NIHR | National Institute for Health and Care Research

Journals Library

Check for updates

Health Technology Assessment

Volume 29 • Issue 65 • December 2025

ISSN 2046-4924

Evaluating whole genome sequencing for rare diseases in newborn screening: evidence synthesis from a series of systematic reviews

Karoline Freeman, Jacqueline Dinnes, Bethany Shinkins, Corinna Clark, Inès Kander, Katie Scandrett, Shivashri Chockalingam, Aziza Osman, Naila Dracup, Rachel Court, Furqan Butt, Cristina Visintin, James R Bonham, David Elliman, Graham Shortland, Anne Mackie, Zosia Miedzybrodzka, Sian M Morgan, Felicity Boardman, Yemisi Takwoingi and Sian Taylor-Phillips

DOI 10.3310/DJRF1124

NIHR | National Institute for Health and Care Research

Health Technology Assessment

Check for updates

Extended Research Article

Evaluating whole genome sequencing for rare diseases in newborn screening: evidence synthesis from a series of systematic reviews

Karoline Freeman[1*] Jacqueline Dinnes[2] Bethany Shinkins[1] Corinna Clark[1] Inès Kander[1] Katie Scandrett[2] Shivashri Chockalingam[1] Aziza Osman[1] Naila Dracup[1] Rachel Court[1] Furqan Butt[1] Cristina Visintin[3] James R Bonham[4] David Elliman[5] Graham Shortland[3] Anne Mackie[3] Zosia Miedzybrodzka[6] Sian M Morgan[7] Felicity Boardman[1] Yemisi Takwoingi[2] and Sian Taylor-Phillips[1]

[1]Warwick Medical School, University of Warwick, Coventry, UK
[2]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[3]UK National Screening Committee, London, UK
[4]Sheffield Children's NHS Foundation Trust, Sheffield, UK
[5]Great Ormond Street Hospital, London, UK
[6]Institute of Medical Sciences, University of Aberdeen, Aberdeen, UK
[7]All Wales Genetics Laboratory, Institute of Medical Genetics, University Hospital of Wales, Cardiff, UK

[*]**Corresponding author** K.Freeman@warwick.ac.uk

Published December 2025
DOI: 10.3310/DJRF1124

This report should be referenced as follows:

Freeman K, Dinnes J, Shinkins B, Clark C, Kander I, Scandrett K, *et al*. Evaluating whole genome sequencing for rare diseases in newborn screening: evidence synthesis from a series of systematic reviews. *Health Technol Assess* 2025;**29**(65). https://doi.org/10.3310/DJRF1124

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 4

A list of Journals Library editors can be found on the NIHR Journals Library website

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 4 and is ranked 30th (out of 174 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal

Manuscripts are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This article

The research reported in this issue of the journal was commissioned and funded by the Evidence Synthesis Programme on behalf of NICE as award number NIHR159928. The protocol was agreed in June 2023. The draft manuscript began editorial review in July 2024 and was accepted for publication in February 2025. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Abstract

Background: Newborn screening using whole genome sequencing is being evaluated in numerous projects across the world, including Genomics England Limited's Generation Study. It presents considerable challenges for policy advisors, not least, given the logistics of simultaneously evaluating the evidence for the suggested 200 rare genetic conditions. The 'genotype-first' approach has the potential for harms through overdiagnosis, and benefits are uncertain.

Objective: To assess different approaches to evaluating whole genome sequencing for newborn screening to inform the development of a robust method of evaluation for informing policy decisions.

Methods: We approached the objective with systematic review methods for a sample of five conditions (considering gene penetrance, expressivity, accuracy and effectiveness of whole genome sequencing and effect of earlier treatment) (search inception to November 2023), evaluated the National Institutes of Health [US] Clinical Genome Resource (ClinGen) as an alternative evidence source for the five conditions and we compared this to a review of genomic studies of newborn screening cohorts reporting penetrance for pathogenic variants of any paediatric condition (search inception to February 2024). We undertook a methodological review of economic evaluations of whole genome sequencing/whole exome sequencing (search inception to January 2024) and explored public views on evaluating whole genome sequencing.

Data sources: MEDLINE (Ovid), EMBASE (Ovid), Web of Science, Science Citation Index (via Clarivate), the Cochrane Library (via Wiley), cost-effectiveness analysis registry and American Economic Association electronic bibliography.

Actionability reports and scores from the Clinical Genome Resource website (downloaded 30 April 2024).

Results: The traditional review approach identified 268 studies reporting the genetic spectrum of individuals with the five conditions or benefits of earlier, symptomatic treatment. No evidence on the penetrance and expressivity or the accuracy or effectiveness of whole genome sequencing in newborns was identified. A review of 200 conditions would take a team of five reviewers 23 years to complete. Clinical Genome Resource reviews were available for four or five conditions. All four 'actionability' ratings disagreed with the findings of our reviews. Our review of 14 genomic studies of newborn screening cohorts found insufficient information to allow individual highly penetrant pathogenic variants for any condition to be identified. None of the 86 economic evaluations of whole genome sequencing or whole exome sequencing were set in a screening context. Some micro-costing studies are available that could help understand the resource use and costs associated with whole genome sequencing. Following a series of patient and public involvement meetings, attendees appreciated the uncertainties of whole genome sequencing. A wider stakeholder perspective is needed to inform policy decisions.

Limitations: Although we only examined five conditions in depth, the consistency in lack of data suggests that our conclusions are robust.

Conclusions: The systematic review approach for evaluating whole genome sequencing of newborns identified a paucity of high-quality evidence. Extending the review to all 200 conditions is not feasible. Currently, the use of existing genome resources and review of genomic studies of newborn screening cohorts are not viable alternatives. The cost-effectiveness of whole genome sequencing in a newborn screening context is unknown.

Future work: Large-scale collaborative research is required to evaluate the short- and long-term harms, benefits and economic implications of whole genome sequencing for screening newborns. We propose a staged approach to evaluation, considering only conditions with pathogenic variants with high penetrance to minimise harm from overdiagnosis.

Study registration: This study is registered as PROSPERO CRD42023475529.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme (NIHR award ref: NIHR159928) and is published in full in *Health Technology Assessment*; Vol. 29, No. 65. See the NIHR Funding and Awards website for further award information.

Contents

List of tables

List of figures

List of supplementary material

Report Supplementary Material 1 Characteristics of the five conditions under review

Report Supplementary Material 2 Excluded studies

Report Supplementary Material 3 Included studies for Q2 (prevalence of genetic variants in those with biochemical or biochemical and clinical features of each condition)

Report Supplementary Material 4 Data extraction of 14 studies included in the review of genomic studies of newborn cohorts reporting penetrance for pathogenic variants

Report Supplementary Material 5 Data extraction of 86 studies included in the review of cost-effectiveness evaluations of WGS and WES

Supplementary material can be found on the NIHR Journals Library report page (https://doi.org/10.3310/DJRF1124).

Supplementary material has been provided by the authors to support the article and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer-reviewed.

The supplementary materials (which include but are not limited to related publications, patient information leaflets and questionnaires) are provided to support and contextualise the publication. Every effort has been made to obtain the necessary permissions for reproduction, to credit original sources appropriately, and to respect copyright requirements. However, despite our diligence, we acknowledge the possibility of unintentional omissions or errors and we welcome notifications of any concerns regarding copyright or permissions.

Glossary

Biallelic Referring to both alleles of a gene.

Clinical actionability Level of evidence about pathogenicity, penetrance and expressivity of a genetic variant and the extent to which interventions can be used to mitigate the effect of the disease.

Clinical Genome Resource (ClinGen) Resource funded by the National Institutes of Health that centralises information about the clinical relevance of genes and variants.

Digenic Referring to two genes.

Dosage sensitivity Refers to the impact of changes in gene dosage (the number of copies of a gene) on an organism's phenotype.

Exon Coding sequence of DNA within a gene that is retained during the ribonucleic acid splicing process before translation into protein.

Expressivity Degree to which a trait/condition is expressed in individuals with a particular genetic variant.

Founder effect Phenomenon where a small group of individuals establishes a new population, leading to reduced genetic diversity compared to the original population. The presence of an allele at an unusually high frequency in an isolated population.

Gene-disease validity Level of evidence supporting the association between a specific gene and a disease/phenotype.

Gene dosage Number of copies of a particular gene present in a cell/organism and hence related to the amount of gene product.

Generation Study United Kingdom study run by Genomics England Limited in partnership with the National Health Service to sequence the genome of newborn babies in the United Kingdom and screen for 200 rare genetic conditions.

Genomics England Limited Company set up by the United Kingdom Department of Health and Social care to run the 100,000 Genomes Project. It is now overseeing the Generation Study.

Germline variant A variant within germ cells that can be passed on to offspring.

Intron Non-coding sequence of DNA within a gene that is removed during the ribonucleic acid splicing process before translation into protein.

Monoallelic Referring to a single allele of a gene.

Mosaicism Refers to the presence of multiple populations of cells with different genotypes in a single individual due to variants occurring in some cells during development.

Pathogenicity (variant pathogenicity) A variant is classified as pathogenic if evidence confirms that it causes disease based on variant interpretation. Information for variant interpretation and variant prioritisation is generally based on family, clinical and case control studies, which are enriched for etiological cofactors, which means penetrance in these cohorts is overestimated. Pathogenicity annotations used in the screening context have got implications for specificity, as 'known' pathogenic variants have lower penetrance in population-based studies. Degree to which a variant impacts the function of a gene; > 99% probability of pathogenicity is defined as pathogenic variant and 90–99% probability of pathogenicity is defined as likely pathogenic variant.

Penetrance Proportion of individuals with a given genetic variant, who display the traits/condition.

Private variants Gene variants that is specific to an individual or family and not commonly found in the wider population.

Proband Individual in a family who is first identified as having a particular genetic condition.

Sporadic variants Gene variants that arise spontaneously in an individual without being inherited from a parent.

Variant Change/alteration in the DNA sequence that may affect how a gene functions. Variants can be benign, pathogenic or of unknown significance.

Variant annotation Variant annotation is the process of assigning functional information to a DNA alteration such as the effect on protein structure.

Variant interpretation Variant interpretation is the process of drawing direct links between individual variants and a disease phenotype for decisions on reporting. This requires expert interpretation and literature review and consideration of context [other genetic factors, environmental factors, ancestry, sex and type of phenotype (symptomatic disease) under consideration]. Guidelines for variant interpretation, for instance by American College of Medical Genetics and Genomics, are available to standardise the interpretation of variant pathogenicity and categorisation into pathogenic, likely pathogenic, uncertain, likely benign and benign.

Variant of unknown significance Variant for which there is insufficient information to determine its impact on disease risk/health.

Variant prioritisation Variant prioritisation is the process of filtering variants identified through sequencing using bioinformatic tools to identify variants most likely linked to a disease phenotype. More advanced tools include information on predictions about effect on protein structure, conservation (conserved sequences across all vertebrates), constraint (gene regions intolerant to loss of function variants), variant frequency, mode of inheritance and gene–disease associations.

Whole exome sequencing Technique for sequencing all the protein-coding regions of genes (exons) in a genome.

Whole genome sequencing Technique for sequencing the entire DNA (genome) of an individual.

List of abbreviations

α-AASA	alpha aminoadipic semialdehyde
CI	confidence interval
CEA	cost-effectiveness analysis
ClinGen	Clinical Genome Resource
ClinVar	clinical variation
CNV	copy number variant
DBS	dried blood spot
DNA	deoxyribonucleic acid
EconLit	American Economic Association electronic bibliography
FAOD	fatty acid oxidation disorders
FGF-23	fibroblast growth factor 23
FH	family history
fHLH	familial haemophagocytic lymphohistiocytosis
G6PDD	glucose-6-phosphate dehydrogenase deficiency
GEL	Genomics England Limited
GS	genome sequencing
HLH	haemophagocytic lymphohistiocytosis
HPLC	high-performance liquid chromatography
hRB	heritable retinoblastoma
HRQoL	health-related quality of life
HSCT	haematopoietic stem cell transplant
HTA	Health Technology Assessment
IEM	inborn errors of metabolism
IMD	inherited metabolic disease
IQ	intelligence quotient
LRT	lysine reduction therapy
MCADD	medium-chain acyl-CoA dehydrogenase deficiency
MeSH	medical subject heading
MLPA	multiplex ligation-dependent probe amplification
NBS	newborn blood spot
NGS	next-generation sequencing
NHS EED	NHS Economic Evaluation Database
NICU	neonatal intensive care unit
NIH	National Institutes of Health [US]
NSC	National Screening Committee
PCR	polymerase chain reaction
PDE	pyridoxine-dependent epilepsy
PHEX	phosphate-regulating neutral endopeptidase
PICO	population, intervention, control/comparison, outcome
Pi/D	oral phosphate and calcitriol (active vitamin D)
PKU	phenylketonuria
PLPB	pyridoxal phosphate-binding protein
PPIE	patient and public involvement and engagement
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QALY	quality-adjusted life-year
RB	retinoblastoma
RCT	randomised controlled trial
REC	retinoblastoma variant effect class
ROBIS	Risk of Bias in Systematic Reviews
RQ	research question
SCI	Science Citation Index
SD	standard deviation
SGP	Scottish Genomes Partnership
SMA	spinal muscular atrophy
SNV	single nucleotide variant
VUS	variant of unknown significance
WES	whole exome sequencing
WGS	whole genome sequencing
WoS	Web of Science
XLHR	X-linked hypophosphataemic rickets

Plain language summary

The government has funded a research project that tests newborns' entire DNA sequence (whole genome sequencing) for over 200 rare conditions, before babies appear to be ill. This could be a new screening programme. Policy-makers need a way to assess, for all 200 conditions, whether it improves lives and is a good value for the National Health Service.

Whole genome sequencing detects variations in our DNA. We all have many variations. They make us unique and only a few cause harm. It is difficult to predict true disease from a genetic finding. A baby may have a variant that looks harmful, but that does not mean they will develop the associated disease. Using whole genome sequencing to detect harmful variants could provide health benefits. Whole genome sequencing may also cause uncertainty and anxiety in parents and perhaps harm babies by giving them unnecessary treatment.

We assessed:

1. examining evidence from research studies to assess whole genome sequencing separately for 200 conditions by sampling 5
2. measuring uncertainty when predicting any condition from genetic findings
3. using an online resource with information on genetic conditions as evidence.

We also looked for, but could not find, studies that weighed up the cost and health implications of implementing whole genome sequencing in screening.

We met five times with parents, expectant parents and charity representatives to explore the challenges of whole genome sequencing. Challenges include communication, consent, data security, privacy and uncertainty.

We found, that:

- there is insufficient evidence for the five conditions to inform policy
- it would take 23 years to assess 200 conditions
- studies using whole genome sequencing only report the number of variants detected, not how well these predict disease
- the online resource does not have good quality evidence for screening policy.

None of the three approaches is currently useful for evaluating the benefits and harms of whole genome sequencing. We first need to collect more evidence.

Scientific summary

Background

In 2021, Genomics England Limited (GEL) launched its Generation Study of whole genome sequencing (WGS) to screen for over 200 rare diseases in 100,000 newborns to explore its potential for an expanded UK newborn screening programme. This presents a number of new challenges for policy advisors.

Multiplex testing has been available for years. Tandem mass spectrometry, currently used in the newborn blood spot (NBS) screening programme, can detect dozens of statistical abnormalities in the blood spot. Each condition included on the NBS programme and any potential candidate conditions have been assessed in turn. However, there is pressure to assess all the conditions that might be found with WGS at once.

The genotype-first approach has the potential for harms, and in some cases, may be more uncertain than more traditional methods. Not everyone with a pathogenic variant will develop symptomatic disease (incomplete penetrance), and symptoms caused by the same genetic variant can vary in severity among affected people (expressivity).

A cost-effectiveness analysis of WGS in newborn screening will be needed for a policy decision, but screening for potentially hundreds of conditions with a single test will require a different methodological approach than one that focuses on a single condition.

Finally, the use of WGS for newborn screening presents several ethical challenges. The majority are common to all screening programmes (anxiety, informed choice and penetrance), but there are some that are more pressing or likely in this programme. For example, some of the genetic variants might only be of significance later in life, and there are implications for the relatives if a variant is found and there is considerable commercial interest in secondary uses of the data which will not benefit participants directly.

We, therefore, aimed to (1) assess different evidence sources and approaches to evidence synthesis, (2) review methods for evaluating cost-effectiveness and (3) collate views of the public on the main challenges of WGS to inform an approach to assessing WGS for newborn screening in the future.

Objectives

1. To undertake a series of five systematic reviews covering a stratified (by burden and cost of the intervention) random sample of rare diseases to establish the evidence base per condition and to provide a reference case for comparison with alternative review approaches. The reviews addressed six questions mapped to the UK National Screening Committee (NSC) criteria on penetrance and expressivity, the proportion of children with disease who carry gene variants, test accuracy, effectiveness of earlier treatment, effectiveness and benefits and harms of WGS.
2. To explore the utility of the existing online resource Clinical Genome Resource (ClinGen) to provide evidence on the actionability of rare paediatric genetic diseases in order to evaluate it as a potential evidence source for the UK NSC.
3. To undertake a review of genomic studies of newborn screening cohorts reporting penetrance of pathogenic variants to explore the feasibility of identifying highly penetrant pathogenic variants that could be considered for a screening programme.
4. To produce a methodological overview of existing published economic evaluations and costing studies of WGS or whole exome sequencing (WES).
5. To explore patient and public views about the introduction of WGS for newborn screening.

Methods

Review of five conditions

A stratified random sample of five conditions was reviewed. Stratification was based on a range of scenarios that might reasonably have an impact on the UK NSC's recommendations relating to WGS for newborn screening. The five conditions were:

a. pyridoxine-dependent epilepsy (PDE)
b. heritable retinoblastoma (hRB)
c. X-linked hypophosphataemic rickets (XLHR)
d. familial haemophagocytic lymphohistiocytosis
e. medium-chain acyl-CoA dehydrogenase deficiency (MCADD).

Data sources

MEDLINE (via Ovid), EMBASE (via Ovid), Science Citation Index (SCI) (via Clarivate) and the Cochrane Library (via Wiley) from inception to November 2023.

Study eligibility criteria were defined for each review question and included the following:

Population

This included the below:studies of newborn screening cohorts, or
studies of newborns and children (≤ 18 years) with clinical, or biochemical and clinical features of the five conditions.

Exposure/intervention

This included the below:presence of pathogenic variants in the relevant gene(s) detected by sequencing
eligible interventions relevant to the screening context, with 'early' intervention defined separately for each condition.

Target condition

This included the below:clinically or clinically and biochemically defined disease.

Outcomes

These included the below:
measures of disease-specific morbidity and mortality
any health-related health outcomes that could be measured across conditions
any harms or other benefits from WGS.

We produced a narrative synthesis of studies.

Exploring Clinical Genome Resource as an evidence source

We searched the ClinGen database on 19 February 2024 for each of the genes included in the review of five conditions, and we tabulated actionability scores and evidence levels comparatively against the evaluation from GEL and against our assessment using the UK NSC criteria.

Review of genomic studies of newborn screening cohorts reporting penetrance of pathogenic variants

Data source: MEDLINE (via Ovid), EMBASE (via Ovid), SCI (via Clarivate) and the Cochrane Library (via Wiley) from inception to January 2024.

Study eligibility criteria: Studies of unselected newborns sequenced for any rare condition with outcomes of penetrance or an approximation.

We produced a narrative synthesis of our findings.

Methods for review of cost-effectiveness evaluations of whole genome sequencing and whole exome sequencing

Data source: MEDLINE (Ovid), EMBASE (Ovid), cost-effectiveness analysis registry, Web of Science (WoS) and American Economic Association electronic bibliography (EconLit) from inception to February 2024 and hand searches of identified systematic reviews.

Study eligibility criteria: Economic evaluations, clinical trials and Health Technology Assessments reporting costs of WGS or WES in human health care.

Evidence synthesis: a general narrative synthesis of the methodological approaches adopted will be reported. We will also focus on two specific methodological questions:

1. How were the costs associated with WGS and WES estimated?
2. What comparators were included in each study?

Consideration of the public voice in the evaluation of whole genome sequencing

Eight members of the public attended five 2-hour virtual meetings between 15 January and 21 May 2024. Meetings were deliberative and pre-defined topics related to WGS were explored: harms and benefits, genetic uncertainties, systematic review findings and the role of patient and public involvement and engagement (PPIE) in future reviews.

Themes of participants' views were narratively synthesised.

Results

Review of five conditions

Extrapolating the traditional approach to 200 conditions

We screened 19,689 titles and abstracts for the 5 traditional reviews, of which 1348 were selected for full-text assessment (range 55–449 per condition). A total of 268 studies were eligible for inclusion across the 5 reviews (range 31–78). No evidence was identified for the four review questions that required studies to be conducted in newborns. Overall, the five traditional reviews yielded very little of the evidence required by the UK NSC. Considering the time taken to identify and select the evidence, and extrapolating to a review of 200 conditions, we could expect as many as 787,560 unique records, 53,920 full texts to be screened and 8840 studies to be reviewed and synthesised, which is estimated to take a team of 5 reviewers 23 years to complete.

Evidence on the genetic spectrum in children with disease

Two hundred and sixty studies (range 26–73) were included, which reported the genetic spectrum in children with the five conditions. The proportion of children testing positive on sequencing varied for each condition by:

1. definition of disease from broadest (symptomatically defined) to narrowest (genetically defined) category
2. the testing strategy (type of test, number of genes and extent of gene sequencing and additional genetic testing to supplement sequencing)
3. the extent of 'pre-screening' using biochemical and clinical markers.

At variant level, studies provided data on the proportion of novel variants and type of variants but very little information on the severity of disease for specific variants. The large number of novel variants present a challenge to sequencing newborns as their pathogenicity is difficult to ascertain.

Evidence on early versus late treatment

Twenty-two studies (range 1–9) reported the outcomes of early versus late treatment. No study was designed to compare treatment effectiveness in screen-detected versus symptomatically detected children. Definitions of early and late treatment varied and relied on study authors' definitions. The evidence base pointed towards some benefit in

early treatment. However, the quality and volume of the evidence were low because of the definition of early versus late, the type of study, the number of participants and the number of studies available. Therefore, there was insufficient evidence to clearly judge the effect.

Learnings from the five traditional reviews

A single approach to reviewing five conditions was not feasible due to differences in the conditions' characteristics, treatment and aim of screening. For instance, each search strategy was developed individually. Disease-specific categories were needed to organise studies by the population subtype because of differences in the availability of biochemical tests, the number of disease groups with overlapping symptoms and whether conditions could only be defined genetically. The definition of early versus late depended on whether the relevant intervention was preventative, curative or for symptom management, whether an early intervention phase could be defined and whether conditions were progressive or presented following a trigger. A review of 200 conditions would require 200 individual reviews; however, some learning may be transferable between reviews of similar conditions, which we could not explore with the 5 conditions.

Exploring Clinical Genome Resource as an evidence source for the United Kingdom National Screening Committee

Four of the five conditions reviewed (PDE, hRB, XLHR and MCADD) had a paediatric actionability report available on ClinGen in March 2024. However, no information on variant classification in terms of pathogenicity was available for any of the genes.

Comparison of our assessment of the five conditions using the UK NSC criteria with the ClinGen scores of actionability alongside GEL's decisions to include genes on their gene list was complicated. The overall decision on actionability differed for four/four conditions between ClinGen and our assessment using the UK NSC criteria and for five/five conditions between GEL's assessment and our assessment.

It would be inappropriate for the UK NSC to base decisions on potential screening programs on the actionability reported in ClinGen without further assessment.

Review of genomic studies of newborn screening cohorts reporting penetrance of pathogenic variants

Fourteen studies reported experiences with gene sequencing in newborns, of which five provided information that approximated penetrance by reporting some clinical follow-up after a sequence positive test. The number of included genes ranged from 134 to 954 across the five studies and the number of newborns sequenced ranged from 127 to 29,989. Gene selection and variant interpretation varied across studies.

The proportion of babies designated as screen positive from these studies ranged from 1.7% to 9.7%. Half of the positive screens were for conditions not included on conventional newborn screening panels in the study countries (USA and China). However, the clinical significance of a large majority of 'positive screen' results on sequencing (83.3–100%) was unknown, so we do not know if detecting and reporting these was overdiagnosis of clinically insignificant disease, misdiagnosis of disease or early detection of late-onset disease.

Follow-up ranged from 2 months to > 5 years. Penetrance was approximated by the number of confirmed cases after clinical follow-up. For all genes considered together, penetrance ranged from 1.6% after follow-up of 24–48 months to 58.4% after up to 3-year follow-up. The studies did not provide sufficient evidence to understand penetrance for any genetic variant because:

- The number of infants with a specific condition displaying a range of variants was too low.
- Infants with confirmed genetic disease received management that precludes the estimation of penetrance and expressivity for cases without symptomatic confirmation of disease.
- Clinical follow-up was not long enough to include all childhood-onset cases.

Overall, there was little agreement on what genes should be considered in newborn screening, no indication of how to interpret discordant results from NBS programmes and genetic screening and evidence of overdiagnosis. The studies

demonstrated unequivocally that if WGS was to be introduced without further research, it would cause significant problems.

Review of cost-effectiveness evaluations of whole genome sequencing and whole exome sequencing

Eighty-six studies were included in the review. None of them focused on the use of WGS or WES in a screening context. Under half of studies (*n* = 39/86, 45%) were full economic evaluations, of which only 10 were cost–utility studies, that is, studies which estimated the cost per quality-adjusted life-year (QALY). Most evaluations focused only on the costs and outcomes associated with the diagnostic pathway, avoiding the complexity of capturing the impact of a diagnosis on patient management. Two-thirds of the included studies reported a costing perspective; of which one-third [29/86 (36%)] adopted a broad healthcare system perspective, 15 a specific health system perspective, 8 a patient perspective and 5 a societal perspective. Only seven studies (8%) adopted a lifetime horizon. Of the studies that included a comparator (78/86, 91%), 44 (56%) explicitly stated that the comparator was the current standard of care testing, consisting of a broad range of tests. Different assumptions were made in terms of which tests would no longer be needed following the incorporation of WES or WGS in the diagnostic pathway.

Consideration of the public voice in the evaluation of whole genome sequencing

The group largely supported WGS for newborn screening. As meetings progressed and the complexities were explored, however, views became more nuanced; for example, one participant mentioned that they now were 'sitting on the fence a bit'.

Participants identified a wide range of benefits and harms and broadly felt that the benefits outweigh the harms. Key harms they were concerned about ranged from personal (anxiety) to societal (strain on health services). Key benefits included saving lives and avoidance of a diagnostic odyssey.

The process would have benefited from having more time to develop and discuss ideas. For a future review, it would be beneficial to increase the diversity of viewpoints.

Conclusions

A traditional systematic review approach to evaluating WGS of newborns is unfeasible, and we were unable to identify an acceptable alternative way to evaluate WGS for newborn screening in a single mechanism. Cost-effectiveness evidence for WGS has only focused on symptomatic populations to date. Our review highlights the main evidence gaps and informs the direction of future research efforts.

We propose research undertaken in large joined-up collaborations to produce the evidence that is needed for policy advisors before an evaluation of WGS is feasible. This may include a co-ordinated international approach to collecting penetrance data for pathogenic variants with a clear treatment plan. This could be followed by a staged approach of evaluation considering only those of the 200 conditions for screening that have pathogenic variants with very high penetrance.

Study registration

This study is registered as PROSPERO CRD42023475529.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme (NIHR award ref: NIHR159928) and is published in full in *Health Technology Assessment*; Vol. 29, No. 65. See the NIHR Funding and Awards website for further award information.

Chapter 1 Introduction

Rare diseases are a group of disorders that are characterised by their relatively low prevalence in the population and are typically defined in the UK as affecting < 1 in 2000 individuals.[1] There are approximately 7000 rare disorders, which, when combined, affect around 6% of the population in the Western world.[2] About 80% of rare disorders are thought to have a genetic cause.[2] If a disorder is caused by a variation in a single gene, the disorder is termed as a monogenic disorder. For only half of the estimated 5000 monogenic disorders, the underlying genes are known.[3] The identification of disease-causing variation in the known genes is challenging. Every human individual is believed to have up to 5 million variants compared to the reference genome.[4] This variation is responsible for the human diversity, and only a small fraction of variants affects human health.

Variants in genes associated with human health can be benign or pathogenic or be of unknown significance, meaning that the link between genetic variation and clinical phenotype is unknown. For pathogenic variants, the link between genotype and phenotype is known but can be incomplete such that only a proportion of individuals with the pathogenic variation may develop symptoms (incomplete penetrance). These symptoms can further vary in severity even among affected family members (expressivity). This makes it difficult to predict disease even if the underlying genetic cause is well characterised.

Disorders caused by genetic variants can present at any time, from birth [e.g. phenylketonuria (PKU)] to much later in life (e.g. Huntington's disease). It is estimated that there are about 600 childhood-onset conditions for which there is a potential intervention,[4,5] and newborn screening aims to identify the disorders that benefit from pre-symptomatic detection. Early diagnosis enables surveillance and early intervention when available, which can significantly improve outcomes, and it is particularly important in conditions with rapid progression or that cause irreversible damage. Current approaches to screening for rare disorders as part of the UK national newborn blood spot (NBS) screening programme have been successful in both their capacity for detection of infants with rare conditions and in the wide uptake of the screening programme by the public. However, the number of diseases currently screened for in the UK is limited to nine conditions, and screening is based on biochemical markers. In recent years, whole genome sequencing (WGS) as a first-line screening test has emerged as a possible tool for an expansion of the NBS screening programme to identify the genetic disease.

Genetic testing aims to identify gene variants to help confirm or rule out genetic disorders (diagnosis) or establish the likelihood of a person developing and passing on a genetic disorder in the future (screening). Genetic testing can target single variants or single genes either alone or as part of a gene panel, or can consist of whole exome sequencing [WES, looking at all coding portions of the deoxyribonucleic acid (DNA)] and WGS (looking at both coding and non-coding regions of the DNA). WGS can potentially detect every variation in a genome, testing for hundreds of genetic diseases at the same time.

In 2021, Genomics England Limited (GEL) launched its Generation Study using WGS to sequence 100,000 newborns to screen for over 200 rare diseases to test its potential for an expanded UK NBS programme.[6] Recruiting of pregnant women began in the first half of 2024. This poses several challenges to the UK National Screening Committee (UK NSC), who advise the four UK governments on screening-related questions, including the addition of new conditions to the NBS screening programme.

No methods exist for the evaluation of hundreds of conditions identified by one test. Multiplex testing has been available for years, and tandem mass spectrometry (MS/MS), currently used in the NBS programme, can detect dozens of statistical abnormalities in the blood spot. The approach to assessing each of the conditions that could be detected on the blood spot has been to look at each condition, in turn allowing a thorough evaluation of the evidence around testing, early treatment and harms and benefits of screening and an assessment of the evidence against the 20 UK NSC screening criteria.[7] However, there is pressure to assess all the conditions that might be found with WGS at once, but a condition-by-condition approach for 200 conditions, with many thousands of gene variants, may not be feasible or effective, considering the limited evidence base for rare genetic diseases.

The 'genotype-first' approach has the potential for harms in the form of overdiagnosis and overmedicalisation, particularly for conditions with pathogenic variants that have low penetrance and for conditions with a high number of variants of unknown significance. The wider psychological and societal impact of genetic testing on such a large scale is also unknown. Most healthy, adult, UK Biobank participants were found to have one or more rare non-synonymous variants when a panel to look at > 500 disease genes was used.[8] In healthy adults, these can be assumed to be benign. Similar findings in asymptomatic newborns would be concerning, demonstrating the difficulties in variant interpretation. To assess pathogenicity (e.g. > 99% probability of pathogenicity for pathogenic variants and 90–99% probability of pathogenicity for likely pathogenic variants), variants need to undergo a complex interpretation process, considering the prevalence of the variant, inheritance pattern, variant type and their predicted effects as well as observed gene–disease associations.[9] However, not all pathogenic variants cause disease in all individuals who carry the variant. Penetrance is a measurement of the relationship between a genotype and phenotype. For rare diseases, the existing evidence and context for variant interpretation comes mainly from family and clinical studies, which means that penetrance is generally overestimated and the reported pathogenicity does not equate with penetrance.[10] Moving to the screening context, knowledge of penetrance of known pathogenic variants in the general population is key for understanding the proportion of newborns who are likely to benefit from detection at screening and those who are likely overdiagnosed.[11] An evaluation of WGS for newborn screening should be able to select variants with high penetrance for a low-risk screening programme or identify a variant annotation approach that is effective in filtering out harmless variants. We do not know how informative a review focused on penetrance outcomes in newborn sequencing would be.

A comprehensive health economic evaluation of the impact of screening for 200 conditions simultaneously is unlikely to be feasible, while addressing the question on a condition-by-condition basis is unlikely to be cost-effective. WGS may detect conditions for which high-risk, expensive treatments are available or for which no treatment is available. If more diseases are included in the newborn screening programme, adequately resourced referral pathways must be in place, which will present a challenge for resource allocation in healthcare systems. There is a resource trade-off between the early genetic diagnosis and intervention for less sick or asymptomatic individuals and resource-intensive diagnostic odysseys and later treatment when a rare disease presents symptomatically.[12] Typically, an initial step in a cost-effectiveness evaluation is to map out the respective clinical pathways for the intervention and the comparator(s), which is likely to consist of the current clinical practice. Costs and health outcomes are then estimated based on the resources used and outcomes of each pathway. Mapping the full clinical pathways, that is, diagnosis and treatment, for such a wide range of conditions where the patient is asymptomatic, is unlikely to be feasible. A different methodological approach is needed for the economic evaluation screening for potentially hundreds of conditions with a single test.

Finally, the use of WGS for newborn screening presents several ethical challenges. These include consent, data ownership and psychological implications. Since newborns cannot provide consent, the NBS screening programme in the UK relies on obtaining informed consent from parent(s)/guardian(s). However, obtaining fully informed consent for WGS could be challenging and parents may not feel that they can turn down screening.[13] As such, seeking consent for WGS, where large panels of conditions are screened for simultaneously, in the same manner may be inappropriate, and there is the question whether repeat consent is required for stored genetic data and whether and when consent should be transferred from the parent to the child. Beyond the consent issues, the possibility of retaining genomic data, linking results to health records and re-evaluating data throughout a person's life raise important questions about privacy and confidentiality, data ownership and secure data storage.[14] This may have psychological implications, and an understanding of how the public and new parents feel about WGS, making decisions about taking part, data security, communication of results, access to treatment and impact on the child and the whole family need to be understood.

