivity and specificity), incidence of TYR1, management of TYR1, incidence of condition-specific sequelae, treatment of sequelae, mortality, resource use and costs associated with screening, and utility values associated sequelae.

2.5.1 Pathway of live-born babies

The model required information relating to the percentage of babies being screened on day 5, the proportion of babies undergoing cascade testing because of a previously affected sibling, of babies detected symptomatically before 14 days of life and of babies who eluded screening. These values are reported in Table 1. We assumed that the probability of parents who declined screening to be equal to the probability of babies who would elude screening.

Table 1: Model assumptions: Pathway of live-born babies in England

Parameter	**Value (%)**	**Source**
NBS screening with or without addition of TYR1 screening (day 5)	96.5%	NBS screening programme in the UK: Data collection and performance analysis report 2016 to 2017 (England data)[7]
Cascade testing	0.001800%	Unpublished data collected by clinical expert and published data by Bartlett et al.[8] for the West Midlands, number of births per year in the West Midlands from the "West Midlands Newborn Screening Laboratory Annual Report",[9] corrected for the higher incidence of TYR1 in the West Midlands than in the whole of England.[10]
Symptomatic presentation (≤14 days)	0.000042%	
Elude screening	3.498158%	NBS screening programme in the UK: Data collection and performance analysis report 2016 to 2017 (England)[7] minus proportion being cascade tested or presenting symptomatically ≤ 14 days.
NBS, newborn blood spot; TYR1, tyrosinaemia type 1		

2.5.2 Test accuracy

Test accuracy parameters used in the economic model were estimates of the sensitivity and specificity of the two screening strategies. Sensitivity and specificity estimates for each screening strategy and its sources are shown in Table 2. More details of our approach are provided in Appendix 1.

Table 2: Test accuracy to detect TYR1

Parameter	Value	95% confidence interval	Source
Screening tests			
MS/MS measurement of SUAC (proposed TYR1 screening protocol)			
Sensitivity	1.00000	0.903996, 1.00000	Summary estimates derived from the 2016 NSC review[3; 5] combined with the two additional studies our update clinical effectiveness review (11 studies in total)
Specificity	0.999983	0.999974, 0.999988	
MS/MS measurement of phenylalanine and second-tier tyrosine (current PKU screening protocol)			
Sensitivity	0.250000	0.004258, 0.946393	Estimated from unpublished data obtained from clinical expert
Specificity	0.999865	0.999737, 0.999931	Estimated from 2016/2017 performance data of NBS screening programme in the UK[7]
Confirmatory test (diagnostic protocol)			
Sensitivity	1.00	-	Assumption
Specificity	1.00	-	
MS/MS, tandem mass spectrometry; NBS, newborn blood spot; NSC, National Screening Committee; PKU, phenylketonuria; SUAC, succinylacetone; TYR1, tyrosinaemia type 1			

2.5.3 Confirmatory protocol

Definitive diagnosis of TYR1 is made using a confirmatory diagnostic protocol. In the model, we assumed that babies with a positive test result on the index test and those with clinical symptoms suggestive of TYR1 would receive a confirmatory diagnosis before initiation of treatment with diet and nitisinone. Here, although we assumed that the confirmatory protocol is 100% accurate, the index tests are not 100% specific, so there will be babies with false positive test results. The use of the confirmatory protocol should correctly confirm that babies with a false positive results do not have TYR1, and should therefore reduce the anxiety caused by the index test,[11] and eliminate the need for treatment.

2.5.4 Birth prevalence of TYR1 and method of detection

The model includes the different methods used to identify babies/children with TYR1 in the UK, with each requiring information about the TYR1 prevalence. Figure 4 and Figure 5 show the flow diagrams of a hypothetical UK birth cohort with or without TYR1 screening added to the current NBS screening programme.

- In a hypothetical UK cohort of 679,106 live-born babies per year (number of live births in England and Wales in 2017[12]), we would expect around seven TYR1 cases per year (assumed UK incidence 1:100,000 similar to the incidence observed in the West Midlands region excluding Birmingham).[10]
- We assumed that 0.0018% of live-born babies (approximately 12 babies per year) would have received cascade testing for TYR1 due to a previously affected sibling, of which 25% would have TYR1 (approximately three cases identified by cascade testing per year).
- Very rarely (assumed 0.000042% of live-born babies, which equals to one baby every 3.5 years) would present with TYR1-specific symptoms before 14 days of age, we assumed that 50% would have TYR1 confirmed (approximately one TYR1 case every seven years).
- We assumed that approximately 23,756 (~3.5%) newborns per year would elude NBS screening. The birth prevalence of TYR1 in this group was estimated to be approximately 1:189,000 (derived from birth prevalence in total UK birth cohort[10] excluding cases detected by cascade testing or cases with TYR1-specific symptoms before 14 days of age). Therefore, one unscreened TYR1 case would be detected symptomatically every eight years.
- The remaining 655,337 newborns (assumed NBS screening coverage 96.5%) would undergo NBS screening. We assumed that the current PKU screening programme would detect one of the remaining four TYR1 cases per year (assumed 25% sensitivity), while three TYR1 cases would be missed and present later in life with symptoms.
- The proposed universal TYR1 screening programme would detect all four TYR1 cases (summary sensitivity 100% from 11 studies) with no TYR1 cases presenting symptomatically after 14 days of age.

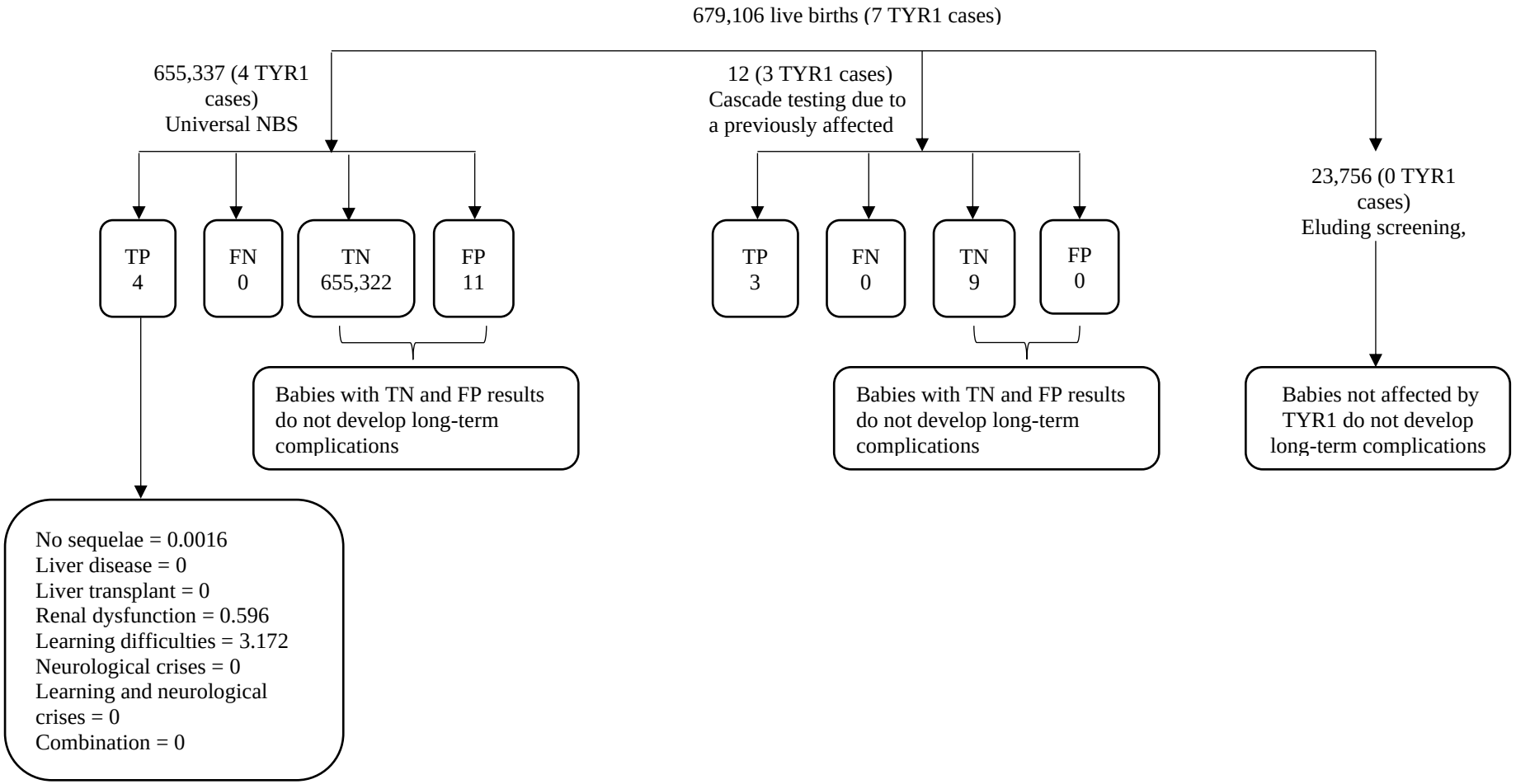

Figure 4: Flow diagram for a hypothetical cohort of 679,106 live births in England and Wales – universal NBS screening for TYR1

FN, false negative; FP, false positive; NBS, newborn blood spot; TN, true negative; TP, true positive; TYR1, tyrosinaemia type 1

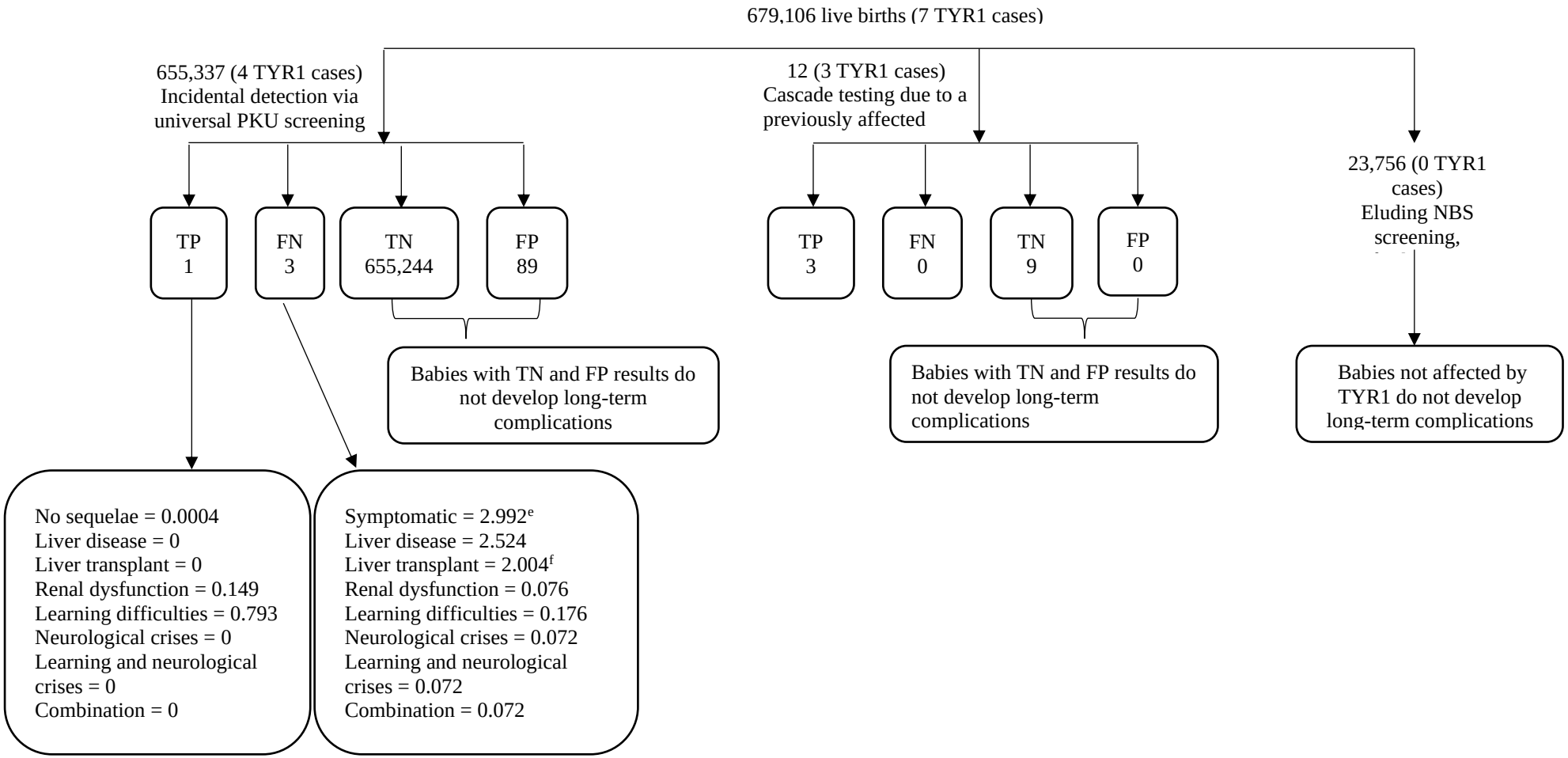

Figure 5: Flow diagram for a hypothetical cohort of 679,106 live births in England and Wales – current approach without universal NBS screening for TYR1

FN, false negative; FP, false positive; NBS, newborn blood spot; PKU, phenylketonuria; TN, true negative; TP, true positive; TYR1, tyrosinaemia type 1

Model inputs / assumptions for Figure 4 and Figure 5:

a) Cohort based on 679,106 live births in England and Wales in 2017 (ONS). TYR1 birth prevalence in England and Wales assumed to be 1:100,000, so there should be ~7 TYR1 cases per year. Due to rounding and omitting the one baby every 3.5 years (0.28 baby per year) that has TYR1-specific symptoms before 14 days of age, the numbers do not add to 679,106, but 679,105.

b) Uptake of NBS screening 96.5%; eluding screening 3.498%; cascade tested due to affected siblings 0.0018%; presenting with symptoms <14 days 0.000042%.
The assumed 0-1 babies per year presenting with TYR1-related symptoms before 14 days of age are not presented in this flow diagram.

c) Babies with previously affected siblings have a 25% chance of having TYR1.

d) Test accuracy of screening and diagnostic tests for TYR1 detection:
Proposed screening test using SUAC as primary marker: sensitivity 100%, specificity 99.9983%;
Screening test using phenylalanine with second-tier tyrosine (current PKU screening protocol): sensitivity 25%, specificity 99.9865%;
Diagnostic protocol: sensitivity 100%, specificity 100%.

e) Proportion of people who develop symptoms suggestive of TYR1 within the 10-year time period having had a false negative result.

f) Proportion of people undergoing a liver transplant from those with a liver disease.

2.5.5 Incidence of condition-specific sequelae and transition probabilities

In the model, we included health states for the following complications and long-term sequelae: liver disease, renal dysfunction, learning difficulties, neurological crises, and a combination of learning difficulties and neurological crises. Information on the incidence of long-term complications was obtained from the published evidence identified by the 2016 UK NSC review[3] and from related post-hoc analyses.[4] From this information, annual transition probabilities were estimated to simulate the number of people entering each health state by assuming a constant rate of events. Information and any assumptions in deriving these transition probabilities used in the model, are summarised in Table 3 and Appendix 2

Table 3: Four and six-month transition probabilities by complication (central estimates)

Long-term complication	**Screen detected**		**Symptomatically detected**		**Source**
	4-month transition probability	**6-month transition probability**	**4-month transition probability**	**6-month transition probability**	
Liver disease	0	0	0.118	0.172	Larochelle et al.,[13] "Post hoc 1" analysis (Screen detection vs symptomatic detection, all with direct nitisinone initiation) as in Geppert et al.[4]
Liver transplant in cases with liver disease	0	0	0.012	0.018	McKiernan et al.[14]; information on follow-up time in screen-detected cases extrapolated from Bartlett et al.[8]
Renal dysfunction	0.002	0.003	0.004	0.005	Mayorandon et al.;[15] data for "Renal dysfunction".
Learning difficulties	0.010	0.016	0.008	0.012	Mayorandon et al.[15] (frequency/odds ratio for "psychomotor impairment")
Neurological crisis	0	0	0.003	0.005	Larochelle et al.[13]: "Post-hoc 3" analysis (Screen detection, direct nitisinone initiation vs screen detection, 1–12 months delayed nitisinone initiation) as in Geppert et al.[4]
Combination of learning difficulties and neurological crises	0	0	0.003	0.005	Assumed to be equal to the transition probability for neurological crises
Combination of long-term complications	0	0	0.003	0.005	Assumption to be equal to the transition probability for neurological crises

2.6 Resource use and costs

The resource use and costs included are those that are directly incurred by the NHS and PSS. Costs were required for the index tests, diagnostic protocol, treatment with nitisinone, costs associated with the treatment of babies and children with TYR1, and costs for the treatment of long-term sequelae. All costs are presented in 2017/18 prices and after the first year, both costs and benefits were discounted at a rate of 3.5% per annum. Costs obtained from the literature through systematic reviewing were uprated to 2017/18 prices where necessary using the Hospital and Community Health Services (HCHS) index from Unit Costs of Health and Social Care 2018.[16]

2.6.1 Cost of index tests

Drops of blood are collected from the neonate by heel prick and infused on special screening cards in order to screen for inborn errors of metabolism. Screening cards are sent to laboratories to be analysed. In the model, we used the cost of identifying PKU as a proxy for the cost of an incidental finding of TYR1. The cost of £2.70 per sample is based on using MS/MS to screen for PKU (and other conditions), inclusive of labour, capital and consumables (expert opinion, 2019). With MS/MS, other disorders can be screened for without the need to collect additional blood samples from the neonate.[11] Hence, extending the current screening programme to include screening for TYR1 by measurement of SUAC can be conducted simultaneously using existing equipment, and without adding to the run time. Adding TYR1 to the current NBS screening programme was estimated to cost an additional 10p per baby tested which includes the cost of using 'home brew' or 'in-house' reagents and an internal standard (expert opinion, 2019). We therefore used a cost of £2.80 per sample for the proposed, expanded NBS screening programme.

2.6.2 Cost of diagnostic protocol

We used the cost (£257, uprated to 2018 prices using the HCHS pay and price indices)[16] of the confirmatory protocol for PKU from Pandor et al.[11] as a proxy for the cost of diagnosing TYR1. Pandor and colleagues derived this cost based on the additional resource use needed associated with laboratory staff and consumables, arranging referral and providing advice for diagnosing babies with PKU.

2.6.3 Treatment with nitisinone

It was assumed that as soon as a definitive diagnosis for TYR1 is made, treatment is initiated. Some babies will commence treatment pre-symptomatically and others through symptomatic presentation. Treatment is in the form of nitisinone and dietary intervention, with the latter including information about dietetic management of TYR1. Based on expert opinion, it was assumed that children up to the age of 10 years would receive the oral suspension of nitisinone, and that those older than 10 years

would receive nitisinone in the form of capsules. We also assumed that the average dosage of nitisinone treatment is 1 mg per kg body weight per day. This regimen was combined with average age and gender-specific body weights from the UK-WHO growth charts[17] to derive unit costs for nitisinone treatment. Cost per year for nitisinone treatment was obtained from clinical expert opinion, which is based on a 4mg/1ml oral suspension, with a 90ml oral sugar-free suspension costing £1,692 (eMC Dictionary of Medicines and Devices Browser). In Table 4 we present the annual costs for treating people with TYR1 using nitisinone.

Table 4: Annual costs for treating people with TYR1 using nitisinone

Age	Weight (kg)	Daily dose (1mg/kg)	Formulation	Cost per year
Neonate	3.5	3.5 mg (0.87ml)	4mg/1ml Oral Suspension	£5,970
1 month	4.3	4.3 mg (1.1ml)	4mg/1ml Oral Suspension	£7,549
2 month	5.4	5.4 mg (1.3ml)	4mg/1ml Oral Suspension	£8,921
3 month	6.1	6.1 mg (1.5ml)	4mg/1ml Oral Suspension	£10,293
4 month	6.7	6.7 mg (1.7ml)	4mg/1ml Oral Suspension	£11,666
6 month	7.6	7.6 mg (1.9ml)	4mg/1ml Oral Suspension	£13,038
1 year	9	9 mg (2.3ml)	4mg/1ml Oral Suspension	£15,783
3 years	14	14 mg (3.5ml)	4mg/1ml Oral Suspension	£24,017
5 years	18	18 mg (4.5ml)	4mg/1ml Oral Suspension	£30,879
7 years	23	23 mg (5.8ml)	4mg/1ml Oral Suspension	£39,800
10 years	32	32 mg (3x10mg & 1x2mg)	Capsules (Available as 2mg, 5mg and 10mg)	£30,797
12 years	39	39 mg (3x10mg & 1x5mg & 2x2mg)	Capsules (Available as 2mg, 5mg and 10mg)	£38,507
14 years	50	50 mg (5x10mg)	Capsules (Available as 2mg 5mg and 10mg)	£47,040
Adult Male	68	68 mg (6x10mg & 4x2mg)	Capsules (Available as 2mg, 5mg and 10mg)	£66,741
Adult Female	58	58 mg (5x10mg & 4x2mg)	Capsules (Available as 2mg, 5mg and 10mg)	£57,333

2.6.4 All other resource use and costs

Detailed information about resource use and costs for the management of babies with TYR1 including inpatient stay, outpatient visits, contact with healthcare staff including dietician, testing, diet and costs associated with long-term conditions is presented in Table 5 and summarised in Appendix 3.

Table 5: Resource use and unit costs

Resource use	Cost	Source
Neonatal admission (HRG code)		
• Inborn Errors of Metabolism with CC Score 0-2 (Elective inpatient stay) (KC04B)	£572.03	NHS reference costs 2016/17[18]
• Neonatal Critical care, High dependency care (XA02Z)	£909.81	NHS reference costs 2016/17[18]
• Neonatal Critical care, Normal care (XA05Z)	£429.17	NHS reference costs 2016/17[18]
Paediatric admission		
• Average cost per stay	£2,880.00	PSSRU 2018[16]
Outpatient visits		
• Paediatric consultant-led outpatient attendance	£201.00	PSSRU 2018[16]
• Paediatric non-consultant-led outpatient attendance	£151.00	PSSRU 2018[16]
• Adult outpatient attendance	£134.00	PSSRU 2018[16]
Staff costs		
• Dietician	£86.00	PSSRU 2018[16]
• Health visitor	£59.11	PSSRU 2018[16]
Tests		
Bloods		
• Blood gases	£5.89	Homerton University Hospital[19]
• Full blood count	£10.33	BCH
• Coagulation (PT, PTT, fibrinogen)	£8.36	BCH
• Liver function tests (Bilirubin, AST, ALT, alkaline phosphatase, GGT, albumin)	£4.55	BCH
• Urea and electrolytes, creatinine, calcium, phosphate	£3.79	BCH
• Amino acids (quantitative, tyrosine, phenylalanine)	£118.00	BCH
• Alpha-fetoprotein	£6.10	Homerton University Hospital[19]
• Glucose and ammonia	£0.82	BCH
• Iron and ferritin, vitamins A, D, E, folate and vitamin B12, micronutrients (selenium, zinc, copper)	£75.44	North East Lincolnshire CCG[20] + BCH
Urine		
• Glucose	£4.17	NICE guidance NG45[21]
• Amino acids	£36.00*	BCH
• Tubular re-absorption of phosphate	£0.76	BCH
• Calcium/creatinine ratio	£0.73	BCH
• Urine acidification	£4.17	NICE guidance NG45[21]
• Albumin, protein, β2-microglobulin	£0.74	BCH
• Organic acids	£52.00*	BCH
Imaging		
• Liver imaging: ultrasound	£115.39	Homerton University Hospital[19]
• Renal imaging: ultrasound	£115.39	Homerton University Hospital[19]
• Liver imaging: MRI or CT	£515.14	Homerton University Hospital[19]
Other tests		
	£360.79	

Resource use	Cost	Source
• Developmental evaluation / neuropsychological assessment • Bone mineral density • Eye examination • Molecular genetics	£82.29 £29.19 £617.28*	NHS reference costs 2016/17[18] NHS reference costs 2016/17[18] Olson et al, 2013[22] BCH
Diet costs • Prescription charges • Annual cost of diet	 £8.80 £8,887.99	 UK NHS prescription authority[23] Belanger-Quintana 2012[24]
ALT, alanine aminotransferase; AST, aspartate aminotransferase; BCH, Birmingham Children Hospital; CCG; clinical commissioning group; CT, computed tomography; GGT, gamma glutamyl transferase; NG, national guideline; NHS, National Health Service; MRI, magnetic resonance imaging; NICE, National Institute for Health and Care Excellence; PSSRU, Personal Social Services Research Unit; PT, prothrombin time; PTT, partial prothromboplastin time		

* One-off tests

2.7 Utility values

The utility values applied in the base-case model were mainly obtained from the cost-effectiveness analysis undertaken by Tiwana and colleagues.[25] Briefly, the authors estimated the cost-effectiveness of expanding the newborn screening programme compared with the previous standard screening for inborn errors of metabolism in the USA. Estimates associated with a false positive screening result, as well as long-term sequelae, were obtained from the literature. Although it was not explicitly stated by the authors, it was assumed that babies identified with TYR1 via early/late detection had the same utility values. This was considered a reasonable assumption, so was applied in our model. The current model includes health states for children who may have neurological crises, as well as learning difficulties. Here, we assumed a utility value of 0.82 for these children. Additionally, we assumed a utility value of 0.30 for children who may have a combination of sequelae. The utility values used in the model are listed in Table 6.

Table 6: Utility values included in the model

Health state	Utility values	Range	Source
False positive screen result	0.97	0.95-0.99	Tiwana et al., 2012[25]
Liver disease	0.20	0.10-0.30	
Liver transplant	0.67	0.58-0.74	
Renal dysfunction	0.67	0.58-0.74	
Learning difficulties	0.79	0.59-0.84	
Neurological crises	0.84	0.70-0.85	
Treatment without complications	0.90	0.85-0.95	
Neurological crises and learning difficulties	0.82	0.72-0.90	Assumptions
Combination of sequelae	0.30	0.10-0.50	
Overdiagnosis	-0.09	-	

In addition to utility values associated with these complications, we included age-related disutilities which take the effect of ageing on a person's health-related quality of life into consideration. These age-related disutilities have been derived using an algorithm, which estimates the general population utility scores as a function of age and gender.[26]

The base-case model excludes disutility associated with overdiagnosis. Overdiagnosis can be defined as the diagnosis of a condition, which would not have resulted in symptoms or harm to the patient during their lifetime if it had gone unrecognised.[27] There is very limited information on overdiagnosis, however, the Quebec cohort included an individual in the non-nitisinone group (n=28) who received no treatment at all but did not appear to develop any complications throughout their childhood.[13]

Mortality

Two types of mortality were included in the model, death following liver transplant and death from other causes. General population mortality was obtained from the Office for National Statistics (ONS)[28] and an average of the mortality rate for males and females was used in the model. We applied a 14% risk of mortality following liver transplant.[14] It was assumed that babies with liver disease have a 30% increased risk of death compared to the general population. Additionally, it was assumed that all babies/children with untreated TYR1 would die by the age of 10 years.

2.8 Outcomes

Three different outcome measures were used in the analysis: cases of TYR1 correctly identified by screen-detection, life-years, and quality-adjusted life years (QALYs). For life-years, we used a value of 1 where people were alive and 0 when they had died. To derive QALYs, in each cycle, a utility pay-off is assigned based on the health-state occupied. For each strategy, the sum of the QALYs is derived over the model time horizon.

2.9 Analysis

The economic analysis was undertaken from the perspective of the NHS and PSS. The results of the analysis are presented in terms of an incremental cost-effectiveness ratio (ICER), expressed as cost per additional screen-detected TYR1 case, cost per life-years gained (LYG) and cost per QALY gained. The Markov component of the model had different cycle lengths to reflect monitoring of babies. The cycle lengths were: 4-monthly for the first year of life, then 6-monthly. Cost-effectiveness was assessed over a lifetime horizon, and all costs incurred and benefits accrued were discounted at 3.5% per annum in line with recommended guidelines.[29]

2.10 Probabilistic sensitivity analysis

We undertook probabilistic sensitivity analysis (PSA) to determine the joint uncertainty in the key model input parameters of prevalence, sensitivity and specificity, costs and utility values. We undertook the PSA based on the outcome of cost per QALY gained only. In the PSA, each model parameter is assigned a distribution, reflecting the amount and pattern of its variation, and cost-effectiveness results are calculated by simultaneously selecting random values from each distribution. This process was repeated 10,000 times to give an indication of how variation in the model parameters leads to variation in the ICERs. The distributions used in the PSA are presented in Appendix 5. We also calculated the probability that each strategy is the most cost-effective at different willingness-to-pay (WTP) thresholds per QALY gained in line with NICE guidance.[29]

2.11 Scenario analyses

Tyrosinemia type I is a rare condition and there was significant uncertainty in several model inputs. We addressed this through scenario analyses for different values for each variable, and a sensitivity analysis using an optimistic and a pessimistic value for a range of variables simultaneously. A number of scenario analyses were undertaken using our base-case model:

- Changing the discount rate applied to both costs and benefits.
- Including disutility values associated with overdiagnosis. We assumed that 3% of screened babies would be overdiagnosed.
- The cost of the screening test.
- Using different age cut-offs of survival for children with untreated TYR1.

Discount rates

NICE guidelines for health technology assessments suggest applying the standard discount rate of 3.5% for both costs and benefits.[29] However, for public health interventions the NICE guidelines recommend applying a discount rate of 1.5% for both costs and benefits. NICE also recommends undertaking scenario analyses by applying the 1.5% discount rate when the benefits are substantial in restoring health and sustained over a long period of time. Therefore, we examined the effects of changing the discount rates accordingly.

Cost of screening for TYR1

The base-case analysis includes an additional cost of £0.10 for extending the current screening programme per baby screened. Based on clinical expert opinion, the additional cost of testing when using a separately purchased commercial kit is likely to be £1.16 per baby tested. A scenario analysis was run using this additional cost to explore the impact to the base-case ICER.

Including disutility values associated with overdiagnosis

In this scenario analysis, we included a disutility of -0.09 associated with overdiagnosis. The assumption here is that 3% of screen-detected babies are overdiagnosed (estimated from the non-nitisinone group reported by Larochelle et al. where one TYR1 case without any treatment among 28 [3.6%] TYR1 patients born prior nitisinone availability did not seem to experience any complications for about 12 years [13]).

Using different age cut-offs of survival for children with untreated TYR1

The base-case assumes that babies with a false negative screening test result who have not developed symptoms suggestive of TYR1 by 10 years of age would die. This scenario assumes that children with untreated TYR1 would not survive beyond five years.

2.12 Value of information

Value of information (VOI) analysis was undertaken to provide a framework for analysing uncertainty within the economic model by estimating the expected costs associated with imperfect information when deciding between alternative strategies. This could be characterised as uncertainty.[11] The extent to which analysts can be confident that based on the current information the best strategy is being adopted and the intended benefits would be obtained. Reducing this uncertainty could potentially lead to an alternative strategy being adopted. The value of this additional information depends on how much this additional information will reduce the uncertainty.[11] A key VOI measure is the expected value of perfect information (EVPI), which represents the monetary value of obtaining perfect information to eliminate uncertainty in the overall decision-making process and for key parameters.[30] If the costs of obtaining further information exceeds the EVPI, there is little justification for undertaking further research.[30]

3 Results

For this first part of the results section, we present the preliminary draft results including scenario analyses which were sent to the UK National Screening Committee in September 2019 and before feedback was received from the experts at the Tyrosinaemia workshop which was held on 22nd July 2020.

We present the preliminary deterministic results in Table 7 through to Table 9 based on the outcomes cost per additional screen-detected TYR1 case, cost per LYG and cost per QALY gained, respectively. We based the estimations on a hypothetical UK birth cohort of 100,000 newborns, of which we assumed that 96,500 (96.5%) would undergo NBS screening. The model comprising two stages provides a quantitative framework to link the diagnostic accuracy of universal screening compared to no universal screening to the short-term costs (costs associated with screening and diagnosing TYR1) and benefits (number of TYR1 cases identified) and the long-term costs and health

outcomes by expressing the results in terms of LYG and QALYs gained. Additionally, we present the results for the scenario and sensitivity analyses based on the outcome cost per QALY gained. Subsequently, probabilistic sensitivity analysis and scenario analysis results are presented for the outcome cost per QALY gained.

3.1 Preliminary deterministic base-case results

3.1.1 Cost per additional screen-detected case of TYR1 (short-term time horizon)

Table 7 presents the estimates of the number of children with TYR1 that would be screen-detected in the UK if universal screening for TYR1 using MS/MS measurement of SUAC was added to the current NBS screening programme compared to the current practice for a hypothetical screening cohort of 96,500 newborns (96.5% of 100,000 live births in the UK). In the group undergoing NBS screening, we assumed a TYR1 birth prevalence of ~1:189,000 (corresponding to 0.5105 TYR1 cases). The remaining 0.4895 TYR1 cases that would be expected in the UK per 100,000 live births would not benefit from a change in the NBS screening approach as they would be detected in both strategies by cascade testing (0.45 cases), detected symptomatically before 14 days of life (0.021 cases) or would elude screening and present symptomatically after 14 days (0.0185 cases). These results show that the current practice not including universal TYR1 screening was the least costly strategy, with expected costs (for screening tests and follow-up diagnostic testing of screen-positive babies) of approximately £264,410 and 0.1276 TYR1 cases detected per 96,500 screened newborns, while the proposed strategy with added universal TYR1 screening would detect all 0.5105 TYR1 cases at an estimated cost of £271,165. This indicates that for each additional screen-detected TYR1 case in the expanded NBS screening programme, an additional £17,642 would be incurred on screening and diagnostic tests. This refers to the cost of testing only, and not the costs of treatment.

Table 7: Deterministic results based on expected costs and number of screen-detected TYR1 cases per 96,500 screened newborns

Strategy	**Expected total costs (£)**	**Incremental costs (£)**	**Number of screen-detected TYR1 cases**	**Incremental number of screen-detected TYR1 cases**	**ICER (£) per additional screen-detected TYR1 case**
No universal screening for TYR1	264,410	-	0.1276	-	-
Universal screening for TYR1	271,165	6,755	0.5105	0.3829	17,642
ICER, incremental cost-effectiveness ratio Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

3.1.2 Cost per life-years gained (LYG) (life-time horizon)

The results presented in Table 7 could be considered as short-term impact expected from the additional diagnosis of TYR1 through screen detection. In Table 8, we now included all costs for the detection of TYR1

(including costs of NBS screening and follow-up diagnostic testing in screen-positive babies, costs of cascade testing, and costs of diagnostic testing in babies with symptomatic presentation before 14 days of life and those who eluded screening and present later with TYR1-related symptoms) as well as the long-term costs of TYR1 management and treatment in relation to the expected life-years in a hypothetical UK cohort of 100,000 live-born babies. The results show that in this cohort, universal screening for TYR1 is expected to yield 0.50 more life-years at an additional cost of £214,712 compared to the current strategy without universal screening for TYR1, which corresponds to an ICER of approximately £424,200 per LYG.

Table 8: Deterministic results based on the expected costs and life years per 100,000 live births

Strategy	**Expected total costs (£)[a]**	**Incremental costs (£)[a]**	**Life-expectancy discounted**	**Incremental life years [a]**	**ICER (£) per LYG**
No universal screening for TYR1	2,132,413	-	27.702524	-	-
Universal screening for TYR1	2,347,125	214,712	27.702529	0.50	£424,219
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; LYG, life years gained Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

The only slight gain in life-years is likely due our assumption that the life expectancy is the same for screen-detected cases and symptomatically detected cases without liver disease. Babies identified pre-symptomatically through screen-detection or cascade testing are not at risk of liver disease/liver transplant, so there is no increased risk of mortality that is associated with babies pre-symptomatically detected. Therefore, the observed benefit in life-years will come from the avoided false negative screening test results in babies undergoing universal TYR1 screening (estimated to be 0.3829 per 100,000 live-born babies) resulting in a smaller number of symptomatically detected TYR1 cases with long-term complications like liver disease (including liver cancer) and liver transplantation.