We aimed to assess different evidence sources and approaches to evidence synthesis, review methods of costing and collate views of parents with children with rare diseases and expectant parents on the main challenges of WGS to inform an approach to assess WGS for newborn screening in the future. To that effect, we undertook five traditional evidence reviews covering five conditions to establish the evidence base and to provide a reference case for comparison with alternative review approaches. We evaluated the use of the existing resource Clinical Genome Resource (ClinGen) as an evidence source on actionability for the genes associated with the five selected conditions, where 'actionability' refers to the level of evidence about pathogenicity, penetrance and expressivity of a genetic variant and the extent to which interventions can be used to mitigate the effect of the disease.[15] We undertook a focused review on genomic studies in paediatric screening populations reporting penetrance that may allow the identification

of gene variants with high penetrance and expressivity for a screening program that maximises benefits and minimises harms. We conducted a review of economic evaluations of WGS and WES to better understand the methodology employed to date to evaluate these tests. Finally, we explored the patient and public views on questions relating to evaluating and communicating WGS in newborns.

Chapter 2 Methods

The methods for conducting this review were predetermined and published on PROSPERO (registration number CRD42023475529).

Review of five conditions

The traditional approach to evaluating the evidence for a new screening programme is to conduct a thorough systematic review and cost-effectiveness analysis (CEA) of the test and condition to inform discussions around the extent to which UK NSC screening criteria are met. This traditional review approach was not considered to be feasible for each of the 200 rare diseases included in GEL's Generation Study.[6] Instead, a series of five systematic reviews covering 5 of the 200 rare diseases was undertaken to establish the evidence base per condition and to provide a reference case for comparison with alternative review approaches.

Objectives

Six key objectives aligned with evidencing UK NSC criteria were identified:

1. to identify the penetrance and expressivity of different gene variants associated with each condition in untreated infants/young people up to 18 years old
2. to identify the proportion of infants/young people up to 18 years with biochemical or biochemical and clinical features of each condition carrying the genetic variants known for the conditions
3. to evaluate the diagnostic accuracy (clinical validity) of gene sequencing for each condition
4. to evaluate the effectiveness of earlier intervention (treatment or surveillance) for each condition, or, if comparative data on early versus late intervention are unavailable, to separately evaluate the effectiveness of treatment in screen detected cases and following clinical presentation
5. to evaluate the effectiveness of WGS for newborn screening for each condition in terms of disease-related morbidity and mortality
6. to identify any harms of WGS for newborn screening for each condition and to identify any additional benefits beyond those afforded by earlier intervention.

The relevant UK NSC criteria per question are outlined in *Table 1*.

Selection of five conditions for review

A range of scenarios that would be likely to impact the UK NSC's advice on the implementation of WGS for newborn screening were identified. These scenarios were primarily driven by the nature of the intervention(s) that could be introduced on detection of the condition, that is:

1. widely available and relatively low-cost treatment that carries a low risk of harm (e.g. adverse effects) for the patient and wider family, for example, vitamin therapy
2. intervention centred around long-term surveillance to allow earlier detection of the clinical manifestation of the condition with the associated anxiety and costs both in terms of resource use and time required to attend appointments, for example, regular outpatient appointments to assess biochemical or developmental changes
3. ongoing long-term high-cost treatment, potentially with more significant side effects and requiring hospital visits and long-term monitoring, for example, high cost or new drugs with unpleasant or unknown side effects

TABLE 1 Key questions for the evidence summary and relationship to UK NSC criteria

Key question[a]	Related UK NSC criteria[b,16]
Question 1: What is the penetrance and expressivity of different gene variants associated with each condition?	1. The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious or treatable disease
	3. If the carriers of a variant are identified as a result of screening, the natural history of people with this status should be understood, including the psychological implications
Question 2: What is the prevalence of genetic variants in those with biochemical or biochemical and clinical features of each condition?	1. The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious or treatable disease
Question 3: What is the diagnostic accuracy (clinical validity) of gene sequencing for each condition?	4. There should be a simple, safe, precise and validated screening test
	5. The distribution of test values in the target population should be known and a suitable cut-off level should be defined and agreed
	8. If the test is for a particular variant or set of genetic variants, the method for their selection and the means through which these will be kept under review in the programme should be clearly set out
Question 4: What is the effectiveness of earlier vs. later intervention (treatment or surveillance) for each condition?	9. There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual when compared with usual care. Evidence relating to wider benefits of screening, e.g. those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened, then the screening programme should not be further considered
	10. There should be agreed evidence-based policies covering which individuals should be offered interventions and the appropriate intervention to be offered
Question 5: What is the effectiveness of WGS for newborn screening for each condition?	11. There should be evidence from high-quality RCTs that the screening programme is effective in reducing mortality or morbidity. Where screening is aimed solely at providing information to allow the person being screened to make an 'informed choice' (such as Down syndrome or cystic fibrosis carrier screening), there must be evidence from high-quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual being screened
Question 6: What are the harms of WGS for newborn screening for each condition and any additional benefits beyond those from earlier treatment?	13. The benefit gained by individuals from the screening programme should outweigh any harms, e.g. from overdiagnosis, overtreatment, false positives, false reassurance, uncertain findings and complications

RCT, randomised controlled trial.

a Questions 1–3 can be conceptually related to test accuracy if one considers 'index test positive' to mean the presence of pathogenic genetic variants. Penetrance can be considered as akin to positive predictive value (PPV), i.e. the proportion of those identified to have the pathogenic variants of interest who develop the condition. Question 2 is equivalent to half of a two-group test accuracy study focusing only on cases and not controls, i.e. what proportion of those who have the condition are identified as having the pathogenic variants of interest. Test accuracy studies will additionally consider those who do not have the condition of interest so that calculation of true negatives is possible, which cannot be achieved by a combination of questions 1 and 2.

b From: Criteria for a population screening programme – GOV.UK (www.gov.uk), UK National Screening Committee, www.gov.uk/government/publications/evidence-review-criteria-national-screening-programmes/criteria-for-appraising-the-viability-effectiveness-and-appropriateness-of-a-screening-programme

4. potentially curative interventions carrying high short-term risks and costs to NHS but long-term lower impact on both NHS, patients and their families, for example, stem cell transplantation
5. where existing screening and treatment pathways exist such that the impact of WGS would be incremental, for example, hearing screening using Automated Auditory Brainstem Response test and WGS.

Genomics England Limited provided a shortlist of 27 monogenic conditions, which (1) were considered to meet their four principles (GEL score 1, judgement in July 2023)[17] and (2) have a relatively high prevalence, within the context of rare disease. Following exclusion of conditions previously reviewed and not recommended for screening by the UK

NSC, one condition with an intervention falling under each of the five scenarios above was selected at random. The final list consisted of the following five conditions:

a. Pyridoxine-dependent epilepsy (PDE): PDE is a neurological condition resulting from an enzyme deficiency, which causes the accumulation of metabolites which inactivate pyridoxine. This pyridoxine depletion causes intractable neonatal seizures that become recurrent and prolonged if left untreated. Treatment with high-dose pyridoxine (vitamin B6 supplementation) reduces the incidence and severity of seizures.
b. Heritable retinoblastoma (hRB): hRB is cancer of the eye caused by disrupted function of the tumour suppressor protein retinoblastoma (RB). hRB occurs where both copies of the faulty gene *RB1* are present on conception (either inherited or occurring sporadically, referred to as 'germline' or 'constitutional' variants). Early intervention for hRB centres around regular ophthalmologic surveillance from birth to allow earlier identification and treatment.
c. X-linked hypophosphataemic rickets (XLHR): XLHR is an endocrine condition caused by loss of function in the phosphate-regulating neutral endopeptidase (*PHEX*) protein, which ultimately leads to hypophosphataemia (low phosphate levels) manifesting as rickets. Calcitriol and oral phosphate can be used as preventive treatments; the only available curative treatment being monoclonal antibody burosumab (currently not licensed for < 1-year-olds).
d. Familial haemophagocytic lymphohistiocytosis (fHLH): fHLH is an immune deficiency caused by malfunction of the perforin/granzyme cytotoxic pathway, leading to a proliferation of lymphocytes and overactive macrophages, which ultimately lead to infiltration and damage of organs, including the bone marrow, liver, spleen and brain. fHLH is usually activated by infection and presents as an acute illness. Active disease can initially be managed using chemoimmunotherapy with allogenic haematopoietic stem cell transplant (HSCT) providing curative treatment.
e. Medium-chain acyl-CoA dehydrogenase deficiency (MCADD): MCADD is a metabolic condition, currently included in the UK NBS screening programme. MCADD is caused by inactivity or deficiency of the MCAD protein, which affects proper functioning of the liver and can cause metabolic crises during periods of prolonged fasting or increased energy demands. MCADD is not curable, but it is managed preventively through dietary advice to avoid fasting and strict feeding regimens and through the provision of an emergency regimen (glucose polymer feed) to be used during illnesses.

See *Report Supplementary Material 1* for full details of the five conditions.

The evidence was reviewed by the condition with all genetic variants considered rather than the pre-specifying gene variants to be reviewed. This approach is not quite the same as WGS, which aims to detect particular variants and therefore does not screen for conditions.

Reviews were undertaken using a rapid evidence assessment approach, producing an Evidence Summary as described in the UK NSC guidance on the evidence review process.[7]

Identification and selection of studies

Search strategy

Search strategies were developed for each condition by an Information Specialist. The searches were developed in a test database [MEDLINE (Ovid)] and were informed and refined through a series of scoping searches, checks of a proportion of results from these searches and iterative discussions between the Information Specialist (ND), project lead (KF) and members of the reviewing team (IK, JD and SC). The full process of search development is documented in *Appendix 1*.

The following databases were searched from inception to November 2023 (see *Appendix 2* for exact dates and full search details): MEDLINE (via Ovid), EMBASE (via Ovid), Science Citation Index (SCI) (via Clarivate) and the Cochrane Library (via Wiley). No date, language or study-type filters were applied. Search results were managed using EndNote 20 [Clarivate Analytics (formerly Thomson Reuters), Philadelphia, PA, USA] and systematically deduplicated using the University of Leeds method.[18]

Inclusion criteria

Study eligibility criteria were defined for each review question, and these are summarised below. Full details are reported in *Table 2* to allow the comparison of similarities and differences across review questions.

TABLE 2 Eligibility criteria for each review question

PICO category condition	Eligibility criteria according to RQs			
Population	**RQs 2, 4**		**RQs 1, 3, 5, 6**	
PDE	*Clinical*: intractable seizures that respond to pyridoxine treatment *Biochemical*: multiple biomarkers (mainly, α-AASA, pipecolic acid) in the urine, blood or cerebral spinal fluid *Gene*: *ALDH7A1*		1. Newborn babies with no symptoms or known FH of the five selected conditions (i.e. newborn screening cohorts)	
hRB	*Clinical*: positive eye examination (indirect ophthalmoscopic examination with scleral indentation) usually following dim or absent red reflex testing *Biochemical*: *N/A* *Gene: RB1*		2. Sibling or family studies or other approximations	
XLHR	*Clinical:* bone deformity, dental abscesses, stunted growth and bone and joint pain *Biochemical:* low serum phosphate, high urine phosphate *Gene: PHEX*			
fHLH	*Clinical:* fever, enlarged spleen and liver, lymphadenopathy, an array of neurological symptoms *Biochemical:* high ferritin, abnormal cell counts, disturbed liver function markers *Gene: PRF1*, *UNC13D*, *STX11*, *STXBP2* (fHLH)			
MCADD	*Clinical*: hypoglycaemic episodes characterised by seizure and metabolic decompensation (vomiting, coma and death) *Biochemical*: C8 acylcarnitine levels ≥ 0.5 µmol/l and C8:C10 ratio ≥ 1.0 *Gene*: *ACADM*			
Intervention	**RQs 1, 2, 3**	**RQ 4 intervention**	**RQ 4 timing of intervention**	**RQs 5, 6**
PDE	Presence of pathogenic variants in the relevant gene(s) was detected by direct sequencing using any technology	Supplements of pyridoxine or vitamin B6; LRT	Early treatment or surveillance of 'gene-positive' newborns or an approximation thereof	Screening with WGS
hRB		Regular ophthalmoscopic examination usually accompanied by red reflex testing		
XLHR		Supplements of oral phosphate, active vitamin D, monoclonal antibody burosumab		
fHLH	n	Chemoimmunotherapy and allogeneic HSCT		
MCADD		dietary advice		

continued

TABLE 2 Eligibility criteria for each review question (*continued*)

PICO category condition	Eligibility criteria according to RQs			
Comparator	**RQ 1, 2, 3**	**RQ 4 intervention**	**RQ 4 timing of intervention**	**RQ 5, 6**
All conditions	No comparator	As for 'Intervention' above	Late treatment or surveillance of symptomatically detected newborns, children or young people up to the age of 18 or an approximation	1. no WGS screening 2. current practice, for example NBS screening 3. no comparator
Outcomes	**RQ 1, 2, 3**	**RQ 4**	**RQ 5**	**RQ 6**
All conditions	RQ 1, 2, 3: Presence of the condition (clinical or clinical and biochemically defined) in those with the genetic variant of interest	Disease-specific morbidity and mortality	Health-related health outcomes that could be measured across conditions (e.g. quality of life, time to diagnosis and intervention and mortality)	Harms or other benefits from WGS in newborns

α-AASA, alpha aminoadipic semialdehyde; FH, family history; LRT, lysine reduction therapy; PICO, population, intervention, control/comparison, outcome; RQ, research question.

Population

For the evaluation of penetrance and expressivity (question 1), diagnostic accuracy of gene sequencing (question 3) and the effectiveness and harms and benefits of WGS (questions 5 and 6), studies of newborn babies with no symptoms or known family history (FH) of the five selected conditions (i.e. newborn screening cohorts) were eligible for inclusion. Where evidence in newborns was limited, sibling or family studies or other approximations were considered.

For the evaluation of the prevalence of different genetic variants (question 2) and the evaluation of the effectiveness of earlier intervention (question 4), studies conducted in newborns, children and young people up to age 18 years with clinical, biochemical or genetic indicators of the five conditions were eligible. The clinical and biochemical definitions for each condition, where applicable, are provided in *Table 2*. Intervention studies (question 4) ideally included participants who were asymptomatic for the condition of interest ('early' intervention) and those treated following clinical presentation ('later' intervention).

Exposure/intervention

For review questions 1–3, the exposure was defined as the presence of pathogenic variants in the relevant gene(s) for each of the five conditions detected by direct sequencing using any technology [WGS, WES, next-generation sequencing (NGS), Sanger sequencing, etc.]. The identification of a genetic variant as pathogenic or 'likely pathogenic' is a complex process; however, established classification systems, such as the one developed by the American Board of Medical Genetics,[9] are often used. We relied on study authors' classification of pathogenicity.

For review question 4 (effectiveness of earlier intervention), eligible interventions were defined separately for each condition (*Table 2*). 'Early' intervention was ideally defined as the treatment or surveillance of newborns identified as positive for pathogenic variants of the relevant gene(s) (any test) for each condition or an approximation thereof, for example, treatment or surveillance of second or third siblings of the original proband who carried the same pathogenic variants. 'Late' intervention was considered to occur in symptomatically detected (without screening) newborns, children or young people up to the age of 18 years for each of the five conditions or an approximation (anything else available). Where evidence was limited, any author-defined 'early' versus 'late' intervention was accepted.

For review questions 5 and 6 (effectiveness and harms from WGS), the intervention was screening using WGS.

Comparator

No comparators were considered for review questions 1–3.

For review question 4, the comparators were the same as the interventions per condition, introduced at a later time point in the disease, that is, following symptomatic detection of the conditions ('late' as defined by the study authors).

For review questions 5 and 6, the comparator strategy was no screening using WGS, comparison with current practice (which in the UK is NBS screening for nine different rare diseases using methods such as MS/MS) or no comparator.

Outcomes

For review questions 1 (penetrance and expressivity) and 3 (diagnostic accuracy of sequencing), the 'outcome' is the target condition and reference standard used to measure the target condition. For each question, the target conditions were clinically, or clinically and biochemically, defined conditions. For review question 2 (proportion of different genetic variants in those with the conditions of interest), the population and target conditions are the same (i.e. those with the conditions of interest). Each combination of biochemical and clinical features should be considered as the 'reference standard' for detecting the presence of disease and calculation of diagnostic accuracy (review question 3).

For review question 4 (effectiveness of earlier intervention), eligible outcomes were measures of disease-specific morbidity and mortality.

For review question 5 (effectiveness of WGS), any health-related health outcomes that could be measured across conditions (e.g. quality of life, time to diagnosis and intervention and mortality) were eligible.

For review question 6 (harms or other benefits from WGS in newborns), harms included effects associated with false-positive results, overdiagnosis (including identification of variants of uncertain significance), ethical issues, parental or proband anxiety, referral to surveillance pathway, missing management pathways, data storage and adverse events from treating asymptomatic newborns. Potential benefits included greater certainty (doctors and patients), reduced anxiety, fewer investigations, appropriate surveillance or management plan (therapeutic yield), earlier diagnosis and earlier treatment.

Study designs

All the conditions selected for this review are considered to be rare; therefore, we expected a relative lack of evidence for each of the six review questions. We did, however, rank study designs in the order of priority for each review question. Systematic reviews were eligible for all six review questions.

For review questions 1 and 3, we considered the following types of primary study:

- observational studies of newborn screening population without treatment and follow up to disease
- observational studies of any screening population with treatment and follow up to disease with matched comparator (or with no comparator)
- observational studies of screening of which only gene variant positives are included
- case-control studies (review question 3 only)
- sibling studies
- case series (i.e. more than one case or family).

For review question 2, any observational studies reporting sequencing results for individuals with the conditions of interest were eligible. A minimum of four cases with the condition was required.

For review questions 4–6, randomised controlled trials (RCTs), non-randomised studies, before-and-after studies and other cohort studies were eligible.

Exclusion criteria

Papers that fulfil the following criteria were excluded:

- studies of people older than 18 years at diagnosis
- studies of non-hereditary forms of the five conditions
- qualitative studies
- studies that provided insufficient information for the assessment of methodological quality/risk of bias
- studies reporting outcomes not relevant to our review questions
- studies where > 10% of the sample did not meet our inclusion criteria and are not reported separately
- study reports not available in the English language
- single case studies (studies of one case or one family; however, we reported the number of case studies per condition)
- letters, reviews, editorials, communications, conference abstracts and other grey literature, publications that contained no numerical outcomes data.

Selection of studies

Titles and abstracts of records identified by the searches were screened by one reviewer. A random 20% sample of records were screened independently in duplicate by a second reviewer, and any records with any uncertainty over inclusion (coded ‘Maybe’ by either reviewer) were discussed and a consensus decision was reached. The full publications of all records selected were obtained and assessed by one reviewer, with a random 20% sample assessed independently by a second reviewer. Disagreements were resolved by consensus, or through discussion with a third reviewer.

Data collection and analysis

Data extraction

All data extractions were extracted into a piloted electronic data collection form. Data were extracted by one reviewer and checked by a second reviewer. Any disagreements were resolved by consensus or discussion with a third reviewer.

Assessment of methodological quality

a. At the outset, we planned to carry out quality appraisal using dedicated tools for each study type, for example, using the Risk of Bias in Systematic Reviews (ROBIS) tool for systematic reviews,[19] the Quality in Prognostic Studies tool[20] or tools developed by the Joanna Briggs Institute.[21]
b. Ultimately, the majority of reports included were of relatively small series of patients with the conditions of interest. The decision was, therefore, taken to use a single tool designed for the appraisal of case series and case reports, tailored to the review question.[22] Use of a single tool allowed an overall picture of study quality across study types and conditions.

Quality assessment was conducted by one reviewer with all assessments checked by a second reviewer.

Methods for synthesis

We planned to employ an order of priority approach to synthesis, providing a narrative synthesis of studies that employ the highest priority design available with an accompanying tabulation of key details from studies using lower priority designs (e.g. study design, countries in which the studies took place, sample sizes and key outcomes) (see Protocol for full details, PROSPERO: CRD42023475529).

Due to the paucity of studies using higher priority study designs and the lack of available data for review questions 1 (penetrance and expressivity), 3 (diagnostic accuracy of sequencing), 5 and 6 (effectiveness and harms or other benefits from WGS in newborns), an alternative approach to synthesis was adopted.

Studies providing data for review question 2 (prevalence of different genetic variants in those with the condition of interest) were classified into subgroups according to the definition of disease using clinical and/or biochemical measures (broadest to narrowest), and the largest most representative study from each subgroup was synthesised. For any subgroups where a single study was insufficient to cover the breadth of tests (e.g. numbers of genes included) and population, more than one study was selected for synthesis. Data concerning study design, country in which the study took place, definition of the condition, number of participants, genes tested and genetic tests used, gene frequency and types of variant identified were presented. Available data on expressivity in those with the condition were also presented. A subset of data items was reported for the remaining studies that were considered relevant for the review question 2 (see *Appendix 3*, *Table 16*).

Studie providing data for review question 4 (evidence on earlier vs. later treatment) were classified into subgroups according to the definition of 'early' versus 'late' treatment that was used. Studies using definitions most closely related to the review question were synthesised. For completeness, case reports (typically, family or sibling studies with four or fewer patients) reporting a comparison of early versus late treatment were reported in *Appendix 5*, *Table 31*. Data concerning study design, country in which the study took place, definition of the condition, number of participants, definition of early and late treatments, outcome measures, time point of measurement and results were presented. A subset of data items was reported for the remaining studies considered relevant for the review question 4 (see *Appendix 3*, *Table 19*).

Exploring Clinical Genome Resource as an evidence source

The existing online research ClinGen was explored as an evidence source for the clinical actionability of rare paediatric genetic diseases (i.e. the risk of variants is known to be high, and intervening will prevent/mitigate disease), and findings

were compared to the traditional reviews of the five conditions PDE, hRB, XLHR, fHLH and MCADD. The main aim was to assess the resource as a potential evidence source for the UK NSC, considering that traditional reviews for over 200 conditions are unlikely to be feasible.

Research question (RQ): What is the utility of ClinGen as an evidence source for the clinical actionability for the genes associated with the 5 selected conditions and scaled up to 200 conditions?

Background to Clinical Genome Resource

The or ClinGen is a National Institutes of Health [US]-funded, open access and centralised resource that defines the clinical relevance and actionability of genomic variants. ClinGen's action group developed practical methods to identify genetic disorders with the greatest clinical utility when detected in previously undiagnosed adults[23] and adapted the methods to the paediatric context.[24] Working groups use a standardised protocol to produce summary reports and semi-quantitative metric scores for childhood-onset rare genetic disorders. The methods to produce summary reports are adapted from work by Goddard *et al.*[25] to guide decisions about returning incidental findings. The method provides a transparent, systematic, evidence-based process for identifying and quality-rating of evidence. The evidence is reviewed by an expert panel that applies a semi-quantitative metric based on Berg *et al.*[26] to score the overall clinical actionability of gene variants. Each topic is scored independently by multiple members, and the scores are then discussed using consensus for assigning a single actionability score.

The ClinGen focuses on four main curation activities: gene–disease validity (pathogenic variants in the gene clearly cause disease), dosage sensitivity (loss or gain of the gene results in disease), variant pathogenicity (categorisation of variants in the gene into benign, uncertain and pathogenic) and clinical actionability. As of March 2024, the database reported 6357 curated variants across 100 genes (not restricted to paediatric-onset diseases) and included 144 paediatric actionability reports.[24] Conditions/genes have to meet a set of minimum requirements to be reviewed. These include:

1. Guidelines on an intervention relevant to an undiagnosed paediatric population exist (focusing on disease prevention, lowering the clinical burden and improving clinical outcomes and not including 'personal utility', reproductive decision-making and 'ending the diagnostic odyssey').
2. At least one variant in the gene should have moderate-to-high penetrance (40% penetrance, relative risk = 2), or no information on penetrance is available.
3. The health condition is significant.

Actionability reports determine the clinical actionability of secondary findings in paediatric patients undergoing clinically indicated diagnostic testing. Although the paediatric protocol states that elements relevant for population-based screening decisions may be captured, there is a lack of consideration of systems-based practice and availability of population-scale follow-up. Actionability ratings from ClinGen are therefore insufficient for recommending screening in asymptomatic cohorts.[24]

Paediatric summary reports summarise the evidence on gene–condition pairs under four dimensions: severity of disease, likelihood of disease (similar to penetrance), effectiveness of intervention and nature of intervention. The dimensions are scored from 0 to 3 (3 being best for actionability) based on the evidence, and the evidence is rated for the likelihood and effectiveness dimensions (poor evidence to substantial evidence). The scores are summarised across the four dimensions to provide an overall score for each outcome–intervention pair. The scores and the evidence are taken into consideration for the final assertion on actionability. The evidence review allows for a pragmatic approach allowing non-systematic and expert based references, because it is recognised that this is the most commonly available evidence for rare genetic disorders.

Clinical Genome Resource information for the five conditions reviewed

We searched the ClinGen database on 19 February 2024 for each of the genes included in our review of five conditions and identified reports from the paediatric actionability working group. For each of the five conditions, we extracted the scores for condition–intervention pairs where available, the overall assertion, the evidence provided underlying the scores and the references cited. We tabulated the scores and evidence level comparatively against the evaluation

from GEL and against our assessment against the UK NSC criteria. We used the upper quintile of the ClinGen score as the cut-off for actionability as recommended by the Locus-Variant Binning Committee, who developed the transparent semiquantitative metric for evaluating the clinical actionability for pathogenic variants.[26]

We performed a narrative synthesis of our assessment.

Review of genomic studies of newborn screening cohorts reporting penetrance for pathogenic variants

Knowledge of penetrance is important for the decision about which pathogenic variants of genes associated with childhood diseases should be reported for action following detection by sequencing. This is particularly important in the screening context to maximise the number of babies who are likely to benefit from detection at screening and those who are likely overdiagnosed (detection rate of pathogenic variants is not sufficient). The main aim of this focused review was to identify studies reporting penetrance as an outcome following sequencing in the newborn screening setting for any paediatric condition – to explore the feasibility of identifying highly penetrant pathogenic variants that could be considered for a screening programme (i.e. minimising harm from reporting variants of unknown or uncertain significance).

RQ: What is the penetrance or actionability of pathogenic/likely pathogenic variants of rare genetic child-onset diseases identified in newborn screening populations using WGS?

Identification and selection of studies

Searches were developed iteratively in a single database (MEDLINE via Ovid). The final search combines the concept of newborn screening with either WGS, WES, penetrance, actionability, sequencing or allele frequency.

The following databases were searched from inception to January 2024 (see *Appendix 2* for exact dates and full search details): MEDLINE (via Ovid), EMBASE (via Ovid), SCI (via Clarivate) and the Cochrane Library (via Wiley). No date, language or study-type filters were applied. Records were exported to EndNote and systematically de-duplicated using a process based on the University of Leeds method.[18]

Eligibility criteria

This approach relies on the availability of information on penetrance/expressivity and generalisability of findings to the screening context. We, therefore, only included studies of newborn screening populations that report as a minimum the penetrance (or an approximation) of gene variants linked to rare genetic diseases with childhood onset.

Studies that satisfied the criteria listed in *Table 3* were included.

Papers that fulfilled the following criteria were excluded:

> Studies of populations other than newborns, studies on populations at risk or with symptoms, studies where WGS is the second-tier test, qualitative studies, studies only reporting variant frequency without an estimation of penetrance, studies only reporting carrier frequency, studies that provide insufficient information for assessment of methodological quality/risk of bias, studies reporting outcomes not listed in our inclusion criteria, studies where > 10% of the sample did not meet our inclusion criteria and were not reported separately, articles not available in the English language, single case studies, letters, reviews, editorials, communications, conference abstracts and other grey literature, publications that contain no numerical outcomes data.

Review strategy

Titles and abstracts of records identified by the searches were screened by one reviewer. A second reviewer independently assessed a random 20% sample of the titles/abstracts plus records labelled as unclear by the first reviewer. Disagreements were resolved by consensus. Full-text articles were independently assessed against the

TABLE 3 Study eligibility criteria

Population	Newborn babies without symptoms or known FH of rare genetic diseases
Target condition	Any rare genetic condition with childhood onset of symptoms
Exposure/ intervention	Screening using WGS or WES Step down: direct sequencing using panel tests or single gene testing in newborns without symptoms or known FH
Comparator	Order of priority (comparator is only needed for approximation of penetrance using allele frequency): 1. No comparator necessary 2. Randomisation WGS vs. standard care 3. Contemporaneous cohort of matched/random newborns without genetic screening 4. General population of newborns without genetic screening 5. General population of healthy adults
Outcomes	For comparator 1: a. Follow-up without treatment to clinical presentation Step down: b. Follow-up to/comparison to biochemical tests (e.g. conventional NBS test), indicating presence of disease processes c. Categorisation into levels of penetrance based on existing literature d. Categorisation into pathogenic/likely pathogenic based on existing classification systems For comparators 2–5: Variant frequency matching the inheritance pattern of the relevant condition in screened and comparator to approximate penetrance
Study designs	Order of priority: Observational study of WGS of asymptomatic newborns without reporting to parents (no treatment) and follow-up to clinical features (or reporting results, but no treatment until symptomatic or reporting results and then later stopping treatment to determine whether necessary) Observational study of WGS with follow-up testing (e.g. biochemical test) Observational study of WGS follow-up to symptom onset despite treatment Randomised trial of WGS (with results reported and treatment) vs. standard care to determine variant frequency and approximate gene penetrance Comparative studies using comparators detailed above to determine variant frequency and approximate gene penetrance
Language	English language

inclusion/exclusion criteria by two reviewers. Disagreements were resolved by consensus. Records rejected at full-text stage are listed in *Report Supplementary Material 2* with reasons for exclusion.

Data extraction strategy

Data were extracted into a piloted electronic data collection form by one reviewer and were checked by a second reviewer.

Assessment of methodological quality

Methodological quality of each study was assessed based on study design, as no appropriate tool was identified to assess bias in single-arm cohort studies.

Methods for analysis/synthesis

We tabulated and narratively synthesised information on the proportion of sequencing positive cases, agreement with conventional screening results and the proportion with confirmed disease on clinical follow-up (penetrance) for each condition separately and for all conditions combined per study. We considered positive genetic screening outcomes as unconfirmed if within the studies' type and length of follow-up, the condition could not be confirmed with

confirmatory testing and/or clinical follow-up. We provided our learning from this review for the context of WGS in newborn screening.

Review of cost-effectiveness evaluations of whole genome sequencing and whole exome sequencing

Based on a quick search, we could not find any cost-effectiveness evaluations of WGS or WES specifically for newborn screening. We included studies evaluating the use of WES because there is likely to be an overlap in the methodological challenges faced when considering their cost-effectiveness. To help inform future cost-effectiveness evaluations of using WGS for screening, the objectives of this review were to provide: (1) an overview of the contexts in which the cost-effectiveness of WGS or WES has been evaluated to date, (2) a high-level summary of the methodological approaches adopted in these evaluations, (3) an in-depth appraisal of how health-related quality of life (HRQoL) has been measured and incorporated into cost–utility studies and (4) a summary of the cost components included in micro-costing studies. The motivation for providing an overview of the contexts in which WGS or WES has been evaluated to date is that, if focusing on WGS or WES as a means to detect a disease earlier, the approach adopted to capture the benefits and harms of earlier diagnosis and potentially treatment may be applicable when considering a screening context.

We kept our search strategy and inclusion criteria broad to ensure that any WGS or WES studies conducted in a screening context would be identified.

Objective: To produce a methodological overview of published economic evaluations and costing studies of WGS or WES in any context or population.

Search strategy

Searches for cost or economic evaluations were conducted in the following databases in February 2024:

- MEDLINE (Ovid)
- EMBASE (Ovid)
- CEA registry
- Web of Science (WoS)
- American Economic Association electronic bibliography (EconLit)

Pre-print sources were not searched.

Our search terms were broadly based on those used in a previously published systematic review of WGS/WES cost-effectiveness studies.[27] Searches combined concepts relating to sequencing analyses, including, but not limited to, terms for WGS and WES and (1) costing studies, (2) budget impact studies, (3) economic evaluations and (4) economic models. A previous scoping review, Nurchis *et al.*[28], was identified, which focused specifically on Health Technology Assessments (HTAs) of WGS. Targeted searches for HTAs were not included in the Schwarze *et al.*[27] review, so we added targeted searches of MEDLINE, EMBASE and International HTA from 2022 (i.e. post the searches conducted by Nurchis *et al.*)[28] to ensure identification of any additional HTA reports that our original search may have missed. Searches were restricted to English language and humans. The search development methods can be found in *Appendix 1* and the full search details are reported in *Appendix 2*.

Studies included in the Schwarze *et al.* systematic review[27] and the Nurchis *et al.* HTA scoping review[28] were assessed against our inclusion criteria. References of relevant systematic reviews were checked for any additional primary studies that were not identified by our search.

Inclusion and exclusion criteria

The inclusion and exclusion criteria for the review (*Table 4*) broadly replicate those used in the Schwarze *et al.* review, apart from the outcome inclusion criteria.[27] This review had a broader scope than our review in that it also included

TABLE 4 Inclusion and exclusion criteria

	Inclusion criteria	Exclusion criteria
Population	Studies in human health care	Studies which focused on populations with suspected communicable diseases
Intervention	WGS and WES (any platform)	Targeted sequencing
Comparator	Any comparator or no comparator	
Outcomes	Costs	Included health outcomes only
Study design/ publication type	HTAs, costing studies, budget impact analyses, economic evaluations (partial and full), evidence-based guidelines, clinical trials	Conference abstracts

eight studies which only focused on health outcomes (even though the searches targeted economic evaluations). Because our inclusion criteria were restricted to studies that included costs as an outcome, we excluded these eight studies. We also excluded studies that focused on using WGS or WES to diagnose or monitor communicable diseases (e.g. bacterial infections). The reason for this was that the role of sequencing in this context was often very different compared to using it to diagnose a rare disease in an individual. For example, the motivation for sequencing was often to monitor for outbreaks and/or detect treatment resistance – potentially independent of the health of the individual concerned. The timing of the test is crucial in this context, and samples from multiple individuals were often batched together. We also excluded conference abstracts because word count restrictions meant that the studies could not be reported in sufficient methodological detail to facilitate meaningful evidence synthesis.

Data were not extracted from included literature reviews, but the studies included in the reviews were checked against our database search results to ensure that we added any eligible studies that had been missed.