3.1.3 Cost per quality-adjusted life years (QALYs) gained (life-time horizon)

In Table 9 we included all costs for the detection of TYR1 (including costs of NBS screening and follow-up diagnostic testing in screen-positive babies, costs of cascade testing, and costs of diagnostic testing in babies with symptomatic presentation before 14 days of life and those who eluded screening and present later with TYR1-related symptoms) as well as the long-term costs of TYR1 management and treatment in relation to QALYs in a hypothetical UK cohort of 100,000 live-born babies. In this cohort, adding TYR1 screening to the current NBS screening programme is expected to yield 3.6 more QALYs than the current strategy without TYR1 screening at an incremental cost of approximately £215,000, which equates to an ICER of approximately £58,800 per QALY gained.

Table 9: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected QALY discounted	Incremental QALY [a]	ICER (£) per QALY gained
No universal screening for TYR1	2,132,413	-	26.67387	-	-
Universal screening for TYR1	2,347,125	214,712	26.67391	3.6	58,848
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

3.2 Preliminary deterministic scenario and sensitivity analyses

3.2.1 Scenario analysis: Impact of higher TYR1 screening costs on the cost per additional screen-detected case of TYR1

Including the cost of a commercial kit to screen for TYR1 would lead to a cost of £3.86 for each baby tested in the proposed, expanded NBS screening programme. The results show that the expected total costs for universal screening increased, while the expected total costs for the current strategy without universal TYR1 screening and the expected number of screen-detected TYR1 cases remained constant across both strategies.

Table **10** shows that increasing the costs of the screening test from £2.80(increase by 10p) to £3.86 (increase by £1.06) increases the ICER from approximately £17,600 in the base-case analysis (Table 7) to approximately £259,700 per additional screen-detected case of TYR1.

Table 10: Scenario analysis (higher TYR1 screening costs) based on the expected costs and number of TYR1 cases detected per 96,500 screened newborns

Strategy	Expected total costs (£)	Incremental costs (£)	Number of screen-detected TYR1 cases	Incremental number of screen-detected TYR1 cases	ICER (£) per additional screen-detected TYR1 case
No universal screening for TYR1	274,000	-	0.1276	-	-
Universal screening for TYR1	373,455	99,455	0.5105	0.3829	£259,741
ICER, incremental cost-effectiveness ratio; TYR1, tyrosinaemia type I Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

3.2.2 Scenario analyses: Impact on the cost per QALY gained

Table 11 shows the results of all analyses performed scenario in a hypothetical UK cohort of 100,000 live-born babies taking into account all costs related to TYR1 detection, treatment and management in relation to the expected QALYs. These results show that changing the discount rate from 3.5% to 1.5% per annum for both costs and benefits, and the use of a commercial kit to test for TYR1 had the greatest impact on the ICERs, increasing the ICERs by approximately 55% and 48%, respectively.

Table 11: Summary of the scenario analyses results based costs and measures of effect in QALYs in 100,000 live births

Screening strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected QALYs discounted	Incremental QALYs [a]	ICER (£) per QALY gained
Base-case results					
No universal screening for TYR1	2,132,000	-	26.67387	-	-
Universal screening for TYR1	2,347,000	215,000	26.67391	3.6	58,848
Changing the discount rate applied to both costs and benefits (1.5% per annum)					
No universal screening for TYR1	3,440,000	-	44.3180	-	-
Universal screening for TYR1	3,923,000	483,000	44.3181	5.3	91,133
Increasing the costs of the screening test (£3.86 per tested baby)					
No universal screening for TYR1	2,132,000	-	26.67387	-	-
Universal screening for TYR1	2,449,000	317,000	26.67391	3.6	86,885
Using a different age cut-off (five years) of survival for children with untreated TYR1					
No universal screening for TYR1	2,132,000	-	26.67387	-	-
Universal screening for TYR1	2,346,000	214,000	26.67391	3.7	58,490
Including disutility (-0.09) associated with overdiagnosis					
No universal screening for TYR1	2,132,000	-	26.67387	-	-
Universal screening for TYR1	2,347,000	215,000	26.67391	3.6	59,298

[a] Values have been multiplied by 100,000
ICER, incremental cost-effectiveness ratio; QALYs, quality adjusted life-years; TYR1, tyrosinaemia type 1

This scenario analysis includes disutility associated with overdiagnosis. Here we assumed that 3% of screen-detected cases would have been overdiagnosed. These results showed that the expected total costs remained unchanged across both strategies, with a slight decrease in the expected QALYs yielded, which equated to an ICER of approximately £59,300 per QALY gained.

Appendix Table 2 and Appendix Table 3 present the results in terms of cost per QALY gained for the scenario analyses that considers all of the optimistic and the pessimistic estimates for the probabilities of long-term complications following pre-symptomatic detection and symptomatic detection of TYR1, respectively. The optimistic estimates resulted in an ICER of approximately £109,700 per QALY gained, while the pessimistic estimates resulted in an ICER of approximately £38,100 per

QALY gained. These results show that the transition probabilities for long-term complications, in particular for liver disease impacts on the results. Further explanation of the observed results is provided in Appendix 6 and Appendix 7.

3.2.3 One-way sensitivity analysis results

Deterministic sensitivity analysis results were conducted by varying key model input parameters by ±50% of the base-case values used in the model to see the impact on QALYs, and presented in the form of a tornado diagrams (see Figure 6). We increased and decreased the number of false positive and false negative results by ±50% to derive the lower and upper values for sensitivity and specificity of the screening tests. Sensitivity analysis results show which individual parameter has the greatest impact to the base-case ICER of approximately £58,800 per QALY gained ; that is, which parameter is the key driver of the costs-effectiveness for the comparison between universal screening for TYR1 versus no universal screening for TYR1. These results show that varying the cost of the different testing strategies (TYR1 screening using SUAC as marker and incidental detection by PKU screening when measuring Phe) had the greatest impact to the results based on cost per QALY gained.

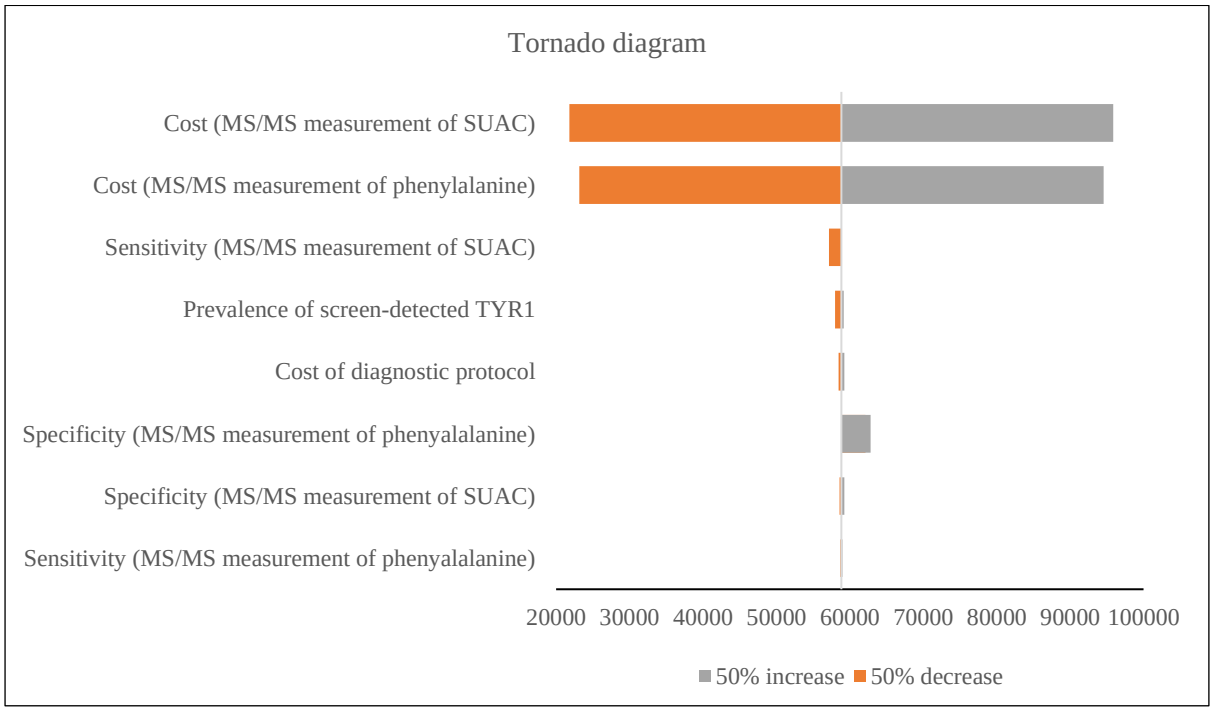

Figure 6: Tornado diagram for the impact of a ±50% change in individual parameters on the ICER per QALY gained

3.2.4 Probabilistic sensitivity analysis (PSA) results

The probabilistic results are presented in the form of a scatterplot and its corresponding cost-effectiveness acceptability curves (CEAC). Figure 7 presents the results of the 10,000 runs of the simulations, the scatterplot shows that there is some variation in the incremental QALYs but less variation for the incremental costs. This result is expected as majority of the babies screened would

not undergo any further testing; hence, not incurring any further costs. Figure 8 shows the probabilistic results presented in the form of a CEAC, which shows the probability that an intervention is cost-effective at different willingness-to-pay thresholds per QALY gained. At the £30,000 per QALY gained threshold adding universal TYR1 screening to the current NBS screening programme has a very low probability of being cost-effective. At the £100,000 per QALY threshold the expanded NBS screening strategy has a probability of one of being cost-effective compared to current practice without universal TYR1 screening.

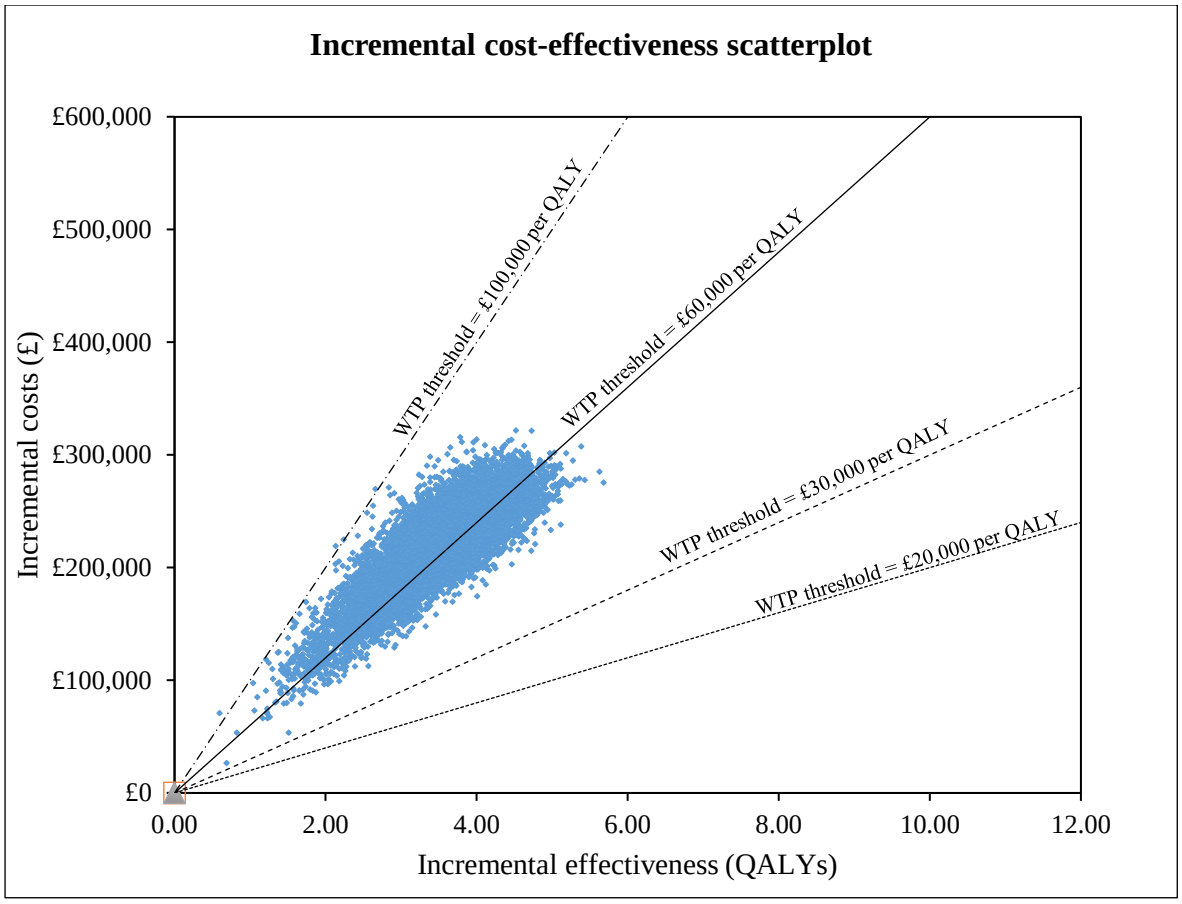

Figure 7: Incremental cost-effectiveness scatterplot for the comparison between universal screening versus no screening

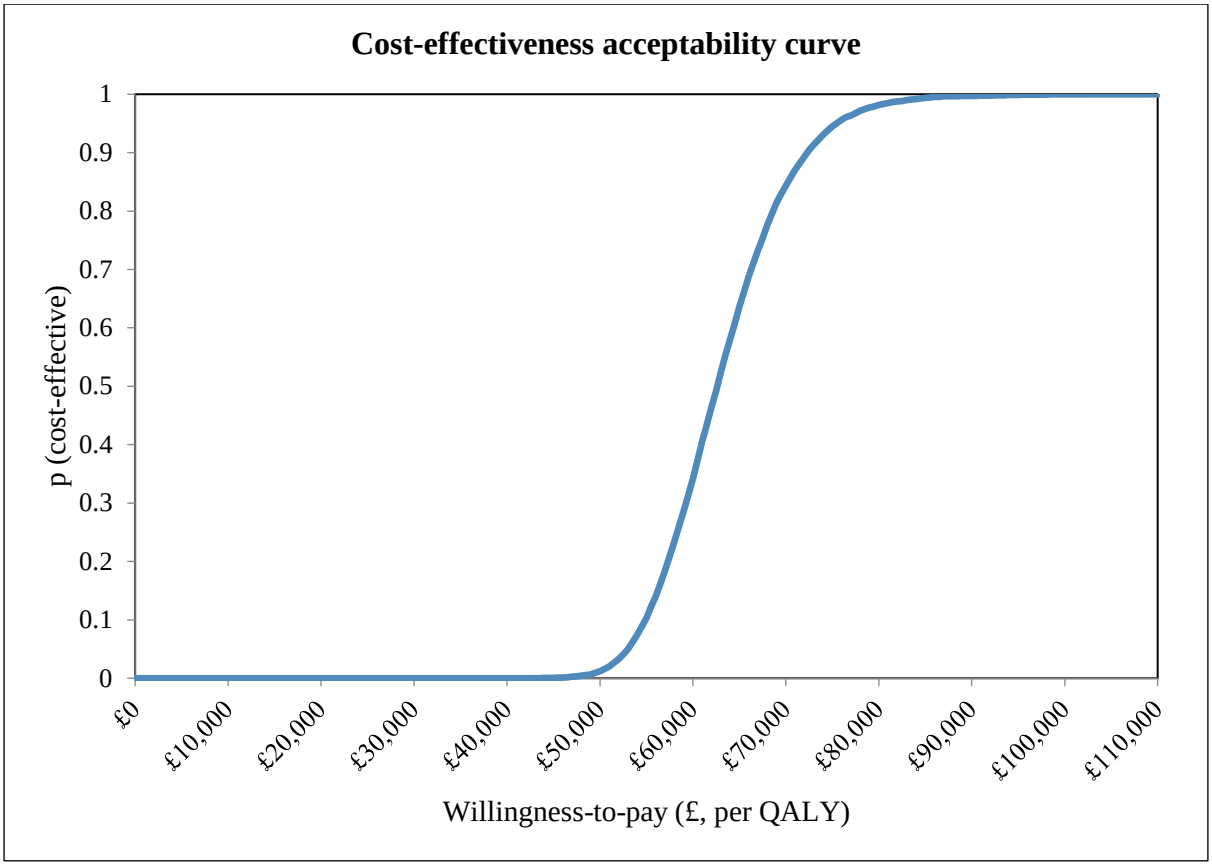

Figure 8: Cost-effectiveness acceptability curve for both strategies at different willingness-to-pay thresholds

3.3 Final results

Following on from the workshop in July 2020, further additional analyses were undertaken, and a final workshop took place on 6th May 2021. These additional analyses are the final base-case results and sensitivity analyses.

3.3.1 Changes to the preliminary base-case analysis

In this section, we report the results of each change individually, then assess the impact to the ICER by making all changes simultaneously.

a) Re-examine the rate of learning difficulties in the screen detected cases using the same rate as in symptomatically detected: 4-month transition probability (0.008) and 6-month transition probability (0.012).

Table 12: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected QALY discounted	Incremental QALY [a]	ICER (£) per QALY gained
No universal screening for TYR1	2,132,334	-	26.67387	-	-
Universal screening for TYR1	2,346,995	214,661	26.67391	3.7	57,948
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

Table 12 shows that changing the transition probabilities does not change the magnitude or the direction of costs or QALY differences and the ICER is very similar (£58,000) to the previous base-case (£58,800 per QALY gained).

b) Revise the cost of nitisinone to more recent costs (estimated to be 25% of costs in the model, Mary-Anne Preece provided accurate and up-to-date costings)

Table 13: Annual costs for treating people with TYR1 using nitisinone

Age	Weight (kg)	Daily dose (1mg/kg)	Formulation	Cost per year	Cost under tender
Neonate	3.5	3.5 mg (0.87ml)	4mg/1ml Oral Suspension	£5,970	£10,152
1 month	4.3	4.3 mg (1.1ml)	4mg/1ml Oral Suspension	£7,549	£10,152
2 months	5.4	5.4 mg (1.3ml)	4mg/1ml Oral Suspension	£8,921	£10,152
3 months	6.1	6.1 mg (1.5ml)	4mg/1ml Oral Suspension	£10,293	£10,152
4 months	6.7	6.7 mg (1.7ml)	4mg/1ml Oral Suspension	£11,666	£11,506
6 months	7.6	7.6 mg (1.9ml)	4mg/1ml Oral Suspension	£13,038	£12,859
1 year	9	9 mg (2.3ml)	4mg/1ml Oral Suspension	£15,783	£15,566
3 years	14	14 mg (3.5ml)	4mg/1ml Oral Suspension	£24,017	£23,688
5 years	18	18 mg (4.5ml)	4mg/1ml Oral Suspension	£30,879	£30,456
7 years	23	23 mg (5.8ml)	4mg/1ml Oral Suspension	£39,800	£39,254
10 years	32	32 mg (3x10mg & 1x2mg)	Capsules (Available as 2mg, 5mg and 10mg)	£30,797	£24,141
12 years	39	39 mg (3x10mg & 1x5mg & 2x2mg)	Capsules (Available as 2mg, 5mg and 10mg)	£38,507	£29,456
14 years	50	50 mg (5x10mg)	Capsules (Available as 2mg 5mg and 10mg)	£47,040	£38,365
Adult Male	68	68 mg (6x10mg & 4x2mg)	Capsules (Available as 2mg, 5mg and 10mg)	£66,741	£50,526
Adult Female	58	58 mg (5x10mg & 4x2mg)	Capsules (Available as 2mg, 5mg and 10mg)	£57,333	£42,853

Table 14: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	**Expected total costs (£)[a]**	**Incremental costs (£)[a]**	**Expected QALY discounted**	**Incremental QALY [a]**	**ICER (£) per QALY gained**
No universal screening for TYR1	1,831,555	-	26.67387	-	-
Universal screening for TYR1	1,997,449	165,894	26.67391	3.6	45,469
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

Table 13 shows the updated costs for nitisinone and Table 14 provides the ICER using these updated costs. As shown in Table 14 , the incremental QALYs do not change, however the incremental costs have fallen, and this generates an ICER of approximately £45,500 per QALY gained.

c) We propose using a utility value of 0.70 for neurological crises, learning difficulties 0.84 and combined learning difficulties and neurological crises 0.70.

Table 15: Utility values included in the model

Health state	**Utility values**	**Range**	**Source**
False positive screen result	0.97	0.95-0.99	Tiwana et al., 2012
Liver disease	0.20	0.10-0.30	
Liver transplant	0.67	0.58-0.74	
Renal dysfunction	0.67	0.58-0.74	
Learning difficulties	0.70	0.60-0.80	Assumption
Neurological crises	0.84	0.70-0.85	
Treatment without complications	0.90	0.85-0.95	Tiwana et al., 2012
Neurological crises and learning difficulties	0.70	0.60-0.80	Assumption
Combination of sequelae	0.30	0.10-0.50	
Overdiagnosis	-0.09	-	

Table 16: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	**Expected total costs (£)[a]**	**Incremental costs (£)[a]**	**Expected QALY discounted**	**Incremental QALY [a]**	**ICER (£) per QALY gained**
No universal screening for TYR1	2,132,413	-	26.67386	-	-
Universal screening for TYR1	2,347,125	214,712	26.67389	3.3	64,097
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

Table 15 lists of the updated utility values and Table 16 provides the ICER with the changes to the utility value. As we can see, the cost differences do not change, however there is slight fall in QALY

differences, and this means that the ICER is slightly more (£64,000) than the previous base-case (£58,800 per QALY gained).

d) Based upon a reagent cost of £0.20/baby screened and a staff cost of £0.40/baby screened and assuming that no additional equipment is needed then the total cost per baby screened will be £0.60/baby screened.

Table 17: Deterministic results based on expected costs and number of screen-detected TYR1 cases per 96,500 screened newborns (assumed 96.5% screening uptake)

Strategy	**Expected total costs (£)**	**Incremental costs (£)**	**Number of screen-detected TYR1 cases**	**Incremental number of screen-detected TYR1 cases**	**ICER (£) per additional screen-detected TYR1 case**
No universal screening for TYR1	263,937	-	0.1276	-	-
Universal screening for TYR1	319,029	55,092	0.5105	0.3829	143,912
ICER, incremental cost-effectiveness ratio Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

Table 18: Deterministic results based on the expected costs and life years per 100,000 live births

Strategy	**Expected total costs (£)[a]**	**Incremental costs (£)[a]**	**Life-expectancy discounted**	**Incremental life years [a]**	**ICER (£) per LYG**
No universal screening for TYR1	2,132,413	-	27.702524	-	-
Universal screening for TYR1	2,395,375	262,962	27.702529	0.50	525,924
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; LYG, life years gained Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

Table 19: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	**Expected total costs (£)[a]**	**Incremental costs (£)[a]**	**Expected QALY discounted**	**Incremental QALY [a]**	**ICER (£) per QALY gained**
No universal screening for TYR1	2,132,413	-	26.67387	-	-
Universal screening for TYR1	2,395,375	262,962	26.67391	3.6	72,073
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

Table 19 shows that increasing the cost of the screening for the reagent and staff cost from £0.10 per baby to £0.60 per baby, the overall QALY differences remain the same, but the cost differences

increase from approximately £214,000 to £263,000 and in turn the ICER is higher (£72,000 per QALY gained) than the previous base-case (£58,800 per QALY gained).

3.3.2 Final base-case results

Table 20 and Table 21 show the final base-case results when all four changes listed above are applied simultaneously.

Table 20: Deterministic results based on the expected costs and life years per 100,000 live births

Strategy	**Expected total costs (£)[a]**	**Incremental costs (£)[a]**	**Life-expectancy discounted**	**Incremental life years [a]**	**ICER (£) per LYG**
No universal screening for TYR1	1,831,475	-	27.702524	-	
Universal screening for TYR1	2,045,569	214,094	27.702529	0.50	422,998
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; LYG, life years gained Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

Table 21: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	**Expected total costs (£)[a]**	**Incremental costs (£)[a]**	**Expected QALY discounted**	**Incremental QALY [a]**	**ICER (£) per QALY gained**
No universal screening for TYR1	1,831,475	-	26.67386	-	-
Universal screening for TYR1	2,045,569	214,094	26.67390	3.5	61,756
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

Table 21 shows that adding TYR1 screening to the current NBS screening programme is expected to yield 3.5 more QALYs than the current strategy without universal TYR1 screening at an incremental cost of approximately £214,100, which equates to an ICER of approximately £61,800 per QALY gained. These incremental costs and QALY differences are very similar to the previous base-case analysis and the ICER has only increased slightly.

3.3.3 *Probabilistic sensitivity analysis results*

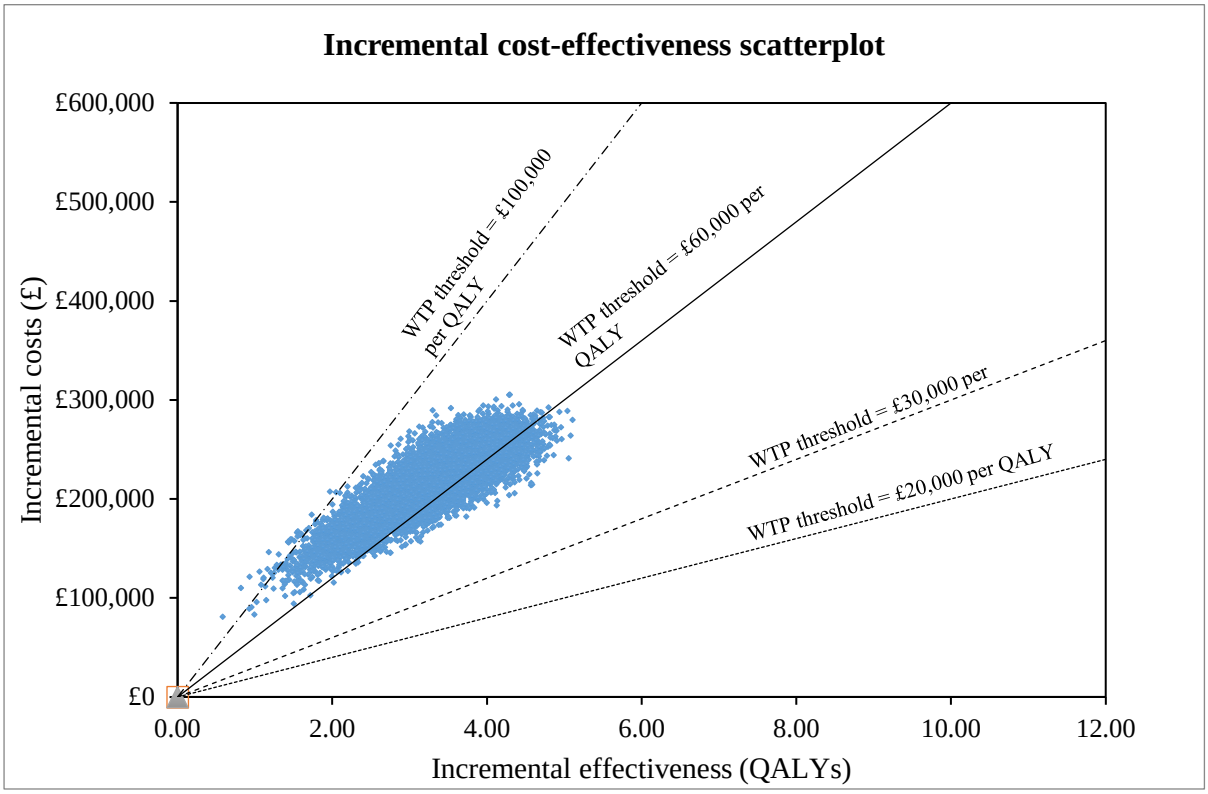

Figure 9: Incremental cost-effectiveness scatterplot for the comparison between universal TYR1 screening versus no universal TYR1 screening
WTP, Willingness to pay; QALY, Quality adjusted life years.

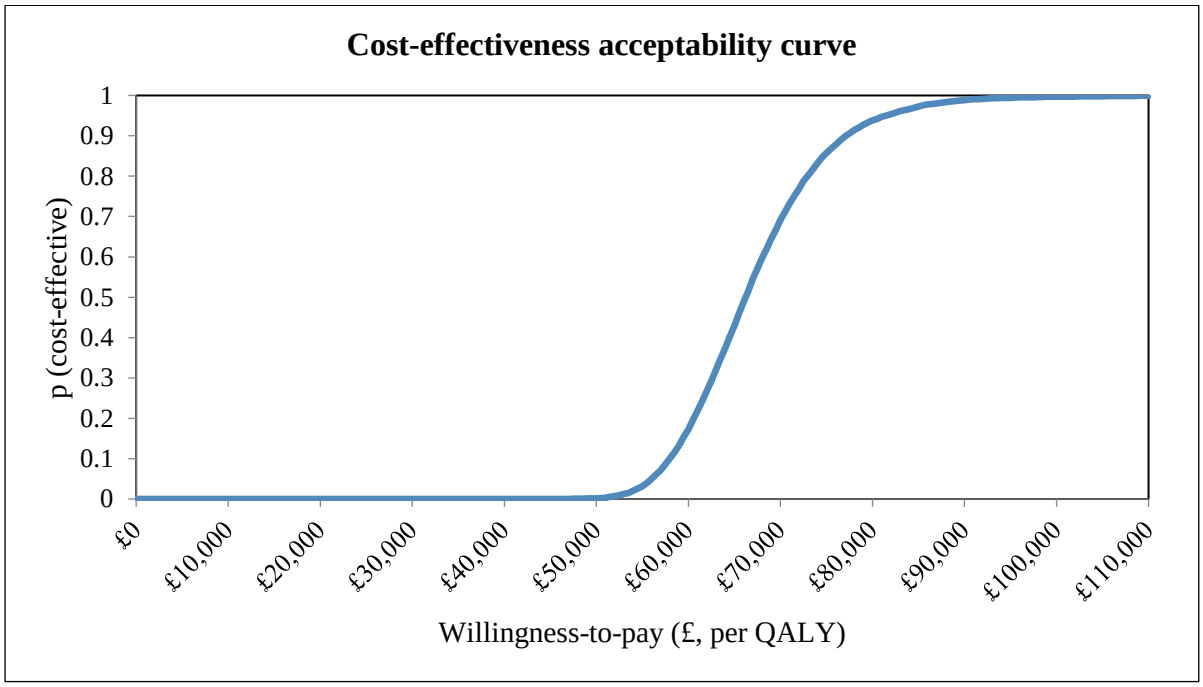

Figure 10: Cost-effectiveness acceptability curve for screening at different willingness-to-pay thresholds

The probabilistic results are presented in the form of a scatterplot which shows that there is some variation in the incremental QALYs but less variation for the incremental costs (Figure 9) and its

corresponding cost-effectiveness acceptability curves (CEAC) which shows the probability that an intervention is cost-effective at different willingness-to-pay (WTP) thresholds per QALY gained. At a WTP threshold of £20,000 per QALY gained, adding universal TYR1 screening to the current NBS screening programme has a zero probability of being cost-effective. However, for orphan drugs in ultra-rare disease settings (prevalence 1/40,000) the threshold used by NICE is very different.

The lifelong impact of the intervention on the patient determines the upper limit of the threshold in these settings. This varies from £100,000 per incremental QALY for interventions that deliver less than 10 QALYs to the patient in their lifetime to £300,000 for interventions that deliver more than 30 incremental QALYs to the patient in their lifetime.

Base-case deterministic results show that adding TYR1 screening to the current NBS screening programme compared to the current strategy without universal TYR1 screening resulted in an ICER of approximately £61,800 per QALY gained, which is below the acceptable threshold for orphan drugs in ultra-rare diseases. Results from the PSA show that at a willingness-to-pay threshold of £100,000 per QALY adding TYR1 screening to the current NBS screening programme has a probability of 1 for being cost-effective.

3.3.4 Scenario analysis results based on final base-case

e) Sensitivity analysis on the number of TYR1 cases detected via cascade testing (reduced from approximately 45% to 30%)

Table 22: Deterministic results based on expected costs and number of screen-detected TYR1 cases per 96,500 screened newborns (assumed 96.5% screening uptake)

Strategy	**Expected total costs (£)**	**Incremental costs (£)**	**Number of screen-detected TYR1 cases**	**Incremental number of screen-detected TYR1 cases**	**ICER (£) per additional screen-detected TYR1 case**
No universal screening for TYR1	263,937	-	0.1577	-	
Universal screening for TYR1	319,029	55,092	0.6308	0.4731	116,439
ICER, incremental cost-effectiveness ratio Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

The model originally estimated costs and benefits under the assumption that 45% of TYR1 patients are detected by cascade testing and would therefore not benefit from the introduction of universal TYR1 NBS screening. With less TYR1 cases identified through cascade testing, and more TYR1 cases identified through NBS screening, the ICER equates to approximately £116,400 per additional screen-detected TYR1 case (see Table 22) and the cost per QALY gained is approximately £59,300 (see Table 23).

Table 23: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected QALY discounted	Incremental QALY [a]	ICER (£) per QALY gained
No universal screening for TYR1	1,991,253	-	26.67382	-	-
Universal screening for TYR1	2,242,843	251,590	26.67386	4.2	59,289
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

f) Re-examine the transition probabilities for liver disease in symptomatically detected cases (perhaps have a fixed probability of 90% at 6 months)

Table 24: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected QALY discounted	Incremental QALY [a]	ICER (£) per QALY gained
No universal screening for TYR1	1,727,974	-	26.67385	-	-
Universal screening for TYR1	1,978,225	250,251	26.67389	4.2	60,227
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

The results in Table 24 show that there is a reduction to the expected total costs across both strategies and the ICER. More children would need to undergo liver transplantation, and if successful, this reduces the need for nitisinone treatment.

g) Plasma amino acid test done annually

Table 25: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected QALY discounted	Incremental QALY [a]	ICER (£) per QALY gained
No universal screening for TYR1	1,805,602	-	26.67386	-	-
Universal screening for TYR1	2,016,746	211,145	26.67390	3.5	60,905
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

If plasma amino acid tests are done annually, the cost differences fall slightly and there is very little change in the ICER (see Table 25).

h) Urine amino acid test not undertaken

Table 26: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected QALY discounted	Incremental QALY [a]	ICER (£) per QALY gained
No universal screening for TYR1	1,831,451	-	26.67386	-	-
Universal screening for TYR1	2,045,531	214,080	26.67390	3.5	61,752
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

If the urine amino acid test is not undertaken, the cost differences fall slightly and there is very little change in the ICER (see Table 26).

i) Patients with TYR1 are seen in clinic every three months until 2 years of age and then every six months.

Table 27: Deterministic results based on costs and QALYs per 100,000 live births

Strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected QALY discounted	Incremental QALY [a]	ICER (£) per QALY gained
No universal screening for TYR1	1,830,802	-	26.67386	-	-
Universal screening for TYR1	2,044,882	214,080	26.67390	3.5	61,752
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years Exact results have been obtained from TreeAge, but were rounded by the authors and presented.					

The model originally estimated costs and benefits for patients with TYR1 who were seen in clinic every three months until school age (4 years) and then every six months. However, patients with TYR1 are seen every 3 months for the first 2 years, then every 6 months. Applying these changes to the model results in very little change to the ICER (see Table 27).