Screening

Initial screening of titles and abstracts, followed by full-text screening was carried out using Rayyan.[29] The titles and abstracts of records identified by the searches were independently screened by two reviewers (BS and AO). Where any disagreements occurred, the record was taken through to full-text screening. Full texts were assessed against inclusion/exclusion criteria by one reviewer (BS or AO), while 20% was independently checked by the other reviewer. Disagreements at this stage were resolved through discussion. There were some studies where it was unclear whether the test being evaluated was WGS or WES. These were shared with our genetics advisors for confirmation. The reasons for excluding records at the full-text stage were documented.

Data extraction

An electronic data collection form was developed in Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA), piloted and refined. The form included the same study characteristics extracted in the Schwarze *et al.* systematic review (for continuity) but with additional components to capture additional methodological issues relating to costing of WGS/WES pathways and comparator pathways.[27] Studies which produced a comparative analysis of costs and health outcomes were categorised as full economic evaluations. This included cost–utility studies, where the outcome is defined as the cost per quality-adjusted life-year (QALY), or cost-effectiveness studies, where the outcome is defined as the cost per change in a particular outcome. Partial economic evaluations, such as cost–consequence analyses, were defined as studies where costs and outcomes are reported, but they are reported in isolation. Costing studies are those which did not evaluate patient outcomes and just reported costs.

Data were extracted by one reviewer (BS or AO), with a random 20% checked by the other reviewer. Disagreements were resolved by consensus or discussion with a third reviewer.

Critical appraisal

Given that the main focus of this cost-effectiveness review was to develop an understanding of the methodology used, rather than the results of the studies identified, we did not appraise the methodological quality of included studies.

Evidence synthesis

For the studies meeting our eligibility criteria, we first provide a tabular/graphical overview of the key study characteristics, including number of publications by year, continent/country, the type of economic analysis conducted, which sequencing approaches were evaluated (WES, WGS or both) and the target population (grouped broadly by age).

To summarise the contexts in which WGS or WES has been evaluated, we provide a narrative overview of the included studies, structured around the PICO:

- Population: who were the target population, that is age, disease area.
- Intervention: specifics about the timing of the test or aspects relating to how the test was conducted and integrated into the diagnostic pathway.
- Comparator(s): what were the comparator(s) in the studies, if any, and what types of data were used to inform the comparator pathways and associated resource use.
- Outcomes: what was the main outcome of the study, for example, cost per patient, cost per diagnosis, cost per QALY.

We then provide a summary of some specific aspects of the methodological approach adopted, which are of particular interest when considering the economic evaluation of WGS. These include the costing perspective adopted, the time horizon of the analysis and if and what discounting was applied and whether any sensitivity or scenario analyses were conducted to capture and demonstrate uncertainties in the underlying evidence base. The time horizon adopted is important as WGS often has long-term implications, particularly for conditions with lifelong impacts or where early detection can alter long-term outcomes. Given the likely long-term nature of the impacts of WGS (and screening more broadly), the use and choice of a discount rate is also likely to significantly influence the value of future costs and benefits. The costing perspective is important as there are disparities and debate around who is responsible for the immediate and long-term cost consequences associated with WGS. Using WGS at scale may also lead to societal costs/cost savings that require consideration. There is also likely to be a high level of uncertainty in the evidence base underpinning economic evaluations of WGS, and we wanted to assess whether sensitivity or scenario analyses were routinely conducted to capture these uncertainties. There are other methodological factors important to consider when critiquing an economic evaluation, such as characterising heterogeneity across population subgroups or the analytic approach adopted, but we felt the appropriate approach to these are more context-specific and therefore challenging to critique in a review with such a broad clinical focus.

Since the reference standard outcome for a UK NSC economic evaluation is cost per QALY, we present a more detailed methodological review of the included cost–utility studies. Specifically, we focus on how HRQoL has been measured and incorporated into these evaluations.

Additionally, we provide a narrative summary of the costing studies that utilised a bottom-up approach to estimating the costs of WGS. Examining the components included in these costings and the methodological approaches employed offers valuable insights for future evaluations of newborn screening programs incorporating WGS.

Consideration of the public voice in the evaluation of whole genome sequencing

The broad aims of the patient and patient and public involvement and engagement (PPIE) were to:

1. build an understanding of the perspectives and experiences of members of the rare disease community around the key challenges and opportunities of WGS in newborn screening
2. explore views towards, and understandings of, screening programmes and the role of the UK NSC
3. discuss methodological challenges identified during the review process
4. note views on limitations in the evidence base
5. contribute meaningfully to the development of future PPIE in this area.

Recruitment targeted people with differing lived experiences of rare genetic conditions (including parents and adults living with rare conditions), third-sector representatives and prospective parents from the public. Recruitment took place between October and December 2023. Due to the restricted time frame, recruitment was initially through a targeted approach via known contacts at several third-sector organisations and through GEL's PPIE groups. Through these contacts, we recruited the charity representatives (including an adult living with a genetic condition). Four of the parent representatives were recruited through a post on a social media page for families living with rare genetic conditions, and one approached the research team after reading a press release about the study. The public representative had previously been involved in PPIE work related to screening for a rare genetic condition, so they were able to meaningfully contribute without needing extensive guidance/input on genomics, screening and health care.

The group consisted initially of eight members: five women and three men:

- a member of the public, who was also an expectant, and later a new, parent
- an adult living with a rare genetic condition, who was also a Diversity, Equity, Inclusion and Belonging consultant and representative for a charity organisation supporting people living with a genetic condition
- representative from a charity advocating for and supporting people living with genetic conditions
- five parents of children living with a rare genetic condition (ages of their children ranged from 5 years to young adults).

One of the parents withdrew from the group very early in the process due to their caring commitments.

Five (2-hour) virtual meetings took place between 15 January and 21 May 2024 (for discussion topics, see *Table 5*). Meetings 1 and 4 included the involvement of a member of the review team (KF) to present the review and its progress, with the remaining three meetings being independent of the review process. Meetings were recorded to facilitate note-taking. To maximise inclusion, participants who were unable to join any of the meetings were offered the opportunity to catch up by watching recordings of missed meetings (with the permission of the participants present at the meeting) and sending feedback via e-mail.

Meetings were deliberative, to create the knowledge space required for discussion, drawing on relevant evidence and the experience of the participants. Ahead of each meeting, participants were informed of the topic for discussion and were given up to 1 hour of preparatory work in the form of reading from a variety of sources and/or formulating ideas/questions.

TABLE 5 Schedule of PPIE meeting discussion topics

Meeting	Topic
1	Introduction to the group and project Team and PPIE group introductions. Presentation on the purpose and process of the review. Answering questions from the PPIE group on this
2	Harms and benefits Pre-reading and presentation on: background/context to newborn screening, concept of harms and benefits, the role of the UK NSC, how WGS differs from current screening and that different stakeholder groups may have different views. Open discussion on 'What do you think are the key harms and benefits of screening that policy-makers should be taking into account'
3	Genetic uncertainties Pre-reading and presentation: current understanding of genetic uncertainties, including defining terms such as penetrance, expressivity, variants of uncertain significance. Open discussion on the impact of this information on balancing harms and benefits and how to approach these complexities when explaining genetic screening to parents
4	Systematic review findings Pre-reading: executive summary of draft review and open discussion on the methodological challenges and evidence gaps identified during the review process and the conclusions reached
5	Conclusions and role of PPIE in future reviews Open discussion on topics, including: the most important harms and benefits; how these can be captured and traded off; how these harms and benefits can be transformed into evidence to support the decision-making of the UK NSC; how can PPIE be best incorporated into future reviews and who needs to be involved in these conversations

Following each meeting, participants were asked to complete a short evaluation questionnaire. This had two main purposes: (1) to evaluate the format and content of the meeting and (2) to allow participants time to reflect on the content of the meeting and provide any short summary points on this. In the final meeting, which brought together the group's thinking on all of the topics covered, participants were asked to contribute to an online whiteboard (before, during and after the meeting) with their responses to prompts: What do you see as the main or most important? (1) Benefits and (2) harms to WGS for screening in newborns, (3) how should potential harms be prevented/dealt with, (4) how should benefits and harms be balanced against each other, (6) what kinds of evidence should the UK NSC prioritise, (7) what should happen where the evidence is not available, and finally, (8) how should PPIE contribute to this process (including when and who)?

Chapter 3 Results

Using a traditional review approach

This traditional review approach covers a stratified random sample of 5 of the 200 rare diseases included in GEL's Generation Study[6] to establish the evidence base per condition and to provide a reference case for comparison with alternative review approaches.

Workload and implications for scaling up to more than five conditions

Extensive scoping, refinement, testing and running of the electronic searches for the five conditions took a total of 6 weeks. We cannot envisage any way in which this could be simplified or shortened in duration. A traditional review for 200 conditions would require 200 individual searches to be developed.

The average time required to sift 100 abstracts was 40 minutes across five reviewers with varying levels of systematic review experience. A single full-text sift took an average of 5–10 minutes, depending on the question and the condition. *Table 6* reports the approximate time required for searches and sifting activities for the 5 conditions, with extrapolation to 200 reviews of 200 conditions. This does not include time for double sifting, data extraction and quality assessment, discussion of disagreements and categorisation of studies for synthesis purposes, that is, by population for genetic testing (review question 2) and by timing of treatment initiation (review question 4). These time estimates, therefore, only illustrate the extrapolation of two reviewing tasks and do not reflect the time needed for a complete review process. The complete review process took a team of three full-time and two part-time reviewers 7 months to complete. We anticipate that a single, similarly sized review team would need about 280 months or 23 years to review and synthesise the evidence for 200 conditions. The approximate time estimates demonstrate the scale of the effort required to evaluate WGS on a condition-by-condition basis, although some learning may be transferable to conditions similar to those reviewed here.

Volume of evidence for the six review questions for five individual reviews

We sifted a total of 19,689 title and abstracts, of which 1348 were selected for full-text assessment (range 55–449 per condition). A total of 268 studies were eligible for inclusion across the five reviews (range 31–78). *Table 7* summarises the search and eligibility results across conditions, with flow diagrams per condition presented in *Appendix 4*, *Figures 2–6*). The excluded studies and reasons for exclusion are listed in *Report Supplementary Material 2*. Extrapolating to a review of 200 conditions, we could expect as many as 787,560 titles and abstracts and 53,920 full texts to be sifted.

No evidence was identified for four of the pre-specified review questions for any of the five selected conditions (*Table 7*). All four questions were those for which studies were required to be conducted in newborns:

- penetrance of the condition in a newborn cohort where sequencing was the first line screening test (Q1)
- test accuracy of WGS in the newborn screening setting (Q3)
- clinical effectiveness of WGS in newborns (Q5)
- harms or additional benefits from WGS in newborns (Q6).

TABLE 6 Approximate time (in hours) needed for searches and sifting for a traditional review of 200 conditions extrapolated from five individual reviews

	Approx time taken (hour) to develop searches	Approx time (hour) taken to run searches	Approx time (hour) to sift titles/abstract	Approx time (hour) to sift full texts
5 conditions	47.5	21	131	170
200[a] conditions	1900 (253 full time working days)	840 (112 full time working days)	5241 (699 full time working days)	6795 (906 full time working days)

a Numbers multiplied by 40 to give estimated number for 200 conditions.

TABLE 7 Volume of evidence for 5 conditions taking the traditional review approach and an extrapolation to a review of 200 conditions

	PDE	hRB	XLHR	fHLH	MCADD	Total	200 conditions[a]
Titles and abstracts sifted	992	5797	4787	5151	2962	19,689	787,560
Included on title and abstract	170	245	313	449	180	1357	54,280
Included at full text	31	78	42	52	65	268	10,720
Number included per review question							
Question 1:[b] Penetrance	0	0	0	0	0	0	0
Question 2:[c] Genetic spectrum in patients	26	73	39	51	56	245	9800
Question 3:[d] Test accuracy	0	0	0	0	0	0	0
Question 4:[e] Early vs. late treatment	5	5	3	1	9	23	920
Question 5:[f] Effectiveness of WGS	0	0	0	0	0	0	0
Question 6:[g] Harms of WGS	0	0	0	0	0	0	0

a Numbers multiplied by 40 to give estimated number for 200 conditions.
b What is the penetrance and expressivity of different gene variants associated with the five conditions in untreated infants/young people up to 18 years?
c What proportion of infants/young people up to 18 years with (1) biochemical and (2) biochemical and clinical features of the five conditions carry the gene variants known for the conditions?
d What is the diagnostic accuracy (clinical validity) of gene sequencing for each of the five conditions?
e What is the evidence on early [following screen detection or sibling detection (cascade testing)] vs. late (following clinical presentation) treatment?
f What is the effectiveness of WGS for newborn screening for each condition?
g What are the harms of WGS for newborn screening for each condition and any additional benefits beyond those from earlier treatment?

Evidence was identified only for the two review questions focused on individuals with the conditions of interest (*Table 7*), that is:

- the prevalence of genetic variants and genetic spectrum in a paediatric population with the condition of interest (range 29–89 studies) (Q2)
- evidence for early versus late treatment, where 'early' was the closest approximation to management of screen-detected cases that were identified (range 1–9 studies).

Considering that the five conditions are rare diseases, the number of identified records was larger than expected. For some conditions such as PDE, data for the same study participants are likely to have been represented in multiple publications by different authors, but this was not always clearly identifiable within the study reports. Overall, however, the five individual reviews yielded very little evidence of the sort required by the UK NSC.

Findings from the traditional review approach for the five conditions

Gene/variant frequency in patients with condition(s) of interest

Review question 2 looks to identify the proportion of infants/young people up to 18 years with either (1) biochemical or (2) biochemical and clinical features of the five conditions (PDE, hRB, XLHR, fHLH and MCADD) that carry the gene variants known for the conditions. This question can be considered in multiple parts. First and foremost, the idea was to compare the genetic spectrum in children with disease to that in sequencing positive newborns (question 1) to assess the difference of clinical and genetic diseases. Secondly, to identify the proportion of individuals with the conditions who have an underlying genetic cause (that can be identified by WGS) and the proportion of patients who may be missed because of non-genetic cause of the condition or because of non-specific symptoms caused by variation in a different gene. Thirdly, consideration of the genetic spectrum in those with the condition to identify whether recurrent (or 'reported') genetic variants are responsible for large proportions of cases or whether novel variants that can be more

analytically intensive to identify are frequently responsible. Finally, any reported patterns of expressivity associated with different variants or types of variants were considered.

The extent to which identified studies can inform these questions depends on the type of genetic test used and the inclusion of specific genes or genetic variants in the studies as well as the nature of the included target population. The target populations identified did not easily fit into the two pre-specified categories of clinically defined and clinically/biochemically defined disease and extended to genetically defined disease. Disease-specific categories were therefore used, which ranged from clinical symptoms, with or without biochemical markers, to study inclusion based on the presence of a genetic variant (*Table 8*). The categories were based on the expectation that the differences in the definition of disease may have an impact on the genetic spectrum reported.

Table 8 summarises the volume of evidence by disease definition for the five selected conditions. These definitions were condition-specific according to the availability of biochemical markers (not available for hRB), the number of umbrella terms of disease with overlapping symptoms (XLHR, hereditary hypophosphataemia and hypophosphataemia) and whether conditions can only be defined by genetic information (four conditions of fHLH).

Below, we report the largest most representative study from each subgroup per condition; some subgroups have more than one study where results are complementary. Details per study regarding gene/gene variant frequency are reported in *Appendix 5*. An overview of all included studies is included in *Report Supplementary Material 3*. The detailed quality assessment of extracted studies is presented in *Appendix 3*, *Tables 17*, *18* and *20* and summarised below for each condition. Given the differences in populations and variant frequencies in the selected studies resulting from the different disease definitions, synthesising across studies was challenging. We therefore start with a short summary of results across the disease definitions and then describe results in detail per disease definition beginning with the broadest, with a particular emphasis on the disease category that is most relevant to WGS in newborns. Because the frequency of genes and variants identified in a population does not depend on the sequencing method, we did not prioritise studies based on sequencing method used.

Pyridoxine-dependent epilepsy

The three subgroups and studies for PDE were: childhood-onset, pharmacoresistant seizures,[30] clinically or biochemically defined PDE[31] and genetically confirmed PDE (all with variants in *ALDH7A1*).[32,33] Details per study are

TABLE 8 Categorisation of disease per condition and number of studies identified per disease category based on study populations

	PDE	hRB	XLHR	fHLH	MCADD
Disease definitions per condition (n studies)[a]					
Disease definition from broadest (top) to narrowest (bottom)	Childhood-onset pharmacoresistant seizures (*n* = 6)	Any RB (*n* = 55)	Hereditary rickets (*n* = 1)	HLH (*n* = 38)	IEI/IEM (*n* = 2)
	Clinical–biochemical PDEs(*n* = 9)	Sporadic RB (*n* = 6)	Hypophosphataemic rickets (*n* = 20)	fHLH (two or more of *PRF1*, *UNC13D*, *STXBP2* and *STX11*) (*n* = 2)	FAODs (*n* = 1)
	Gene (*ALDH7A1*)-positive PDE (*n* = 11)	Familial/germline/ bilateral RB (*n* = 12)	Hereditary hypophos-phataemic rickets (*n* = 5)	Gene (*PRF1*)-positive (*n* = 6)	ACAD (*n* = 1)
	–	–	XLHR (*n* = 13)	Gene (*UNC13D*)-positive (*n* = 3)	Clinical–biochemical MCADD (*n* = 51)
	–	–	–	–	Gene (*ACADM*)-positive MCADD (*n* = 1)

ACAD, acyl-CoA dehydrogenase deficiency; FAOD, fatty acid oxidation disorders; HLH, haemophagocytic lymphohistiocytosis; IEM/IMD, inborn errors of immunity/inherited metabolic disease.
a The number of disease categories varies from three to five between conditions.

provided in *Appendix 5*, *Table 21*. Three of the four studies were retrospective,[30–32] including one multicentre registry-based study.[32] Two studies exclusively used sequencing to identify variants, two used targeted panels[31,34] and one used WES followed by short- or long-read genome sequencing (GS) for cases not solved by WES.[30] The registry-based study did not report the genetic tests used to identify variants.[32]

Quality assessment

Of the four extracted studies, only Boonsimma *et al.*[30] was considered to have met all quality assessment criteria. This was a single-centre study with clearly defined population and period of recruitment, and the sequencing methods were described in detail and cases and variants were adequately ascertained. The study also broadly indicates the population of patients with unexplained infantile-onset seizures who may benefit from newborn WGS. We had several concerns about the other three studies extracted based on lack of information about the selection of participants, level of detail provided concerning the participants[33] or testing methods used.[31,32] Two of the four studies were considered to report the genetic spectrum of PDE–*ALDH7A1* based only on sequencing techniques.[30,33] It is worth noting, however, that the majority of PDE presents in the first 4 weeks of life such that the window of opportunity for WGS to benefit infants prior to developing seizures is small.

Summary of results

Using the broadest definition of a target population who might be suspected of having PDE, up to 6% could be confirmed as having *ALDH7A1* variants following genetic testing (based on a sample of 103 children with infantile-onset (< 12 months of age) pharmacoresistant seizures.[30] Restricting the population to those with a pyridoxine response increased the percentage detected by WES to 11% (4/35), 75% (3/4) of whom also had elevated biochemical markers for PDE.[31] Studies of patients with clinical and biochemical indicators of PDE have reported higher percentages with biallelic *ALDH7A1* variants of up to 86% (18/21).[35]

The international registry-based study of patients with clinical suspicion of PDE and at least one *ALDH7A1* variant allele[32] provides the most comprehensive current picture of the genetic spectrum of patients with PDE, with data regarding Chinese patients with PDE reported in Jiao *et al.*[33] The majority of patients in both studies had compound heterozygous *ALDH7A1* variants (58.9%, 109/185[32] and 83.9%, 26/31[33]) compared to homozygous variants. Missense variants were the most commonly observed, accounting for 57% of variant alleles[32] and 65% of all variants.[33] Although four individual (missense) variants accounted for 38% (140/367) of all variants identified in Coughlin *et al.*,[32] the majority of identified variants each occurred in a single individual. The occurrence of novel and often 'private' variants was a recurrent phenomenon across the included studies (17/26 variants occurred in a single individual in Jiao *et al.*[33]), with sequencing of one or both parents to confirm pathogenicity of identified variants commonly reported (e.g. in two of six children reported in Boonsimma *et al.*).[30]

Results for each subgroup by definition of disease

Childhood-onset pharmacoresistant seizures

The broadest category defined for PDE is a clinically defined population of children with early-onset (neonatal or at latest infantile-onset) seizures that have been shown to be resistant to one or more standard pharmacological treatments for epilepsy (i.e. refractory or pharmacoresistant). This is the group in whom PDE might be ultimately suspected on clinical presentation. Potentially, hundreds of genes can be associated with epilepsy; however, only a relatively small number (including *ALDH7A1*) have therapeutic implications (i.e. identification of the gene directly informs treatment options). Other (treatable) pyridoxine responsive seizures include pyridoxal phosphate-responsive seizures (resulting from variants in *PNPO*) or pyridoxal phosphate-binding protein (PLPB) deficiency (resulting from variants in *PLPB*).

Boonsimma *et al.* included 103 cases with infantile-onset (age ≤ 12 months) pharmacoresistant epilepsy that was seen or referred for genetic testing at a Thai tertiary care centre.[30] Genes associated with genetic epilepsy syndrome (n = 728) were initially targeted using WES, followed by sequencing for additional selected candidate pathogenic variants in unsolved cases (only one heterozygous variant was identified).

Of 103 cases, 6 (5.8%) were identified as having biallelic variants in *ALDH7A1*; 4 (66%) were identified on initial WES and 2 (33%) required short- (*n* = 1) and long-read (*n* = 1) sequencing to identify two novel variants (all cases and both parents underwent sequencing). Two additional patients were identified as having 'treatable' disorders, one with a biallelic *PNPO* variant and one with a *BTD* variant. An additional 36 patients were identified as having genetic variants that could 'inform' treatment decisions.[36]

A total of eight *ALDH7A1* variants were identified in the six patients, five recurrent and three novel variants [two-thirds were copy number variants (CNVs)]. Of the five recurrent variants, two occurred in more than one patient within the study sample, suggesting a possible founder effect (one was identified in four patients, and one in two patients).

Clinical–biochemically defined pyridoxine-dependent epilepsy

Traditionally, PDE was clinically defined based on seizure recurrence (increase in number or severity) following pyridoxine withdrawal. Increasingly, urinary- or blood-based biomarkers [elevated alpha aminoadipic semialdehyde (α-AASA), piperidine-6-carboxylic acid and pipecolic acid concentrations] are used to define PDE and may help to distinguish it from other pyridoxine responsive seizures such as pyridoxal phosphate-responsive seizures (*PNPO* gene) or PLPB deficiency (*PLPB* gene). As PDE is an autosomal recessive disorder, those affected are usually children of unaffected carriers of the *ALDH7A1* variant such that FH is not a strong indicator of the likelihood of PDE, although more than one sibling in a generation can be affected. Further, there is an increased risk in consanguineous families. It is important to note that this category includes those who have either clinically, or biochemically defined PDE, or both, and this difference in disease definition may have an impact on results obtained from sequencing.

Koul *et al.* included all children (*n* = 35) from a single centre with refractory neonatal or infantile seizures that did not respond to antiepileptic drugs, but later responded to pyridoxine.[31] Only 7.9% (3/35) had elevated biochemical markers for PDE (pipecolic acid and AASA). All probands received 'targeted variant testing in *ALDH7A1*', but no further detail was given to describe the nature of this test. If the initial test was negative, WES was conducted on multiple genes, including *PLPBP*, *PRRT2* and *ALDH7A1*. The full spectrum of genes considered for WES was not reported.

Of the 35 cases, 4 (11%) were identified as having a variant in *ALDH7A1*, including the 3 with biochemical indicators of PDE, but it is unclear whether these variants were detected by the first-line testing strategy or by WES. The number of patients with homozygous and compound heterozygous variants was not reported. Twelve (34%) patients were identified as having a *PLPHP* variant, and 2 (6%) had a variant in the *PRRT2* gene, detected through WES. Specific variants were not reported.

Age at seizure onset and developmental delay were reported separately for patients positive for *ALDH7A1* variant (*n* = 4), those positive for a *PLPBP* variant (*n* = 12) and those with neither of these variants (*n* = 19). Age at seizure onset was lower for those with *ALDH7A1* variants (30 minutes–1 hour), and delayed development was reported in a higher proportion [75% (three-fourths)] compared to the other groups. For those with *PLPBP* variant, seizure onset was between 1 hour and 10 days old and 17% had developmental delay; for those with neither variant, seizure onset was reported at 4 hours–29 months and 37% (7/10) had developmental delay. It is not clear to what extent these differences are a result of selection into the study, and furthermore, based on the small numbers reported, results cannot be considered to be representative of the whole population of patients with clinically and biochemically defined PDE.

Gene (ALDH7A1)-positive pyridoxine-dependent epilepsy

The most tightly defined populations are those defined genetically. PDE–ALDH7A1 refers to PDE resulting specifically from variants in *ALDH7A1*. The primary focus for studies in this group is to characterise the genetic spectrum associated with *ALDH7A1* variants.

Two studies are reported. One reported an international study, including 185 patients with clinical suspicion of PDE and at least one confirmed pathogenic variant in *ALDH7A1*.[32] Participants were recruited from four clinical genetics laboratories that perform clinical testing of *ALDH7A1*, the international registry (which appears to cover North America, Europe, the Middle East and Australia) for PDE[37] and the pyridoxine-dependent seizures patient registry.[38] Specific details of the genetic tests used to identify genetic variants was not reported. The second study, based in China,

included 33 participants from 31 families, 31 with PDE–ALDH7A1 and 2 with PLPB variants.[33] Selection into the study appears to have been based on genetic testing results. A targeted polymerase chain reaction (PCR) panel was used to sequence each exon (1–18) and exon–intron boundary of the *ALDH7A1* gene.

Biallelic variants in *ALDH7A1* were identified in 98% (182/185) of patients,[32] resulting in 367 alleles with variants; 3 patients had only 1 variant allele identified (1.6%). The percentage of patients with compound heterozygous variants was 58.9% (109/185),[32] and 84% (26/31)[33] of the remaining patients were homozygous. Jiao *et al.* reported two (6.5%) patients as homozygous for *ALDH7A1*; however, the supplementary table to the report clearly reports five patients as having the same genetic variant on both alleles, indicating homozygosity (16% of total).[33] The reason for this discrepancy is not clear.

Types of variants were reported slightly differently between the studies; however, some similarities can be observed. In the registry-based study,[32] 209 (57%) of 367 variant alleles were missense [compared to 65% (17/26) of variants identified in Jiao *et al.*[33]], 66 (18%) were splicing errors [compared to 12% (*n* = 3) of variants[33]], 29 (8%) were inDel (insertion and/or deletion of nucleotides), 29 (8%) were single nucleotide variant (SNV) terminations, 18 (5%) were synonymous SNVs and 15 (4%) were CNVs. The remaining variants in Jiao *et al.*[33] were nonsense (8%; 2/26) or deletions (15%; 4/26).

Of the 26 different variants identified in Jiao *et al.*,[33] 9 were recurrent within the study population (7/17 missense and 2/3 splicing site), and 17 (65%) occurred in only 1 patient each. The two most commonly identified variants were observed in 23% (7/31) of all PDE patients (a missense variant) and 19% (6/31) of patients (a splicing site variant), respectively; remaining recurrent variants were observed in between 2 and 4 patients each.

The total number of recurrent variants identified in the study population was not reported by Coughlin *et al.*;[32] however, 4 individual variants accounted for 38% (140/367) of all variant alleles, 1 of which was identified in a quarter of all alleles (94/367, 25.6%). Forty-nine missense variants accounted for the 209 alleles with missense variants; however, the majority (65.3%; 32/49) of these were only identified in a single individual and 17 were recurrent (responsible for 177 of all variant alleles identified).

Heritable retinoblastoma

The three subgroups and studies for hRB were: any RB,[39] sporadic RB[40] and a combined category of familial, bilateral or germline RB.[41] Details per study are provided in *Appendix 5*, *Table 22*. All three studies were retrospective and were conducted at a single centre and all exclusively considered the RB1 gene. None of the selected studies exclusively used sequencing to identify variants.

Quality assessment

Quality assessment raised similar concerns about all three extracted studies for hRB.[39–41] Insufficient detail about the genetic testing methods used led to concerns about the ascertainment of variants by sequencing and about replication or application of results beyond the study. It was not possible to determine the applicability of study results to the review question, as techniques other than sequencing may have been used to determine the genetic status of the patients.

Summary of results across disease definitions

The percentage of RB cases with an identified germline *RB1* variant (i.e. hRB) using various testing strategies varied from 20.5%[40] for the most narrowly defined population (sporadic unilateral RB) to 44.2%[39] for the most broadly defined population (any RB). Both of these studies included participants with germline mosaicism (4.6%[39] and 27.5%[40] of germline cases), which is not always detectable in peripheral blood. Mosaicism is where a percentage of cells in the body carry the variant allele, but others carry a normal copy. Sequencing is often done on DNA extracted from peripheral blood samples, which means patients with mosaicism (especially low-level mosaicism) may not be detected by WGS strategies if the DNA extracted from the blood does not carry the variant allele.[42] Specific techniques, such as use of unique molecular identifiers and NGS are of interest for detecting mosaicism.[43] Results from the largest, most inclusive study[39] demonstrate that, without genetic testing, as few as 10% of RB cases might have been identified

for ophthalmologic surveillance from birth based on known FH of the condition. This study did, however, employ a stringent definition of familial RB.[39]

Salviat *et al.*[39] further demonstrated that as much as 38% of the total population (*n* = 517) had no RB variant identified (negative on germline screening and no tumour tissue available for testing). While it is likely that the majority of the 517 patients had somatic RB, a small proportion may have germline RB that was not detected by the genetic testing strategy. In Hulsenbeck *et al.*,[41] for example, 3 of the 821 patients at the centre were excluded as genetic data revealed no *RB1* variant and high *MYCN* (a different gene) amplification, suggesting a different genetic pathway to the development of RB that may not be identified using WGS.

In two studies that reported the variant types, the most common were nonsense and frameshift variants.[39,41] The percentage of those with bilateral RB was reported as higher in those with nonsense variants in both of these studies.[39,41]

Results for each subgroup by definition of disease

Any retinoblastoma

Retinoblastoma is a childhood-onset cancer of the eye caused by biallelic variants in the *RB1* gene. RB can be either heritable (hRB) or somatic. hRB occurs where a variant on one of the alleles is present from conception (either inherited from a parent or occurring sporadically) and is therefore present in every cell of the body (germline), and the second variant occurs within the cells of the eye at some point after conception (this can occur prenatally or at any time point after birth). Around 40% of all RB are heritable. Somatic RB occurs when both variants occur within the cells of the eye such that the variants can only be identified by genetic testing of the tumour tissue as opposed to testing peripheral blood samples. Patients with bilateral RB are frequently assumed to have hRB, the majority of which occurs sporadically; only approximately 10% of all cases of RB have a known FH of the disease (familial RB). Studies that include 'any RB' provide the best estimate of the percentage of RB patients who would benefit from newborn WGS.

Salviat *et al.*[39] included 1371 consecutive RB cases (including bilateral and unilateral, familial and non-familial) who successfully completed genetic counselling at a single centre. Multiple genetic screening methods were used dependent on the year of testing, but these included various combinations of: denaturing HPLC, quantitative multiplex PCR of short fluorescent fragments, methylation-restriction PCR, multiplex ligation-dependent probe amplification (MLPA), comparative genomic hybridization, Sanger sequencing and NGS. The promoter region and all exons with their flanking intronic sequences were screened. The identified pathogenic variants were classified as germline (identifiable in blood and therefore present from the point of conception) or somatic (occurring at any point following conception and therefore only identifiable in tumour tissue) and then as associated with the presence or absence (complete loss) of RB protein. Where tumour tissue was available, this was screened first to identify the *RB1* variants, with peripheral blood then tested for the identified variants to determine the germline (heritable) status (*n* = 293). The remaining patients with no tumour tissue available underwent germline screening (of peripheral blood) only.

Of the original 1404 eligible participants, 118 (8.4%) had a known FH of RB (defined as families with at least two germline carriers of a *RB1* pathogenic variant). Of those who completed the study, 44.2% (606/1371) were found to have a germline *RB1* variant (hRB), including 497 with bilateral RB and 109 with unilateral RB. Germline mosaicism was identified in 28 of 606 (4.6%). Of the 765 patients with no germline variant identified, 248 (32.4%) were identified as having somatic RB (biallelic variants in tumour tissue only) and 517 (67.6%) were negative on germline screening (no tumour tissue available). Of those with germline variants identified, the majority (561/606, 92.6%) were identified by the germline screening strategy, and 7.4% (*n* = 45) were identified through the first-line tumour screening strategy. The most common variants (comprising 77.2% of identified germline variants) were nonsense (222, 36.6%), frameshift (140, 23.0%) or out-of-frame splice variants (110, 18.2%), all of which are associated with a loss of RB protein (total of 537/606 identified germline variants were associated with loss of RB protein).

Among those with hRB (germline variants) (*n* = 606), the incidence of bilateral RB was higher in patients with variants associated with a loss of RB protein (84.2%; 452/537) compared to those with variants with no loss of RB protein (65.2%; 45/69) ($p = 0.01$). Germline variants associated with a loss of RB protein were also associated with earlier mean age at diagnosis of RB ($p < 0.001$), and later stage at diagnosis ($p = 0.047$) compared to variants not associated with the complete loss of RB protein.

Sporadic retinoblastoma

Sporadic RB occurs where there is no FH of disease, and it can be either heritable (the variant is present in the germline and can therefore be passed on) or somatic (occurring only in the tumour). Children who develop sporadic germline RB (hRB) are the population with the greatest potential to benefit from newborn screening with WGS, as regular intensive surveillance can be initiated to allow earlier detection and treatment. This disease definition identifies the additional patients who might undergo surveillance for RB as a result of WGS. It is worth noting, however, that WGS is still of benefit to those with a known FH of RB because genetic testing will identify those who do not need to undergo such intensive surveillance as the *RB1* variant has not been passed on.

The study by Temming *et al.*[40] included 195 patients who presented with unilateral sporadic RB and underwent genetic testing at the request of patients or their legal guardians. Only those with ophthalmological follow-up until age 5 years were eligible for inclusion. One or more of the following methods was used for genetic testing of blood or tumour tissue: analysis of allele loss in tumours, cytogenetic analysis, denaturing HPLC, exon-by-exon sequencing, MLPA, methylation-sensitive PCR, quantitative fluorescent multiplex PCR, quantitative real-time PCR, real-time PCR and single-strand conformation polymorphism. Forty (20.5%) patients were identified as having a germline *RB1* variant (hRB), 29 (72.5%) of whom had a heterozygous *RB1* variant and 11 (27.5%) had germline mosaicism (which can be passed on to offspring of the proband if it is found to be present in the particular germ cell that forms the embryo).[44] Of those with heterozygous germline variants (hRB), 10 (34%) were classified as whole gene deletions, 13 (45%) as premature terminations and 6 (21%) as 'mild' variants.

Of the 195, 9 (4.6%) developed bilateral RB during 5-year follow-up, 8 of whom had a heterozygous RB1 variant (3 with whole gene deletions and 5 with premature terminations) and 1 had germline mosaicism.

Heritable/bilateral retinoblastoma

Heritable RB includes cases who have a FH of RB (familial or inherited germline variant), and those whose germline variant in *RB1* occurs sporadically at the time of conception (sporadic germline variant). Most patients with bilateral RB have hRB such that studies frequently defined study eligibility as familial or bilateral RB as a proxy for identifying cases of germline hRB. Consideration of studies in this category provides an indication of the most commonly found types of variant; however, in many studies, multiple genetic tests are used to characterise the genetic spectrum and it is not always possible to identify those variants that would be most easily identified on WGS.

Hulsenbeck *et al.*[41] included 287 cases with RB from a total population of 815 patients (342, 42% of whom were identified as having a germline variant). Patients with confirmed heterozygous pathogenic constitutional *RB1* variants (hRB) who had not previously undergone ophthalmological screening for familial RB were included in the study report. Those with mosaicism and non-*RB1* variants were excluded. Multiple genetic screening methods were used, some of which included sequencing or sequential testing. A retinoblastoma variant effect class (REC) was developed to classify the identified pathogenic variants according to their effect on RB protein structure and quantity (i.e. extent of loss in RB protein), from REC-I (largest effect) to REC-V (smallest effect).

The most common variants (comprising 98.6% of identified variants) were nonsense or frameshift variants (REC-I) (199, 69.3%), whole *RB1* gene deletions (REC-II) (39, 13.6%) and missense or in-frame SNVs (REC-III) (45, 15.7%). Of those with whole *RB1* gene deletions (n = 39), 27 (69%) were identified as also having deletion of the *MED4* gene (hypothesised as being associated with lower penetrance of RB).[45]

The percentage of bilateral RB was highest in those with nonsense or frameshift variants (REC-I) (186, 93.5%) compared to those with REC-II (30, 76.9%) and REC-III (36, 80.0%). Age at diagnosis was lowest in patients with REC-I variants [median: 7.3 months (range: 0.2–48.0)], followed by REC-II [10.3 months, (0.4–40.9)] and REC-III [11.6 months (0.9–45.1)] variants.

X-linked hypophosphataemic rickets

The XLHR is one form of hereditary rickets (see *Report Supplementary Material 1*). Clinical symptoms of rickets are not specific to XLHR but are characteristic for a much broader range of conditions. Some conditions can be distinguished based on biochemical markers, while others require information on the inheritance pattern or genetic testing. The

proportion of rickets patients identified by sequencing the *PHEX* gene will therefore depend on whether disease is defined clinically, biochemically or genetically. The findings of studies on *PHEX* frequency in children with the disease defined in four different ways are summarised in *Appendix 5*, *Table 23*.

Overall, six studies were selected across four different categories of hereditary rickets. One study was identified and described for the broad hereditary rickets category,[46] and two studies were included for each of hypophosphataemic rickets[47,48] and hereditary hypophosphataemic rickets[49,50]. The largest study in these two categories only tested for the *PHEX* gene, while the second largest study tested for additional genes, thus providing information about the *PHEX*-negative cases, which could either be undetected *PHEX* cases or caused by different genes. One study was selected for the XLHR category.[51]

Four studies were conducted at a single centre,[46–49] while one study included patients from all paediatric hospitals in Norway[50] and one study was an international registry study.[51] Three studies were prospective[46–48] and three were retrospective.[49–51] Sequencing techniques were employed to identify variants in the three prospective studies, while both sequencing and MLPA were used in two retrospective studies. The registry study of XLHR patients did not specify the genetic test(s) used.[51]

Quality assessment

The quality assessment using the Murad tool raised concerns across the different dimensions assessed. For selection of patients in both Gaucher *et al.* and Jacob *et al.*, it is unclear whether the included participants were representative of all eligible patients.[46,48] The ascertainment of patients in Rafaelsen *et al.* is unclear.[50] Except for Marik *et al.*, all studies did not describe the cases in sufficient detail to enable other investigators to replicate the research.[46–51]

In three studies, concerns regarding the applicability of study findings to WGS was low because the reported genetic spectrum was based on sequencing that more closely resembles WGS in the screening application,[46–48] while concerns were high in the two studies that used MLPA in addition to sequencing[49,50] and unclear in one study where 'genetic testing' was not further specified.[51]

Summary of results

The range of genes identified varied across the different categories of rickets. The frequency of *PHEX* varied from 33% in the most broadly defined population[46] to 89.7% in the most narrowly defined population.[51] 'Pre-screening' of the population using biochemical testing resulted in a greater proportion of patients with *PHEX* in both the hypophosphataemic rickets category[48] and the hereditary hypophosphataemic rickets category.[49]

The extent of test negatives (no confirmed pathogenic variant identified) varied from 3%[49] to 21.1% in Rafaelsen *et al.*,[50] which was largely due to the testing strategy used [number of genes considered, extent of sequencing, sequencing method (e.g. WGS vs. Sanger sequencing) and additional testing]. The same study sequenced the untranslated 3-prime region as well as exons and intronic regions in recognition that XLHR may be caused by variants in the regulatory region of the messenger ribonucleic acid.[50] MLPA was used in both Del Pino *et al.*[49] and Rafaelsen *et al.*[50] to overcome limitations of sequencing to detect deletions and duplications. Neither study reported the proportion of variants detected by MLPA in addition to those detected by sequencing.

The XLHR is characterised by many different variants precluding the investigation of expressivity of specific variants, with one study reporting that there is no clear genotype–phenotype link.[50] Furthermore, a great proportion of variants in each study was novel, which presents challenges for the confirmation of variant pathogenicity.

Results for each subgroup by definition of disease

Hereditary rickets

This category is an umbrella term of any type of hereditary rickets (excluding nutritional rickets). It consists of two main types: (1) vitamin-D-dependent rickets (low phosphate levels secondary to vitamin D deficiency) and (2) hypophosphataemic rickets (low phosphate is the primary defect and rickets is therefore vitamin-D-resistant). Each type consists of several conditions caused by a series of genes with similar clinical symptoms. Some have distinct biochemical markers.

Jacob *et al.*[46] conducted a comprehensive assessment to identify the genotypic spectrum of rickets in 10 Indian families, with 10 patients suspected of having hereditary rickets. All 10 patients had symptom onset in childhood; however, 2 patients did not receive a diagnosis until early adulthood. Exome sequencing identified variants in six different genes, including 3 patients (3/10, 33%) with a *PHEX* variant. The results revealed three known truncating variants, c1482 + 5G > C, c1586_1586 + 1del and c.58C > T. The c1586_1586 + 1del variant resulted in a more severe phenotype compared to the other two variants. Other implicated genes included *CYP27B1* (*n* = 3 patients), *CYP2R1*, *VDR*, *SLC34A3* and *SLC2A2* (all one patient each). No cases had an unidentified genetic cause.

Hypophosphataemic rickets

In hypophosphataemic rickets, low serum phosphate due to renal losses is the primary defect, which is either mediated by the hormone fibroblast growth factor 23 (FGF-23) (raised FGF-23 levels) or is independent of FGF-23 (normal FGF-23 levels). Measuring levels of FGF-23 can aid the distinction between the two types. There are 15 genetically distinct disorders that are grouped into FGF-23-dependent and FGF-23-independent hypophosphataemic rickets.[52]

Gaucher *et al.* analysed the *PHEX* gene in 118 families, including 56 familial and 62 sporadic cases using classical sequencing. Sequencing covered all 22 exons, intronic regions and the region at the 3-prime end, which is not translated into a protein but serves regulatory processes.[48] The study was conducted at a single centre but encompassed a multiethnic population, comprising individuals of European, North African, Caribbean and Asian backgrounds. The inclusion criteria were based on low serum phosphate and ratio of tubular maximum reabsorption rate of phosphate to glomerular filtration rate (TmP/GFR) levels, bone deformities and radiological evidence of rickets. *PHEX* variants were found in 78% (93/118) of probands. The 93 variants comprised 78 different variants of which 60 (77%) were novel. Variant types included nonsense (28%), frameshift (30%), splice-site variant (23%) and missense variant (19%).

Some uncertainty regarding the pathogenicity of novel *PHEX* variants was noted. One patient with a novel c.1206A > G variant (a single nucleotide replacement) also exhibited a second missense variant. In two other cases, both patients with the c.505G > A variant (another single nucleotide replacement) harboured additional variants, one involving an insertion leading to frameshift and the other a deletion leading to frameshift. The findings emphasise the significant role of *PHEX* in X-linked dominant hypophosphataemic rickets and suggest that family members should be screened when a *PHEX* variant is found in a sporadic case. Additionally, when a missense variant is detected, a search for another *PHEX* variant should be conducted.

The study's approach to sequence only the *PHEX* gene meant that a large proportion of *PHEX* negative cases remained unexplained. For 3/25 negative cases, a reason was proposed. Missing PCR samples for two exons in one case led to the conclusion that one patient had a large deletion. One patient was subsequently diagnosed with a different type of hypophosphataemic rickets, and one patient was diagnosed with tumour-induced osteomalacia with secondary hypophosphataemia.

Marik *et al.*[47] screened 66 consecutive Indian patients with refractory hypophosphataemic rickets using WES. Patients were characterised by a lack of healing despite treatment with cholecalciferol and had lower-than-normal phosphate levels for their age. The mean age of onset of symptoms was 22.5 ± 14.3 months; 24/66 (26.4%) patients had a confirmed *PHEX* variant, and 40/66 (60.6%) patients had a confirmed variant in a different gene. The remaining two cases had a variant of unknown significance (VUS) which was classified as negative and for whom genetic testing could not confirm the diagnosis. All 24 *PHEX* variants were different and 13/24 were novel.

In the context of expressivity, one patient carried two *PHEX* variations [c.2048T > A; *p*.(Leu683His) and c.2071-1G > C], whereas her mother, who was clinically mildly affected, had only one *PHEX* variation [c.2048T > A; *p*.(Leu683His)]. This difference may explain the variability in disease severity between them.

Hereditary hypophosphataemic rickets

Hereditary hypophosphataemic rickets is marked by increased FGF23 activity, leading to hypophosphataemia due to renal phosphate wasting. Genes associated with hereditary hypophosphataemic rickets include *PHEX*, *FGF23*, *ENPP1*, *DMP1* and *FAM20C*, which are clinically and biochemically similar but follow different inheritance patterns.

Del Pino *et al.* included 96 patients diagnosed with hereditary hypophosphataemic rickets, of whom 42 underwent molecular testing of *PHEX* by Sanger sequencing and MLPA to detect gene deletions and duplications.[49] The condition was characterised by the typical presence of combination of clinical, laboratory and radiographic findings. Deleterious sequence alterations or large deletions in the *PHEX* gene were identified in 85.7% (36/42) of patients. The remaining six patients were not genetically confirmed.

Rafaelsen *et al.*[50] investigated 28 Norwegian children with hereditary hypophosphataemic rickets from 19 families. Inclusion was based on serum phosphate levels below the age-dependent reference range combined with tubular reabsorption rate of phosphate (not due to hyperparathyroidism). Sanger sequencing and MLPA analysis (to look for deletions and insertions in sequencing negative patients) of the *PHEX* gene was conducted. This was followed by Sanger sequencing of *FGF23*, *DMP1*, *ENPP1KL* and *FAM20C* successively in *PHEX*-negative patients. Overall, 13/19 (68.4%) probands were identified with *PHEX* variants. The 13 variants were all different and none were novel. There was one variant each in *FAM20C* and *SLC34A3*; 4/19 (21.1%) had no confirmed variant. Exploration of the effect of different types of variants (missense vs. nonsense) revealed no difference in clinical outcomes.

X-linked hypophosphataemic rickets

X-linked hypophosphataemic rickets is the most common form of hereditary hypophosphataemic rickets. It is caused by variants in the *PHEX* gene, which is inherited in an X-linked dominant fashion. Genetic testing or knowledge of a FH with a typical X-linked dominant inheritance pattern can support the diagnosis of XLHR.

Ariceta *et al.*[51] is a registry-based study, including multinational data of 579 participants with XLHR. XLHR diagnosis was based on the clinical judgement of an XLH-treating expert physician (FH, clinical, radiological and biochemical findings) and/or by genetic testing. However, genetic testing was not required for patient registration; 282 of children underwent genetic testing, which was not further described. A total of 89.7% (253/282) children tested were identified with a variant in the *PHEX* gene. There was a small number of patients with different genetic disease who were incorrectly registered, including four with a variant in *FGF23*, one in *SLC34A3* and seven where the gene was not specified; 17/282 (6.0%) had no variant confirmed. The study only reported genetic findings at the gene level. While symptoms were not reported by *PHEX* variants, the study concluded overall that children with XLHR diagnosis despite early detection and treatment did not do too well.

Familial haemophagocytic lymphohistiocytosis

Haemophagocytic lymphohistiocytosis (HLH) is not a single disease but a syndrome that is associated with several heritable and non-heritable conditions. Symptoms and biochemical markers are, therefore, non-specific and do not aid in the differential diagnosis, making a population of patients with fHLH difficult to define based on the clinical and biochemical characteristics. Historically, HLH was divided into primary (early onset, genetic condition) and secondary HLH (later onset, secondary to underlying medical condition such as cancers, infections or autoimmune disorders) using the main underlying trigger of symptomatic disease to define subgroups. More recently, the boundary between primary and secondary HLH has blurred with a better understanding of the complexity of the syndrome (discovery of new genes involved, differences in severity, genetic involvement in secondary HLH and detection of digenic disease). It is more accepted now that primary HLH is an artificial and ill-defined category, which was reflected in the published studies and could not be adopted here.

The fHLH is used to describe a subset of primary HLH disorders caused by biallelic variants in the four genes *PRF1*, *UNC13D*, *STX11* and *STXBP2*. Distinguishing between the four conditions is not feasible using clinical symptoms or biochemical markers, and this relies on genetic testing. The three categories and representative studies were, therefore, HLH, more broadly,[53] fHLH encompassing the four genes *PRF1*, *UNC13D*, *STX11* and *STXBP2*[54] and any one of the four conditions, of which we identified studies for three of the four genes.[55–57] The findings of the studies are summarised in *Appendix 5*, *Table 24*.

Four of the studies were retrospective and multicentre studies,[53,54,56,57] while Amirifar *et al.* was a systematic review.[55] The test was well defined in one of the five studies, which specified different sequencing methods (Sanger sequencing, NGS and WES) for patients with different indications based on biochemical assays.[54]

Quality assessment

The quality assessment of Amirifar *et al.*[55] was conducted using the ROBIS-2 tool for systematic reviews,[19] while other studies were evaluated using a modified Murad tool. We observed that the four primary studies lacked clear or sufficient information on patient selection.[53,54,56,57] The studies provided limited details regarding the inclusion of patients from different centres and did not report the time frames of recruitment. Amirifar *et al.*[55] presented an unclear risk of bias due to insufficient details on the review's conduct, study selection and data synthesis.

Summary of results

In patients with HLH symptoms that met the HLH diagnostic criteria of the Histiocyte Society, a genetic diagnosis of fHLH was established in one-third of the patients.[53] This means, a third of clinically defined HLH patients could be detected by sequencing the four fHLH genes; 7% of patients would be missed because they carried variants in different genes, 3.5% would be missed because the genetic cause could not be identified, 10% would be missed because they had a monoallelic disease and 46% would be missed because they had a non-genetic disease. The number of patients with monoallelic disease identified in three out of five studies led to the suggestion in one study that there is a gene–dosage effect, which means that fHLH can no longer be regarded as a simple recessive disease.[53] This needs to be considered in the interpretation of sequencing results and is further complicated by the occurrence of digenic disease (monoallelic variants in two of the four genes). However, the risk of disease in patients with monoallelic disease is currently unknown. Across all four conditions, there is some indication that homozygosity is associated with earlier onset and more severe disease[54] and severity is linked to the type of variant[55] but that there is no clear genotype–phenotype link, as siblings with the same genotype displayed different phenotypes and different ages of disease onset.[53] Overall, the studies suggest that there are a few common variants that are linked to particular ethnicities, highlighting the need to understand the genetic disease in a broad spectrum of patients, which is applicable to the screening setting before considering the implementation of sequencing as a screening tool in a diverse population.

Results for each subgroup by definition of disease

Haemophagocytic lymphohistiocytosis

Haemophagocytic lymphohistiocytosis is a hyperinflammatory syndrome generally defined by the diagnostic criteria recommended by the Histiocyte Society including symptoms of fever, splenomegaly, cytopenia, elevated cytokines and haemophagocytosis.[58]

Cetica *et al.*[53] is an Italian registry-based study that analysed 500 HLH patients over 25 years. The multiethnic study included patients of southern European, Eastern European, African, Asian and Hispanic origin. In 426 patients who underwent sequencing, a genetic diagnosis was possible in 171/426 (40.1%) of patients, while 43/426 (10.1%) had a monoallelic disease and 197/426 (46.2%) were believed to have a non-genetic disease; 15/426 (3.5%) cases with assumed genetic disease were missed. Of 171 patients with a genetic diagnosis, 141 (82.5%) had fHLH with *PRF1* and *UNC13D* variants, accounting for 131/141 (92.9%) of fHLH cases. The 69 patients with *PRF1* and the 62 patients with *UNC13D* biallelic disease carried 34 and 37 different variants, respectively. A small number of variants were common occurring in up to 19 patients. The data on disease onset in sibling pairs revealed that there is no clear genotype–phenotype relationship. In 9/26 sibling pairs, disease onset varied up to 17 years, while in 1 pair, one sibling developed the disease at 6.7 years, while the other remained unaffected at 25 years. The number of patients with monoallelic disease indicates that HLH likely results from both genetic predisposition and exogenous triggers.

Familial haemophagocytic lymphohistiocytosis

Familial HLH is an artificial category of four conditions caused by variations in four distinct genes. Patients with different fHLH subtypes are clinically and biochemically similar.

Shabrish *et al.*[54] investigated 101 molecularly confirmed fHLH patients, of whom 98 were under the age of 18 years, over 10 years, from 20 referral centres in India. 86/98 (87.8%) had biallelic disease; 12/98 (12.2%) patients had a monoallelic disease and would be considered test negative on sequencing. The *PRF1* and *UNC13D* variants accounted for 70/86 (81.4%) of cases with biallelic disease. Molecular analysis revealed that missense variations were the most common type of variation in all four genes. The number of different variants was significant (25 different variants in 34 patients with *PRF1* and 28 different variants in 23 patients with *UNC13D* fHLH). Patients with homozygous variants

across all four genes had an earlier disease onset (median 10 months) compared to those with compound heterozygous variants (median 3 years).

Single gene/variant

The narrowest category consists of patient populations that had confirmed fHLH caused by one specific gene, while patients with confirmed variants in other genes were excluded. The studies' aim was to characterise the variant spectrum in this tightly defined patient population.

Trizzino *et al.*[57] included 124 patients with confirmed biallelic *PRF1* disease and a median age of disease onset of 3 months from six different international centres. They detected 63 different variants: 11 nonsense, 10 frameshift, 38 missense and 4 in-frame deletions; 15/63 variants were novel. The most common single variant was a missense variant in 32 patients. Specific *PRF1* variants were strongly linked to Turkish, African American and Japanese ethnic groups. Patients with two disruptive variants had a younger age at onset than patients with missense variants only.

The systematic review by Amirifar *et al.*[55] analysed clinical features, immunologic data and genetic findings from 57 articles covering 322 patients with *UNC13D* variant with a median age of onset of 6 months; 269/322 (83.5%) had biallelic disease, 50/322 (15.5%) had monoallelic disease, and for 3 patients, this information was not reported. Missense variations were the most common type of variation. Severe features appeared to be associated with a homozygous genotype and missense variants. Splice-site errors and compound heterozygosity were more prevalent in patients with mild features.

Pagel *et al.*,[56] a multinational study with patients included mainly from Germany and Turkey, included 37 patients with confirmed biallelic *STXBP2* variants. One of the 37 patients was 19 years of age at the time of diagnosis. Nine novel variants were reported. Variants included nine different missense variants, four different splice-site variants and several small deletions or insertions. Three variants were seen in more than 5 patients; 13/37 patients carried one of two splice-site variants affecting exon 15. The exon 15 splice-site variant was associated with mild disease and an atypical disease course. These patients often experienced chronic, recurrent episodes with long periods without HLH symptoms, and their reactivations typically responded to steroids-only treatment or underwent spontaneous remission.

Medium-chain acyl-CoA dehydrogenase deficiency

The MCADD is part of a group of conditions called inborn errors of metabolism (IEM). MCADD symptoms overlap with those of multiple IEM. Differential diagnosis can be achieved by determining the disease-specific biomarker profile, relevant enzyme activity levels and the underlying genetic variant. Because MCADD is on the current NBS screening panel, MCADD can be defined based on screening outcomes as positive on the initial MS/MS screening test and as positive on confirmatory serum and urine tests.

The five disease definition categories for MCADD (from broadest to narrowest) and representative studies were: IEM/inherited metabolic disease (IMD),[59] fatty acid oxidation disorders (FAOD),[60] ACAD deficiencies,[61] clinical–biochemical MCADD[62,63] and gene (*ACADM*)-positive MCADD.[64] The two studies in the clinical–biochemical MCADD category included one that detected cases through NBS screening (genetic spectrum in MS/MS positives and in follow-up positives)[62] and one that diagnosed patients clinically or biochemically following family screening.[63]

Details per study for each of the five disease definitions considered are provided in *Appendix 5*, *Table 25*. The genetic testing method was not reported in two studies,[59,63] three used a sequencing technique[60,61,64] and one used sequential testing.[62] All six studies were conducted retrospectively.

Quality assessment

Three of the studies met all the quality assessment criteria.[60–62] We had concerns regarding the selection criteria for sequencing for the study by Touw *et al.*,[64] and reporting was considered inadequate in the study by Mesbah *et al.*[63] Applicability to the review question was unclear in two studies, because the testing method was not reported.[59,63]

Summary of results across disease definitions

Five of the six studies described above reported the experience of using genetic confirmatory testing in national newborn screening programmes. The percentage of patients with a genetic MCADD diagnosis varied from 15%[59,60] to 19%[61] in populations with a broad disease definition and to 85% in a population of patients with biochemically confirmed MCADD.[62] As biochemical testing is readily available for MCADD, the latter category is the most relevant to consider. Genetically confirmed disease was less common in those with only an initial positive NBS screening test for MCADD compared to those confirmed on second-tier or follow-up biochemical testing.[61] Testing for only two common *ACADM* variants had a low yield in initially NBS test MCADD-positive babies [8/511 (1.6%) biallelic and 157/511 (30.7%) monoallelic],[62] however, this strategy could not be fully evaluated because results were not reported for those with biochemical MCADD that was confirmed on follow-up testing.

Results from the study by Mesbah *et al.*,[63] which included patients with a clinical diagnosis of MCADD, suggest that 18.0% (almost one in five) of MCADD cases cannot be detected through genetic screening. However, it is important to note that the sample size was small (with only 17 cases) and the genetic testing method used was not reported, so it is not clear how applicable these results are to WGS.

Of the studies that presented variant frequency, the 985A > G variant was found to be the most common.[59,62–64] In the study restricted to genetically confirmed cases of MCADD, 61.8% were homozygous with this variant.[64] However, all four studies were from countries with populations of mainly Caucasian origin. The only study from Asia did not report this variant in any of their three genetic MCADD cases. *ACADM* variants present little heterogeneity in the studies (17 genotypes in 68 genetically confirmed MCADD cases); however, this may be misleading as ethnicity appears to affect the genotypes detected.[64]

Little information regarding expressivity was presented in the included studies. Generally, the evidence points towards some variants causing 'milder' disease based on residual enzyme activity studies, and it appears that those were newly detected in the screening context,[64] highlighting that detecting genetic variants in a screening context is not the same as finding them in a diagnostic context.

Results for each subgroup by definition of disease

Inborn errors of immunity/inherited metabolic disease

Inborn errors of metabolism, also known as IMD, includes a group of approximately 600 conditions that are individually rare but collectively common. They are hard to diagnose, given the non-specific symptoms that many affected patients' experience.[65] IEM can be caused by variants in different genes that affect the same metabolic pathway at different stages, resulting in different conditions with similar symptoms. Overall, IEM are very heterogeneous, resulting in groups of disorders affecting different metabolic pathways with different epidemiology, presentation and heritability.

Martin-Rivada *et al.*[59] reported details for 224 Spanish newborns who underwent genetic testing following biochemical indication of an IEM as part of the national screening programme. The original cohort included 902 consecutive newborns with an initial abnormal NBS test result. The molecular genetic testing method was not reported; however, 30 different genes were included. Of the 224 babies, 222 (99.1%) were diagnosed with a genetically confirmed IEM; 2 participants were considered to have biochemical hyperphenylalaninemia (no variants identified). In the wider group of initial NBS-positive babies, this equates to 24.6% (222/902) with genetically confirmed IEM. Of those with a genetic variant, 19.3% (43/222) were identified with a variant in the *ACADM* gene. Among these 43 children, 14 different genotypes were identified. The 985A > G variant made up 70% of all alleles (60/86) in 43 newborns; 22 newborns were homozygous and 16 were compound heterozygous for this variant. The remaining five cases were compound heterozygous for other variants. Only one patient (homozygous for 985A > G) showed symptoms of MCADD before newborn screening results were available.

Fatty acid oxidation disorder

Fatty acid oxidation disorders are a particular group of IEM caused by variants of the genes associated with the metabolic pathway of fatty acids in the mitochondria. Symptoms overlap and different FAOD are identified by their specific acylcarnitine (fatty acid metabolites) profiles and can be confirmed by gene sequencing.

Maguolo *et al.*[60] included 30 Italian patients with FAOD; 20 were infants diagnosed following NBS screening and 10 were clinically diagnosed with a mean age at onset of 29 years and therefore excluded from review. Five of the 20 infants were biochemically positive for MCADD. Sequencing was conducted using a custom-designed FAOD panel, including 15 genes. Biallelic MCADD was identified in 15% of newborns (3/20) or 60% (3/5) infants with biochemical MCADD; while 2 infants carried monoallelic ACADM variants and would be classed as negative on genetic testing. Fifteen patients (75%) had different FAOD subtypes (5/15 with genetically confirmed disease; 1/15 with monoallelic disease and 10/15 without genetic information). Of six *ACADM* variants in those with biallelic disease, five were different. One infant was homozygous for the 985A > G variant and two infants were compound heterozygous carrying a total of four different variants. All three infants with biallelic *ACADM* variants had a residual MCAD enzyme activity of < 5% associated with severe disease.

Acyl-CoA dehydrogenase deficiencies

The FAOD subgroup of ACAD deficiencies are caused by variants in 11 genes, with *ACADM* being one of them.

Wang *et al.*[61] reported genetic testing results for 20 newborns with ACAD confirmed on diagnostic biochemical testing from a cohort of 83 newborns with an initial positive screening result for ACAD deficiency on NBS screening. Out of the 20 newborns, 4 had biochemically confirmed MCADD. High-throughput sequencing and Sanger sequencing was used with a wider IEM panel of 306 genes. Three of the 20 (15%) newborns with confirmed ACAD were identified with biallelic compound heterozygous variants in the *ACADM* gene. In the wider group of babies with an initial positive screening test, this would be 3.6% (3/83). One newborn was test negative on sequencing due to a monoallelic variant. Therefore, genetic testing identified three-fourths of patients with biochemical MCADD. The three biallelic MCADD patients carried five different variants, of which two were novel. None of the Chinese newborns carried the 985A > G variant common in European cohorts.

Clinical–biochemical medium-chain acyl-CoA dehydrogenase deficiency

This disease definition includes those with a diagnosis of MCADD, either confirmed by clinical characteristics or by biochemical testing. Ideally, we would have considered the two categories separately to investigate the impact of different disease definitions on the genetic spectrum of disease. However, studies tended to include a mix of patients or poorly defined the populations included.

We included two studies for this disease definition category, one of which included patients identified with MCADD through a NBS programme.[62] The other study in this subgroup included children who were clinically diagnosed with MCADD or detected through family screening and presumably confirmed biochemically.[63]

Nichols *et al.*[62] included 511 newborns with NBS octanoylcarnitine (C8) levels ≥ 0.3 µmol/l who were subsequently referred for molecular genetic testing using a sequential sequencing method. First-tier testing specifically aimed to identify two of the most common variants prevalent in the USA c.985A > G and c.199T > C, second-tier testing used full *ACADM* sequencing in those with at least one variant or those without variant detected on first-tier testing but repeat C8 levels of at least 0.4 µmol/l. Mesbah *et al.*[63] included 17 children younger than 18 years of age with clinically diagnosed MCADD; 4 were diagnosed via family screening and 2 via post-mortem. The genetic testing method was not reported. Both studies solely considered the *ACADM* gene.[62,63]

The percentage of patients positive (biallelic *ACADM* variants) on first-tier screening for two variants initially was 1.6% (8/511 with MS/MS-positive screening test) in Nichols *et al.*[62] A further 157/511 (30.7%) of newborns were monoallelic, and 83/511 (16.2%) had neither variant on first-tier testing. In 20 newborns with positive clinical follow-up, sequencing of the full *ACADM* gene revealed 17 (85%) with biallelic *ACADM* variants and 3 with monoallelic variants classified as test negative. Mesbah *et al.*[63] reported genetic test results of 14 of 17 included patients, 11 of whom had biallelic *ACADM* variants; and 3 cases were missed by sequencing as only one variant was identified. The most common variant in these two studies from the USA and Ireland was 985A > G (12/17 and 11/11,[62] either homozygote or compound heterozygote, respectively).[62,63] Nichols *et al.*[62] reported 13 different variants of which 5 were novel. They reported that the c.199Y > C/c.134A > G genotype resulted in 'mild' MCADD.

Gene (ACADM)-positive medium chain acyl-CoA dehydrogenase deficiency

This disease definition includes those with confirmed *ACADM* variants and provides useful information on variant frequency.

The study by Touw *et al.*[64] included 68 children from the Dutch newborn screening programme with confirmed variants in the *ACADM* gene by sequencing all exons and adjacent intron regions. Seventeen different genotypes were reported, of which 7 were novel. Most of the children (42/68, 61.8%) were homozygous for the 985A > G variant; a further 20 were compound heterozygous, including the 985A > G variant. The authors categorised genotypes into 'classic' (previously recognised in clinically confirmed cases, n = 53) and 'variant' (genotypes not previously recognised in clinically confirmed cases, n = 15), and they reported median residual MCAD enzyme activity of 0% for the former and 25% for the latter group. This may support the theory that screening identifies disease with milder MCADD due to genotypes not recognised in clinical cases.

Evidence on early versus late treatment

Question 4: What is the evidence on early [following screen detection or sibling detection (cascade testing)] versus late (following clinical presentation) treatment?

In the absence of RCTs investigating the outcomes of pre-symptomatic versus symptomatic treatment, we defined early versus late treatment for the five selected conditions based on the natural history of the disease (e.g. age of symptom onset and progressive vs. relapsing/remitting conditions) and the type of available treatment/management (e.g. preventative vs. symptom management). *Table 9* summarises the natural history for the five conditions that informed our definitions of early versus late. *Table 10* summarises the management strategies for the five conditions, and *Table 11* details out our definitions of early versus late based on the available treatment studies. The process illustrates that the different condition–treatment pairs require individual definitions of early versus late.