3.4 Value of additional research

Results from the probabilistic sensitivity analysis showed that at the £30,000 willingness-to-pay (WTP) threshold per QALY, expanding NBS screening to include TYR1 was not cost-effective. However, at a WTP threshold of £100,000 per QALY the expanded NBS screening programme was cost-effective (see Figure 8). In these analyses, distributions were placed around key model input parameters to reflect the joint parameter uncertainty. The PSA was performed within the EVPI framework, whereby different thresholds (£30,000 and £100,000) were used with the PSA results to establish whether there is justification to undertake further research to reduce uncertainty. These

results showed that at both thresholds, there would be little additional value in undertaking further research for all parameters within the economic model.

3.5 Flow diagram

The results are presented diagrammatically in the form of a flow diagram in Figure 4 and Figure 5. Figure 4 depicts a flow diagram with a hypothetical birth cohort of 679,106 live born babies and the numbers of TYR1 cases identified through the proposed universal screening approach using SUAC levels as a marker and babies identified through cascade testing due to a previously affected sibling(s). Figure 5 shows the flow of the same hypothetical 679,106 newborns using the current screening approach with numbers of TYR1 cases identified as incidental findings via routine PKU screening when measuring phenylalanine as primary marker (with second-tier tyrosine measurement) and babies identified through cascade testing due to a previously affected sibling(s). Following detection, the two flow diagrams show the proportion of babies who developed long-term complications. Assuming a UK birth prevalence of 1:100,000, we would expect seven babies with TYR1 in 679,106 live births. Due to expected numbers of less than one per year, we simplified the two flow diagrams by not displaying the TYR1 cases presenting with TYR1-related symptoms before 14 days of age (assumed 0.28 babies per year). We also depicted zero TYR1 cases presenting symptomatically after eluding NBS screening (instead of the assumed 0.14 TYR1 cases per year). These simplifications (and rounding issues) lead to individual numbers only adding up to a total of 679,105 newborns, and not 679,106 newborns.

Figure 4 shows that with universal TYR1 screening, three of the seven TYR1 cases in this hypothetical birth cohort would be identified through cascade testing due to previously affected siblings, with the majority of babies developing learning difficulties. The remaining four TYR1 cases would be identified through universal screening using SUAC levels as a marker, with 11 babies having a false positive result. With respect to developing long-term complications, the model suggests that majority of these babies would develop learning difficulties.

In the current no universal TYR1 screening strategy (Figure 5), again we would expect three TYR1 cases to be identified through cascade testing due to previously affected siblings, with the majority of babies developing learning difficulties. The remaining four cases would undergo the current NBS screening approach including PKU screening. This would identify one TYR1 case, while three babies would have false negative results and 89 babies would have false positive results. With this strategy there are higher number of false negative results and false positive results. The model assumes that a vast majority of the babies with false negative results would be identified symptomatically within 10 years, and there will be an increase in the number of babies receiving a confirmatory diagnostic test. Majority of the babies who had a false negative result and were later identified through symptomatic

detection, further developed liver disease and required a liver transplant. In comparison to the expanded NBS screening strategy, a lower proportion of people developed learning difficulties. Symptomatically detected TYR1 patients were at risk of neurological crises, neurological crises and learning difficulties and a combination of long-term complications. It should be noted that the long-term outcomes of TYR1 cases identified by cascade testing are the same in both strategies.

In summary, universal screening using SUAC as a marker identifies more cases of TYR1 pre-symptomatically (7/7 compared to 4/7 using rounded numbers), fewer false positive results, but might lead to more people developing learning difficulties. Conversely, less TYR1 cases are identified pre-symptomatically as incidental finding during PKU screening using as primary marker phenylalanine (1/7), with 3/7 being identified through symptomatic detection. As a result, these babies are at risk of developing liver disease, and might require a liver transplant.

Assuming a population of 655,000 screened newborns, our model estimates that introducing universal screening for TYR1 will detect an extra three cases which would have otherwise been detected symptomatically around six-months of age, costing approximately an additional £1.4million per year. Overall, we estimate that this will give an extra 24 quality adjusted life years compared to the current practice of phenylketonuria testing. This was largely through improvements in quality of life through avoiding liver disease. The model estimates 511 fewer false positive results per year, but with uncertainty in test accuracy estimates underpinning them. We are very unsure of some of the values included in the economic model, so the results should be interpreted with caution. The model estimated more cases of learning disability, through the earlier administration of nitisinone, but the scientific community remains unsure whether the learning difficulties are as a result of nitisinone treatment or tyrosinaemia itself.

4 Discussion

This economic analysis used a decision analytical model to evaluate the cost-effectiveness of adding TYR1 screening to the current NBS screening programme against the current practice without universal TYR1 screening. Under the current assumptions and given the paucity of some of the clinical information required for the model, careful interpretation of these results is required.

Considering screen-detected cases only, the results showed that in 96,500 screened newborns, an expanded NBS screening programme including TYR1 screening is likely to detect an additional 0.3829 cases of TYR1 compared to current practice without universal TYR1 screening, with an ICER of approximately £17,600 per additional screen-detected case of TYR1. The ability to correctly identify a specific condition/disease is an important consideration when deciding to implement a screening programme. However, this outcome measure alone does not consider the long-term costs or benefits of the screening programme. Costs of the management and treatment of the condition and its long-term complications as well as the impact of diagnostic errors (false positive and false negative results) must also be considered in the analysis.[11]

The economic model was constructed to capture the impact of false positive and false negative results, as well as the treatment of TYR1 and its long-term complications. The base-case results showed that an expanded NBS screening programme including TYR1 screening is slightly more expensive and is expected to yield more QALYs compared to the current approach without universal TYR1 screening, with an ICER of approximately £61,800 per QALY gained. This is above the NICE recommended willingness-to-pay thresholds for standard conditions[32] but within the accepted threshold for treatment of rare or 'orphan' diseases. The threshold for screening of rare diseases is unclear to the authors. Of note, there is little difference in the expected life-years between the strategies; but this difference can be seen in terms of the health-related quality of life. Screen-detected babies would avoid liver disease and liver transplant and could potentially have a better health-related quality of life. However, it should be noted that the model estimated an increase in the rate of learning disability and renal dysfunction in keeping with cases detected through current practice. This incremental gain in expected QALYs leads to a reduction to the ICER. The results from the probabilistic sensitivity analysis show that at a willingness-to-pay threshold of £100,000 per QALY, an expanded NBS screening programme including TYR1 screening has a probability of one of being cost-effective compared to current practice not including universal TYR1 screening.

Additional analyses in the form of scenario analyses and sensitivity analysis were undertaken, and these results showed that changes made to the costs of the screening tests for both strategies had the greatest impact on the ICER. We compared our results with those of previous economic evaluations identified through our literature review (see Appendix 9 for further details). One common result

across analyses was the minimal incremental life-years gained between strategies, which shows some agreement between economic analyses. However, the results in terms of cost per LYG showed that there was agreement between our analyses and the analyses by Cipriano et al.,[39; 41] with high costs (CAN$ 309,400) for each additional LYG.

4.1 Implications for practice

4.2 Strengths and limitations of the economic analysis

Strengths

- For the current screening approach, sensitivity and specificity of the current PKU screening protocol to identify TYR1 cases were derived from information obtained for DBS phenylalanine levels of babies with TYR1 on day 5. We identified the proportion of cases that could have been identified using the current PKU screening protocol using an unpublished local dataset.
- We presented detailed resource use and costs associated with management of people with TYR1.
- The model considers both the short-term impact of screening (i.e. identifying additional cases of TYR1 through screen-detection) and future costs and benefits of identifying a greater proportion of TYR1 cases pre-symptomatically.

Limitations

- We assumed that utility values for screen-detected and clinically detected babies were the same. To our knowledge there are no published studies that suggest differences in utility weights to use for these two different groups.
- In the model, we assumed that there is a probability of children developing both learning difficulties and neurological crises. We have assumed a utility value of 0.70 for this health state.
- The model included long-term complications associated with TYR1 or its treatment. The incidence of these complications was obtained from the literature.[8; 13; 15; 33] For some long-term complications, it was difficult to decipher the events that occurred in people who were screened-detected and those in babies who were clinically diagnosed. Additionally, assumptions were made about length of follow-up. Furthermore, we assumed that these events occurred at a constant rate over time and this might have led to an over- or underestimation of these events.
- The economic analysis was undertaken from the viewpoint of the NHS and PSS. A more comprehensive societal perspective could have been undertaken to capture the costs borne by

patients and their families, thus including costs associated with loss of productivity, carers' costs and out-of-pocket expenditures such as travel costs.

- There was very poor quality data in many model input parameters, so results should be interpreted with caution.

4.3 Future research recommendations

Even though the value of information analysis indicated that expanding the current NBS screening programme for TYR1 is unlikely to change the conclusion, more accurate model data inputs need to be collected (i.e. birth prevalence and methods of detection of TYR1 in the UK; probabilities for long-term complications in pre-symptomatically and symptomatically detected cases; resource use and costs of diagnosing and treating TYR1; and on utilities, and harms/side-effects of treatments for TYR1), in order to help refine and provide a more accurate cost-effectiveness estimate.

5 References

1. Weinstein MC, O'Brien B, Hornberger J *et al.* (2003) Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* **6**, 9-17.

2. Public Health England (2017) Guidance. UK NSC: evidence review process. (accessed 24 May 2018)

3. Geppert J, Taylor-Phillips S, Stinton C *et al.* (2016) *Newborn Blood Spot Screening for Tyrosinaemia Type 1 in the UK - External review against programme appraisal criteria for the UK National Screening Committee (UK NSC)*. https://legacyscreening.phe.org.uk/tyrosinaemia.

4. Geppert J, Stinton C, Freeman K *et al.* (2017) Evaluation of pre-symptomatic nitisinone treatment on long-term outcomes in Tyrosinemia type 1 patients: a systematic review. *Orphanet J Rare Dis* **12**, 154.

5. Stinton C, Geppert J, Freeman K *et al.* (2017) Newborn screening for Tyrosinemia type 1 using succinylacetone - a systematic review of test accuracy. *Orphanet J Rare Dis* **12**, 48.

6. Public Health England (2017) *NHS Newborn Blood Spot Screening Programme: A laboratory guide to newborn blood spot screening for inherited metabolic diseases.*

7. Newborn Blood Spot Screening Programme in the UK (2018) Data collection and performance analysis report 2016 to 2017. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/709367/Newborn_blood_spot_screening_data_collection_and_performance_analysis_report_2016_to_2017.pdf (accessed 11 March 2018

8. Bartlett DC, Lloyd C, McKiernan PJ *et al.* (2014) Early nitisinone treatment reduces the need for liver transplantation in children with tyrosinaemia type 1 and improves post-transplant renal function. *Journal of Inherited Metabolic Disease* **37**, 745-752.

9. National Health Service (NHS): Birmingham Women's and Children's NHS Foundation Trust West Midlands newborn screening laboratory annual report 2016-2017. https://bwc.nhs.uk/download.cfm?doc=docm93jijm4n2146.pdf&ver=2863 (accessed 15 March 2019

10. Hutchesson AC, Hall SK, Preece MA *et al.* (1996) Screening for tyrosinaemia type I. *Arch Dis Child Fetal Neonatal Ed* **74**, F191-194.

11. Pandor A, Eastham J, Beverley C *et al.* (2004) Clinical effectiveness and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review. *Health technology assessment (Winchester, England)* **8**, iii, 1-121.

12. Office for National Statistics (2017) Births in england and Wales: 2017. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsummarytablesenglandandwales/2017

13. Larochelle J, Alvarez F, Bussieres JF *et al.* (2012) Effect of nitisinone (NTBC) treatment on the clinical course of hepatorenal tyrosinemia in Quebec. *Molecular genetics and metabolism* **107**, 49-54.
14. McKiernan P (2017) Liver Transplantation for Hereditary Tyrosinaemia Type 1 in the United Kingdom. *Advances in Experimental Medicine & Biology* **959**, 85-91.
15. Mayorandan S, Meyer U, Gokcay G *et al.* (2014) Cross-sectional study of 168 patients with hepatorenal tyrosinaemia and implications for clinical practice. *Orphanet Journal of Rare Diseases* **9**.
16. Curtis LA, Burns, Amanda (2018) Unit Costs of Health and Social Care 2018. Project report. University of Kent.
17. Royal College of Paediatrics and Child Health (January 2013) UK-WHO growth charts - 0-4 years. https://www.rcpch.ac.uk/resources/uk-who-growth-charts-0-4-years (accessed 12 March 2019 2019)
18. NHS (2018) Reference costs; 2017/2018 reference cost data. https://improvement.nhs.uk/resources/reference-costs/ (accessed 12 March 2019 2019)
19. Homerton BHS trust Homerton NHS tariff 2014-2015 http://www.homerton.nhs.uk/media/174141/1251-private-patient-tariff-for-2014-15.pdf (accessed 15 March 2019)
20. North East Lincolnshire Clinical Commissioning Group (CCG) North East Lincolnshire CCG costs 2013-2014. www.northeastlincolnshireccg.nhs.uk/data/uploads/.../nel-435-attachment-response.doc (accessed 15 March 2019
21. National Guideline Centre (2016) Preoperative tests (update): Routine preoperative tests for elective surgery (NG45). https://www.nice.org.uk/guidance/ng45/evidence/full-guideline-pdf-87258149468 (accessed 12 March 2019 2019)
22. Olson J, Sharp P, Goatman K *et al.* (2013) Improving the economic value of photographic screening for optical coherence tomography-detectable macular oedema: a prospective, multicentre, UK study. *Health technology assessment (Winchester, England)* **17**, 1-142.
23. NHS Business Services Authority (NHSBSA) NHS Prescription Services. https://www.nhsbsa.nhs.uk/pharmacies-gp-practices-and-appliance-contractors/drug-tariff (accessed 13 March 2019
24. Belanger-Quintana A, Dokoupil K, Gokmen-Ozel H *et al.* (2012) Diet in phenylketonuria: a snapshot of special dietary costs and reimbursement systems in 10 international centers. *Molecular genetics and metabolism* **105**, 390-394.
25. Tiwana SK, Rascati KL, Park H (2012) Cost-Effectiveness of Expanded Newborn Screening in Texas. *Value in Health* **15**, 613-621.
26. Ara R, Brazier JE (2010) Populating an economic model with health state utility values: moving toward better practice. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* **13**, 509-518.

27. Kale MS, Korenstein D (2018) Overdiagnosis in primary care: framing the problem and finding solutions. *BMJ (Clinical research ed)* **362**, k2820.
28. Office for National Statistics National life tables: UK. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables (accessed 12 Mach 2019
29. National Institute for Health and Care Excellence (2013) Guide to the methods of technology appraisal 2013 https://www.nice.org.uk/process/pmg19/resources/guide-to-the-processes-of-technology-appraisal-pdf-72286663351237 (accessed 12 March 2019 2019)
30. Thorn J, Coast J, Andronis L (2015) Interpretation of the Expected Value of Perfect Information and Research Recommendations: A Systematic Review and Empirical Investigation. *Medical Decision Making* **36**, 285-295.
31. Public Health England (2015) Criteria for appraising the viability, effectiveness and appropriateness of a screening programme https://www.gov.uk/government/publications/evidence-review-criteria-national-screening-programmes/criteria-for-appraising-the-viability-effectiveness-and-appropriateness-of-a-screening-programme (accessed 18 March 2019
32. National Institute for Health and Care Excellence (2012) Methods for the development of NICE public health guidance (third edition). https://www.nice.org.uk/process/pmg4/resources/methods-for-the-development-of-nice-public-health-guidance-third-edition-pdf-2007967445701 (accessed June 2019
33. McKiernan PJ, Preece MA, Chakrapani A (2015) Outcome of children with hereditary tyrosinaemia following newborn screening. *Archives of Disease in Childhood* **100**, 738-741.
34. Wijburg FA, Reitsma Ch WC, Slooff MJH *et al.* (1995) Liver transplantation in tyrosinaemia type I: The Groningen experience. *Journal of Inherited Metabolic Disease* **18**, 115-118.
35. Organisation for Economic Co-operation and Development (OECD) Prices and purchasing power parities (PPP) data. http://www.oecd.org/sdd/prices-ppp/ (accessed 13 March 2019
36. National Institute for Health and Care Excellence (2015) Costing statement: Challenging behaviour and learning disabilities Implementing the NICE guideline on challenging behaviour and learning disabilities (NG11). https://www.nice.org.uk/guidance/ng11/resources/costing-statement-pdf-70691581 (accessed 12 March 2019 2019)
37. Curtis LA, Burns, Amanda (2015) Unit Costs of Health and Social Care 2015. Project report. University of Kent.
38. Institute of Health Economics Alberta Canada (2016) *Newborn blood spot screening for galactosemia, tyrosinemia type I, homocystinuria, sickle cell anemia, sickle cell/beta-thalassemia, sickle cell/hemoglobin C disease, and severe combined immunodeficiency*. *Alberta STE Report*.
39. Cipriano LE, Rupar CA, Zaric GS (2007) The cost-effectiveness of expanding newborn screening for up to 21 inherited metabolic disorders using tandem mass spectrometry: results from a decision-analytic model. *Value in Health* **10**, 83-97.