TABLE 9 Overview of the natural history of the five conditions

Condition	Natural history/characteristics of condition
PDE	Progressive
	Deficiency of the enzyme α-AASA dehydrogenase, which is involved in the breakdown of lysine in the brain, results in the accumulation of metabolites that inactivate pyridoxine. Pyridoxine depletion causes intractable neonatal seizures that become recurrent and prolonged if left untreated (with pyridoxine). Uncontrolled seizures can ultimately lead to death, but this appears to be less common. Classic PDE usually presents during the neonatal period (i.e. within 28 days of birth) with prolonged seizures that are difficult to control with anti-seizure medication; in 75% of cases, seizures may occur within the first few hours of life. These seizures last for several minutes and involve loss of consciousness, spasticity and convulsions. If untreated, periods of encephalopathy are common (irritability, crying, fluctuating tone and poor feeding). In some cases (up to 30%), affected individuals do not experience seizures until they are 1–3 years old (late-onset PDE). Intellectual disability and developmental delay are often present (around 75% of cases), especially in those with classic PDE
	More than 75% of PDE patients have IDD, which is not thought to be correlated with lack of seizure control. IDD may be due to the accumulation of neurotoxic metabolites associated with α-AASA dehydrogenase deficiency (LRT is targeted at reducing impact of this on outcome)
hRB	Progressive
	Disrupted function of the tumour suppressor protein RB results in cancer of the eye, which can be heritable (either familial or sporadic) or somatic and non-heritable. Bilateral disease is usually considered heritable, but unilateral can also be heritable
	Heritable RB usually occurs at an average of 15 months of age and may be picked up by targeted ocular screening before any symptoms develop if there is FH of the disease. The most common first symptom of RB is leukocoria or visible whiteness of the pupil, which may be noticed in photographs taken using flash photography. Other common symptoms include strabismus (squint), proptosis [protruding eye(s)], glaucoma and hypopyon (presence of pus). If the tumour is large, the eye may become painful and inflamed. High-risk features on presentation (e.g. optic nerve invasion) are more common with increasing age and are associated with poorer outcome. If RB is left untreated, blindness can occur and metastases will most likely develop. Studies have reported detection of RB at birth in some infants

continued

TABLE 9 Overview of the natural history of the five conditions (*continued*)

Condition	Natural history/characteristics of condition
XLHR	Progressive
	Loss of function of the PHEX protein leads to increased FGF23, which consequently decreases renal phosphate reabsorption, which increases urinary phosphate excretion and decreases calcitriol production, leading to hypophosphataemia and other imbalances
	Hypophosphataemia manifests as rickets (i.e. lower leg deformities (from 6 months), waddling gait, progressive lower leg deformities, delayed gross motor development, widening of the distal metaphysis at the wrists and ankles [from age 1 year, dental abscesses/malpositions (3 years+), stunted growth, bone pain and hearing loss (older children)]. Ikegawa *et al.*[66] noted that while lower leg deformities may start at 6 months–1 year, it's typically only recognised when toddlers start walking
	Four criteria are used for clinical diagnosis: hypophosphataemia, increased levels of serum alkaline phosphatase, decreased tubular reabsorption of phosphate and radiological evidence of rickets
	For all four markers to be positive can take up to 6 months and infants develop them inconsistently
	Generally, biochemical markers are detectable before radiological evidence.
	Presentation with symptoms is often later when stunted growth and bowed legs become apparent
fHLH	Condition requiring activation/reactivation to become symptomatic
	Malfunction of the perforin/granzyme cytotoxic pathway leads to a proliferation of lymphocytes and overactive macrophages impeding normal downregulation of immune response macrophages and subsequently to an escalation of the immune response, including abnormal targeting of red blood cells and cytokine storms, leading to anaemia and organ damage
	Hyperinflammatory syndrome with several genetic causes
	Usually, it activates in infancy (during the first year of life in 70%, though later onset is possible and timing of onset may be dependent on fHLH subtype/variant) following infection. The most common symptoms include: fever that does not respond to antibiotics, rash, hepatomegaly, splenomegaly, enlarged lymph nodes and can lead to seizures, unconsciousness and coma. It is fatal if untreated and reactivates when immune system is triggered again
MCADD	Condition requiring activation/reactivation to become symptomatic
	Inactivity or deficiency of the medium-chain acyl-CoA dehydrogenase (MCAD) protein prevents catalysing the beta-oxidation of fatty acids for formation of ketone bodies in the liver, leading to lack of an alternative energy source during periods of prolonged fasting or increased energy demands resulting in metabolic crises
	Individuals present as healthy at birth; first clinical presentation is typically between 3 and 24 months of life (a reduction in overnight feedings can trigger onset of symptoms or common infections in previously asymptomatic individuals), some may present with a metabolic crisis in the neonatal phase before screening results would be available. Others do not present with symptoms until childhood or even adulthood. Most symptomatic cases present from 3 months to 3 years of age
	Under normal conditions, MCADD patients can use alternative glycogen stores elsewhere in the body. A person with MCADD who never experiences low blood sugar would never experience symptoms
	Up to 25% of MCADD-affected individuals will die during their first clinical manifestation
	Clinical presentation: unexplained lethargy, vomiting, altered consciousness, hypoglycaemia, encephalopathy which may progress to seizures and coma
	Not all individuals with MCADD develop such a clinical presentation, whether a severe metabolic crisis or milder symptoms. This causes some confusion as to whether the MCADD phenotype consists of specific clinical signs and symptoms or the biochemical evidence of an enzyme disorder.[67]
	Because signs and symptoms are highly variable, no consistent definition of what is regarded as symptomatic MCADD is used in the literature

IDD, intellectual developmental delay.

TABLE 10 Overview of management strategies for the five conditions

Condition	Treatment category (used to select conditions)	Treatment relevant to screening context	Treatments not considered
PDE	Low cost to both NHS and patient/family	First screening-relevant Rx: guideline recommendations are for pyridoxine initiation following PDE–ALDH7A1 diagnosis to prevent or reduce the severity of further seizures. Diagnosis usually occurs following a pyridoxine trial for intractable seizures or based on the combination of seizure and biochemical indicators of PDE, rather than following a genetic diagnosis. The guideline does not explicitly mention asymptomatic treatment initiation; however, theoretically, asymptomatic treatment should prevent seizures or reduce severity if seizures occur. Expert advice suggests that treatment is not usually initiated asymptomatically because of potential harm from high-dose pyridoxine; those who are identified as gene positive are more likely to be treated on first seizure Second screening-relevant Rx: LRT (lysine-restricted diet and arginine supplementation) aims to reduce the accumulation of potentially neurotoxic metabolites (e.g. pipecolic acid or α-AASA, among others) in order to prevent or reduce developmental delay.[68] LRT can be initiated soon after birth for infants who are not being breastfed as there are lysine-restricted formula milk options Risk of Rx: the most well-known adverse effect of pyridoxine is sensory neuropathy; however, this requires very high doses and seems to have been primarily reported in adults.[69] Clinical expert advises that due to reports of toxicity associated with high-dose Rx in infants/children (e.g. pyridoxine-induced seizures), asymptomatic treatment is unlikely to be initiated even where there is a known *ALDH7A1* variant LRT: intolerance to a lysine-free amino acid formula, or severe adverse effects (nutritional, neurological or other) (diet should be terminated. In the former case, a natural protein-restricted diet may be considered[70]	Some children require a combination of pyridoxine and standard anticonvulsants for adequate seizure control
hRB	Long-term surveillance with the associated anxiety and costs	First screening-relevant Rx: surveillance to identify clinically presenting RB as early as possible (e.g. with fewer tumour foci). (Surveillance is usually based around red reflex testing, which begins after birth and may be repeated every few months until the child is 5 years old. Children with dim or absent red reflex are referred to a specialist ophthalmology service for eye examination under general anaesthetic. In some countries, children with a FH of RB are recommended to have more intensive ophthalmologist surveillance from birth.) Second screening-relevant Rx: for clinically presenting cases, 'early' treatment is where local therapies, for example, cryotherapy or laser treatment, can be used to firstly save vision in the affected eye, and ultimately, saving the eye (prevent enucleation). In some cases, these focal therapies may be used alongside chemotherapy. Focal treatments are considered curative, but a risk of recurrence or development of RB in the other eye remains,[71] particularly for heritable forms, and children with RB will undergo continuous follow-up care Risk of Rx: no obvious clinical risks other than that associated with general anaesthetic; all surveillance programmes carry some risk of harm (e.g. psychological distress), particularly for those who do not go on to develop the condition. Risks associated with focal therapy include retinal thinning and hole formation, retinal detachment, vitreous condensation and vitreous haemorrhage with tumour seeding and cataract[72]	Radical Rx: enucleation, radiotherapy or chemotherapy

continued

TABLE 10 Overview of management strategies for the five conditions (*continued*)

Condition	Treatment category (used to select conditions)	Treatment relevant to screening context	Treatments not considered
XLHR	High- and long-term costs to both patient/family and NHS	First screening-relevant Rx: Pi/D for newborns/asymptomatic patients = supplements to address low levels of phosphate and active vitamin D and prevent symptoms to develop/severity (traditional first-line treatment, can be administered to newborns or before the development of clinical or radiological signs of rickets, started once diagnosed either by biochemical markers, radiographic images or onset of clinical symptoms, ± genetic testing), feasibly at 2–3 months of age based on hypophosphataemia and increased levels of alkaline phosphatase[73] or following genetic confirmation,[74] treatment is preventative Second screening-relevant Rx (but not licenced for < 1-year-olds): burosumab = monoclonal antibody to neutralise FGF23 to prevent low levels of phosphate and active vitamin D (could be given pre-symptomatically as first-line treatment according to clinical advisor, so probably a future treatment), only curative treatment Risk of Rx: toxicity of alfacalcidol (vomiting, diarrhoea, nephrocalcinosis, nephrolithiasis and reduced kidney function)	Pi/D regimen for symptomatic disease Cinacalcet (suppresses parathyroid hormone secretion, which influences serum phosphate concentration, suggested as an adjunct treatment) Vitamin D-only (dated management strategy) 25-hydroxycholecalciferol (=circulating form of vitamin D which is converted to calcitriol the active from. Does not work in vitamin D-resistant rickets because a defect in the conversion of vitamin D(3) to its active 25-hydroxy metabolite is not the cause of XLHR) Paricalcitol (a vitamin D analogue that prevents and treats secondary hyperparathyroidism) Recombinant human growth hormone (never first-line treatment) Corrective surgery (late treatment option when nothing else has worked, not relevant to screening context but could be used as outcome measure for earlier treatments)
fHLH	Short-term high costs to NHS but long-term lower costs to NHS and patients	First screening-relevant Rx: allogenic HSCT (± prophylactic treatment up until HSCT) with matched related (first line) or unrelated (second line) donor. Once engrafted, new stem cells can differentiate into healthy blood cells and restore normal immune response. Includes a pre-transplant conditioning regimen: aims to immunosuppress host to that the graft takes. Conditioning regiments can be myeloablative (full dose, carries higher risks) or 'reduced intensity'. Different drug combinations are used in conditioning regimens, which is usually specified in the studies. Only curative treatment Risk of Rx: Conditioning: drug toxicity HSCT itself: graft vs. host disease, transplant-related mortality [85–90% success rate (alive, cured)] HSCT is the only pre-symptomatic treatment option. No official guidance about pre-symptomatic treatment: risks of HSCT need to be weighed against risk of waiting until activation[75] Second screening-relevant Rx: prophylactic treatment to reduce the chances of disease activation while waiting for HSCT	Chemoimmunotherapy of active disease: induction therapy following the HLH-2004 protocol (HLH-1994) to suppress life-threatening inflammation by targeting the abnormally activated immune cells: 8 weeks of chemoimmunotherapy to control inflammation and then continued therapy until HSCT is available
MCADD	Existing screening and treatment pathways, so impact of WGS would be incremental	Screening-relevant Rx: dietary advice consisting of avoidance of fasting, strict feeding regimens and emergency regimen, with or without low fat diet or dietary supplementation of L-carnitines, preventative measure to avoid low blood sugar levels, triggering decompensation (metabolic crisis) Risk of management: none, but low compliance may be an issue	Carbohydrates given by mouth/IV administration of dextrose solution (treatments for symptoms following a trigger/treatments to reverse catabolism and prevent metabolic crisis)

IV, instrumental variable; Pi/D, oral phosphate and calcitriol; Rx, treatment.

TABLE 11 Definition of early vs. late treatment initiation for the five conditions

Condition	Aim of screening	Definition of 'early'[a]	Definition of 'late'
PDE	To identify PDE in patients before their first seizure. While pyridoxine may not be given pre-symptomatically, it can be given immediately at the time of first seizure without trials of ineffective anticonvulsants	1. Asymptomatic or pre-clinical (i.e. prior to first seizure) 2. After first seizure (if FH present or known to be *ALDH7A1* +ve) 3. After trial of single anticonvulsant	Depends on definition of early 1. Any clinical presentation (a–c below) 2. After trial of one or more anticonvulsants (b or c below) 3. After trial of multiple anticonvulsants (c below)
		The asymptomatic stage ranges from a few hours to days (too short for screening to identify asymptomatic disease), or, less typically, months after birth. There are also reports of unusual fetal behaviour indicating likely antenatal seizures. Some family studies document antenatal treatment with pyridoxine or asymptomatic, prophylactic treatment of second-born siblings; however, treatment is more usually initiated soon after first seizure (if sibling is affected or known to be *ALDH7A1*-positive) or following failure to control seizures with anticonvulsant medication (i.e. within days/weeks of first seizure). Pyridoxine is unlikely to be a first-line treatment for neonatal seizures per se	Clinical presentations, include: a. After first seizure b. After trial of single anticonvulsant c. After trial of multiple anticonvulsants
hRB	To identify the predisposition of RB in patients before cancer becomes symptomatic to allow surveillance for timely management of early signs of cancer to improve prognosis	1. Surveillance from birth, either because gene-positive or known FH 2. Outcomes in patients with known FH (this option presumes at least some will be screen detected, or detected early because of family awareness of the risk of RB) 3. Detection of RB at stage A or B (the earlier the stage at which RB is detected, the greater the chance of local therapeutic options being feasible, or at the very least avoidance of enucleation) 4. RB by age on presentation (e.g. < 1 year) (age at diagnosis is a proxy for earlier detection; however, children can present with more aggressive or advanced tumours even at very young ages)	1. No surveillance, clinical presentation of RB 2. Outcomes in patients with no FH (presume no screening) 3. Detection of RB at stages C, D, E 4. RB by age on presentation (e.g. ≥ 6 months or 1 year)
XLHR	To identify XLHR in patients before rickets signs present and treat hypophosphataemia to prevent manifestation of rickets	1. Asymptomatic or pre-clinical (i.e. before signs of rickets like leg bowing, stunted growth, waddling gait become apparent). Early can mean getting detected based on (1) genetically confirmed XLHR, (2) abnormal biochemical markers and/or (3) radiological signs of rickets through screening or knowledge of FH with relevant inheritance pattern 2. Age < 1 year at treatment start (onset of main clinical symptoms from year 1 onwards) to increase the chance that the early group includes asymptomatic or 'early symptomatic' patients (unless all < 1 year olds are symptomatic indicated by treatment not first-line, i.e. on treatment at enrolment, treatment switching reported or type of treatment not screening relevant)	1. Symptomatic or with clinical signs of rickets (symptoms specified, treatment not first-line, patients < 1 year excluded) 2. Age > 1 year at treatment start

continued

TABLE 11 Definition of early vs. late treatment initiation for the five conditions (*continued*)

Condition	Aim of screening	Definition of 'early'[a]	Definition of 'late'
fHLH	To identify the likelihood of HLH due to a genetic variation before disease activation by infection to consider curative HSCT. Pre-symptomatic treatment aims to minimise the time the hyperinflammatory response can be triggered, which requires chemoimmunotherapy with limited treatment success	1. Pre-symptomatic (= before first disease activation), identified through screening [*PRF1*, *STX11*, *STBPB2* or *UNC13D*-positive/any biochemical screening (i.e. reduced/absent perforin expression; reduced natural killer cell activity)] of either newborns or individuals identified through FH 2. Age < 3 months at HSCT (diagnosis to HSCT in 6–8 weeks should be achievable)	1. After first disease activation (regardless of whether patient is still in active disease or in remission) 2. Age > 3 months at HSCT
MCADD	To identify MCAD deficiency in patients before the first metabolic crisis occurs. Pre-symptomatic detection allows management of the deficiency by avoiding fasting as compared to treating life-threatening metabolic decompensation events	1. Pre-symptomatic, i.e. before first trigger (low blood sugar levels that cannot be compensated) resulting in decompensation, identified either through biochemical screening, genetic screening or a combination of both (neonatally or because of FH) 2. Age < 6 months at start of dietary advice	1. Symptomatic, i.e. with clinical symptoms of active disease 2. Age > 6 months at treatment start

a Option 1 is the preferred option. All subsequent options (2 onwards) are considered a step down and were only taken forward to data extraction if no studies for option 1 existed.

Appendix 5, *Tables 26–30*, summarises the outcomes from studies investigating early versus late treatment as defined by the studies for the five conditions. Case studies comparing outcomes in early versus late-treated siblings are presented in *Appendix 5*, *Table 31*. The detailed quality assessment is included in *Appendix 3*.

Pyridoxine-dependent epilepsy

There were five studies that compared relevant outcomes in children 'treated early' for PDE and children 'treated late' for PDE (see *Appendix 5*, *Table 26*).[68,76–79] Two studies included a series of families[76,77] and three were single-arm studies where patients were recruited retrospectively.[68,78,79] Treatments included pyridoxine monotherapy,[76,77] LRT[78] or a combination of the two.[68,79] The definition of 'early' and 'late' treatments varied between the studies. Jiao *et al.*[77] and Tseng *et al.*[79] defined 'early-treated' as the sibling with the shortest delay in receiving treatment after seizures began. In Bok *et al.*,[76] the impact of treatment given antenatally (or asymptomatically) was compared to symptomatic treatment initiation, and the remaining two studies used age at treatment initiation to differentiate early and late treatments.[68,78] The number of PDE cases in the included studies were low and ranged from 4 (from 2 families)[76] to 60.[68]

None of the studies reported seizure control as an outcome but instead focused on developmental outcomes. Three studies reported intelligence quotient (IQ) scores;[76,78,79] one reported a 'standardised developmental assessment' score (similar to IQ score),[68] and one reported psychomotor development.[77] Other reported outcomes included motor performance[76] and a standardised neurological outcome.[78] The timing of the outcome assessment was not reported in two studies[77,78] and ranged between 4 years (age of one child)[76] and a mean of 15 years of age[79] in the other studies.

Quality assessment

We had methodological concerns regarding the selection of patients for all studies[76–79] apart from the one by Coughlin *et al.*[68] The selection process was not clearly reported in these studies and no recruitment dates were given. The follow-up period was not reported in the study by Strijker *et al.*[78] and, in Jiao *et al.*,[77] one child was only 6 months old at the last follow-up. In addition, the description of the assessment processes was inadequate in the latter study.[77]

Findings

There is some indication of improved developmental outcomes in 'early-treated' children. In studies that reported IQ scores[76,78,79] or a quantitative measure of developmental assessment,[68] scores were higher (suggesting improvement) in the 'early-treated' groups, apart from those in the study by Tseng *et al.*[79] When treated with pyridoxine and LRT, the full-scale IQ score was slightly lower in the 'early-treated' group (76.0 vs. 77.4). There were suggestions of improved psychomotor development[77] and motor performance[76] in 'early-treated' groups; all three children in the 'early-treated' group were assessed as normal compared to only one child assessed as normal in the 'late-treated' group in Jiao *et al.*,[77] and one 'early-treated' child showed a slightly better outcome than their 'late-treated' sibling in the study by Bok *et al.*[76] (walking 4 months earlier).

There was insufficient evidence to determine whether there were any apparent differences as a result of the type of treatment (pyridoxine monotherapy, LRT or a combination of the two).

No studies that looked at outcomes in 'early-treated' patients alone were identified.

Heritable retinoblastoma

Five studies compared relevant outcomes in children who underwent pre-emptive surveillance to allow an early detection of RB ('treated early') compared to those who presented clinically ('treated late') (see *Appendix 5*, *Table 27*).[80–84] All five studies were retrospective and single-armed.[80–84] Two studies[81,82] were multicentre and three were single-centre studies.[80,83,84] The number of included RB cases ranged from 13[84] to 264.[80] Four[81–84] of the five studies included children with a FH of RB, and the fifth study[80] included all cases of RB, but the main comparisons of relevance that were reported were in those with a FH of RB (with or without surveillance).

Four studies[80,82–84] compared outcomes in 'screened' versus 'not screened' at the participant level, one of which[83] further categorised the 'screened' group into intensively screened and screened. In the remaining study by Chantada *et al.*,[81] the income status of the country was used as a proxy for 'screened' versus 'not screened', since it was assumed

that more children in a high-income country will be detected by screening than their counterparts in low- and middle-income countries.

Quality assessment

We had concerns regarding whether the length of follow-up was long enough for outcomes to occur in the study by Moll *et al.*[82] In all studies apart from Rothschild *et al.*,[83] reporting was not sufficient to allow others to replicate the research.

Findings

There was some variability in reported outcomes between studies. In three studies, the mean or median age at diagnosis was lower in screen-detected versus clinically detected RB, suggesting that screening does allow an earlier detection of RB than would happen otherwise (mean 4.9 months vs. 17.2 months,[82] mean 4.7 months vs. 16.7 months[84] and median 0 months for intensively screened, 4 months for screened and 9 months for not screened.)[83] In terms of patient health outcomes, some measure of both ocular survival and survival was reported in three studies on a per participant level.[80,83,84] The largest study[80] using Kaplan–Meier analysis demonstrated considerably higher rates of ocular survival at 1 year for the screened FH group (n = 86) compared to the not screened FH group (n = 178) (83.2% compared to 47.5%), with a smaller difference at 5 years (67.7% compared to 58.2%). In contrast, however, a marginal difference in survival at 1 year (100% vs. 97.3%) had increased slightly by 5 years (93.2% vs. 87.4%), potentially suggesting a benefit from earlier detection.[80]

Similar results were reported by Rothschild *et al.*,[83] with enucleation rates of 0% (0/16) for those intensively screened, 8.7% (2/23) for those 'screened' and 65% (13/20) for those not screened, and mortality rates of 0% for both screened groups and of 5% (1/20) for the not screened group (follow-up time point not reported). The smallest study reported enucleation rates of 0% for those screened (0/5) and 75% for those not screened (6/8), with no deaths (median follow-up of 4.8 years).[84]

The final study reported considerably lower enucleation rates in the USA (25%; 8/32) compared to those in developing countries (71.7%; 43/60) and a considerably higher probability of event-free survival at 5 years {0.92 [standard deviation (SD) 0.05] in the USA compared to 0.81 (SD 0.07) in developing countries}.[81] It is likely that other differences in the delivery of care have contributed to the observed differences in outcomes, and it is not possible to properly attribute these differences to the effect of screening alone.

'Early-only' studies

In addition, we identified three studies reporting on 'early treatment' only, that is identification of RB via surveillance.[85–87] Results from two of the three studies identified high proportions of infants with RB present in the first 1–2 weeks after birth [70% (12/17)[85] and 50% (4/8)[86]]. The third study reported screening of 23 asymptomatic siblings of probands with RB; 13% (3/23) were identified as having active RB on screening at a median of 6 months of age.[87] It is not possible to determine whether the use of surveillance in these studies resulted in 'better' outcomes than would have occurred if they had presented clinically.

X-linked hypophosphataemic rickets

There were three studies that compared the outcomes in children with XHLR who underwent early treatment with those who received treatment later.[50,88,89] All studies implemented standard treatments for rickets {i.e. daily oral phosphate, vitamin D or an analogue of vitamin D (alfacalcidol) [oral phosphate and calcitriol (Pi/D)]} as opposed to the more recently licensed burosomab. No studies using burosumab as first-line treatment were identified. *Appendix 5*, *Table 28* provides the summary of their findings.

All three studies were retrospective and were conducted without control groups. Two studies involved multiple centres[50,89] and one[88] was conducted at a single centre. All three studies used the same definition of early (before 1 year of age) versus late treatment (at or after 1 year of age). In one study, patients in the early treatment group were diagnosed prior to the onset of clinical signs of rickets, while those in the late treatment group were diagnosed after the appearance of clinical symptoms.[88]

Study outcomes included measurements of height and various biochemical parameters such as serum calcium, phosphate, ALP, creatinine, parathyroid hormone and vitamin D3 levels. Makitie *et al.*[88] and Quinlan *et al.*[89] assessed the activity of rickets through radiographic examination. Additionally, Makitie *et al.*[88] made predictions regarding adult height. The evaluation of results occurred at different time intervals across the three studies. Makitie *et al.*[88] measured outcomes at the end of the first year of treatment and before puberty. Rafaelsen *et al.*[50] conducted measurements at each clinic visit and assessed the results at the last recorded consultation. The mean age at the last recorded consultation of the early group was documented as 11.1 years, while the late group had a mean age of 8.4 years. Quinlan *et al.*[89] analysed outcomes at medium treatment durations of 8.5 years and 11.9 years for early and late treatment groups, respectively.

Quality assessment

The evaluation of study quality revealed that Quinlan *et al.*[89] lacked sufficient information on patient selection, while it remained uncertain whether the follow-up duration in Rafaelsen *et al.*[50] was adequate to assess outcomes. The study by Makitie *et al.*[88] satisfactorily addressed all aspects of the quality assessment.

Findings

Study findings suggest that initiating treatment during the early stages of growth moderately enhances outcomes for patients with XLHR. Both Makitie *et al.* and Quinlan *et al.* reported median height closer to the expected average in the early treatment compared to the late treatment groups [SD scores of −0.7 (n = 8) vs. −1.8 (n = 11), p = 0.009;[88] and −0.7 (n = 10) vs. −2.0 (n = 13), p = 0.009[89]]. Additionally, Makitie *et al.* reported pre-pubertal height closer to the expected average in the early (−1.3 SDS; n = 8) compared to the late treatment group (−2.0 SDS; n = 11) (p = 0.054). Treatment effect in the early-treated group (z-score −0.2), compared to the late-treated group (z-score −1.2), was within that expected by chance (p = 0.06) that was observed by Makitie *et al.*[88] By contrast, results from Rafaelsen *et al.*[50] showed the same trend as Makitie *et al.*[88] but were within that expected by chance.

Regarding biochemical parameters, Makitie *et al.*[88] found a similar degree of hypophosphataemia between groups; however, serum alkaline phosphatase levels remained elevated in the late treatment group throughout childhood. Quinlan *et al.*[89] reported no difference in the median levels of serum phosphate or serum alkaline phosphatase between the early and late treatment groups.

In terms of rickets activity, Makitie *et al.*[88] observed more pronounced radiographic signs of rickets in the late treatment group, whereas patients receiving early treatment still displayed significant skeletal rickets changes. Quinlan *et al.*[89] reported similar ricket severity scores in both groups.

'Early-only' studies

In addition, we identified three studies only reporting on early treatment.[73,90,91] Exploration on early treatment indicated that early treatment could enhance metabolism, promote growth and mitigate deformities. However, these findings are not entirely consistent with the findings of the studies reporting a comparison of early versus late treatment.

Familial haemophagocytic lymphohistiocytosis

One retrospective, multicentre study that evaluated outcomes for children who received treatment for fHLH following asymptomatic detection before activation (treated 'early') and children who received treatment following clinical symptomatic detection or activation of fHLH (treated 'late') was identified[75] (see *Appendix 5*, *Table 29*). The study included 32 genetically confirmed sibling pairs/triplets with fHLH. The asymptomatic children were diagnosed following the diagnosis of their sibling. Outcomes included mortality, cause of death and number of patients in complete remission at the end of follow-up. The follow-up period for each patient was different.

Quality assessment

The primary quality concern in Lucchini *et al.*[75] was the lack of detail regarding which international centres provided the data and whether they sent in details for all eligible patients in their care. The patient characteristics, treatment and outcomes were reported adequately, and the follow-up time was sufficient to assess treatment outcomes.

Findings

We present results separately for per-protocol and intention-to-treat populations using a slightly different definition for 'early' and 'late' treatment.

Intention-to-treat population

For the intention-to-treat population, we defined 'early' and 'late' treatments as cases that were asymptomatic and symptomatic at diagnosis regardless of activation before or after treatment initiation. In other words, the 'early'-treated group in the intention-to-treat population included patients who were asymptomatic at diagnosis but with some subsequently experiencing activation either before or shortly after commencing treatment, and this may confound the true effect of 'early' treatment on outcomes.

Mortality was lower among the 'early'-treated group (15% vs. 38%). Six patients [two (8%) in the 'early'-treated group and four (15%) in the 'late'-treated group] died due to transplant complications, which was the most common cause of death. The proportion in complete remission was higher in the 'early'-treated group (81% vs. 62%).

Per-protocol population

Ideally, 'early' treatment would begin in individuals before their first activation. However, individuals who present as asymptomatic may have experienced disease activation before or shortly after the initiation of treatment. Therefore, for the per-protocol population, we defined 'early' and 'late' treatment using four distinct groups, dependent on the activation status. The 'early'-treated group can be split into two categories: asymptomatic patients who were treated with HSCT ± prophylactic treatment and did not activate (group 1, n = 15) and asymptomatic patients who were treated with prophylactic treatment and subsequently activated (group 2, n = 3). The 'late'-treated group was separated into those who were symptomatic and were subsequently treated with active disease protocol ± HSCT (group 3, n = 26) and those who were asymptomatic at diagnosis but experienced disease activation before the start of treatment and thus were treated with active disease protocol ± HSCT (group 4, n = 7). Defining the treatment groups in this way helps to reduce any potential confounding effects that activation may have had on outcomes and allows us to better isolate the impact of 'early' treatment of truly asymptomatic individuals on clinical outcomes.

Mortality was lowest among group 1 (7%) and was similar between the remaining three groups (group 2 33%, group 3 38% and group 4 27%). Note that group 1 reflects those with less severe disease than group 2, since these patients did not activate after treatment initiation. Also, group 2 only included three patients, one of whom died. The most common cause of death was disease progression. Three patients died following transplant complications (four in group 3 and two in group 4). The number of patients in complete remission was highest in group 1 (93%). Proportions were considerably lower in the other three groups (group 2 33%, group 3 10% and group 4 7%). However, group 1 had the shortest median follow-up time, so it is possible that patients in this group subsequently experienced symptoms or other events. One patient in group 2 was lost to follow-up.

Group 3, who were symptomatic at diagnosis, and likely represent those with most severe disease, had the highest proportion of deaths and lowest proportion of patients in complete remission at the end of follow-up. Outcomes in the groups who had experienced activation following an asymptomatic diagnosis either before (group 4) or after (group 2) the start of treatment were similar to group 3. The opposite was true, however, for those in group 1, which indicates that 'early' treatment before activation has been experienced may be effective. Further, only three asymptomatic patients with no previous activation subsequently activated after commencing 'early' treatment. As previously mentioned, this reduces our confidence in results from group 2, but does, however, increase confidence in the conclusion that treatment in asymptomatic individuals with no previous activation may be effective.

Medium-chain acyl-CoA dehydrogenase deficiency

There were nine studies that looked at relevant outcomes in children with MCADD who received early management following asymptomatic detection through screening (treated 'early') and those who received management following clinical symptomatic detection of MCADD (treated 'late') (see *Appendix 5*, *Table 30*).[92–100] All studies, apart from the study by Abdenur *et al.*,[92] were single-arm and all were conducted retrospectively. Three Australian studies had overlapping patient cohorts.[96,98,99] The sample sizes in the included studies ranged from 2[92] to 90.[94,101]

The definitions of 'early' and 'late' treatment varied slightly between the studies. In eight of the nine studies (all apart from Wilson *et al.*[100]), 'early' treatment was defined as management following detection of MCADD through NBS screening. One of the eight studies also included those detected through family screening.[93] In six of the eight studies,[92,94–96,98,99] patients were all asymptomatic, but in two studies, this group also included symptomatic patients.[93,97] In the study with data for only two patients, one patient received dietary management from 5 months following diagnosis by NBS screening, and the other patient received management following symptom presentation.[92] In Wilson *et al.*,[100] 'early' treatment was management following asymptomatic screening due to an affected sibling.

In all nine studies, 'late'-treated was defined as management following clinical presentation of MCADD.[92–100] The study by Anderson *et al.*[94] also included those detected through family screening in the 'late' group, and it is unclear how many patients this group included and whether they were symptomatic or asymptomatic. It is important to bear these differences in the definitions of 'early' and 'late' treatments in mind when interpreting results. Where reported, management strategies included avoidance of fasting,[93,94,97,98,100] carnitine supplementation,[93–95,97,100] various diets[92,94,95,97,98] and sick-day regimens.[98,100]

The reported outcomes were heterogeneous and included mortality/severe episodes, descriptions of patients' clinical statuses, various measures of physical and psychological development, healthcare use and biomarker levels. Results for each of these outcomes are described separately below. Follow-up length was varied, but two studies reported outcomes within the first 4 years of life.[96,98]

Quality assessment

In four studies, we found that cases were not described in sufficient detail to allow for replication of the study or to make inferences, and this was mostly due to the paucity of treatment definition and description.[95,96,98,102] In four studies, the method for selection of patients was unclear.[92,93,95,102] The follow-up times were variable, and this was unknown in one study.[93]

Findings

Mortality/severe episodes

Four studies reported mortality as an outcome.[95,96,98,100] This outcome was assessed at 4 years of age in two studies,[96,98] and follow-up was variable in the other two studies.[96,100] In two studies, the proportion of patients who died by the age of 4 years was lower in the 'early'-treated groups (4% vs. 17% in Haas *et al.*[96] and 4% vs. 19% in Wilcken *et al.*[98] in the early- and late-treated groups, respectively). In the study by Wilson *et al.*,[100] 21% and 17% of children had died by the age of 6 years in the early-treated and late-treated groups, respectively. However, it is important to note that in this study, the definition of early treatment was not those detected by NBS screening, and instead, this group included those diagnosed due to an affected older sibling. No patients died in the study by Gong *et al.*[95] Wilcken *et al.*[98] reported the number of severe episodes by the ages of 2 and 4 years. At both ages, the percentage who had experienced a severe episode was lower in the 'early'-treated groups.

Description of clinical status

Four studies reported the clinical status of patients after varying follow-up periods.[92,93,95,102] Where reported, age at assessment ranged between 24 months[92] and 11 years.[102] No studies assessed patients after a standardised follow-up period. In the study by Li *et al.*,[102] all six children (four 'early'-treated and two 'late'-treated) were assessed as normal. In two studies, the 'early'-treated groups were assessed as normal, but some children in the 'late'-treated groups showed poorer outcomes.[92,95] One of the two studies had only two patients; the late-treated child showed clinical abnormalities, including severe seizure disorder and cerebral palsy, and nasogastric feeding in one child.[92] In the second study, one late-treated patient reported intermittent fasting hypoglycaemia and one experienced hemiplegia due to disease episode.[95] In the study by Alcaide *et al.*,[93] nine (29%) patients were symptomatic in the 'early'-treated group, and both children in the 'late'-treated group were symptomatic.

Physical and psychological outcomes

Various measures of physical and psychological outcomes were reported, and overall, there appeared to be few differences between the 'early'- and 'late'-treated groups. There were no differences in terms of height, weight or

neuropsychological function (within the first 4 years of life) in one study,[96] or in terms of intellectual ability score (assessed at > 4 years of age) in another study.[98] In Wilcken *et al.*,[99] one child (4%) in the late-treated group had a mild intellectual handicap, and two (8%) required extra assistance at school, whereas the children in the 'early'-treated group were all assessed as normal (at 6 years of age or last follow-up).

Healthcare use

Four studies reported some measure of number of hospital/emergency room (ER) visits.[94,96,98,100] The mean [95% confidence interval (CI)] number of hypoglycaemia-related hospital days and ER visits per patient-years was slightly lower in the 'early'-treated group [0.09 (0.03 to 0.15) vs. 0.11 (0.04 to 0.19)] in the study by Anderson *et al.*[94] However, in this study, it is important to note that the 'late'-treated group may have included asymptomatic, screen-detected patients. In two studies, the percentage of patients who had previously been admitted to hospital was lower in the 'early'-treated groups [42% vs. 85% (by age 4)[98]] and 25% versus 36% (variable follow-up).[100] Haas *et al.*[96] reported whether the hospital visits were inpatient, emergency or outpatient within the first 4 years of life. The percentage of children with inpatient stays and emergency room visits was lower in the 'early'-treated group, but the number of outpatient visits was higher in this group. Length of inpatient stay in those admitted was also similar between the 'early'- and 'late'-treated groups.

Biomarker levels

Biomarker levels at diagnosis were reported in Gong *et al.*[95] Mean levels of C6, C8 and C10 were lower in the 'early'-treated children, but there were no differences in the mean C8: C2 or C8: C10 ratios. However, it should be noted that these measures were reported inconsistently.