40. Chilcott J, Bessey A, Pandor A *et al.* (2013) Expanded newborn screening for inborn errors of the metabolism.
41. Auguste P, Geppert J, Taylor-Phillips S *et al.* (2019) Cost-effectiveness of newborn blood spot screening for Tyrosinaemia type 1 using tandem mass spectrometry University of Warwick.
42. Jonsson B (2009) Ten arguments for a societal perspective in the economic evaluation of medical innovations. *Eur J Health Econ* **10**, 357-359.
43. Canadian Agency for Drugs and Technologies in Health (2016) Guidelines for the Economic Evaluation of Health Technologies: Canada. *Discounting and the evaluation of health care programs*. (accessed (May 2019)

6 Appendix

6.1 Appendix 1. Methods and sources for the test accuracy estimates

The values for the universal NBS screening for the TYR1 strategy are based on identifying TYR1 babies using MS/MS measurement of succinylacetone (SUAC) from DBS as estimated by the 2016 UK NSC report[3] and from our updated searches.[5] The current strategy used in the UK to identify babies with TYR1 is through cascade testing of newborns with previously affected siblings, and by incidental detection on NBS screening for PKU. To estimate the probability of PKU screening to detect TYR1, we used an unpublished local dataset from the West Midlands that reported DBS phenylalanine levels of babies with TYR1 on day 5. We then identified the proportion of cases that could have been identified by the current PKU screening protocol (expert communication, January 2019). From this, we derived a sensitivity of 25%. A specificity of 99.987% was estimated using information from the 2016/2017 performance analysis report from the NBS screening programme in the UK.[7] In 2016/2017, 779,688 babies were screened for PKU of which 107 screened positive. If we expect 4-8 cases of TYR1 within the screened newborn population, approximately 25% (see sensitivity estimate) would be detected by PKU screening (1-2 true positive cases), resulting in an estimated 105 false positives and 779,575 true negatives for TYR1 (worst case 107 false positives and 779,573 true negatives, specificity 99.986%). We did not require test accuracy information for cascade testing or those presenting with symptoms suggestive of TYR1, as in the current clinical pathway, these babies receive the diagnostic protocol. In the model, we assumed that the diagnostic protocol used to confirm TYR1 is 100% accurate.

6.2 Appendix 2. Transitions probabilities

First the instantaneous rate of the event was derived using equation 1, where *p* is the probability and *t* is time in years.

$$r = -[\ln(1-p)]/t \qquad \text{equation 1}$$

This rate was converted to the appropriate transition probabilities to be used in the model.

$$p = 1-\exp\{-rt\} \qquad \text{equation 2}$$

The previous 2016 UK NSC review concluded that whilst there is clear evidence that nitisinone is an effective treatment, benefits of early treatment initiation following screen-detection over later treatment initiation following symptomatic detection were not clear.[3] To reflect this uncertainty, we used the identified literature from the 2016 UK NSC review[3] and performed, if applicable, post hoc analyses of published individual patient data[4] to derive a central estimate of the potential benefit of pre-symptomatic detection over later symptomatic detection for each outcome, and an optimistic and pessimistic version. The optimistic version may overestimate the benefit of screening and the pessimistic version may underestimate the benefit of screening. Where there was very little or no data for optimistic and pessimistic estimates we increased or reduced the effect size from the central estimate. For the pessimistic estimate, we increased the number of events in the screen-detected group by 50% and decreased the number of events in the symptomatically presenting group by 50%. For the optimistic estimate, we decreased the number of events in the screen-detected group by 50% and increased the number of events in the symptomatically presenting group by 50%.

The 2016 UK NSC review identified two prospective cohorts (one UK, one Canadian) and one international survey examining the outcomes for people with TYR1 with nitisinone treatment started at different ages.[3] To populate the model, we considered data from the UK cohort,[8; 33] the Canadian Quebec cohort,[13] and the international survey[34] alongside related post-hoc analyses published by Geppert et al.[3] In choosing estimates we prioritised larger studies, UK-based studies, and studies judged to be at lower risk of bias. Where possible, we used the reported individual patient data to calculate estimates so as to reduce bias as follows: where possible we removed cases detected symptomatically before 14 days from the analysis as they cannot benefit from screening and would result in underestimates of benefit if included in the early group. In the UK, TYR1 is typically detected symptomatically in the first six months, so where possible we excluded those TYR1 cases diagnosed in the pre-nitisinone era who did not receive nitisinone within the first year of life from the late treatment group. The international survey by Mayorandan et al.[15] did not provide individual patient data. We were therefore unable to remove early (<14 days) symptomatic cases or cases born in the pre-nitisinone era who did not receive nitisinone until they were older than one year. For this international survey[15] we used the reported age category of TYR1 cases treated before the age of one

month as a proxy for the screen-detected group and the age category of TYR1 cases with treatment initiated between 1-6 months of age as a proxy of the symptomatically detected group. The international survey did not report the follow-up time by age groups at start of nitisinone treatment; we therefore used the reported mean follow-up time for all included TYR1 cases and assumed it was the same in each group, which is unlikely to be true. We then estimated the follow-up time until event or study end date, respectively, by using the provided information of mean age at onset of the complication of interest.

This information was used to derive transition probabilities for the model. Transition probabilities determine the speed (and direction) at which the hypothetical cohort of people move between health states within the Markov model. For example, the central estimate for renal dysfunction, the transition probabilities of 0.002 and 0.003 represent the 4-monthly and 6-monthly probability that pre-symptomatically detected babies would develop renal dysfunction each cycle of the model. Likewise, for babies who were clinically detected, there is a 4-monthly and 6-monthly probability of 0.004 and 0.006, respectively, for developing renal dysfunction in each cycle of the model.

Liver disease

To derive the central estimate, we performed a post hoc analysis of the individual patient data from TYR1 patients in Quebec reported by Larochelle et al.[13] As described in the "Post hoc 1" analyses by Geppert et al.,[4] we compared screen-detected versus symptomatic detection in cases with direct nitisinone initiation after diagnosis. No case out of 24 screen-detected and early (<1 month) nitisinone-treated patients developed detectable liver disease at a mean follow-up of 9.0 years (0 events in 216.1 person-years). Four out of five symptomatically presenting cases had liver disease (hepatic failure or hepatomegaly) at presentation at a mean follow-up time of 2.1 years (4 events in 10.6 person-years).

As an optimistic estimate, we used data from the UK-based study by McKiernan et al.[33] McKiernan et al.[33] reported no cases with liver disease in 12 pre-symptomatically detected TYR1 patients with a mean estimated follow-up time from birth of 7.8 years (follow-up time was extrapolated from individual patient data for the 10 pre-symptomatically detected cases reported in Bartlett et al.,[8] adding one year of follow-up to each case, plus adding two 3-year old cases). All three surviving cases that had presented symptomatically developed liver disease at a mean age at onset of 0.67 years (3 events in 2.0 person-years).

As pessimistic estimate, we reduced the effect size of our central estimate as described above.

Liver transplantation in cases with liver disease

As a central estimate, we used individual patient data from the UK reported by McKiernan et al.[33] In the 12 pre-symptomatically detected cases, there was no case with liver disease and no patient needed a liver transplantation at a mean estimated follow-up time from birth of 7.8 years (median age 8.5 years, range 3-12.5 years). All three surviving cases that presented symptomatically had liver disease. One out of three cases with liver disease required a liver transplantation at the age of 5 months, while the other two patients are clinically normal with compensated liver disease aged 10 and 17, respectively (one event in 27.4 person-years).

To derive the optimistic estimate, we performed a post hoc analysis of the data from TYR1 patients in Quebec reported by Larochelle et al.[13] As described in the "Post hoc 1" analyses by Geppert et al.,[4] we compared screen-detected versus symptomatic detection in cases with direct nitisinone initiation after diagnosis. No case out of 24 screen-detected and early (<1 month) nitisinone-treated patients developed detectable liver disease and none required liver transplantation at a mean follow-up of 9.0 years. Four out of five symptomatically presenting cases had hepatic failure or hepatomegaly at presentation. Of these four patients with liver disease, three required liver transplantation at a mean follow-up time of 5.5 years (3 events in 22.1 person-years).

For the pessimistic estimate, we used the central estimate derived from McKiernan et al.[33] and reduced the number of events in the symptomatically presenting group by 50%.

Renal dysfunction

As central estimate, we used data for "Renal dysfunction" from the international survey by Mayorandan et al.[15] Two out of 37 cases treated before one month of age had renal dysfunction, while 4/45 cases with nitisinone treatment started at 1-6 months had renal dysfunction at a mean follow-up for all 168 patients of 9.1 years (odds ratio 1.7 compared to those treated below the age of one month). Mean age at onset of "renal tubular dysfunction" was reported as 28.4 months. We therefore estimated the mean follow-up time to be 8.7 years in the cases treated before one month and 8.5 years in the cases with nitisinone treatment started between 1-6 months.

For the optimistic estimate, we used the central estimate derived from the international survey by Mayorandan et al.,[15] decreased the number of events by 50% for the screen-detected and increased the number of events by 50% for symptomatically detected cases.

For the pessimistic estimate, we used the central estimate derived from the international survey by Mayorandan et al.,[15] increased the number of events by 50% for the screen-detected and decreased the number of events by 50% for symptomatically detected cases.

Learning difficulties

As central estimate, we used data for "Psychomotor impairment" provided by the international survey by Mayorandan et al.[15] There were 8/37 cases with psychomotor impairment in patients with nitisinone treatment started before one month of age and 8/45 TYR1 cases with psychomotor impairment in patients with treatment started between 1-6 months (reported odds ratio 0.7 compared to those treated below the age of one month). Mean follow-up time for all 168 included patients was 9.1 years. Mean age at onset of "impaired psychomotor development" was reported as 29.4 months. We therefore estimated the mean follow-up time until the event or study end date, respectively, to be 7.7 years in the TYR1 patients with treatment started before one month and 7.9 years in cases with nitisinone treatment initiated between 1-6 months.

For the optimistic estimate, we used the central estimate derived from the international survey by Mayorandan et al.,[15] decreased the number of events by 50% for the screen-detected and increased the number of events by 50% for symptomatically detected cases.

For the pessimistic estimate, we used the central estimate derived from the international survey by Mayorandan et al.,[15] increased the number of events by 50% for the screen-detected and decreased the number of events by 50% for symptomatically detected cases.

Neurological crisis

To derive our central estimate, we performed a post hoc analysis of the individual patient data from Quebec reported by Larochelle et al.[13] As described in the "Post hoc 3" analyses by Geppert et al.,[4] we compared screen-detected patients with direct nitisinone initiation with screen-detected cases with 1-12 months delayed nitisinone treatment. There were no cases with neurological crisis in 24 patients treated before one month of age at a mean follow-up time of 9.0 years (0 events in 216.1 patient-years). One neurological crisis occurred at around 11 months of age (before nitisinone initiation) among 10 screen-detected cases with nitisinone started between 1-12 months of age (mean follow-up time until event or study end date, respectively, estimated to be 11.6 years).

For the optimistic estimate, we used the data from Quebec as published by Larochelle et al. for the early (<1 month) and late (>1 month) nitisinone-treated groups.[13] There were 0/24 cases with neurological crisis in patients treated before one month of age (mean follow-up time 9.0 years), while

6/26 cases had a total of 16 episodes of neurological crisis, all occurring before nitisinone treatment was started (mean follow-up time until event or study end date, respectively, estimated to be 8.8 years).

To derive the pessimistic estimate, we performed a post hoc analysis of the individual patient data from Quebec TYR1 patients reported by Larochelle et al.[13] As described in the "Post hoc 1" analyses by Geppert et al.,[4] we compared screen-detected versus symptomatic detection in cases with direct nitisinone initiation after diagnosis. There were 0/24 cases with neurological crisis in screen-detected patients treated before one month of age (mean follow-up time 9.0 years) and no episode of neurological crisis occurred in the five symptomatically presenting cases with direct nitisinone initiation after diagnosis (mean follow-up time 6.26 years).

Combination of learning difficulties and neurological crises

Assumed to be equal to the transition probability for neurological crises.

Combination of long-term complications

Assumption to be equal to the transition probability for neurological crises**.**

In Appendix Table 1, we derived optimistic and pessimistic transition probabilities to show the impact to the ICERs (cost per LYG and cost per QALY). These estimates were derived mainly from the literature; where information was unavailable, we assumed that the optimistic estimate is equivalent to -50% of the central estimate for screen-detected and +50% for symptomatically detected. Likewise, for the pessimistic estimates we assumed +50% for screen-detected and -50% for symptomatically detected.

Appendix Table 1: Four and six-month transition probabilities by complication

Long-term complication	Screen detected		Symptomatically detected		Source
	4-month transition probability	6-month transition probability	4-month transition probability	6-month transition probability	
Liver disease					
Central estimate	0	0	0.118	0.172	Larochelle et al.,[13] "Post hoc 1" analysis (Screen detection vs symptomatic detection, all with direct nitisinone initiation) as in Geppert et al.[4]
Optimistic estimate	0	0	0.393	0.528	McKiernan et al.[33]; information on follow-up time in screen-detected cases extrapolated from Bartlett et al.[8]

Long-term complication	Screen detected		Symptomatically detected		Source
	4-month transition probability	6-month transition probability	4-month transition probability	6-month transition probability	
Pessimistic estimate	0	0	0.061	0.090	Larochelle et al.,[13] (central estimate): Number of events -50% for symptomatically detected cases.
Liver transplant in cases with liver disease					
Central estimate	0	0	0.012	0.018	McKiernan et al.[33]; information on follow-up time in screen-detected cases extrapolated from Bartlett et al.[8]
Optimistic estimate	0	0	0.044	0.066	Larochelle et al.,[13] "Post hoc 1" analysis (Screen detection vs symptomatic detection, all with direct nitisinone initiation) as in Geppert et al.[4]
Pessimistic estimate	0	0	0.006	0.009	McKiernan et al.[33] (central estimate): Number of events -50% for symptomatically detected cases.
Renal dysfunction					
Central estimate	0.002	0.003	0.004	0.005	Mayorandan et al.[15]; data for "Renal dysfunction".
Optimistic estimate	0.001	0.002	0.006	0.008	Mayorandan et al.[15] (central estimate): Number of events -50% for screen-detected and +50% for symptomatically detected.
Pessimistic estimate	0.003	0.005	0.002	0.003	Mayorandan et al.[15] (central estimate): Number of events +50% for screen-detected and -50% for symptomatically detected.
Learning difficulties					
Central estimate	0.010	0.016	0.008	0.012	Mayorandan et al.[15] (frequency/odds ratio for "psychomotor impairment")
Optimistic estimate	0.005	0.007	0.013	0.019	Mayorandan et al.[15] (central estimate): Number of events -50% for screen-detected and +50% for symptomatically detected.
Pessimistic estimate	0.017	0.025	0.004	0.006	Mayorandan et al.[15] (central estimate): Number of events +50% for screen-detected and -50% for symptomatically detected.

Long-term complication	Screen detected		Symptomatically detected		Source
	4-month transition probability	6-month transition probability	4-month transition probability	6-month transition probability	
Neurological crisis					
Central estimate	0	0	0.003	0.005	Larochelle et al.[13]: "Post-hoc 3" analysis (Screen detection, direct nitisinone initiation vs screen detection, 1–12 months delayed nitisinone initiation) as in Geppert et al.[4]
Optimistic estimate	0	0	0.010	0.015	Larochelle et al.,[13] including all early vs. late nitisinone-treated patients
Pessimistic estimate	0	0	0	0	Larochelle et al.[13]: "Post hoc 1" analysis (Screen detection vs symptomatic detection, all with direct nitisinone initiation) as in Geppert et al.[4]
Combination of learning difficulties and neurological crises					
Central estimate	0	0	0.003	0.005	Assumed to be equal to the transition probability for neurological crises
Optimistic estimate	0	0	0.010	0.015	
Pessimistic estimate	0	0	0	0	
Combination of long-term complications					
Central estimate	0	0	0.003	0.005	Assumption to be equal to the transition probability for neurological crises
Optimistic estimate	0	0	0.010	0.015	
Pessimistic estimate	0	0	0	0	

6.3 Appendix 3. Detailed information on resource use and costs

6.3.1 Inpatient stay

Based on expert clinical opinion it was assumed that babies diagnosed with TYR1 before the age of one would stay in their local hospital for one week before being transferred for a two-week stay at a specialist hospital (we assumed that one week would be spent in high dependency care and the other week in normal care). The cost of stay in the local hospital was based on NHS reference costs for inborn errors of metabolism; the cost of stay in the specialist hospital was based on neonatal critical care costs derived from NHS reference costs.[18]

We assumed that infants (up to 1 year) who were clinically diagnosed with TYR1 would require inpatient stay and would have an average cost per stay of £2,880.[16] In the model, it was assumed that babies up to the age of one year old would require inpatient stay.

6.3.2 Outpatient visits

Based on expert clinical opinion, patients with TYR1 are seen in clinic every three months until school age (4 years) and then every six months. These clinic visits were costed as paediatric consultant-led clinics for patients aged up to 18 years. For 19 years and older, the cost for an adult outpatient clinic was used.[16] These consultant-led clinics are multi-disciplinary and the patient would also see a dietician in these clinics.

Patients attend separate clinics to receive dietetics advice (costed as paediatric non-consultant led clinics).[16] The dietician would cover areas such as overview of condition, first-line dietetic management, practical teaching of making up and adjusting feeds, weaning and guidance on protein exchanges, label reading, excluded and included foods and dietary principles when starting school. On average patients would expect to have an extra seven clinic consultations with a dietician between birth and starting school.

6.3.3 Contact with dietician

During the initial hospital stay, the parents/carers would also see the dietician twice a day for 30 minutes. Based on expert opinion, there is constant contact between the family/carers and the dietician. Telephone calls are twice weekly during the first two months after birth and then weekly during the next four months (lasting on average about 15 minutes each). Between the ages of six months and one year, telephone calls lasts between 30 to 60 minutes each and are twice weekly during months aged 7 and 8 and then weekly during the next four months (months aged 9 to 12). From two years of age until the patient attends school there are telephone calls with the dietician every three months lasting around 15 minutes each. In addition, the dietician would visit the patient at home

during the first month, and when the patient starts school they would also do school visits (we have assumed three visits in total – one visit for when the patient starts reception, infants and juniors).

6.3.4 Other health care resource use

We based the following assumptions about health care resource usage on discussions with our clinical experts:

- Health visitor input: Once the patient is discharged from hospital, they would see the health visitor at a local clinic once a week during the first six months.
- Blood and urine tests: Undertaken when a patient with TYR1 visits the outpatient clinic. Four-monthly during the first year; six-monthly between ages 1 and 11 years; and yearly aged 11 and over except for:
    - Full quantitative amino acid tests undertaken twice yearly and phenylalanine or tyrosine amino acid tests which would be undertaken weekly.
- Imaging: Liver and renal ultrasound imaging six-monthly between one and 11 years and yearly for those aged 11 years and over. Either magnetic resonance imaging or computed tomography liver scanning undertaken yearly.
- Other tests:
    - Molecular genetics testing as a one-off undertaken at birth;
    - An eye examination at age four years;
    - Neuropsychological assessment undertaken ideally when the patient is of 2-3, 4-5, 8-10 and 12-14 years of age; and
    - Bone density assessment undertaken at age four and then every two years.