'Early-only' studies

In addition, we identified nine studies only reporting on early treatment in which MCADD patients were detected through NBS screening or family studies and no patients were detected and treated following symptom onset.[60,103–109] One further study identified five cases admitted to hospital before their NBS test results became available, but outcomes were not reported separately for the NBS screening group.[64]

Conclusions and learning from the review of five conditions

The five traditional reviews yielded insufficient evidence in populations relevant to a screening context to inform a UK NSC decision about WGS in newborns. The lack of evidence on the penetrance and expressivity of the genetic variants for all five conditions does not appear to support the identification of specific conditions or groups of individuals for which WGS may be beneficial in a screening context.

Available evidence about the genetic spectrum of patients with symptomatic or biochemical disease frequently pointed to small numbers of recurrent variants accounting for considerable proportions of cases; nevertheless, large numbers of variants were novel or occurred only in a very small number of cases. The pathogenicity and penetrance of such variants are unclear. We also saw that genetic spectrum can be strongly affected by factors such as ethnicity and the prevalence of consanguinity. Studies often employed a suite of genetic testing methods rather than relying on sequencing alone such that we could not determine the proportion of cases that could have been identified with WGS alone. As such, the results observed are not necessarily transferable to the screening context of apparently healthy newborns.

In terms of identifying benefit from earlier treatment, our results generally reflect acknowledged difficulties in evaluating the effectiveness of interventions for rare conditions. Studies were generally small, with variable definitions of 'early' and 'late' both within and between conditions and with a frequent reliance on an earlier intervention in siblings or where there was a known FH. Only a few examples of asymptomatic or very early initiation of treatment were identified. Outcome measures were often short term, with limited follow-up to identify longer-term patient-relevant outcomes such that any benefits from an earlier diagnosis resulting from newborn screening will be difficult to quantify.

A few aspects were identified that would render certain conditions less likely candidates for a newborn screening programme using WGS:

- The genetic heterogeneity is large and novel potentially pathogenic variants are common, causing a lot of uncertainty which has downstream implications in terms of time needed to determine pathogenicity.
- The types of frequent variations (mosaicism, large deletions, etc.) are not sufficiently captured by WGS and require additional genetic testing.
- The pre-symptomatic phase is likely to be shorter than the time until test results are available for diagnosis.
- An early intervention phase cannot be defined.
- The preferred curative treatment option is not licensed for newborns.
- The curative treatment option carries a high risk of adverse events.
- The available treatment option is for symptom management only.

A single approach to reviewing 5 conditions was not feasible, and a review of 200 conditions would require 200 individual reviews. The five reviews were undertaken by three full-time and two part-time reviewers, and it took 7 months to complete without writing up the findings. A review team of similar size could take as much as 280 months (23 years) to undertake 200 consecutive reviews for 200 conditions. Some learning may be transferable between reviews of similar conditions, shortening certain review processes. For instance, IEM are a group of related conditions that are generally managed with a specific diet or dietary advice, therefore some thinking and decisions may be applicable more widely across several conditions. We could not explore this with the five conditions reviewed, which we selected for a range of treatment scenarios. The five conditions were highly varied in their characteristics, treatment and aim of screening and the review process had to be tailored for each condition. Firstly, each search was developed individually, requiring an understanding of the condition in terms of:

- the presence of a non-genetic version in addition to a genetic version of the condition
- alternative names and aliases
- relevant umbrella terms.

Secondly, categorisation of sequencing studies by the studies' disease definition required condition-specific categories, depending on:

- the availability of biochemical tests (i.e. the definition of disease in biochemical terms is possible)
- the number of disease groups with overlapping symptoms
- whether conditions were only defined genetically
- whether the condition is already screened for.

Thirdly, the definition of early versus late was specific for each condition and depended on:

- whether the relevant intervention is earlier detection or therapeutic and whether the treatment is preventative, curative or management of symptoms
- whether an early intervention phase could be defined
- whether conditions are progressive or present following a trigger
- whether early could be defined in other terms than pre-symptomatic (e.g. early stage of progressive disease).

Exploring Clinical Genome Resource as an evidence source for the United Kingdom National Screening Committee

The ClinGen is an evidence review resource that could be potentially used for the evaluation of WGS for 200 paediatric conditions. We, therefore, assessed the evidence base provided in ClinGen for the five conditions reviewed in our traditional review and compared ClinGen scoring dimensions and scores to the UK NSC criteria and our decisions. We also assessed the alignment of the dimensions and UK NSC criteria to the four principles used by GEL for decisions on gene inclusion. High agreement between criteria from the three resources would mean that decisions from ClinGen and/or GEL could inform UK NSC recommendations in the future.

Clinical Genome Resource as an evidence source for the five conditions in our review

We considered the gene–disease validity, variant classification (level of pathogenicity) and the actionability scores reported by ClinGen. The gene–disease validity has been confirmed in ClinGen for the eight genes of the five conditions considered in our review, and four of the five conditions reviewed (PDE, MCADD, hRB and XLHR) had a paediatric actionability report available on ClinGen in March 2024. However, no information on variant classification in terms of pathogenicity was available for any of the genes. Therefore, the resource has got limited informative value for decision-making on the variant level. The evidence provided on the actionability of the four conditions with a paediatric actionability report is explored in the next chapter (*Comparison of the paediatric reports for pyridoxine-dependent epilepsy, medium-chain acyl-CoA dehydrogenase deficiency, heritable retinoblastoma and X-linked hypophosphataemic rickets from Clinical Genome Resource with our review findings*).

Comparison of the paediatric reports for pyridoxine-dependent epilepsy, medium-chain acyl-CoA dehydrogenase deficiency, heritable retinoblastoma and X-linked hypophosphataemic rickets from Clinical Genome Resource with our review findings

The paediatric actionability reports for PDE, MCADD, hRB and XLHR provide information and scores on four dimensions: severity of disease, penetrance, treatment effectiveness and burden of intervention. These are the most relevant aspects in determining the medical actionability for genetic conditions identified as incidental findings according to ClinGen. *Table 12* summarises the ClinGen scores for the four conditions.

According to ClinGen, PDE, hRB and MCADD are conditions with high actionability (overall score 10 or 11 out of 12) and XLHR is pending an actionability assertion. XLHR has got an overall score of 9, suggesting moderate actionability based on the score alone; however, the Paediatric Actionability Working Group may over-ride the score based on their clinical expertise. At the time of review, the evidence searches for the paediatric reports were 3.5–6 years old (see *Appendix 6*, *Table 32*).

The severity of disease for all four conditions was scored 2 out of 3 (a reasonable possibility of death or major morbidity) and penetrance was scored 3 out of 3 (> 40% chance of serious outcome). A summary of the penetrance findings is provided in *Appendix 7*. The score for penetrance was based on minimal evidence (level C evidence) for PDE, XLHR and hRB (see *Appendix 7*, *Table 33*) and on a single published study that was identified non-systematically through an expert review for MCADD. The evidence on penetrance for PDE and XLHR was provided by GeneReviews without further references to a primary data source (see *Appendix 6*, *Table 32*).[113] For hRB, the penetrance information came from a non-systematic review and a published guideline, in addition to GeneReviews. The penetrance information for MCADD was based on a published study of 81 NBS test-positive newborns who were identified through the Danish NBS screening prgramme.[119] Overall, 7/12 references that form the evidence base for penetrance and expressivity for these four conditions were websites; no evidence cited was based on sequencing data of an unselected newborn population and no additional references to our four systematic reviews were identified (see *Appendix 6*, *Table 32*).

The treatment available for PDE and MCADD was rated as highly effective (score 3 of 3) with minimal risk (score 3 of 3) (*Table 12*). Surveillance for hRB, while highly effective (score of 3), was judged to be associated with moderate risk (score of 2). The treatment for XLHR was classed as moderately effective with moderate risk (both score of 2). The treatments considered for the conditions are summarised based on the ClinGen information given in *Appendix 7*, *Table 33*. The evidence base for the effectiveness of treatment and surveillance was classed as moderate for hRB, MCADD and PDR, where at least one guideline or treatment recommendation was identified (5/8 as websites only, 3/8 published in peer-reviewed journals), and as minimal for XLHR in the absence of a clinical guideline or treatment recommendation. The evidence on the treatment effectiveness of Pi/D for XLHR was based on one non-systematic review, information from the OMIM website and GeneReviews. The main information that ClinGen included from GeneReviews was the findings from the retrospective study looking at Pi/D treatment in patients with XLHR < 1 year of age versus > 1 year of age by Makitie *et al.*[88] included in our review above. One more reference to early versus late management was available for hRB, where the information provided was for screened (family members of probands with hRB) versus probands. This study concluded that surveillance resulted in earlier diagnosis and better outcomes measured by the likelihood of enucleation, eye radiation and visual acuity; however, less than half of the early group actually received surveillance.[120] This study formed part of the treatment recommendation by Skalet *et al.*,[121] cited in the ClinGen paediatric actionability report for hRB. This study was missed by our searches but would not have been included in our review because the definition of early treatment did not meet our inclusion criteria.

TABLE 12 Summary of actionability for the five conditions based on ClinGen, our review and GEL

Condition	ClinGen dimensions (score[a]/level of evidence[b])	NSC criteria[c] (met/not met/rationale)	GEL principles [met (y/n)/ evidence]
The condition			
	Severity: What is the nature of the threat to health to an individual carrying a clearly deleterious allele in this gene?	1. The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious or treatable disease	(A) There is strong evidence that the genetic variant(s) causes the condition and can be reliably detected
PDE	2/–	Met/intractable neonatal seizures associated with intellectual development delay in > 75% cases, and if uncontrolled leading to death, gene–phenotype link has been verified	Y/ClinGen[24]
fHLH	–/–	Met/abnormal immune response leading to fatal multiorgan failure in infancy, gene–phenotype link has been verified	Y/Trizzino 2008[57]–PRF1, Gadoury-Levesque 2020[110]–UNC13D, Al Ahmari 2021[111]–STX11, Gadoury-Levesque 2020[110]
MCADD	2/–	Met/metabolic disease with symptomatic presentation within first 2 years of life, leading to death in 20% cases if undiagnosed, gene–phenotype link has been verified, already part of NBS program	Y/ClinGen[24]
hRB	2/–	Met/malignant neoplasm of the eye with symptoms onset at about 15 months of age, leading to blindness and metastases, gene–phenotype link has been verified	Y/ClinGen[24]
XLHR	2/–	Met/hypophosphataemia leading within first 2 years of life to bone deformity, dental abscesses, stunted growth and bone and joint pain, gene–phenotype link has been verified	Y/Ruppe 2011[112]
	Likelihood of disease: What is the chance that a serious outcome will materialize given a deleterious variant (akin to penetrance)?	3. If the carriers of a variant are identified as a result of screening, the natural history of people with this status should be understood, including the psychological implications	(B) A high proportion of individuals who have the genetic variant(s) would be expected to have symptoms that would have a debilitating impact on quality of life if left undiagnosed
PDE	3/C	Not met/penetrance of variants detected through screening/sequencing is not known, only information on genetic spectrum in children with disease is available	Y/GeneReviews[113]
fHLH	–/–	Not met/penetrance of variants detected through screening/sequencing is not known, only information on genetic spectrum in children with disease is available	Y/GeneReviews[113]
MCADD	3/N	Not met/penetrance of variants detected through sequencing is not known, only information on genetic spectrum in children with disease or with positive NBS test is available	Y/GeneReviews[113]
hRB	3/C	Not met/penetrance of variants detected through screening/sequencing is not known, only information on genetic spectrum in children with disease is available	Y/GeneReviews[113]
XLHR	3/C	Not met/penetrance of variants detected through screening/sequencing is not known, only information on genetic spectrum in children with disease is available	Y/GeneReviews[113]

continued

TABLE 12 Summary of actionability for the five conditions based on ClinGen, our review and GEL (*continued*)

Condition	ClinGen dimensions (score[a]/level of evidence[b])	NSC criteria[c] (met/not met/rationale)	GEL principles [met (y/n)/ evidence]
The intervention			
	Effectiveness: How effective is the selected, specific intervention for preventing or significantly diminishing the risk of harm when initiated during childhood (< 18 years)?	9. There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, e.g. those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened, then the screening programme should not be further considered	(C) Early or pre-symptomatic intervention for the condition has been shown to lead to substantially improved outcomes in children compared to intervention after the onset of symptoms
PDE	3/B	Not met/the evidence-based direction is that there might be some benefit in early treatment; however, the quality and volume of the evidence are low because of the definition of early vs. late, the type of study, the number of participants and the number of studies; therefore, there is insufficient evidence to clearly judge the effect	Y/GeneReviews[113]
fHLH	–/–	Not met/the evidence-based direction is that there might be some benefit in early treatment; however, the quality and volume of the evidence is low because of the definition of early vs. late, the type of study, the number of participants and the number of studies; therefore, there is insufficient evidence to clearly judge the effect	Y/GeneReviews[113]
MCADD	3/B	Not met/the evidence-based direction is that there might be some benefit in early treatment; however, the quality of the evidence is low because of the definition of early vs. late, the type of study and the number of participants; there is insufficient evidence to clearly judge the effect. While MCADD is on the current NBS screening panel, there is no indication that the decision was based on a strong evidence base	Y/GeneReviews[113]
hRB	3/B	Not met/the evidence-based direction is that there might be some benefit in early treatment; however, the quality and volume of the evidence are low because of the definition of early vs. late, the type of study, the number of participants and the number of studies; therefore, there is insufficient evidence to clearly judge the effect	Y/GeneReviews[113]
XLHR	2/C	Not met/the evidence-based direction is that there might be some benefit in early treatment; however, the quality and volume of the evidence are low because of the definition of early vs. late, the type of study, the number of participants and the number of studies; therefore, there is insufficient evidence to clearly judge the effect	Y/NICE HST8 2018[114]
	Nature of the intervention: How risky, medically burdensome or intensive is the given intervention?	10. There should be agreed evidence-based policies covering which individuals should be offered interventions and the appropriate intervention to be offered	(D) Conditions screened for are only those for which the interventions are equitably accessible for all
PDE	3/–	Not met/published consensus guidelines state: 'All newborns with PDE–*ALDH7A1* should be treated with 100 mg/day of pyridoxine supplementation'. Not mentioning asymptomatic treatment specifically.[115] Clinical advice was that pyridoxine should not be given asymptomatically due to reports of toxicity, it is unclear what common practice is	Y/GeneReviews[113]

TABLE 12 Summary of actionability for the five conditions based on ClinGen, our review and GEL (*continued*)

Condition	ClinGen dimensions (score[a]/level of evidence[b])	NSC criteria[c] (met/not met/rationale)	GEL principles [met (y/n)/ evidence]
fHLH	–/–	Met/clinical advice was: 'I think most physicians would agree with pre-emptive HSCT for Perforin, Syntaxin, MUNC13-4 and MUCN18-2 deficiency (although, even for these disorders there are some patients who follow a mild course)'	Y/Clinical Commissioning Policy: Anakinra for HLH for adults and children in all ages[116]
MCADD	3/–	Met/dietary advice is provided following treatment guidance such as …	Y/BIMDG guidelines[117]
hRB	2/–	Met/surveillance as offered currently for patients detected through FH	Y/NHS commissioning board NHS standard contract RB service[116]
XLHR	2/–	Met/clinical advice was to start patients on Pi/D, and published recommendations provide dosage specification for pre-symptomatic children Not met/burosumab is not licensed for children < 1 year of age currently	Y/NICE HST8 2018[118]
Final rating			
	Total score/final assertion		Category[d]
PDE	11 CB/strong actionability	1/8 criteria considered in review were met, overall rating: not met	1
fHLH	–/–	2/8 criteria considered in review were met, overall rating: not met	1
MCADD	11 NB/strong actionability	2/8 criteria considered in review were met, overall rating: not met	1
hRB	10 CB/strong actionability	2/8 criteria considered in review were met, overall rating: not met	1
XLHR	9 CC/assertion pending	2/8 criteria considered in review were met, overall rating: not met	1

CB, cord blood; CC, complexity and comorbidity.
a Score between 0 and 3 (3 is best for actionability), sum of scores gives total score.
b Level of evidence: A substantial, B moderate, C minimal, D poor, N non-systematically identified or expert contributed evidence.
c The rating should be interpreted as a recommendation and can be changed following stakeholder consultation.
d Out of 4: category 1: gene/condition appears to satisfy the four principles, Category 2: unclear whether gene/condition satisfies the four principles – expert input required, Category 3: gene/condition does not satisfy the four principles and is childhood onset, Category 4: gene/condition does not meet the four principles and is adult onset.

Source
Reproduced from Freeman K, Taylor D, Dinnes J, Clark CCA, Kander I, Scandrett K, *et al*. Challenges in evaluating whole genome sequencing for newborn screening: series of systematic reviews and roadmap for evidence generation for policy advisers. *BMJ Medicine* 2025;**4**:e001726. https://doi.org/10.1136/bmjmed-2025-001726. This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for non-commercial use, provided the original work is properly cited. See: https://creativecommons.org/licenses/by-nc/4.0/. The table includes minor additions and formatting changes to the original text.

Comparison of the ClinGen scores of the five conditions to our review findings using the UK NSC criteria and GEL's decisions for gene inclusion is shown in *Table 12*. This was complicated by three aspects. Firstly, the ClinGen scores out of 3 (3 being best for actionability) for the four dimensions do not provide a definitive yes/no assessment precluding an evaluation of the level of agreement for the individual criteria.

Secondly, the criteria for assessment do not map well across ClinGen, the UK NSC and GEL. While ClinGen assessed the severity of disease in the diagnostic context of an incidental finding in an individual, the UK NSC criterion 1 requires the assessor to consider the importance of disease from a public health perspective, considering whether screening for the condition is worth it, whether the natural history of the disease is understood and the link between the risk factor (here a genetic variant) and the disease is known. GEL's principle A focuses more simplistically on the level of evidence that proves the gene–disease link.

The criterion around penetrance differs between ClinGen and GEL's principle B versus UK NSC criterion 3, because the UK NSC criterion focuses on screen-detected variants rather than variants detected in patients with confirmed disease. This means that the studies identified for question 2 in our traditional review only partially address this criterion and do not provide the evidence needed to understand the penetrance in the complete spectrum of genetic disease identified through screening. This explains at least partly the different outcomes of assessment for this criterion.

There are also some important differences in the evaluation of the evidence of treatment effectiveness. While ClinGen requires an available and effective treatment that can be used in pre-symptomatic individuals, the UK NSC and GEL also require evidence on the benefits of earlier treatment compared to treating symptomatic disease. And, while ClinGen considers the burden of the treatment separately in their fourth dimension, GEL's principle D is focused around equity in access to treatment, and UK NSC criterion 10 focuses on the available evidence of a treatment pathway for early treatment. While some criteria map reasonably well across GEL and the NSC, the evidence bar appears to be higher for the UK NSC, which has led to different decisions for five/five conditions concerning the treatment effectiveness. In contrast to the UK NSC, which requires a systematic and comprehensive approach to evidence synthesis, GEL's approach was similar to that used by ClinGen in that a single supporting reference was sufficient to meet GEL's principles for gene inclusion, and the evidence was often based on information provided by ClinGen or GeneReviews. This means that all five conditions met the GEL principles, while the current evidence base for none of the conditions meets all UK NSC criteria. Furthermore, ClinGen and GEL assessed gene–condition pairs against 4 criteria, while the UK NSC requires evidence to a further 16 criteria which are not part of this comparison.

Thirdly, the subjective approach to reach a final decision is not transparent and decisions may not follow from the evidence. GEL and ClinGen heavily rely on the expertise and opinion of an expert clinical panel. The UK NSC, on the other hand, relies on a balanced interpretation of the amount and quality of evidence, which is reviewed during stakeholder consultations.

Overall, it appears that different outcomes in the assessment of the five conditions are due to a combination of (1) different focus in the assessment criteria used (particularly, the individual patient vs. the public health perspective), (2) the expected limited evidence base for these rare conditions and a higher evidence bar used by the UK NSC and (3) the subjective interpretation of the evidence base to reach a final decision. This means that neither the ClinGen scores nor the GEL decisions would be an appropriate proxy for a UK NSC recommendation for the five conditions. And, using ClinGen as an evidence source for a future review of genomic screening of 200 conditions may not be feasible nor appropriate.

Conclusions and learnings from Clinical Genome Resource

The ClinGen is an excellent and important resource of information on medical actionability for incidental findings of monogenic diseases. It produces reports separately for the paediatric context, uses a transparent and standardised approach of evidence collection and a semi-quantitative metric to score the actionability of diseases across four domains. An expert group reviews the decisions and the evidence and formulates an overall assertion on actionability. The process allows an early rule out of conditions that do not meet the evidence threshold. Although this is a rich source of information and potentially a good starting point for discussion around reporting of genetic findings, the resource has got some limitations as an evidence source for the UK NSC:

1. The protocol that the paediatric actionability working groups follow is a generic one, and there is no information on what each group did in their evidence review and how consistent the approach was across groups.
2. The focus is on individuals with incidental findings in the diagnostic setting rather than the public health perspective for a screening programme decision.
3. The reports tend to be quite dated and would require updating considering the fast-moving field of genetics.
4. The evidence bar for actionability in the diagnostic setting is lower than what the UK NSC requires in terms of quantity and quality of evidence for population screening.
5. Knowledge of penetrance is not a requirement for actionability as early rule out is not triggered in the absence of information on penetrance.
6. Treatment effectiveness is judged by the availability of a treatment guideline with an existing intervention for an undiagnosed paediatric population without the requirement of evidence on benefits of early versus late treatment.
7. The number of conditions covered is limited as of March 2024, and no information on variant pathogenicity was available for any of the five conditions of interest.

Considering the relatively low evidence requirement, it may be feasible to explore using ClinGen as a tool to rule out conditions for reporting, which did not meet the requirements for full review, but it is inappropriate for the UK NSC to base decisions on potential screening programs on the actionability reported in ClinGen without further assessment.

Using existing genomic studies of newborn screening cohorts reporting penetrance as an evidence source

The purpose of this review was to explore the feasibility of identifying pathogenic variants of paediatric conditions that are known to have high penetrance and expressivity in an unselected newborn population to be considered for an initial newborn screening programme that maximises benefit and minimises harm. The review focused on studies using sequencing in newborns for childhood-onset disease that reported penetrance or an approximation.

Out of 4970 articles identified, 105 were taken through to full-text sifting. The majority were excluded because the test did not meet the inclusion criteria (indirect sequencing or sequencing was second- or third-line test), or the outcome was irrelevant (focus on carrier frequency). See *Appendix 4*, *Figure 7* for the study flow diagram and *Report Supplementary Material 2* for the list of excluded studies and their reason for exclusion. There were 14 studies that reported experiences with gene sequencing in newborns which were reported in 16 references.[122–137] Green *et al.*,[122] Green *et al.*[123] and Ceyhan-Birsoy *et al.*[124] all report on the BabySeq project, which will be cited as Green *et al.* from hereon.[122] Full data extractions for the 14 studies are reported in *Report Supplementary Material 4*. Four studies sequenced genes for a single condition[134–137] and were not considered for synthesis, and 10 studies sequenced genes for ≥ 74 conditions.[122,125–133] Five of the 10 studies did not report any follow-up time and therefore could not be used to approximate penetrance. These five studies compared sequencing data to follow-up test results,[126] conventional newborn screening with clinical review,[130] categorised sequencing outcomes into levels of penetrance based on existing literature[132] or categorised sequencing outcomes into pathogenic/likely pathogenic based on existing classification systems with[131] or without clinical review.[128,133] Information on penetrance could be inferred for the remaining 5 of the 10 studies because they reported some clinical follow-up after a positive sequencing test.[122,125,127–129] These five studies are synthesised below and summarised in *Tables 13* and *14*, with one study reporting results for a prospective and a retrospective cohort.[125] The five studies originated from two countries, the USA (n = 2[122,127]) and China (n = 3[122,127]). While the US studies used WGS,[122,127] the Chinese studies used gene panel tests[125,128,129] The number of included genes ranged from 134 to 954 and the number of newborns sequenced ranged from 127 to 29,989. Gene selection and variant interpretation varied across studies. In four out of five studies, the country's newborn screening programme acted as a starting point,[125,127–129] and two out of three Chinese studies stated that they considered conditions on the recommended unified screening panel, which details the mandatory conditions for newborn screening in the USA in addition to their own.[128,129] However, the final scope of included conditions differed. Green *et al.* (2023) had the widest scope, including a small number of adult-onset conditions.[122] The level of detail provided on variant selection also varied. In general, as a minimum, studies reported consideration of the American College of Medical Genetics and Genomics guidelines and/or ClinGen/clinical variation (ClinVar) for the classification of variants from pathogenic

TABLE 13 Characteristics of studies reporting results of genetic sequencing of newborns in the screening setting

Study reference/ country	Study design/quality	Population *N*/ characteristics	Screening method	Number of genes/ conditions	Method of estimating penetrance
Green 2023, Green 2022, Ceyhan-Birsoy 2019;[122–124] (BabySeq), USA	RCT (randomisation to conventional care with or without GS at Brigham and woman's hospital) Only the GS arm is reported	127 apparently healthy infants/159 healthy and critically ill newborns	GS	954 genes from categories A and B prioritising paediatric disorders	Clinical follow-up for 3–5 years following disclosure of results
Hao 2022 (retrospective cohort);[125] China	Retrospective, single-arm, multicentre study Babies were recruited from eight women and children's hospitals nationwide in China	11,484 babies	Targeted gene panel (NESTS)	465 causative genes for 596 early-onset, relatively high incidence and potentially actionable severe diseases	Clinical follow-up at children's current age (2–29 months)
Hao 2022 (prospective cohort);[125] China	Prospective, single-arm, single-centre study Pregnant women were recruited from Shunyi Women and Children's Healthcare Hospital of Beijing Children's Hospital from October 2018 to June 2019	3923 newborns	Targeted gene panel (NESTS)	465 causative genes for 596 early-onset, relatively high incidence and potentially actionable severe diseases	Clinical follow-up (phone call follow-up) was done at ages ranging from 24 to 28 months (February 2020 to May 2021)
Bodian 2016;[127] USA	Single-arm study Neonates recruited to two research studies at Inova Fairfax Hospital in Virginia during 2011–4	1696 pre-term and full-term neonates unselected for newborn screening -related disorders	WGS	163 genes of disorders included or under discussion for inclusion in US newborn screening programs (including 65 genes relating to disorders screened by blood-based newborn screening in Virginia)	Comparison to • Blood-based newborn screening with repeat screening if indicated • Clinical diagnoses extracted from EHRs (some follow-up)
Chen 2023;[128] China	Prospective single-arm study Newborns were recruited from eight newborn screening centres in China from 21 February 2021 to 30 December 2021	29,989 newborns with DBS specimen 3–7 days after birth and parental consent	Targeted gene panel sequencing	142 related genes (128 conditions), 43/128 diseases were included in the biochemical panel	Clinical follow-up (recall or phone calls until diagnosis or censoring on 5 July 2022) Diagnosis was based on confirmatory tests
Huang 2022;[129] China	Retrospective, single-arm, enriched study Randomly selected newborns with a negative initial conventional newborn screening result and newborns with a positive newborn screening result, who received repeat MS/MS analysis, born between January 2017 and December 2019	4986 newborns	Newborn genetic sequencing panel	134 genes of 74 inborn disorders with severe impact on children's health (41/74 are on conventional newborn screening panel)	Comparison to conventional newborn screening and clinical manifestation (some follow-up reported)

DBS, dried blood spot; GS, genomic sequencing; NESTS, NEwborn Screening with Targeted Sequencing.

TABLE 14 Summary of findings from studies of genetic sequencing of newborns in the screening setting

Study reference	Number of cases detected by sequencing/ total sequenced (%)	Agreement with conventional newborn screening	Additional cases/ conditions detected	Information on penetrance
Green 2023, Green 2022, Ceyhan-Birsoy 2019;[122–124] (BabySeq), USA	10/127 (7.9%)	N/A	N/A	2/10 (20%) confirmed on clinical follow-up at 55.5 and 60.1 months, on active management
				8/10 (80%) unconfirmed on clinical follow-up between 37.6 and 60.3 months (no phenotype as variants predict risk for future disease, 7/8 with predicted moderate penetrance and 1/8 with predicted high penetrance)
Hao 2022 (retrospective cohort);[125] China	902/11,484 (7.9%) 488/902 lost to follow-up	NR	NR	50/414 (12.1%) confirmed on follow-up at age 2–29 months 364/414 (87.9%) unconfirmed at age 2–29 months (without obvious related phenotype)
Hao 2022 (prospective cohort);[125] China	381/3923 (9.7%)	NR	NR	6/381 (1.6%) confirmed on clinical follow-up at age 24–28 months, 4/6 with active management
				375/381 (98.4%) unconfirmed on clinical follow-up at age 24–28 months (follow-up may be too short for childhood-onset disease)
Bodian 2016;[127] USA	33/1969 (1.7%) (19/33 conventional newborn screening conditions)	1/19 WGS positive/conventional screening positive (agreement)	14/33 additional conditions not on the newborn screening panel (14/14 unconfirmed)	2/33 (6.1%) confirmed by clinical diagnosis (1/2 newborn screening negative, affected by condition not medically actionable, healthy without follow-up time specified, 1/2 newborn screening positive, with positive follow-up test and active management, healthy without follow-up time specified)
		18/19 WGS positive/conventional screening negative (only detected by sequencing)		17/33 (51.5%) unconfirmed by clinical follow-up (newborn screening negative, healthy at > 1 year of age)
		3 WGS negative/conventional screening positive (missed by sequencing – not included in total of 19)		14/33 (42.4%) unconfirmed by newborn screening (not on newborn screening panel, no clinical data presented)
Chen 2023;[128] China	797/29,989 (2.7%)	445 were positive for conditions on both panels and either positive on conventional newborn screening, genetic screening or both	434/797 were positive for disorders screened solely by genetic screening (395 unconfirmed)	402/797 (50.4%) confirmed on follow-up testing and follow-up to median age of 1.2 years (39/402 only detected on WGS, with symptoms and on active management) 395/797 (49.6%) unconfirmed by follow-up testing and follow-up to median age of 1.2 years

continued

TABLE 14 Summary of findings from studies of genetic sequencing of newborns in the screening setting (*continued*)

Study reference	Number of cases detected by sequencing/ total sequenced (%)	Agreement with conventional newborn screening	Additional cases/ conditions detected	Information on penetrance
		343/445 conditions were positive on both conventional and genetic newborn screening (agreement)		Genetic reason for unconfirmed conditions: 359 (X-linked inheritance) females (unconfirmed genetic G6PDD) 1 male with XXY karyotype (unconfirmed G6PDD)
		20/445 conditions were conventional screening-negative/genetic screening-positive (only detected by sequencing)		1 male with c.1386G > T located on exon 12 which was duplicated (unconfirmed genetic G6PDD)
		82/445 conditions were conventional screening-positive/genetic screening-negative (missed by sequencing – not included in total of 797)		1 homozygous deletion of exon 7 in SMA1, but 3 copies of the SMA2 gene (acting as genetic modifier in unconfirmed genetic SMA)
				33 cases may be attributed to childhood onset or mild phenotypes
Huang 2022;[129] China	113/4986 (2.3%)	56/113 conditions were conventional screening-positive/genetic screening-positive (agreement, including 20 not confirmed by clinical manifestation)	12/113 additional conditions not on conventional newborn screening panel (including 10 without clinical manifestation)	36/113 (31.9%) confirmed by repeat MS/MS and clinical manifestation
		45/113 conditions were conventional screening-negative/genetic screening-positive (only detected by sequencing)		30/113 (26.5%) confirmed by clinical manifestation (26/30 newborn screening positive/NGS-positive; mild phenotype in female patients, 2/30 newborn screening negative/NGS-positive, clinical diagnosis and symptom onset at 14 months and 3 years of age, 2/30 not on newborn screening panel, clinical diagnosis, disease onset at 14 months and 3 years of age)
				47/113 (41.6%) unconfirmed (27/47 not on newborn screening panel, no clinical data or no clinical manifestation, 16/47 newborn screening positive/NGS-positive, slightly elevated biochemical markers, no intervention, variant may be causing mild phenotype, 4/47 newborn screening negative/NGS-positive, follow-up testing negative)

EHR, electronic healthcare records; G6PDD, glucose-6-phosphate dehydrogenase deficiency; N/A, not applicable; NR, not reported; SMA, spinal muscular atrophy.

to benign. Four of the five studies clearly stated that only pathogenic and likely pathogenic variants were reported.[122,125,128,129]

The findings of the five studies are reported in *Table 14*. The proportion of sequencing positive babies ranged from 1.7%[127] to 9.7%.[125] Comparison to conventional newborn screening revealed that genetic screening and conventional screening are complementary, each detecting and missing different cases. Genetic screening has the potential to identify a significant proportion of additional conditions that are not on the conventional screening panel in the USA and/or China. Two of the three studies with a comparison to conventional screening reported that over half of the positive screens were made up of conditions not available on the conventional screening panel.[127,128] However, 83.3–100% of these could not be confirmed clinically within the studies' follow-up, so we do not know if detecting these was an overdiagnosis of clinically insignificant disease or misdiagnosis of disease or early detection of later-onset disease. In Chen *et al.*,[128] 359/395 (90.9%) of these unconfirmed additional cases were female babies with the X-linked inherited condition G6PD. This represents an example where sequencing performs differently in males and females.

The great proportion of unconfirmed cases can be partially explained by the short follow-up, which ranged within a retrospective study from 2 months to 29 months[125] and from a median of 1.2 years[128] to > 5 years[122] in the prospective studies. Overall follow-up was insufficiently described, variable within studies and too short to capture the disease onset for all included conditions. The short follow-up had an impact on the estimates of penetrance because cases developing clinical symptoms after follow-up will be unaccounted for. This impact would differ for early-onset and later-onset childhood conditions. Therefore, the identification of variants in healthy newborns does not exclude a pathogenic role for these variants. Considering the short follow-up, penetrance, approximated by the number of confirmed cases after clinical follow-up for all genes considered, ranged from 1.6% after a follow-up of 24–48 months[125] to 58.4% after up to 3 years of follow-up.[129] While some of the studies were large enough to report a significant number of sequencing positive cases, these large studies did not report the findings on clinical outcome by gene variant for all cases to enable an estimate of penetrance on the variant level. One study reported 16 cases of PKU, all of which were compound heterozygotes with c.158G > A.[129] All were weakly positive on conventional NBS testing due to slightly elevated Public Health England levels at regular intervals. However, none of the infants received any intervention, and the study authors concluded that the variant may be causing a mild phenotype. More such data are needed from sufficiently large studies to enable penetrance estimates to be derived for individual variants. However, the majority of variants reported in the studies appear to be single occurrences precluding the estimation of penetrance.

Clinical management was considered where appropriate in all five studies following a confirmed diagnosis including for 'mild' and subclinical cases on confirmatory testing. While clinical management in the studies was considered confirmatory of the positive sequencing result, early intervention also means that penetrance cannot be estimated for all cases from these studies. It is unknown whether symptoms would have developed even without treatment in confirmatory testing of positive cases.

All five studies were single arm, either retrospective or prospective cohorts. One study used an enriched study design; however, this was not reflected in the proportion of sequencing positive cases.[129] Applicability may be a concern as conditions and variant frequencies vary across geographical regions (even within countries) and three out of five studies were undertaken in Chinese populations. Sequencing negatives were not followed up apart from one study, which reported that 82/445 newborns who received both genetic and conventional screening were missed by targeted gene panel sequencing.[128] While this information is needed to evaluate the test accuracy of WGS, it does not contribute to penetrance estimates.

Conclusions and learnings from existing newborn sequencing studies

In comparison to our traditional review of individual conditions, the focused review of studies reporting penetrance of any gene/condition pair in the newborn screening setting took around 4 weeks to develop, and it produced a robust search strategy which can be re-run in the future without consideration of different conditions. The search picked up all publications from current genomics projects in newborns that were known to us.

Overall, the studies identified highlight that GS for newborn screening is still in its infancy and cannot be implemented without further research. There was a lack of consensus on which genes to include, no indication of how to interpret discordant results from NBS programmes and genetic screening, uncertainty over which test is most appropriate for which condition, evidence of overdiagnosis and a large number of cases that could not be confirmed within the studies' follow-up resulting in interventions, routine surveillance and regular follow-up of uncertain benefit. The studies do not lend themselves to determine a variant threshold for individual genes, that is, the number and types of variants to include in a screening programme that focuses on detecting highly pathogenic and highly penetrant variants with low risk of producing harm. This is because:

- The number of infants with a specific condition displaying a range of variants is too low even in large studies so variant frequencies cannot be estimated.
- Infants with confirmed genetic disease received management, which precludes estimation of penetrance and expressivity for cases without symptomatic confirmation of disease.
- Clinical follow-up was not sufficiently long to include all childhood-onset cases.

The studies, however, give some insight into aspects of WGS not identified from our review of five conditions, including challenges and promises of WGS of newborns, such as:

- The potential complementary role of WGS to traditional newborn screening programmes.
- Levels of agreement with conventional newborn screening.
- Conditions for which conventional screening may be better (e.g. 4/4 NBS test-positive congenital hypothyroidism infants missed by sequencing[129]).
- Conditions where regional differences may exist (e.g. G6PDD and PKU in China[126]).
- Potential differences in PPV between sexes (e.g. G6PDD[128]).
- Proportion of infants positive on sequencing with two recessive variants, where variants are in *cis*gender (i.e. in the same copy of the gene leaving the other gene copy intact) [7/55 (12.7%)[128]], therefore highlighting the potential role of parental testing in cases with two heterozygote variants to determine the phase.
- The impact of including analysis of CNVs [in one study, sequence analysis detected 33/47 (70.2%) at-risk genotypes, and CNV analysis contributed an additional 14/47 (29.8%),[132] highlighting potential limitations of WGS].
- Some indication of turnaround times for WGS [mean turnaround time 1.5 times longer for WGS vs. exome gene panel testing (56 days vs. 37 days),[132] sequencing to issuing formal report was within 11 days,[125] turnaround time of newborn WGS 16 weeks–24 weeks[130]].

In addition to the included studies described above, the review also identified two studies that reported findings that approximate penetrance.[138,139] In these studies, variants of a group of childhood-onset conditions were identified, and the number of healthy adults with disease-associated genotypes with these variants in a general population database were determined.[138,139] Gold *et al.* concluded that the false-positive rate of genomic screening of newborns for diseases treatable with HSCT was 0.04% based on 59/141,456 healthy adults, with implicated genotypes using variants of 127 genes of severe childhood-onset conditions treatable with HSCT.[138] Breilyn *et al.* found that clinically relevant variants according to ClinVar in *ACADS* (gene associated with short-chain acyl-CoA dehydrogenase deficiency) were not associated with the evidence of metabolic disease in a large and ancestrally diverse adult population (2035/30,000 healthy adults with implicated genotypes).[139] These findings could imply that the variants of the 127 genes of severe childhood-onset conditions treatable with HSCT may be suitable for newborn screening, while variants in *ACADS* investigated may not be suitable. These studies may contribute useful insights on the penetrance and expressivity in a future review if the studies meet the following inclusion criteria:

- population database should be of healthy adults rather than newborns
- variants considered should be of genes associated with childhood-onset disease
- outcomes should be number of adults with implicated genotype (not simply allele frequency)
- outcomes should be reported on variant level.

Results for review of cost-effectiveness evaluations of whole genome sequencing and whole exome sequencing

Search results

Following the searches of databases and registries, the titles and abstracts of 2325 records were screened, of which 226 records were identified as potentially meeting the eligibility criteria and were flagged for full-text review [see the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram in *Appendix 4*, *Figure 8*]. Seventy-one of these studies were judged to meet the eligibility criteria.

Nineteen of the records identified at the full-text stage process were literature reviews. The reference lists of all 19 reviews were checked to identify whether there were any potentially relevant records missed by our search or the original Schwarze *et al.* review.[27] All included records in 13 of the 19 reviews had already been identified, but there were 6 reviews that included 9 additional studies that had not been identified. Of these, two studies were eligible for inclusion.

Of the 28 costing studies included in the Schwarze *et al.* systematic review,[27] 13 met our pre-defined eligibility criteria. Four studies were excluded because they were conference abstracts.[140–143] One study focused on the use of WGS for a communicable disease.[144] Six studies were included in the Schwarze *et al.* review as partial economic evaluations, but cost was not included as an outcome in the methods or results sections of the studies; cost was just mentioned briefly at some point in the paper.[145–150] One study was categorised as a partial economic evaluation, but we could not find any mention of costs in the paper.[151] A further study was excluded because the focus of the analysis was on the cost associated with identifying incidental findings, and the main analysis excluded the costs associated with WES and WGS.[152]

Two literature reviews were included in the Schwarze *et al.* review; we did not find any additional papers to assess for eligibility in these reviews.[27,153]

No additional papers were eligible from the Nurchis *et al.* scoping review.[28]

In total, there were 86 studies included in the review.

Study characteristics

Table 15 provides an overview of the characteristics of the studies included in this review. The full data extraction is included in *Report Supplementary Material 5*. Most of the included studies were conducted in high-income settings such as Europe (*n* = 27/86, 31%), North America (*n* = 28/86, 33%) or Australia (*n* = 20/86, 23%). Of those conducted in Europe, around half (*n* = 13/27) were from the Netherlands and four from the UK. The earliest study included in our review was published in 2014, and there has been an increase in the number of studies published on this topic since 2017 (*Figure 1*).

Population

None of the included studies focused on the use of WGS or WES specifically in a screening context; all studies focused on symptomatic populations or the cost of WGS or WES more generally. Over three-quarters of the included studies included newborns or children in their target population (*n* = 70/86, 81%), reflecting the early onset and presentation for most of the conditions targeted by WES and WGS.

The studies evaluated the use of WES and WGS for a wide range of conditions with a genetic component. The description of the target diseases was very broad for many of the papers, particularly those that focused on newborns or children. Terms including ‘genetic disorders’, ‘monogenic disorders’, ‘variety of conditions’, ‘rare diseases’, ‘mendelian disorders’ and ‘mitochondrial disorders’ were used, demonstrating the broad spectrum of possible genetic diseases potentially discoverable by WES and WGS in contexts where there is very little indication of a clear diagnosis for an individual. There were some common recurring themes such as neurological or neurodevelopmental disorders, intellectual disability or developmental delay (*n* = 20/86, 23%), cancer (*n* = 12/86, 14%), congenital anomalies (*n* = 6/86, 7%), different types of epilepsy (*n* = 4/86, 5%) and autism spectrum disorder (*n* = 3/86, 3%).

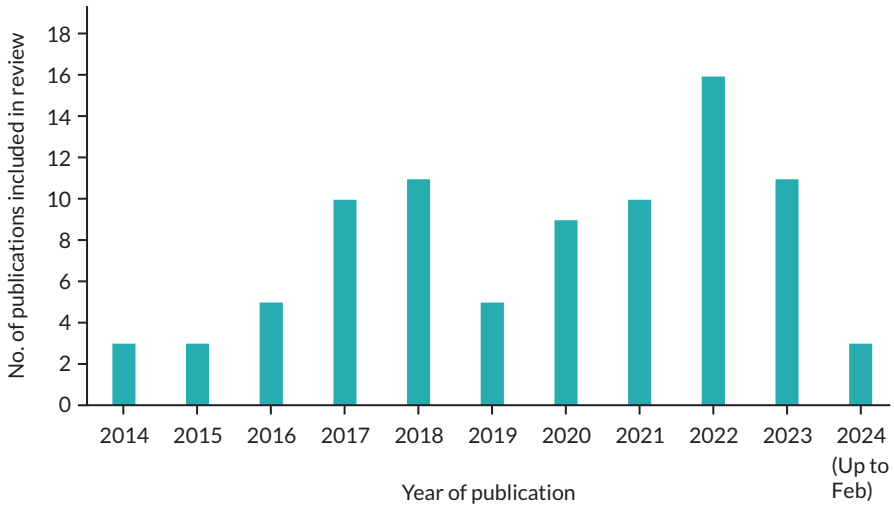

FIGURE 1 Number of studies included in the review by publication year.

TABLE 15 Characteristics of included studies

Characteristic	Category	Number of studies (%)
Type of economic evaluation	Full economic evaluations	
	Cost–utility	10 (12)
	Cost-effectiveness	29 (34)
	Partial economic evaluation	
	Cost–consequence analyses	29 (34)
	Costing study	18 (21)
Study setting	Australia	20 (23)
	Europe	27 (31)
	Hong Kong and China	4 (5)
	International	1 (1)
	Middle East	3 (3)
	North America	28 (33)
	South America	1 (1)
	Southeast Asia	2 (2)
Sequencing approach	WES	48 (56)
	WGS	22 (26)
	Both WES and WGS	16 (19)
Target population	Prenatal	2 (2)
	Prenatal and children	1 (1)
	Newborns or children	50 (58)
	Adults	5 (6)
	Children and adults	19 (22)
	Not specified/no study population	9 (10)

Intervention

Over half of the studies focused on WES as the intervention (*n* = 48, 56%), with just over a quarter focusing on WGS (*n* = 22/86, 26%) and 16 studies (19%) focusing on both tests. Around a third of the studies explored the impact of putting WES or WGS testing earlier or later in the diagnostic pathway (*n* = 29/86, 34%). In some of these studies, WES was already an established part of the diagnostic pathway under evaluation, and the question focused on whether it should be moved to later in the pathway (i.e. more targeted screening to reduce costs) or moved to earlier in the pathway to reduce the need for other diagnostic tests and potentially arrive at a diagnosis earlier.

There were also some studies (*n* = 10/86, 12%) which focused specifically on rapid WES or WGS, which is different to standard sequencing in that it has a much shorter turnaround time, potentially increasing the clinical utility of the test by returning actionable results quicker. All these studies focused on children and newborns and were looking for rare diseases, suspected genetic diseases or 'diseases of an unknown cause'.

Comparator(s)

Of the studies that included a comparator (*n* = 78/86, 91%), 44 (56%) explicitly stated that the comparator was current standard of care testing. This typically consisted of a broad range of tests, some focusing specifically on more targeted genetic testing, and others also encompassing a wide range of imaging, biochemical and biopsy tests. In some studies, these were very clearly broken down by the individual test. Different assumptions were made in terms of which tests would no longer be needed following the incorporation of WES or WGS in the diagnostic pathway, typically dependent on the proposed timing of the different tests.

Many of these studies used cohort data to underpin their analyses (*n* = 58/78, 74%), capturing the healthcare resource use and, where applicable, outcomes for the cohort. Of these, over two-thirds (*n* = 39/58, 67%) used the same cohort of individuals to inform both their intervention and comparator(s). These studies relied on a single cohort who had received either the intervention or comparator, making assumptions about what would have happened to those individuals had they received the opposite. For example, the cohort may have received current standard of care and assumptions were made about what would have happened (usually which diagnostic tests could have been avoided) had the intervention been available. Some studies (*n* = 18, 23%) had two distinct cohorts, where one had received the intervention and one had received the comparator. Sixteen studies were modelling exercises, so the cohorts for the intervention and comparator(s) had been simulated based on aggregate data obtained from the literature.

Outcome

Of the 86 studies included in this review, only 10 were cost–utility studies and defined the outcome of their analysis as cost per QALY. Over a third of the included studies were cost-effectiveness evaluations (*n* = 29/86, 34%), that is, they reported the cost for a change in a specific health outcome. All of these studies focused on the number of diagnoses, additional diagnoses or diagnostic yield as the health outcome of interest. The costing studies included (*n* = 18, 21%) tended to report either the total cost per patient or the cost per test.

Methodology

The costing perspective was not specified in 33 of the included studies (38%). Where it was reported, many adopted a broad healthcare system perspective (*n* = 29/86, 36%) or a more specific health system perspective such as a hospital perspective, a clinical genetics service perspective or a laboratory perspective (*n* = 15). Some studies considered costs to the patient or payer (*n* = 8) or a broader, societal perspective (*n* = 5). The costing perspective was largely driven by whether the cost of WES or WGS testing would fall to a public healthcare system or to the individual. Of those adopting a societal perspective, one study included costs associated with the caregiver time spent in neonatal intensive care unit (NICU); post NICU informal care and direct non-medical costs, for example caretaking, modifications to home and special education.[154] Another study estimated the parental time lost from accompanying their child to medical visits, which was valued using the human capital approach using hourly adult wage.[155]

Over half of the included studies (*n* = 50, 58%) did not state the time horizon of their analyses. Only seven studies (8%) adopted a lifetime horizon, with eight (9%) studies including costs over a 1-year time period or less, typically to only capture costs associated with the diagnostic testing pathway. Where time horizons of > 1 year were explicitly implemented (*n* = 24, 28%), 17 studies (71%) stated that they applied a discount rate to costs beyond the first year.

Only two-thirds of the included studies (*n* = 54/86, 63%) conducted scenario or sensitivity analyses to explore the impact of parameter or structural uncertainty in the economic evaluation.

Cost–utility studies

Cost–utility studies, that is, economic evaluations which produce estimates of the incremental cost per QALY, are typically the preferred economic evaluation method for HTA in the UK. Of the 86 studies included in this review, only 10 were cost–utility studies. All of these studies have been published since 2019. Three of these evaluations focused on individuals with non-small cell lung cancer.[156–158] One study focused on children with prenatally diagnosed non-immune hydrops fetalis and fetal effusions.[159] One study focused on adults and children with suspected monogenic kidney disease.[160]

Crawford *et al.*[154] evaluated the cost-effectiveness of early exome sequencing in critically ill babies who are admitted to NICUs and are suspected to have mitochondrial disorders compared to current typical care. They used a decision tree–Markov hybrid model to estimate the costs and QALYs over a lifetime horizon. The model was populated using data from the published literature, expert opinion and the Paediatric Health Information System database in the USA. The model simplified the health outcome to two broad health states: within the NICU and post-NICU. Given the unavailability of quality of life metrics from the patient perspective, published utility values, based on standard gamble responses from parents with experience of having a child in the intensive care unit, were used as a proxy for the utility of a newborn with severe mitochondrial disease in a level III/IV NICU and post-NICU.

Sanford Kobayashi *et al.*[161] evaluated the cost–utility of rapid WGS in critically ill children admitted to paediatric (not neonatal) intensive care units. Their analysis was based on a cohort of 38 children ranging from 4 months to 17 years old, where rapid WGS had resulted in a molecular diagnosis in 17 of the 38 children. Eight of these 17 children had a change in clinical care, and QALY savings were estimated by Delphi Consensus for 2 of these children. One of these children had autoimmune polyendocrinopathy syndrome, and vaccination for encapsulated organisms was assumed to reduce the risk of mortality. The other child was diagnosed with Factor XIIIA deficiency, and the initiation of prophylactic Factor XIII replacement decreased the risk of repeat central nervous system bleeds and associated mortality and neurologic complications. Cost savings were estimated for four of the eight children based on the change in clinical management.

Stark *et al.*[162] report a cost–utility analysis based on the same cohort of individuals used to underpin a previously published CEA, which has also been included in our review. They collected data from a cohort of 80 infants aged 0–2 years, presenting with multiple congenital abnormalities and dysmorphic features, or other features strongly suggestive of monogenic disorders, from a single tertiary paediatric centre in Australia.[163] In their cost–utility analysis, they analysed the cost-effectiveness of informative and uninformative results on continuing diagnostic investigation, changes in management, cascade testing in first-degree relatives and parental reproductive planning and outcomes. The time horizon over which costs and outcomes are estimated is unclear, but the longest time frame mentioned is 18 months. Schofield *et al.*[164] build on this analysis, adopting a 20-year time horizon, which the authors justify as being adequate for examining the long-term cost-effectiveness of WES without projecting beyond reasonable certainty in the outcomes of rare genetic diseases. Across both studies, health utility values were assigned based on parent-reported preferences for health states using utility values derived from another study.

Lavelle *et al.*[165] evaluated the use of WGS and WES in children with suspected genetic conditions. They modelled two separate patient populations, infants (< 1 year old) who are critically ill and children (aged < 18 years) who were not critically ill but with suspected genetic conditions. They conducted two analyses, one which produced the estimates of cost per additional diagnosis and another which produced the cost per QALY over a lifetime horizon. To achieve the latter analysis, they made a range of assumptions about the proportion of children who receive a change in clinical management following a diagnosis and the proportion who improve following this change in management. Children were assigned costs and QALYs based on their assumed level of long-term disability (none, mild, moderate and severe). The parameters for the model were taken from the literature.

Cost-effectiveness evaluations

Over a third of the included studies were cost-effectiveness evaluations (*n* = 29/86, 34%), that is, they reported the cost for a change in a specific health outcome. Twelve of these studies did not report the time horizon over which costs and outcomes were estimated. Seven studies estimated costs and outcomes over ≤ 3 years, or the time horizon was described as the diagnostic trajectory. Most studies focused only on the costs associated with arriving at a diagnosis. Two studies had a lifetime horizon.[166,167] These two studies were led by the same researchers in Italy, and both explored the cost-effectiveness of WGS versus WES and standard testing in paediatric patients with suspected genetic disorders. The key difference between the studies was that one used a decision tree modelling approach, and the later study used a Markov model approach. Both models were developed using the same data from a cohort of 870 paediatric patients who underwent testing in Rome. Clinical effectiveness was measured in terms of the number of diagnoses. Costs included diagnostic tests as well as management and therapeutic procedures. The authors explain that, although a cost–utility analysis is the preferred economic evaluation method in Italy, there were no follow-up data for the cohort and therefore QALY impacts from change in clinical management were not available. Both studies concluded that WGS sequencing would be cost-effective compared to WES and conventional testing.

Cost–consequence analyses

Around a third of the included studies (*n* = 29/86, 34%) were cost–consequence analyses, that is, they reported cost and health outcomes separately and did not compare results to a cost-effectiveness decision threshold based on the acceptable amount of incremental cost per improvement in a health outcome. Most of these studies also focused on outcomes relating to the number of diagnoses, but some also focused on outcomes such as time to diagnosis and change in clinical management.

Costing studies

Eighteen were costing studies and did not report health outcomes. Nine of these studies were micro-costing studies, where each individual resource use associated with conducting WGS or WES has been measured and reported. These types of studies are particularly useful in the absence of a standard tariff cost for a test in public healthcare settings.

One of these studies was conducted as part of the Scottish Genomes Partnership (SGP) study, where each step of conducting a trio-based WGS is described from clinical assessment and recruitment to participant feedback of the WGS result. Data on testing histories from 259 probands referred to 1 of Scotland's 4 regional genetics centres underpinned the analysis. Activities at regional genetics clinics, the regional genetics laboratories and the sequencing centre were all micro-costed. Costs to GEL were estimated based on the charges. For tests performed in Scotland, unit costs provided by NHS Scotland's genetics laboratories were assigned. For tests conducted in other UK laboratories, unit costs were sourced from these laboratories. The authors identified which section of the standard genetic testing pathway WGS is intended to replace, a section of the pathway estimated to cost £1841 (2018 prices) on average. This cost varied considerably by and within disease, however, and the authors provide a useful breakdown of the costs by condition. The total cost of WGS was estimated to be £6625 per trio.

There was one other UK-based micro-costing study, Schwarze *et al.*,[27] who estimated the cost of Illumina-based WGS in a UK National Health Service laboratory for a cancer case and a rare disease trio case. Cost data were collected for all steps in the sequencing pathway, and sensitivity analyses were conducted to identify key cost drivers. The estimated average cost for a rare disease WGS trio was £7050.11 (2016 values), with sequencing alone costing £4659 per case and bioinformatics and reporting costing £677. Interestingly, the cost for WGS for a cancer case and a rare disease case were quite similar, and the small difference was largely because the cancer case was based on two samples (tumour and germline), whereas the rare disease case was a trio (parents and proband).

Chapter 4 Conclusions

In summary, none of the studies identified evaluated WGS or WES in a screening context and, even when looking at WGS or WES in other clinical contexts, only 10 of the 86 included studies in the review were cost–utility analyses, which is the preferred outcome for the UK NSC cost-effectiveness evaluations. The majority of the included studies focused their time horizon or analyses solely on the diagnostic process, rather than incorporating any changes in patient management and the downstream cost and health consequences of these changes. Our review did identify some valuable micro-costing studies, which provide a detailed breakdown of the exact resource use required for WGS testing (see data extraction in *Report Supplementary Material 5* for details of cost components for each study). These are likely to be useful for any future economic evaluations of WGS for newborn screening; however, there is likely to be additional infrastructure requirements to facilitate WGS testing at this scale to consider and possibly some economies of scale.

The public voice on evaluating whole genome sequencing

Patient and public involvement and engagement group reflections on the topic discussions

At the outset, the group largely supported WGS for newborn screening. There was some diversity of views underpinning this in terms of the way that participants viewed potential harms and benefits, but overall, the benefits were seen to outweigh any harms. As meetings progressed and the complexities were explored, views became more nuanced, for example, one participant mentioned that they now were 'sitting on the fence a bit'. By the final meeting, there was still a sense of general positivity around the potential for WGS in newborns, but the group also articulated as many harms as benefits in the whiteboard exercise, illustrating their considered approach to balancing harms and benefits. Benefits identified included saving lives, improving outcomes and quality of life, a reduction in unnecessary and potentially invasive testing while searching for a diagnosis, information for future reproductive choices and enabling vigilance for the onset of symptoms. Harms identified included parental anxiety and unnecessary worry, the possibility of uncertainty rather than clarity resulting from test results, concerns around education of what screening results mean (both among the public and health professionals), questions of who owns the data and who they are shared with, the strain on health and support services and concerns around whether doctors will treat children as aggressively if the condition is thought to have a poor prognosis.

Reduction of harms focused on the provision of adequate support for parents and children (and children transitioning to adulthood), including emotional support, clear clinical pathways for diagnosis and treatment for babies identified as having a 'condition suspected' result, along with education on what screening results mean (including the possibility of uncertainty and the difference between a screening result and a diagnosis). Several group members reflected on poor experiences with receiving genetic diagnoses that led to harm, and in some cases, trauma, and they expressed the view that these issues with the current system must be addressed before any expansion of screening. Under the headings of 'potential harms' and 'what should happen when the evidence is not available?', several group members mentioned the option of targeted screening being a preferable approach, where results are only disclosed if they meet a threshold of confidence in the result. In these circumstances, it was suggested that the other data gathered could be used for research but that families should be notified if/when more became known about any uncertain findings.

Evaluation of the patient and public involvement engagement process

All participants engaged and participated in the process; they freely shared their views and experiences and openly listened to, and respected, the views of others. Participants responded that taking part was a positive experience. Some found it difficult to attend meetings due to other commitments (including caring commitments), and this led to one person withdrawing from the group. It also meant that at least some of the group members were missing from every meeting.

Reflecting on the process, one area that will require careful consideration for future iterations is the planning of the timescale for PPIE. Recruitment of PPIE representatives took several weeks, which meant that the time available for meetings became condensed. With a largely established PPIE group, this would be less of a concern for future

reviews involving PPIE, but if this process were to be initiated again, more time should be allowed for recruitment and introductions to key concepts. The process would have benefited from having more time to develop and discuss ideas. A lot of complex concepts were discussed over a short time frame and, despite the group having a relatively high level of familiarity with genetics, some participants commented that they would have liked to have more time to continue discussing particularly challenging topic areas. When establishing a new PPIE group, it is important that participants are given time to share their stories and build rapport with other group members (and the team), which needs to be built into the process. The group built rapport quickly, helped by their sense of shared experiences, as most were parents to a child (or adult child) with a rare genetic condition – although, their journeys through these experiences (and the conditions their children lived with) were very different. In PPIE groups with less shared experience, early rapport-building activities and careful chairing of the group are likely to be required.

Producing the documentation for pre-reading also required time and careful consideration to pitch the information at the correct level while also presenting a broad and balanced overview of the topics. The importance of not overburdening participants also had to be taken into account. Most of the documentation provided to the group was 'bespoke' and summarised information from the literature to ensure it was at the right level. This was also necessary to ensure that participants were not overwhelmed, and reading could be done within the hour allowed for meeting preparation. Again, for future iterations, the importance of carefully curating the materials provided to the group should not be underestimated.

The evaluation questionnaires were not only a useful source of evaluation data on the process, but these also proved helpful to allow group members to express views they may not have had the time for, or felt comfortable to express, in meetings. The comments on the forms (which included questions on whether there was anything that surprised them, or particularly stuck in their minds following a meeting) gave an insight into how their views were developing over time.

Comments from the group on the integration of PPIE into future reviews during the whiteboard exercise included the widening of perspectives to include other stakeholders, such as more charity representation, adults with later-onset conditions and children and young adults living with rare conditions. Some also commented on the inclusion of medical professionals, ethicists and representatives from commercial entities (insurance companies and pharmaceutical companies) in debates about particular topics. It was felt that the inclusion of members of the general public, while important, could be challenging in terms of recruitment and finding people with interest in this area and that they would need more background information and training in the subject area.

Chapter 5 Discussion

Main findings

The five traditional reviews yielded little of the evidence necessary for decision-making by the UK NSC. This is belied by the large scale of the five reviews which took 7 months for the identification, selection and review of published studies and comprised sifting nearly 20,000 titles and abstracts and 1348 full texts. The 221 included studies addressed only two of the six review questions. No studies addressed the questions on penetrance or test accuracy. No studies addressed the questions on the clinical effectiveness and harms of genetic screening. In fact, we did not identify any studies that reported on WGS in the newborn screening setting for any of the five conditions. Between one and nine studies across each of the five reviews reported observations on the early versus late treatment, where 'early' tended to only approximate screen-detected children and no study was designed to compare the treatment in screen-detected versus symptomatically detected children. There was some indication that earlier is better across all conditions, but studies were observational, small and we had concerns over their quality mainly around reporting and insufficiently clear definitions of disease.

Between 26 and 89 studies across the 5 reviews reported the genetic make-up of patients with suspected or confirmed disease for the 5 conditions. These studies were of interest for two reasons. Firstly, we wanted to compare the variant spectrum in children with clinical disease with that in asymptomatic children detected by genetic screening to gauge how applicable variant information from children with disease is to the screening context. The example of MCADD illustrated that the variant spectrum and, consequently, the variant frequency shifts when moving from sequencing symptomatic children to sequencing those with biochemical risk factors (positive NBS test), suggesting that studies in children with disease have limited applicability to screening[168] and genetic variants believed to be pathogenic can be very poorly predictive of symptomatic disease when detected by population screening. However, no studies were identified that used sequencing in unselected newborns. We were, therefore, unable to compare the variant spectrum in unselected and symptomatic newborns. This means that the review provided no information on the variant frequency, variant interpretation or penetrance in screening populations or on how much this differs from populations with confirmed disease (additional to the shift from symptomatic to biochemical disease).

Secondly, we wanted to determine the proportion of children with disease, which can be detected by sequencing to assess how many patients may be missed by newborn sequencing. This was complicated by the fact that the studies included different populations of patients, illustrating that disease can be defined in different ways. Before we can determine the detection rate of WGS in screening (i.e. the proportion of patients with disease identified through sequencing), a clear definition of disease is needed. This is not straightforward as the disease can be defined symptomatically or biochemically using various thresholds. Many diseases have similar clinical and biochemical characteristics and can only be diagnosed genetically. For instance, determining the proportion of children with XLHR caused by variants in the *PHEX* gene is complicated by the fact that XLHR has no disease-specific symptoms or biochemical markers. Defining XLHR symptomatically includes additional conditions caused by different genes, which affects the proportion detectable by sequencing the *PHEX* gene. This illustrates that measuring the detection rate of WGS in this way is impractical.

Our review identified other challenges for WGS for the screening of newborns. Firstly, the type of test had an impact on the genetic outcomes reported. This included, for instance, the scope of sequencing, that is, the number of exons covered and whether intron/exon boundaries were considered; the use of additional tests (e.g. MLPA to detect deletions and insertions), as studies generally used exhaustive genetic testing to identify genetic causes in clinically affected patients, and year of testing (since there is a rapid change in test development with the addition of genes and variants over time). This highlights potential concerns over the applicability of some of the study findings to WGS in the screening setting and limitations of WGS as a stand-alone test. Secondly, ethnicity had an impact on the number and type of variants identified.[168] This limits the generalisability of some study findings to the UK context and highlights the importance of considering an ethnically diverse population in the evaluation of WGS for newborn screening and an evaluation of the potential impact of WGS on inequity and inequality in screening. Thirdly, we excluded several studies

that focused on specific subgroups of patients. While unrepresentative to the review question, they identified particular challenges. For instance, patients with mosaicism are not identifiable by WGS, patients with digenic disease would be classified as negative and patients with atypical symptoms may not be identified if knowledge of their variants is limited. Finally, PDE and XLHR had a high number of novel and private variants for which pathogenicity can be difficult to ascertain. These and conditions with similarly high numbers of private variants may pose a challenge to WGS in screening because of potentially high numbers of variants of uncertain (or unknown) significance.

The ClinGen is a useful resource for information on the medical actionability of rare monogenetic diseases. Four of the five conditions reviewed have currently got an actionability report available on ClinGen. When compared to the UK NSC criteria, we identified two main limitations. Firstly, the availability of information on penetrance is not a requirement for the inclusion of a condition on ClinGen. Secondly, the benefits of early versus late treatment are not a focus of the ClinGen review. In addition, the evidence bar for a positive inclusion decision is lower than that for the UK NSC recommendations. As a result, it would be inappropriate for the UK NSC to base decisions for potential screening programs on the actionability reported in ClinGen without further assessment.

Our review of genomic studies of newborn screening populations provided information on international experiences with whole genome and panel testing for screening and illustrated that genetic screening is still in its infancy. Synthesis of the evidence revealed a lack of consensus on which genes to include, no indication of how to interpret discordant results from NBS programmes and genetic screening, uncertainty over which test is most appropriate for which condition, evidence of overdiagnosis and a large number of cases that could not be confirmed by the studies' follow-up processes resulting in interventions, routine surveillance and regular follow-up of uncertain benefit. All of these would cause problems if WGS was implemented prematurely. The studies did not provide the information on variant penetrance and expressivity, which would be needed to determine a variant threshold for individual genes (number and types of pathogenic variants) to select variants for the evaluation for a screening programme that maximises benefits and minimises harm from overdiagnosis. One study reported 31 cases of PKU, of which 15 were caused by 13 different variants which were all confirmed by clinical presentation. This heterogeneity precluded penetrance estimates to be determined for individual variants even in this large study. The remaining 16 cases were clinically unconfirmed and involved a common variant that was believed to cause mild disease. This may represent one example where study information could be used to exclude a variant from consideration for reporting.

The three main approaches explored do not offer an immediate solution to the evaluation of WGS for newborn screening in the future. Our reviews highlight many evidence gaps and challenges for implementation, which may be a contributing factor to the wide variation between studies in decisions of which variants are reported to parents as potentially clinically significant. We conclude that information on the penetrance of pathogenic variants selected for inclusion in genomic studies in the screening context should be one focus of future research efforts.

Our review, unsurprisingly, identified no economic evaluations of WGS in a screening context. All the evaluations focused on individuals with suspected genetic diseases and explored the cost, cost consequences or cost-effectiveness of WES or WGS compared to the standard of care diagnostic testing or to different types of sequencing (rapid vs. non-rapid) or positions in the diagnostic pathway. Many of the studies based their evaluations on data collected from reasonably small (all < 1500) cohorts of individuals, either receiving the intervention or current standard of care or both. Most evaluations did not adopt a lifetime horizon for capturing costs and health outcomes, as recommended, and most were limited to outcomes and costs for the diagnostic trajectory rather than including costs and outcomes associated with changes in clinical management and the consequences of this change in management. Very few studies conducted a cost–utility analysis. This is likely due to a number of reasons. Firstly, WGS can produce a wide range of genetic information, affecting multiple conditions and health outcomes simultaneously. Where this was an issue in the included cost–utility studies, very broad classifications of health states and assumptions around outcomes were made to assign utility values. However, these broad classifications may oversimplify the complexity of the impact of WGS on HRQoL. Secondly, the long-term nature of the impacts of WGS on health outcomes means that the HRQoL needs to be measured or estimated over extended periods of time. Collecting these more detailed data over long periods of time can be challenging and resource-intensive. Finally, there is also the issue of estimating the causal effect of WGS on health outcomes, which would typically require a RCT design, which can be difficult to implement effectively for rarer diseases due to the large sample sizes needed to achieve an adequate statistical power.

Our update of the Schwarze *et al.*[27] systematic review did identify a number of high-quality micro-costing studies, where a granular account of all the resource use associated with WGS or WES had been recorded and costed. These types of studies are incredibly useful as they can be updated as unit costs change over time, and they also help to understand which aspect of the testing pathway is driving the cost. One of the studies identified in the review was published as part of the SGP study, and we have been made aware that similar activity is underway but not yet published by GEL. We would encourage future UK micro-costing studies to be undertaken, particularly when conducted in a routine clinical practice setting, rather than research setting, to help underpin future UK-based cost-effectiveness evaluations of WGS. It would be useful to have these available across a range of clinical contexts to understand more fully when and why costs may differ.