6.3.5 Diet

Patients with TYR1 need to have a protein-restricted diet which is low in both tyrosine (Tyr) and phenylalanine (Phe). During the first six months babies should have either breast and/or formula milk according to national guidelines. In addition, babies with TYR1 would also have special formula milk. We have assumed that during the first six months babies would use one tub of special formula milk a week (allowing for wastage) and that this cost would be borne by the NHS (1 item per week on prescription costs £8.80).[23] When babies are aged between six months and one year, they also have this formula milk but we have assumed that the tub would last two weeks as babies are now in the process of being weaned.

From six months to two years of age, patients would require a daily Phe-free and Tyr-free protein substitute. This comes in the form of a gel and is available in sachets. It would be available to patients via a prescription. (One box contains 24 sachets, and we have assumed that each box is one item and this costs the NHS £8.80).

Estimating the annual costs of special diet for patients is difficult. Many factors make quantification of this cost calculation highly uncertain. In a recent study, authors estimated an annual cost of a low-protein diet for the UK and nine other countries using data on costs and reimbursement for special dietary foods used in the management of PKU.[24] We used purchasing power parities[35] and the HCHS index to uplift their findings to current prices[16] and used these to provide annual costs in the model for patients up to 16 years of age (as this cost would be borne by the NHS and not the patient).

6.3.6 Cost associated with long-term sequelae

Costs were derived for treatment of liver disease (including liver cancer), liver transplant, renal dysfunction, learning difficulties, neurological crisis, for a combination of learning difficulties and neurological crises, and a combination of all long-term complications.

Liver transplant

The costs of liver transplant includes the liver transplantation surgery, inpatient stay, x-rays and ultrasounds, drugs, blood tests, parenteral nutrition, and staff time, including surgical and anaesthetic team costs. The total cost of liver transplant was estimated at approximately £144,600. This cost was provided by Birmingham Children's Hospital (see Appendix 4). In the model, it was assumed that there was an increased risk of death following transplant (see section below on mortality). Additionally, it was assumed that children who had successfully received a liver transplant would discontinue dietary and nitisinone treatment.

Learning difficulties/neurocognitive problems

We assumed that there were no extra costs associated with interventions for neonates between 0-2 years. We assumed that babies from the ages of 2-4 years, would receive an additional four visits by a health visitor costing £59.11 each.[16] We assumed that children from the ages of 4-16 years with learning difficulties would require a school-based intervention estimated at £147.62 per child per year. This includes teacher training, a programme coordinator and materials. This cost was obtained from the National Institute of Health and Care Excellence's (NICE) costing statement: Challenging behaviour and learning disabilities (NG11),[36] then inflated to current prices and used as a proxy.

Neurological crises

We assumed that children with neurological crises would require one visit to a neuropsychologist at £48.88 per hour every six months. This cost was obtained from Unit Costs of Health and Social care (2015)[37] and uprated to current prices.[16]

Combination of long-term complications

It was assumed that patients with a combination of sequelae each received tests/treatments for each long-term complication.

6.4 Appendix 4. Liver transplant costs from Birmingham Children's Hospital (BCH)

	Per episode	Units	Total cost £
Liver transplant costs – recipient			
Transplantation of liver			£53,059
HDU on ward (5-7 days)	932	5	£4,660
Chest x-ray	60	8	£480
Ultrasounds	98.54	8	£788.32
CT Scan	184.61	2	£369.22
Drugs (during and after transplant)			£15,000
BCH inpatient PN			£4,046
Paediatric Intensive Care Unit	2,515	5	£12,575
Daily Blood Tests			£7,298
Central Line insertion & Removal			£2,764
Consultant charges			
Medical team			£15,000
Anaesthetic team			£6,000
Surgical team			£15,000
Histopathology team			£2,000
Radiology Team			£4,350
PICU Team			£1,250
Total liver transplant costs	**£144,640**		

6.5 Appendix 5. Parameters included in the probabilistic sensitivity analysis

Parameter	Base-case value	Distribution
Costs		
Expanded NBS screening including TYR1	£2.80	Gamma
Current NBS screening programme not including universal screening	£2.70	Gamma
Diagnostic protocol	£257	Gamma
Cost of liver transplant	£144,640	Gamma
Proportions screened		
MS/MS screening inclusive of TYR1 (day 5)	96.5%	Dirichlet
Cascade testing	0.001800%	
Symptomatic presentation (≤14 days)	0.000042%	
Elude screening	3.4981584%	
Test accuracy		
Specificity of MS/MS measurement of SUAC	0.999983	Beta (278923.1, 5.63)
Sensitivity of MS/MS measurement of phenylalanine and second-tier tyrosine	0.25	Beta (2.2, 6.7)
Specificity of MS/MS measurement of phenylalanine and second-tier tyrosine	0.999865	Beta (163848.09, 3.28)
Utility values		
False positive screen result	0.97	Beta (242.5, 7.5)
Liver disease	0.20	Beta (150, 600)
Liver transplant	0.67	Beta (87.6, 43.1)
Renal dysfunction	0.67	Beta (87.6, 43.1)
Learning difficulties	0.79	Beta (30.7, 8.2)
Neurological crises	0.84	Beta (75.9, 14.5)
Treatment without complications	0.90	Beta (121.1, 13.5)
Neurological crises and learning difficulties	0.82	Beta (51.66, 11.34)
Combination of long-term complications	0.30	Beta (5.6, 13)

6.6 Appendix 6. Scenario analyses considering optimistic estimates for developing long-term complications

In Appendix Table 2, the results are presented in terms of cost per QALY based on a hypothetical cohort of 100,000 newborns. This scenario analysis considers all of the optimistic estimates for long-term complications following pre-symptomatically detected and symptomatically detected TYR1. These results show that the expanded NBS screening strategy including TYR1 is more costly and expected to yield 3.3 more QALYs, with an ICER of approximately £109,700 per QALY gained. For the optimistic estimate, there is an increase in the 4- and 6-month transition probabilities for liver disease, likewise for liver transplant. There is a reduction in the expected mean costs for the current NBS screening approach without universal TYR1 screening as there are more TYR1 cases with liver disease and hence more liver transplants. Following a successful transplant, babies discontinue nitisinone treatment, which causes the decrease in treatment costs required; thus, leads to a reduction to the expected mean costs for the no universal screening strategy.

Appendix Table 2: Deterministic results for 100,000 newborns based on QALYs (optimistic)

Strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected mean QALY (per newborn)	Incremental QALY [a]	ICER (£)
No universal screening for TYR1	1,686,000	-	26.67399	-	-
Universal screening for TYR1	2,050,000	363,000	26.67402	3.3	109,711
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALYs, quality adjusted life-years					

6.7 Appendix 7. Scenario analysis considering pessimistic estimates for developing long-term complications

In Appendix Table 3, the results are presented in terms of cost per QALY based on a hypothetical cohort of 100,000 newborns. This scenario analysis considers all of the pessimistic estimates for developing long-term complications following TYR1 and its treatment. These results show that the proposed NBS screening strategy with added TYR1 screening is more costly and expected to yield 3.8 more QALYs, with an ICER of approximately £38,100 per QALY gained. Using the pessimistic transition probabilities estimates, less symptomatically detected TYR1 patients are developing liver disease and thus receiving liver transplant. Hence, more TYR1 patients in the current NBS screening approach (not including TYR1 screening) would remain on the costly nitisinone treatment. This is reflected in the higher expected mean costs in the no universal screening strategy resulting in lower incremental costs and therefore a lower ICER compared to the base-case results.

Appendix Table 3: Deterministic results for 100,000 newborns based on QALYs (pessimistic)

Strategy	Expected total costs (£)	Incremental costs (£)	Expected mean QALY (per newborn)	Incremental QALY	ICER (£) per QALY gained
No universal screening for TYR1	2,343,000	-	26.67396	-	-
Universal screening for TYR1	2,487,000	144,000	26.67399	3.8	38,137
ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life-years					

6.8 Appendix 8. Exploratory results

The results presented in this section are based on comparing the proposed, expanded NBS screening approach (including universal screening for TYR1) versus the current approach (no universal screening for TYR1 but universal NBS screening for PKU), by considering those babies who have been identified through screen detection only and excluding other modes for identifying babies with TYR1 (i.e. through cascade testing, symptomatic presentation or elude screening). It assumes that 100% of people would be screened as opposed to the 96.5% in the base-case (as shown in Table 7, Table 8 and Table 9). The assumed birth prevalence in this cohort was 1:100,000. The exploratory results are reported for the outcomes number of TYR1 cases correctly identified via screen-detection (Appendix Table 4), life-years gained (Appendix Table 5) and QALYs gained (Appendix Table 6) per 100,000 screened newborns. It can be seen in these exploratory analyses that the incremental results are similar to those presented in the base-case analysis. Any differences are a result of the additional costs and benefits for including other modes of detecting TYR1 (e.g. cascade testing).

Appendix Table 4: Deterministic results based on number of TYR1 cases detected per 100,000 screened newborns

Strategy	Expected total costs (£)	Incremental costs (£)	Screen-detected TYR1 cases	Incremental screen-detected TYR1 cases	ICER (£) per additional screen-detected TYR1 case
No universal screening for TYR1	274,000	-	0.25	-	-
Universal screening for TYR1	281,000	7,000	1.00	0.75	£9,500
ICER, incremental cost-effectiveness ratio Exact results have been obtained from TreeAge, but were rounded by the authors and presented					

Appendix Table 5 shows that, when all 100,000 newborns would be screened and there would be no TYR1 detection by other modes, the expanded screening approach including universal TYR1 screening is more costly and more effective than the no screening strategy and has an ICER of approximately £417,700 per LYG.

Appendix Table 5: Deterministic results based on life years gained per 100,000 screened newborns (100% of the hypothetical birth cohort)

Strategy	Expected mean costs (£)[a]	Incremental costs (£)[a]	Expected mean LY (per newborn)	Incremental LYG [a]	ICER (£) per LYG
No universal screening for TYR1	1,326,054	-	27.702642	-	-
Universal screening for TYR1	1,755,225	429,170	27.702652	1.00	417,710
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; LYG, life years gained Exact results have been obtained from TreeAge, but were rounded by the authors and presented					

In Appendix Table 6, the results are presented in terms of cost per QALY. These results show that, when all 100,000 newborns would be screened and there would be no TYR1 detection by other modes, the universal screening strategy for TYR1 is more costly and expected to yield 7.2 more QALYs, with an ICER of approximately £59,300 per QALY gained.

Appendix Table 6: Deterministic results based on QALYs gained per 100,000 screened newborns (100% of the hypothetical birth cohort)

Strategy	Expected total costs (£)[a]	Incremental costs (£)[a]	Expected mean QALY	Incremental QALY [a]	ICER (£) per QALY gained
No universal screening for TYR1	1,326,054	-	26.674027	-	-
Universal screening for TYR1	1,755,225	429,170	26.674100	7.2	59,306
[a] Values have been multiplied by 100,000 ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life-years Exact results have been obtained from TreeAge, but were rounded by the authors and presented					

6.9 Appendix 9. Comparison of previous model-based economic analyses with our economic model

This short note focuses on the comparison between the economic evidence identified in the systematic review,[38] the study by Cipriano and colleagues,[39] the ScHARR report from 2013 on the cost-effectiveness of expanding newborn blood spot screening for five conditions (not including TYR1)[40] and our current economic analysis. Here, we highlight and discuss the research questions, methods, model inputs, assumptions, analyses and results of the four economic analyses, which assessed the cost-effectiveness of expanding the current newborn blood spot programme to include screening for TYR1 or other conditions compared to current practice. Brief summaries on the conduct of each study can be found in Appendix Table 7.

Research question

All studies clearly stated their research question, which in general was to estimate the cost-effectiveness of expanding the current newborn bloodspot (NBS) programme to screen for inborn errors of metabolism (IEM) compared to current practice. Of the four analyses, two studies included screening for TYR1 among other IEM,[38; 39] one analysis evaluated NBS screening for five IEM not including TYR1[40], and one analysis focussed on identifying TYR1 alone.[41]

The intervention in the economic analysis published by the Institute of Health Economics (IHE) in 2016 and our current analyses was universal NBS screening for TYR1 using MS/MS measurement of succinylacetone. The economic model by Cipriano et al.[39] evaluated universal NBS screening using

MS/MS measurement of tyrosine,[39] and the ScHARR report from 2013 evaluated expanding the universal newborn blood spot screening with MS/MS for five conditions [maple syrup urine disease (MSUD), homocystinuria (HCU), isovaleric acidaemia (IVA), glutaric aciduria type 1 (GA1), long-chain hydroxyacyl-CoA dehydrogenase deficiency (LCHADD).

Each intervention was compared to the current methods used to identify babies with IEM. The three studies that included identifying TYR1 all used a no universal screening approach for TYR1 as comparator, which included incidental finding of TYR1 through universal newborn screening for phenylketonuria (PKU).[38; 39; 41] It can be seen that the research question set out in the ScHARR report[40] does not include screening for TYR1, and this should be borne in mind when comparing between these analyses.

Methods

A key feature in economic evaluation is the cost and benefit comparison between alternative interventions. The perspective refers to the scope of the evaluation undertaken, and has implications about the costs and benefits included within the analysis.[11] Therefore, it is essential to clearly state the perspective/viewpoint of the analysis. Three studies stated the viewpoint/perspective of their analysis. The current analysis was conducted from the NHS and PSS perspective, while Cipriano and colleagues undertook their analysis from the societal perspective, and the IHE from the payer's perspective. The current analysis included health and social care costs directly incurred by the health care provider for the screening programme, symptomatic detection of TYR1, and treatment of long-term complications, which is in line with the NICE guidelines.[29] The evaluation by Cipriano and colleagues took a broader societal perspective, which in theory should include costs that fall on the society such as productivity losses and other indirect costs such as travel or caring costs. However, it was not clear about the resource use and costs that had been included when this perspective was adopted. Undertaking an analysis from the societal perspective can be considered as being broad/inclusive, thus capturing costs and benefits other than those directly incurred if we adopt a healthcare viewpoint, for example. Excluding relevant costs and benefits in an economic analysis may lead to inefficient allocation of scarce resources.[42] However, careful consideration should be taken when identifying resource use information associated with productivity losses, and the methods used to quantify these indirect costs.[11]

All analyses concluded at a lifetime horizon, with costs and benefits being discounted. Discounting takes into account the impact of time on the value of costs incurred and benefits accrued over time. In the current analysis discounting was based on the 3.5% per annum on costs and benefits, Cipriano and colleagues used a 3% per annum discounting for costs and benefits, while IHE and ScHARR discounted at 5% and 1.5% per annum, respectively. It can be seen that each analysis used different discounting rates in their base-case, which implies that society places different weights on consuming

a product now rather than delaying that consumption in the future. Hence, the higher the discount rate, the higher the value of current consumption is compared to future consumption. Two analyses were undertaken within a UK setting but used different discount rates in their base-case analysis. However, it should be noted that in a scenario analysis the current analysis used a 1.5% discount rate for both costs and benefits. The remaining two analyses were undertaken within a Canadian setting but again used different discount rates, with Cipriano and colleagues using 3% and IHE using 5% per annum discount rate. Current Canadian recommendations for undertaking economic evaluation of health technologies suggests using a discount of 5% per annum for both costs and benefits, and presenting scenario analysis results based on discounting at 0% and 3%.[43]

All analyses used an economic model to compare the cost-effectiveness of universal screening versus no universal screening. Two analyses used a decision tree structure, whilst the other two used a decision tree with Markov nodes to show the lifetime experience of people with TYR1. Decision tree structures have some advantages over the Markov model by being easy to populate and analyse. However, decision trees may not be the most appropriate to handle events that reoccur and can become unwieldly with the addition of more pathways. Additionally, using a decision tree structure can create too much static in the model, by assuming that the probability of events occur at discrete time points in the model and this would have been better captured by using a Markov model. Using a model comprising a decision tree with Markov nodes may be more appropriate by capturing the short term costs and long term benefits of screening. The two analyses that included a Markov model used different cycle lengths. Our analysis uses 4-monthly cycles in the first year then 6-monthly cycles, and the IHE used 3-monthly cycles to reflect when people are followed-up with TYR1.

Appendix Table 7: Summary characteristics of the compared economic evaluations on expanding the newborn blood spot screening programme for TYR1 or other inborn errors of metabolism.

Study name	**Auguste et al[41]**	**Cipriano et al, 2007[39]**	**IHE report, 2016[38]**	**ScHARR report, 2013[40]**
Study title	Cost-effectiveness of newborn blood spot screening for Tyrosineamia type 1 using tandem mass spectrometry	The cost-effectiveness of expanding newborn screening for up to 21 inherited metabolic disorders using tandem mass spectrometry: Results from a decision-analytic model	Newborn blood spot screening for galactosemia, tyrosinemia type I, homocystinuria, sickle cell anemia, sickle cell/beta-thalassemia, sickle cell/hemoglobin C disease, and severe combined immunodeficiency	Expanded newborn screening for inborn errors of the metabolism – health economics
First author	P Auguste	L Cipriano	A Chuck	J Chilcott
Co-authors	J Geppert; S Taylor-Phillips; C Stinton; S Johnson; A Clarke; H Mistry	A Rupar; G Zaric	C Yan; A Waye; I Akpinar	A Bessey; A Pandor; S Paisley
Source of publication	Not published	Value in Health. 2007; 10:83-97	Institute of Health Economics, 2016	School of Health and Related Research, 2013
Language	English Language	English Language	English Language	English Language
Publication type	Report	Original article	Report	Report
Inclusion criteria/study eligibility/PICOS				
Population (and subgroups)	Newborn	Newborn	Newborn	Newborn
Intervention(s)	Universal newborn blood spot screening for TYR1 using MS/MS measurement of succinylacetone	Universal newborn blood spot screening for TYR1 alone with MS/MS measurement of tyrosine	Universal newborn blood spot screening for TYR1 with MS/MS measurement of succinylacetone	Universal newborn blood spot screening with MS/MS for 5 conditions: maple syrup urine disease (MSUD), homocystinuria (HCU), isovaleric acidaemia (IVA), glutaric aciduria type 1 (GA1), long-chain hydroxyacyl-CoA dehydrogenase deficiency (LCHADD)
Comparator(s)	No universal screening for TYR1 (current approach: NBS screening programme for 9 conditions including PKU,	No universal screening for TYR1 (existing Ontario NBS screening programme screened for PKU via Guthrie bacterial inhibition	No universal screening for TYR1 (existing Alberta NBS screening programme screened for 17	No universal screening (existing NBS screening programme tests babies for five

Study name	Auguste et al[41]	Cipriano et al, 2007[39]	IHE report, 2016[38]	ScHARR report, 2013[40]
	MS/MS technology already implemented, cascade testing of babies with previously affected siblings)	assay and for hypothyroidism, MS/MS technology not implemented yet)	conditions, 13 of which screened by MS/MS, including PKU)	disorders including PKU, MS/MS already implemented)
Outcome(s)	Life year gained; QALYs gained	Life-years gained	Life-years gained	QALYs gained
Study design	Model-based cost-effectiveness analysis	Cost-effectiveness analysis	Model-based cost-effectiveness analysis	Model-based cost-effectiveness analysis
Setting and location	England, UK	Tertiary academic hospital in Ontario, Canada	Alberta, Canada	Six screening centres in England, UK
Methods				
Study perspective	NHS and PSS	Societal	Payer	Not explicitly stated
Time horizon	Life-time horizon	Life-time horizon	Life-time horizon	Life-time horizon
Discount rate	3.5% per annum for both costs and benefits	3% per annum for both costs and benefits	5% per annum for both costs and benefits	1.5% per annum for both costs and benefits
Measurement of effectiveness	Number of cases correctly identified via screen detection; life years gained; QALYs	Number of cases (various conditions and diseases) correctly diagnosed	Diagnostic accuracy of MS/MS in screening for TYR1; Nitisinone treatment effectiveness	Number of cases detected; QALYs
Measurement and valuation of preference based outcomes	Utility values were based on a study by Tiwana et al; age-related utility values from the UK population norms	Not applicable	Not applicable	Utility values based upon EQ-5D and from literature; age-related utility values from the UK population norms
Resource use and costs	Resource use and costs associated with the index tests, diagnostic protocol, treatment with nitisinone, diet costs, costs associated with the treatment of babies and children with TYR1 such as inpatient stay and outpatient visits, contact with healthcare staff and costs for the treatment of long-term sequelae (liver disease, liver transplant, renal dysfunction, learning difficulties, neurological crises)	Resource use and costs associated with the screening test (equipment, staff expenses, reagents and consumables), average health care costs by age, hospitalisations, additional health care services, maintenance diagnostics, non-dietary treatments (haemodialysis, liver transplant, kidney transplant), dietary treatments, pharmaceutical treatments, educational and social services	Resource use and costs associated with hospitalisation and other health services (including dieticians and genetic counselling), physician services, medication, diet, laboratory services for implementation, screening test, confirmatory tests, and follow-up testing, software implementation. Costs were also included for the educational and social services for people with mental disability	Resource use and costs including the cost of diagnosis (contact with healthcare staff, including consultant paediatricians, specialist nurses and dieticians, and costs of blood tests), regular appointments with consultant paediatricians, specialist nurses and dieticians, costs of blood tests, cost of dietary supplements, cost of sequelae and intercurrent illness, health and social care costs in respect to neurocognitive outcomes.

Study name	Auguste et al[41]	Cipriano et al, 2007[39]	IHE report, 2016[38]	ScHARR report, 2013[40]
				There are no marginal costs associated with expanding the programmes. The programme costs include the preparation and distribution of pre-screening information, lab screening costs, maintaining the website and cost of making the referral.
Currency, price date and conversion	Costs are in 2017/18 prices in UK pounds sterling. Costs from other time periods were uprated using the UK Hospital and Community Health Services index from UCHSC 2018.	Costs in 2004 prices in Canadian dollars; Costs from other time periods were adjusted using the Canadian Health Care Price Index.	Costs in 2015 prices in Canadian dollars; Costs from other time periods and other countries were converted to CAN$ using the purchasing power parity.	Costs are in 2011/12 prices in UK pounds sterling.
Model type	Decision tree and Markov model with 4-monthly cycles in first year and then 6-monthly cycles	Decision tree	Markov Model with 3-monthly cycles	Decision tree
Assumptions	• Babies are screened on day 5. • National protocol for confirmation of TYR1 is 100% accurate. • Diagnostic results to confirm TYR1 are available soon after testing. • Babies commence treatment for TYR1 (nitisinone and diet) soon after diagnosis. • All screen-detected with a confirmed diagnosis do not show any signs or symptoms related to TYR1, but may develop long-term complications.	• All diseases are grouped into 3 levels of severity: (1) neonatal, classical, severe, or early onset of the disease; (2) later-onset, chronic, or milder forms; (3) mild variations that would not be detected or treated without screening. • Patients in the first 2 categories would eventually be diagnosed clinically. • Patients in the 3rd category have same life expectancy at birth as members of the general population. • All initially positive screening results are confirmed with a	Assumptions specific to TYR1 • 1% of screen positive people receive follow-up testing and genetic confirmation. • Expert group provided proportions for people with learning/ language difficulties and time elapsed before being diagnosed for people without screening and delayed detection.	No assumptions specific to TYR1 as this was not one of the conditions covered.

Study name	Auguste et al[41]	Cipriano et al, 2007[39]	IHE report, 2016[38]	ScHARR report, 2013[40]
	• Average dose of nitisinone treatment is 1mg per kg body weight per day. • People are 100% compliant with diet and nitisinone treatment.	2nd MS/MS analysis before the patient is contacted. • Positive results from MS/MS testing are confirmed with other technologies before a final diagnosis is made. • While waiting for confirmation, all patients receive treatment.		
Results				
Study parameters	Parameters included: sensitivity and specificity of index tests and confirmatory tests; birth prevalence; incidence of condition-specific sequelae and transition probabilities; resource use and costs associated with care; utility values; mortality rates	Parameters included: sensitivity/ specificity and costs associated with taking a sample, annual equipment costs and treatment costs	Parameters included: incidence of the condition, sensitivity and specificity of index tests and confirmatory tests; effectiveness of screening and early detection on the development of sequelae; resource use and costs associated with care; mortality rates	Parameters included: sensitivity and specificity of index tests; prevalence of inborn errors; resource use and costs associated with care; utility values; survival; mortality rates
Incremental costs and outcomes	When comparing universal screening with no screening: 1) the estimated ICER was £18,300 per additional screen-detected case correctly identified. 2) universal screening was expected to yield 0.55 more life-years, with an ICER of approx. £424,200 per LYG. 3) universal screening is expected to yield 3.6 more QALYs, with an ICER of approximately £58,800 per QALY gained.	Results showed that screening compared to no screening is expected to yield an additional 0.0000457 LY at a cost of CAN$ 14.14, which resulted in an ICER of CAN$ 309,400 per LYG.	Results for TYR1 screening showed that screening is expected to yield an incremental gain of 6.0×10^{-5} LY, costing an additional CAN$ 1.90, which equates to approximately CAN$ 31,700 per LYG.	No results specific to TYR1 as this was not one of the conditions covered.

Study name	Auguste et al[41]	Cipriano et al, 2007[39]	IHE report, 2016[38]	ScHARR report, 2013[40]
Characterising uncertainty	One-way sensitivity analyses included: reducing the annual discount rate to 1.5%; including cost of a commercial kit; using different age cut-offs of survival for children with untreated TYR1; analysis including disutility associated with overdiagnosis; and varying key model input parameters by ±50% of the base-case values used in the model to see the impact on LYGs such as the number of false negatives and false positives. Probabilistic sensitivity analysis and expected value of perfect information were also conducted.	One-way sensitivity analysis included: incidence; sensitivity and specificity; costs, inclusion of legal costs, life expectancy, discount rate (0-9%) and a utility was included for parents during uncertainty diagnosis period.	Probabilistic and one-way sensitivity included varying the incidence rates on the number of cases detected. PSA were undertaken but not reported. Scatterplots were presented but no cost-effectiveness acceptability curves.	Uncertainty was presented by an expected value off perfect information analysis for the 5 conditions.
Discussion				
Study findings	Universal screening for TYR1 is both more costly and more effective than no universal screening.	The average cost-effectiveness for PKU and 15 diseases is less than CAN$ 100,000 per LYG; the marginal cost of adding TYR1 is CAN$ 309,400 per LYG.	TYR1 screening alone compared to no TYR1 screening has shown to be cost-effective.	Screening for the 5 conditions (TYR1 was not included) are estimated to be potentially cost saving and result in increased quality of life compared to no screening.
Limitations	• Assumed utility values for screen-detected and clinically detected were the same. • Model long-term complications were obtained from the literature. In some studies, it was difficult to decipher the events that occurred in people who were screened-detected and those babies who were clinically diagnosed.	• Relied on costing information from a large tertiary academic hospital (London Health Sciences Centre [LHSC]). • Resource use and costs do not include physician costs, and a fee for service compensation method was assumed. • They did not consider the potential need for additional investments in infrastructure or	• Cost inputs reflect incremental costs and not absolute costs. • Clinical and epidemiological information were obtained from multiple sources and differ across population and ethnic groups. • Costs associated with infrastructure requirements, capital equipment, nursing and	• None noted in report.

Study name	Auguste et al[41]	Cipriano et al, 2007[39]	IHE report, 2016[38]	ScHARR report, 2013[40]
	• Additionally, assumptions were made about length of follow-up and we assumed that these events occurred at a constant rate over time. • Economic analysis did not take into account societal costs and effects.	human resource training outside of the MS/MS laboratory.	other staff training were not considered. • Potential psychosocial harms associated with false positive results were not included. • Authors assumed that costs of late hematopoietic stem cell transplantation (HSCT) was three times higher that early HSCT. • The cost attribution analysis is conducted from an overall perspective, which is not the same as information generated from a detailed local-level costing exercise.	
Generalisability	Some inputs were derived from unpublished data from the West Midlands region, so may not be representative to the UK	Generalisability may be limited by the resource use and unit costs and the societal perspective adopted. Furthermore, the number of births in this region and the incidence of inborn errors may be not be similar to other countries.	The authors suggested that *"the transferability of the evidence base to the Alberta setting is uncertain, as the value in terms of both health outcomes and costs are ultimately dependent on local epidemiology, clinical practice, system capacity, and costs."*	None noted in report.
Other				
Source of funding	Study was funded by the UK NSC.	Study was funded by the Natural Science and Engineering Research Council of Canada.	Study was funded by a financial contribution from Alberta Health through the Alberta Health Technologies Decision Process, the Alberta model for health technology assessment and policy analysis.	Not stated
Conflicts of interest	None declared	None declared	The authors claim no competing interest.	Not stated

Study name	Auguste et al[41]	Cipriano et al, 2007[39]	IHE report, 2016[38]	ScHARR report, 2013[40]
Comments	• Key inputs based on the best available evidence; however, some inputs were derived from unpublished data from the West Midlands region, so may not be representative to the UK. • It is assumed that people with TYR1 who do not develop liver disease have the same life expectancy as the general population. People who develop liver disease are at increased risk of death. • Same utility values used for people who are screen-detected and those who presented symptomatically. • Increased probability of developing liver disease in those who were symptomatically detected; hence more people are at risk of requiring a transplant. Following successful transplant, it was assumed that people would discontinue nitisinone treatment. • Model assumed that all babies requiring a transplant received the procedure.	• Unclear what is meant by 'early diagnosis'. • Model does not consider the 2% of babies who missed screening. • Unclear if social services costs are borne by the payer. • Utility value were included but results are in terms of LYG and not QALYs. • No probabilistic sensitivity analyses undertaken. • Setting was an academic teaching hospital, with a societal viewpoint. Resource use may not be generalizable because there may be higher levels of monitoring and adherence. • Specifics of 'clinical diagnosis' is not discussed in detail. • Authors have not presented results based on a payers' perspective.	• Probabilistic sensitivity analyses were undertaken but not reported. • Scatterplots were presented but no cost-effectiveness acceptability curves. • Analyses could have benefited from a tornado diagram to show the impact of varying key input parameters, as well as identifying the key drivers of the ICER.	• No analysis specific to TYR1 as this was not one of the conditions covered. • Model details were not comprehensive.

ICER, Incremental cost-effectiveness ratio; LYG, Life years gained; MS/MS, Tandem mass spectrometry; NHS, National Health Service; PKU, Phenylketonuria; PSS, Personal social services; QALYs, Quality adjusted life years; TYR1, Tyrosinaemia type 1

Model inputs

All models required clinical information, as well as resource use and costs information; all of which were obtained from different sources. Clinical information was required for test performance, incidence/prevalence of TYR1, as well as the incidence of long-term complications. The current analysis derived sensitivity and specificity of universal NBS screening and incidental finding of TYR1 cases using PKU testing from information obtained from our clinical expert. The information obtained represented TYR1 cases identified through PKU testing born in the West Midlands region from 1982-2014.

Over time, TYR1 cases were at risk of developing long-term complications, which included liver disease, kidney disease, learning/language difficulties, and a combination of conditions, as seen in the current analysis and the IHE report[38]. There were some differences noted. First, the current model further included risk of developing neurological crises, and a combination of neurological crises and learning/language difficulties. Second, the current analysis derived 4-month and 6-month transition probabilities [pre-symptomatically-detected (universal screening for TYR1, incidental finding on routine PKU screening or cascade testing due to affected sibling)] to show the speed at which TYR1 cases may develop these complications. However, the IHE reported the proportion of TYR1 cases who are likely to develop these complications over the model time horizon, presenting proportions for 'with screening and early detection', and 'without screening and delayed detection.' It should be noted that in parts the same published sources[13; 15] were used to derive this information; however, the interpretation of the information was different between analysts. The current analysis assumes that these long-term conditions occurred at a constant rate over time for the entire time horizon, while it is unclear on the rate of developing these events in the IHE report. Third, the analysis performed by the IHE[38] assumed that the mortality rate of 'screen and early detection' TYR1 cases is the same as the general population, with an increased rate of mortality in people with clinical detection of TYR1. In the current analysis, TYR1 cases identified through newborn screening and those presenting symptomatically had the same risk of mortality as the general population. However, people who developed liver disease (irrespective of how/when TYR1 was diagnosed) had an increased risk of mortality.

Resource use and cost

Each economic analysis clearly outlined the resource use and costs considered in their respective economic evaluation. Quantifying resource use and valuing these resources are key to any economic evaluation, which should be in line with the perspective/viewpoint of the analysis. We note the similarities between analyses as resource use and costs associated with screening and diagnosis, management and treatment of TYR1, and treatment of long-term complications were all considered. As expected, there were differences in the resource use (e.g. number of hospitalisations, length of

appointments, pharmaceutical treatments including nitsinone treatment), and thus costs included in these analyses. Of note, there were additional pharmaceutical treatment (Arginine, cysteine, and sodium-phenylbutyrate) costs considered by Cipriano and colleagues that were not considered in the current analysis. Also, the additional resource use and costs for education and social services (standard education, additional and advanced class support, and living in an institutional or assisted living facility) were not considered in the current analysis. If relevant/applicable, excluding these costs from the current analysis could potentially underestimate the costs associated with learning difficulties, for example. Based on the current results, a high proportion of people who had a true positive result developed learning difficulties. Another difference noted is lifetime costs included Cipriano and colleague's analysis, while the current analysis' resource use and costs are based on the cycle length.

There were similarities and differences noted between studies, which could potentially lead to over/underestimating resource use and costs in these analyses.

Assumptions

All analyses made assumptions to have a workable/executable model. Given that identifying TYR1 was not a condition of interest in the ScHARR report, the assumptions made were related to other conditions of interest. Appendix Table 7 presents the assumptions made in each of the four compared economic analyses. Of note, there were clear differences in the assumptions made in each analysis. First, Cipriano and colleagues grouped conditions by severity levels, with the third group (mild variations that would not be detected or treated without screening) assuming to have the same background mortality as the general population. In the current analysis, we have not grouped by severity levels. However, we have assumed that in people without liver disease there is no increased risk of mortality for people living with TYR1 compared to people without TYR1. Second, Cipriano and colleagues included additional costs for babies who received a second MS/MS following an initial positive test result, which was not included in the current analysis or the analysis performed by the IHE. Re-testing (with the index test) babies with an initial positive result aims to reduce the number of false positive test results. However, the authors have not elaborated on the clinical pathway if the second result was negative, i.e. is there an underlying assumption that these babies are all true negative. Third, other assumptions were made in the current analysis that were not made in the other three analyses.

Summary

Two economic analyses compared universal screening using SUAC as a marker versus no universal screening to identify TYR1 in newborns, one economic evaluation assessed the cost effectiveness of universal TYR1 screening using tyrosine as marker, while one analysis focussed on the neonatal

screening for inborn errors of metabolism, excluding TYR1. There were noticeable differences in the illustrative model structures, which lead to some differences in the model inputs required and assumptions made, and this may have potentially lead to discrepancies in the results. One common result across analyses was the minimal incremental life-years gained between strategies, which shows some agreement between studies. However, the results in terms of cost per LYG showed that there was agreement between our analyses and the analyses by Cipriano et al.,[39; 41] with high costs (CAN$ 309,400) for each additional LYG. Conversely, the IHE results showed that the costs for each additional LYG was low (CAN$ 31,700) and this could have been a result of the analysis being undertaken from the payer's perspective and/or the use of a 5% annual discount rate on costs and benefits.