In this review, standard of care testing that frequently included some sort of genetic testing was often listed as a comparator. Costs associated with standard of care varied widely between individuals, demonstrating the importance of developing large, heterogeneous cohorts of individuals to truly capture the breadth of resource use. Standard of care will be very difficult to define when evaluating WGS in a screening context unless there are clear comparative data on which conditions would have been detected later, when and using which diagnostics. Standard of care in terms of clinical management was rarely included in the economic evaluations in this review, and it is likely to add another layer of complexity, especially when trying to elucidate the benefits and harms associated with earlier diagnosis.

Contribution to existing knowledge

There are no previous systematic reviews of the benefits and harms of WGS of newborns and so our review is novel. We have not identified any other group who is approaching the question from the same perspective. The International Consortium on Newborn Sequencing is a consortium of international genomic projects aiming to bring together, exchange and harmonise ideas and efforts to responsibly implement newborn sequencing.[169] This is a useful source of information, but the focus of the consortium is implementation rather than evidence synthesis for policy advisors. While the consortium website portrays certainty of health benefits from newborn sequencing, our review highlights the uncertainties and evidence gaps that could result in harms if WGS is implemented prematurely.

There are currently at least 15 genomic projects underway internationally aiming to sequence over 400,000 newborns.[169] The study by Downie *et al.*[170] compared gene lists across six of these genomics projects and reported substantial differences in genes' lists due to different starting points, with some prioritising the clinical validity of the gene–disease association, while others prioritise treatability. We have observed similar disagreement in gene lists due to different processes and priorities used across the genomic studies included in our review of penetrance in newborn screening populations. This lack of consensus may reflect the uncertainty of what to report to parents as potentially significant, which is based on current evidence. The gene selection process is complex and multifactorial. It requires evidence to assess whether the condition is monogenic, whether the genotype/phenotype link is established by identifying pathogenic variants and the penetrance and expressivity of these variants. Evidence on pathogenicity, penetrance and expressivity is limited for general population cohorts. Because there is less proof of pathogenicity for likely pathogenic variants, there is no consensus on reporting likely pathogenic variants in screening, while VUSs are generally not reported. It is important to understand and evaluate the underlying gene and variant selection processes that underlie the decision on what to report to parents, and our work on this could be expanded in the future to a full systematic review to understand the different approaches to gene and variant curation and to synthesise the commonalities and differences to inform a robust and unified process.

The ClinGen resource represents one open access effort to develop and implement standardised, evidence-based methods to characterise the clinical actionability of genetic variations. However, our review found that the evidence bar for inclusion is low, the focus around treatment is clinical and therefore on treatability rather than benefits of early versus late treatment; and, while, some aspects are relevant to the screening context (e.g. treatment for pre-symptomatic children), focus is on the secondary findings in individuals in the diagnostic setting where the definition of an important condition may be different to the public health perspective of the screening setting.

Strengths, challenges and limitations

Strengths

To the best of our knowledge, this is the first attempt to identify an approach to assess the benefits and harms of WGS for newborn screening for policy advisors. This was a huge systematic review applying both traditional and novel approaches, and it has the potential to guide policy-makers towards a new approach for both synthesising and producing the evidence required to make evidence-based policy decisions in this complex area. We consulted widely with policy advisors and clinicians during protocol development to frame the right RQs for this review. We used a stratified random sample of conditions for our traditional review and published our protocol on PROSPERO. We consulted with clinicians specialising in the five disease areas, geneticists, clinical and laboratory advisors to the UK NBS programme and members of the UK NSC throughout the review process to ensure we deliver what is required by decision-makers and undertook an extensive piece of work to consider the public and patient views on this ethically challenging topic. While all these aspects ensured a high-quality review, we learned from several challenges that can inform subsequent reviews.

Challenges

Reviewing five complex, very different conditions at the same time was challenging. It is impossible to gain the knowledge and insight for five conditions in a similar time frame, as we would normally have to review a single condition. We attempted to work as one team for consistency in decision-making but had to focus on two to three conditions each. We attempted to use one review approach across all conditions, but we quickly recognised that the conditions required individual considerations. A review of 200 conditions would ideally require 200 review teams or an extensive time for a smaller number of teams to ensure that the rigour and subject knowledge can be developed for each condition. Issues around consistency would need to be factored in and addressed if several review teams undertake the evidence syntheses.

The outcomes presented here represent a snapshot of the evidence in this fast-progressing field. Gene and variant lists are constantly changing and developing in light of new evidence. New variants are identified, and variant classification changes based on new insights from clinical practice and research. Evaluations and policy decisions of 200 conditions once completed will not be up to date for long and will require a system that allows continuous evaluation.

The review highlights the challenge of defining disease because we were thinking across genetic, biochemical and clinical diseases. It is important to be clear about what aspects of disease we would aim to detect and prevent with WGS. This is linked to the challenging concept of penetrance, which depends on the definition of disease and at what time point we measure the disease. This will determine the required follow-up time of screen-positive newborns for the estimation of penetrance and is likely to be different for different conditions.

A further challenge was the integration of expert clinical advice into the review, and this will be a challenge in the future on how to incorporate clinical knowledge into the evidence base to inform a policy decision. Clinical knowledge was important and nuanced and was complementary to the review findings. The issue is that the information comes from individuals; however, in these rare diseases, it would be challenging to identify a group of experts large enough to establish a consensus. We were advised that (1) targeted gene panels would be more economical than WGS (and others share that view[171]), but this is not currently a question that is being addressed; (2) meaningful time scales from testing to reporting are unlikely to be achievable in the overstretched NHS where current waiting times for WGS results for the diagnosis of symptomatic children often exceed 1 year[172] and (3) storage of huge amounts of data is expensive and may not be accessible in the future, while re-sequencing may be more appropriate than accessing stored data in the future. These are important issues that our current review did not ask or identify from the published evidence. Furthermore, one clinical advisor said that they would never initiate asymptomatic vitamin B6 (pyridoxine) because of treatment-induced seizures, yet this is the early intervention of interest here. This was an important finding but was not backed up by published studies. It was impossible for us to know how common this view is.

Limitations

Our approach of selecting the five conditions aimed to include a range of scenarios that would allow us to explore different challenges that the NHS would have to face with newborn sequencing. This meant that the conditions, the

aim of screening for the conditions, the definitions of early versus late treatment and their underlying genetics were very different. Consequently, we were unable to use one approach to reviewing as planned and concluded that it is unfeasible to use one review approach for 200 conditions. However, we do not know whether this applies to a group of similar conditions, for example, metabolic conditions, where learnings from one review may be transferable to other conditions speeding up the review process. Repeating this approach with a group of similar conditions could shed light on this.

The number of studies reporting the genetic spectrum in children with confirmed disease was unexpectedly high for all five conditions. The studies were heterogeneous in terms of population, test and outcomes and would have been difficult to synthesise. The most appropriate population with disease to determine the detection rate for WGS was difficult to define for XLHR and fHLH because of non-specific symptoms. We, therefore, decided to categorise the studies by definition of the disease and extracted data from the largest study only (or more than one where results were complementary) to present the breadth of the evidence rather than aiming for completeness. The aim was to provide examples of outcomes depending on disease definition that could inform the inclusion and exclusion criteria of a future review for this question. However, this meant that we were unable to draw on all reported findings on variant frequency and expressivity or to highlight aspects where studies may have agreed or disagreed. We are uncertain how useful synthesising all the included studies would have been.

We were unable to undertake a number of explorations detailed in the protocol because the shortfall and type of data meant that they were unfeasible. We wanted to explore reporting of penetrance information from the genomic studies of newborn screening cohorts (1) by subgroups of conditions pre-defined by the studies, (2) by the GEL category the conditions would fall into and (3) combined for the conditions with a top quintile ClinGen score. The aim was to explore the feasibility of determining the variant threshold that corresponds with the most severe phenotype for conditions. However, this was not feasible because (1) categories are at condition/gene level and penetrance is reported on variant level, (2) the penetrance information reported in studies was inadequate numerically and qualitatively, (3) the ClinGen score was not reported on variant level and (4) the number of conditions covered in ClinGen was limited. This means that we were unsuccessful in exploring whether a variant threshold for individual conditions could be used to develop a restricted but safe screening programme.

Studies comparing genetic screening with traditional NBS screening suggest that tests are complementary in terms of cases they detect and miss, but they do not tell us anything about the place of WGS in an existing screening program; that is, how both tests could be integrated, how to interpret contradictory outcomes, whether the sequence of tests should be different for different conditions or whether newborn sequencing should be a completely separate programme. In general, the combination of metabolomics alongside genomics, in whichever order, is likely to add to our understanding of genetic variants and their significance, and considering all research in a future review will help to identify the best screening strategy for individual conditions.

Furthermore, there are several questions that we did not address or where there was insufficient evidence available from the studies we reviewed, but those are fundamental for the evaluation of newborn genetic screening.

1. The differences in variant frequency and penetrance in screened (unselected) and symptomatic cohorts is important to understand, but data are lacking for most conditions. It is widely accepted that risk estimates for genetic disease from high-risk groups do not translate to the general population as has been shown for cancer predisposition genes in individuals with and without FH.[173] And, in MCADD, a condition that is already screened for by NBS screening with genetic confirmatory testing, the shift from symptomatic disease to biochemical disease definition resulted in a different variant spectrum.[119] In this study, a lower proportion of screen-detected newborns were homozygous for the c.985A > G variant, which is very common in symptomatic children. A significant number of the newborns had genotypes with variants that had not been observed in patients detected clinically and some, like c.199T > C and c.127G > A, were associated with a milder biochemical phenotype. This knowledge is important for all conditions considered for genetic screening and needs to be extended to understanding the variant spectrum in infants designated screen-positive based on genetic testing only.
2. There was insufficient evidence to explore the differences in variant frequency and penetrance by sex, geographical region and ethnicity. Variants of uncertain (or unknown) significance are more common in non-European ethnic

groups because of an under-representation in reference population databases used for variant frequency annotation.[8] Consequently, there is a lack of genotype–phenotype correlations in populations, such as the UK and USA, with significant ethnic diversity. This is a particular problem for whole-population screening and means that it will be difficult to establish an equitable screening programme if ongoing genomic studies do not generate the evidence from an ethnically diverse population.

3. The review did not address the question of whether WGS is the best test for expanded newborn screening compared to WES, panel testing, sequencing combined with other tests or biochemical assays. Many conditions could be tested for using biochemical assays or bespoke genetic tests, which could be superior to WGS.[172] WGS has got limitations and is generally more expensive. It is unable to detect large deletions and needs to be adapted, for instance, to identify copy number diseases like spinal muscular atrophy (SMA) to ensure they are reliably detected. Furthermore, it can also produce inconclusive results as in cystic fibrosis screening.[174] The gene-by-gene and exon-by-exon performance of WGS is improving all the time, but this remains a factor at the time of writing.
4. Test failure of WGS was insufficiently reported by the included studies.
5. We did not explore the impact of sequencing results on clinician behaviour and clinical care.
6. We did not investigate workforce challenges, which will include the number of genetic counsellors needed, and training of staff to correctly interpret WGS results.
7. We did not identify evidence-based pathways for those with rare asymptomatic genetic disease.
8. More work is needed to explore ethical issues around knowing but not reporting certain variants.
9. More work is needed to explore ethical issues of storing blood samples/genetic information, their re-use and who decides what can be done with the data.
10. Existing resource use data associated with WGS need to be adapted/adjusted to account for the additional infrastructure and staff required to deliver WGS at scale, that is, in the context of a national screening programme.
11. Long-term follow-up on cohorts of, ideally, asymptomatic newborns undergoing WGS is needed to understand the implications of testing on patient management in terms of costs and patient outcomes.

Chapter 6 Patient and public involvement and engagement

The PPIE process itself was successful and viewed positively by participants, who all expressed a view that they would be interested in future PPIE work. The group started the process with a broadly positive view of the benefits of WGS for newborn screening, but the more they identified and developed their understanding of the potential harms, the more cautious many of the group became. This process of increasingly critical, or ambivalent, views being expressed as information and discussion increase has also been observed in other groups considering complex topics in the field of genomics and screening.[175] In the final exercise, several expressed the view that a targeted approach to genomic screening may be preferable to high-throughput screens. This was not true of all of the groups, however, and some remained largely supportive of WGS despite potential risks or harms, seeing all forms of knowledge as useful. This divergence in opinions has been highlighted in the literature,[176] where for some, non-actionable or uncertain results can be seen as empowering, whereas for others, concerns about risk and unnecessary anxiety and stigma were identified. Some parents of children with health conditions demonstrate a greater tolerance for uncertainty from screening because they are already experienced in medical uncertainty.[176]

The way in which conflicting perspectives on complex topics in genomics and screening should be prioritised and weighted in research outputs and policy decisions is a debated topic within acceptability research.[177] Further research with collaborative PPIE engagement that can map such diverse views to social characteristics, backgrounds and/or particular lived experiences may prove particularly useful in facilitating a nuanced understanding of the distributions of harms and benefits of genomic screening to inform policy recommendations and implementation. For this reason, it would be beneficial to increase the diversity of viewpoints if this PPIE process was to be repeated. A limitation of our work was that the recruitment strategy did not target participants based on cultural diversity (although there was some), nor did it aim to be reflective of the UK population. This was not achievable, given the timescale. Future work should allocate more resources to recruitment and should recruit participants from outside of the rare disease community to achieve more representation from members of the public without the experience of living with rare conditions. This will be particularly important as most members of the public offered screening will have no, or limited, experience with rare genetic conditions. The group mentioned the potential for difficulty in recruiting members of the public where the relevance of the topic may not be immediately evident to them (especially, if they are not new or expectant parents) and the added time that would be needed to educate public participants on the complexities around genetic screening and diagnoses. Familiarity and factual knowledge around genomics among the public are low[178] and so time would need to be invested in introducing key terms and concepts around genomics, sequencing and screening to enable meaningful contributions.[179] When this is done, research has shown that previously 'genetics agnostic' members of the public can make substantial and rich contributions.[180] Given the backgrounds of members of this group, and their existing familiarity with genetics, this step could be greatly condensed, but future PPIE work should accommodate additional time for information-sharing and the incremental building of knowledge and debate so that a common language for talking about genomic screening can be established.[179] The group also mentioned the value of bringing in specialist contributors/viewpoints for particular topic areas, such as representatives from the insurance industry.

It is important to allow adequate time and consideration to the composition and recruitment of future PPI groups as well as time for new groups to build rapport and trust. It would also be valuable to allow time for the training of participants in PPIE in reviews and give them the resources to codevelop the aims, terms of reference and timetable for their contribution to the work (which was not possible in this rapid review setting). Finally, the importance of investing time and consideration into the development of materials, resources and viewpoints to share with the group ahead of discussions, in an accessible format, should not be underestimated. This is an essential component of creating a deliberative space to allow meaningful contribution and collaboration between the research team and PPIE contributors.

Chapter 7 Equality, diversity and inclusion

The potential lack of adequate representation of people of different ethnic backgrounds in the evidence reviewed and the implications of this for population screening were highlighted above (Limitations). In addition to ensuring that ongoing genomic studies generate evidence for genotype–phenotype correlations from an ethnically diverse population, it is important to ensure an inclusive approach to the recruitment of PPIE contributors and the patients whom they represent.

Recruitment was targeted to people with experience with rare conditions (adults living with a rare condition and parents of children with rare genetic conditions) and advocates working in the area of rare genetic conditions. We were successful in recruiting people with a broad range of experience with genetic conditions. None of the group members with lived experience of rare genetic conditions (parents and adult living with a rare condition) had experience with the same genetic condition as anyone else in the group. Most of the group members had not been involved with research or PPIE previously.

Due to the time constraints on recruitment and the focus on experience with rare conditions, we did not target recruitment by demographic variables (age, socioeconomic status, geographical location, ethnicity and so on). Most group members were female, which may reflect the tendency for mothers to become the primary caregivers when they have a child with 'medical complexities'.[181] There was a diversity in the ages of the children they supported (under 5 years to over 20 years) and therefore the stage of caring for/supporting a child with a rare condition. As indicated in the discussion, for future work (and with more time available), we would wish to expand the viewpoints represented, by, for example, including more public voices, and ensure diversity of background and experiences, including recruiting from underserved groups.

Much consideration was given to the amount, format and content of documentation provided to participants to maximise accessibility and their understanding of concepts and complexities (e.g. terms such penetrance and expressivity). This also included producing a lay summary of the draft report for meeting 4, where the PPIE group had the opportunity to discuss the findings and ask questions of a representative of the review team. Evaluation forms sent after each form checked with the participants that they were happy with the information that had been provided and asked if we could improve on this; feedback for this was entirely positive, and all participants expressed an interest in being involved in future work in this area.

Chapter 8 Impact and learning

This review will inform the UK NSC in their approach to evaluating WGS for newborn screening. The learning from this review is that a traditional review will unlikely be an effective approach for the evaluation of 200 conditions. We propose an alternative pragmatic approach focusing on first addressing the evidence for penetrance in the screening context for variants identified as actionable by current genomics projects like the Generation Study, a critical question with implications for screening benefits and harms for which appropriate quality evidence is currently not available. Penetrance may be defined differently for different conditions. A biochemical confirmatory test may be sufficient for some conditions to estimate penetrance if the link between biochemical and clinical disease is well understood and strong; for others, a follow-up test may not be sufficiently predictive of clinical disease, and, for those conditions without an available confirmatory test, follow-up to clinical symptoms is essential. Where there is good quality evidence for high penetrance, the evidence synthesis could be expanded to other questions relevant to the benefits and harms of screening.

The review may also prompt the UK NSC to discuss evidence requirements for decisions on screening programs for rare and ultra-rare genetic conditions. This will be useful as it will re-focus the discussion on evidence in relation to WGS for newborn sequencing.

Chapter 9 Implications for decision-makers

Currently, there is no evidence supporting the large-scale implementation of WGS of newborns with concomitant simultaneous detection of many conditions. The cost and the balance of benefit and harm are unknown, and implementation would prevent the research required to measure the benefits and harms (see Research recommendations section). Our review of genomic studies of newborn screening cohorts demonstrates unequivocally that introduction of WGS without substantial further research would cause harm and uncertainty.

We found that the traditional approach to evidence synthesis can currently not be applied to WGS of newborns, the quantity of work is very high and the data provided are of insufficient quality and therefore insufficient value. Evidence from clinically detected cases is often not generalisable to screening, where the spectrum of disease differs, and less clinically significant disease types are more common. While there are many unknowns about the benefits and harms of screening, we advocate addressing the accurate measurement of penetrance first and filtering review efforts concerning other aspects, such as benefit of earlier treatment, only in those conditions with variants with sufficient penetrance. This proposal is a pragmatic stepwise approach based on the paucity of data and the need to engage with the broader scientific community to deliver the types of studies which will deliver the required penetrance data (see Research recommendations section).

Another possible approach to evidence synthesis is to use existing gene and variant curation by genomics projects. Decisions on which conditions, genes and variants to report are generally based on the severity of disease, disease onset, penetrance and expressivity of pathogenic variants, treatability and access to treatment. Gene curation processes vary greatly but are all based on a condition-by-condition approach. Our assessment of the GEL approach and ClinGen approach showed that they do not meet the requirements of the UK NSC and are not suitable for adoption. An assessment of approaches more widely may be needed.

Therefore, what is needed is commissioning of carefully designed research to generate new evidence with a focus on penetrance data in a screening setting for genes previously selected by gene curation processes. Future review efforts can focus on the narrower questions of penetrance and expressivity of pathogenic variants in large screening cohorts first, with additional questions about earlier treatment benefits only in those with sufficient penetrance and expressivity. Future large-scale WGS testing and screening programmes should be established in ways that support robust evaluation, research and assessment of clinical effectiveness to determine immediate and longer-term costs, benefits and harms.

To date, economic evaluations of WGS have focused on symptomatic populations and have focused predominantly on the costs and outcomes associated with the diagnostic process itself. Studies that have attempted to estimate cost per QALYs associated with WGS/WES testing, of which there are few, were either focused on very specific heath conditions where it is easier to define distinct health states, or have applied very crude quality of life and survival estimates to extremely broad health states. Creating a model which accounts for all the conditions is highly likely to be unfeasible. It may be possible to group some of the conditions by their biological pathways or clinical manifestations, providing a mechanism for identifying more homogenous health states to which you could apply more precise HRQoL, survival and cost estimates. There will inevitably be a high degree of uncertainty associated with any model developed for this evaluation question, driven by an inability to capture the whole decision problem rather than parameter uncertainty or model structure limitations. Methods to interrogate this type of uncertainty and understand the risk of making the wrong decision (i.e. recommending/not recommending WGS for newborn screening) in this context would be valuable to support policy decision-making.

Chapter 10 Research recommendations

Whole genome sequencing of newborn babies is a complex intervention with interdependent factors (such as gene selection, variant pathogenicity, penetrance, expressivity and benefits of earlier treatment) contributing to the balance of benefits and harm. There is a paucity of evidence around these factors for the screening context, and we have very low levels of knowledge about the balance of benefit and harm. We have highlighted many research gaps that span the complete evaluation process of WGS for newborn screening and propose possible research approaches to address these. We do not propose an order of priority for these recommendations, as all the research is needed to inform future policy decisions about WGS. It is imperative that this research is undertaken as part of large joined-up and possibly international collaborations to produce the evidence that is needed to thoroughly assess the benefits and harms of WGS for newborn screening for rare and ultra-rare conditions.

The research recommendations are as follows with justifications outlined below:

- a systematic review of gene and variant selection approaches reported in newborn genomic studies
- large studies to establish the penetrance of pathogenic variants in the newborn screening population [either cohort studies, healthy adult biobank studies, i.e. looking for absence of variants believed to be highly pathogenic in childhood or studies of newborn dried blood spots (DBSs)]
- clinical effectiveness studies comparing newborn genetic screening with no screening using rare disease registries
- comparative test accuracy studies of gene panel tests, WES, WGS and expanded NBS testing
- UK micro-costing studies to help underpin future cost-effectiveness evaluations of WGS
- broader research about public perception and understanding of WGS, views on risks and uncertainty and acceptability and accuracy of WGS in different populations and ethnic groups.

A systematic review of gene and variant selection approaches

A systematic review of studies reporting gene and variant selection approaches will help in the understanding of prioritisation and reporting decision of genomic projects. This may aid in the formulation of a consensus, best practice gene curation approach for newborn genetic screening and inform the evaluation of GEL's gene list and classification of pathogenicity of included variants as part of the future evaluation of WGS for newborn screening.

Large studies to establish the penetrance of pathogenic variants in the newborn screening population

We recommend addressing the key question of penetrance in a screening population and produce the evidence needed for an evaluation. The evaluation of the subsequent interdependent factors can then focus on conditions only where there is good evidence of high penetrance for at least some pathogenic variants. Large research studies implementing newborn screening using WGS, and reporting what they detect (i.e. diagnostic yield) and offering treatment, do not provide the penetrance and expressivity data required for decisions about whether to implement. This is because if variants are reported to parents as a positive diagnosis in their newborns and they are offered treatment, and they go on to have a good health outcome, there is no way of knowing whether treatment was curative or whether they would never have had any symptoms or effects and did not require treatment. So, it is unknown if they benefited or were harmed. Current newborn genomic projects report results for variants to parents, whose babies will receive clinician-guided management. This treatment provision precludes the estimation of penetrance from the study. If studies were to choose to report fewer pathogenic variants, only those where there is existing good evidence that penetrance is high, that would provide the evidence that policy-makers need on the penetrance of other pathogenic variants (which can be accurately measured through follow-up to symptomatic disease in the absence of diagnosis and treatment). Another key challenge in this area is that even with a very large sample of 100,000 newborns, the power to analyse results by the condition or variant is extremely low due to the rarity of most relevant conditions.

To understand penetrance, the following combinations of research studies could be employed:

1. The ideal studies would be large cohort studies where either screening with WGS is given to newborn babies, without reporting results to parents or very few conditions of well-evidenced penetrance and expressivity are

reported to ensure more good than harm. Other potential conditions/genes should not be reported to participants unless they present symptomatically to measure the penetrance and clinical significance of pathogenic variants and therefore the benefits and overdiagnosis harms of revealing the test results. This would enable the accumulation of evidence about penetrance for conditions/variants, which could be assessed on an ongoing basis for addition to the study/programme, if there is sufficient evidence of more good than harm. This may generate the evidence needed to allow a gradual or stepped implementation in the future based on levels of evidence on pathogenicity and penetrance/expressivity, burden and cost of available treatment, availability of confirmatory tests and disease onset. However, these studies would raise considerable ethical questions and would only be possible with parents' complete understanding, agreement and consent. There would be concern that once parents were asked whether they want to participate in a trial of WGS without telling them what you find, none of them will be in equipoise.

2. The existing large genomic screening cohort studies can produce evidence on the penetrance for variants which are not included on their panels for reporting and treatment. This may provide some useful data because of the large variation in gene lists between studies. However, a note of caution should be applied because some of the rationale for different inclusion of variants may be differences between populations, which would mean we would be primarily measuring the penetrance for variants that are of lesser importance in that population and penetrance measures may not be generalisable to the UK population. A further complication is that good quality follow-up to confirmed disease is necessary, such as using robust disease registries, which is not always available.
3. Genetic information from healthy adult cohorts, such as the UK Biobank and worldwide datasets like GNOMAD, could be used to identify low penetrance variants which are not suitable for inclusion on a screening panel. If large numbers of healthy adults have a variant believed to be associated with childhood-onset disease, then it will not be a highly penetrant variant suitable for use in newborn screening. However, these studies are less useful for identifying which variants to include because they will exclude people with the disease who have died before reaching adulthood.
4. Whole genome or exome sequencing of stored DBS samples could provide excellent data on penetrance if the samples are of sufficient quality for sequencing and if there is a robust system of follow-up to ascertain symptomatic disease status. An example of this approach has been successfully applied in California, where Adhikari *et al.*[182] performed WES on DBS samples and achieved exomes comparable to exomes from fresh blood in 1090/1416 samples. However, only samples from IEM-affected and MS/MS false-positive samples were included and WES results were compared to follow-up testing precluding the estimates of penetrance for the screening context. The challenges of doing these studies at scale are the cost and linking to good phenotype data.

Clinical effectiveness studies using rare disease registries

For evidence on the clinical effectiveness of genetic screening in newborns, future research could explore the role of rare disease registries. A recent study from Germany evaluated the effectiveness of genetic newborn screening for SMA by comparing the outcomes in screen-detected and symptomatically detected patients (asymptomatic vs. symptomatic treatment start) within the same healthcare system.[183] This was feasible because two pilot projects for genetic SMA newborn screening were performed in Germany in two federal states before its nationwide implementation in 2021. This represents an excellent example of how a study can be designed within a rare disease registry. However, registry data are not collected to address specific RQs and limitations will be common. Registry studies rely on the high quality of data collection and reporting. The registry studies identified in our reviews were of low quality because none was comparative, no information on the type of test was provided and the information on the definitions of disease was insufficient.

Comparative test accuracy studies

Comparative (test accuracy) studies of gene panel tests, WES, WGS and expanded NBS testing are required to provide the evidence on which type of test is most promising for newborn screening. A technology-centric approach (WGS for all conditions) has been identified as inappropriate for population screening[172] and the best-suited test needs to be identified for each condition under consideration. For instance, WES has been shown to have insufficient sensitivity and specificity to replace MS/MS for IEMs in general, but effectiveness varied among individual IEMs.[182]

United Kingdom micro-costing studies

The UK micro-costing studies are needed to help underpin future cost-effectiveness evaluations of WGS, particularly when conducted in a routine clinical practice setting, rather than a research setting and across a range of clinical contexts, to understand more fully when and why costs may differ.

Studies on public perception, views and acceptability

Finally, there is a need for broader research about public perception and understanding of WGS and whether parents would still be in equipoise once they are fully informed. Further research should aim to map diverse views on complex topics in genomics and screening and associated views on risks and uncertainty to social characteristics, backgrounds and particular lived experiences to facilitate a nuanced understanding of the distributions of harms and benefits of genomic screening to inform policy recommendations and implementation. Future research also needs to investigate the acceptability and accuracy of WGS in different populations (including island populations who are genetically diverse from mainland UK) and ethnic groups, including those where consanguinity is common, and any resultant ethical and equity issues. Population genomic testing experience is skewed towards Western Europeans and has resulted, for instance, in poorer prediction of polygenic risk scores in non-European populations.[172] Concerns over deepening health disparities with WGS are warranted.

Chapter 11 Conclusions

A traditional approach to systematic reviewing for WGS of newborns is unfeasible, and the review does not reveal a new way to evaluate WGS for newborn screening in a single mechanism. There are two reasons for this: (1) there was insufficient evidence on each condition to allow a conventional UK NSC assessment and (2) the variations in penetrance, natural history, test accuracy and effectiveness of treatments mean that an aggregate is not informative. Our review highlights the main evidence gaps and informs the direction of future research efforts.

We propose a series of possible research approaches undertaken in large joined-up collaborations to produce the evidence that is needed for policy advisors before an evaluation of WGS is feasible. This may include a co-ordinated international approach to collecting penetrance data. This could be followed by a staged approach of evaluation considering only pathogenic variants with a very high penetrance for screening.

Additional information

CRediT contribution statement

Karoline Freeman (https://orcid.org/0000-0002-9963-2918): Conceptualisation, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – reviewing and editing.

Jacqueline Dinnes (https://orcid.org/0000-0003-1343-7335): Conceptualisation, Investigation, Methodology, Supervision, Writing – original draft, Writing – reviewing and editing.

Bethany Shinkins (https://orcid.org/0000-0001-5350-1018): Conceptualisation, Investigation, Writing – original draft, Writing – reviewing and editing.

Corinna Clark (https://orcid.org/0000-0002-1077-9383): Investigation, Writing – original draft, Writing – reviewing and editing, patient and public involvement contributions (lead).

Inès Kander (https://orcid.org/0000-0002-0634-3616): Investigation, Writing – original draft.

Katie Scandrett (https://orcid.org/0000-0001-6111-2805): Investigation, Writing – original draft.

Shivashri Chockalingam (https://orcid.org/0000-0003-0294-6281): Investigation, Writing – original draft.

Aziza Osman (https://orcid.org/0009-0004-2718-0205): Investigation.

Naila Dracup (https://orcid.org/0000-0001-9372-4654): Investigation, Writing – original draft.

Rachel Court (https://orcid.org/0000-0002-4567-2586): Investigation, Supervision, Writing – original draft.

Furqan Butt (https://orcid.org/0009-0002-8393-3502): Investigation.

Cristina Visitin (https://orcid.org/0009-0006-4250-2924): Conceptualisation, Methodology, Writing – reviewing and editing.

James R Bonham (https://orcid.org/0009-0004-9485-6754): Conceptualisation, Methodology, Writing – reviewing and editing.

David Elliman (https://orcid.org/0000-0002-8128-2825): Conceptualisation, Methodology, Writing – reviewing and editing.

Graham Shortland (https://orcid.org/0000-0002-8994-8468): Conceptualisation, Methodology, Writing – reviewing and editing.

Anne Mackie (https://orcid.org/0009-0002-9975-594X): Conceptualisation, Methodology, Writing – reviewing and editing.

Zosia Miedzybrodzka (https://orcid.org/0000-0003-2890-8136): Writing – reviewing and editing.

Sian M Morgan (https://orcid.org/0000-0002-4943-653X): Writing – reviewing and editing.

Felicity Boardman (https://orcid.org/0000-0002-3268-6276): Methodology, Writing – reviewing and editing.

Yemisi Takwoingi (https://orcid.org/0000-0002-5828-9746): Conceptualisation, Funding acquisition, Methodology, Writing – reviewing and editing.

Sian Taylor-Phillips (https://orcid.org/0000-0002-1841-4346): Conceptualisation, Funding acquisition, Methodology, Supervision, Writing – original draft.

Acknowledgements

We are grateful for the input of the seven members of the PPIE group, who included Kerri Pearce, Matt Howard-Murray, Frances Othen-Wales, Luis Canto E Castro and Helen Bates. The authors would also like to thank Dr Helen Jenkinson, Dr Austen Worth, Dr Emma Footit and Professor Raja Padidela, for their specialist clinical advice on the rare conditions, and Professor Aileen Clarke for her critical review of the manuscript.

Data-sharing statement

No new data have been created in the preparation of this article and therefore there is nothing available for access and further sharing. All queries should be submitted to the corresponding author.

Ethics statement

This report concerns secondary research, for which ethics approval is not required.

Information governance statement

This project did not involve the handling of personal information.

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at https://doi.org/10.3310/DJRF1124.

Primary conflicts of interest: Karoline Freeman, Jacqueline Dinnes, Bethany Shinkins, Corinna Clark, Inès Kander, Katie Scandrett, Shivashri Chockalingam, Aziza Osman, Naila Dracup, Rachel Court, Furqan Butt, Felicity Boardman, Yemisi Takwoingi and Sian Taylor-Phillips received funding from the National Institute for Health and Care Research Evidence Synthesis Programme (ESG_HTA_NIHR159928) to undertake this work. Karoline Freeman received funding from the NIHR (NIHR130107, NIHR202220, NIHR30237). Jacqueline Dinnes and Yemisi Takwoingi were supported by the Birmingham Biomedical Research Centre. Bethany Shinkins is a member of the UK NSC. Corinna Clark and Felicity Boardman received funding from Novartis Gene Therapies to explore the public acceptability of newborn screening for SMA in the UK and from the Public Health England for the evaluation of Next Generation Sequencing for CF Newborn Screening. Cristina Visitin is a research lead at the UK NSC. James R Bonham is President of the International Society of Neonatal Screening and member of the UK NSC Fetal Maternal Child Health group. David Elliman is clinical advisor to the UK NSC and Chair of the UK NSC Blood Spot Task Group. Graham Shortland is Vice Chair of the UK NSC. Anne Mackie is the Director of Programmes for the UK NSC and member of the ESP Evidence Synthesis Groups Funding Committee 2022-present. Zosia Miedzybrodzka is co-chair of the Scottish Clinical Genomics Forum and a committee member of the Clinical Genetics Society, the Scottish Strategic Network for genomic medicine, the NHSE rare disease test evaluation working group and board member of the Royal Colleges Genomics Advisory Board and member of the Diagnostic Technologies and Screening Panel, HTA National Stakeholders Advisory Group 2015–23. Sian M Morgan has nothing to declare. Felicity Boardman received funding from GEL for a process and impact evaluation, MRC/NIHR

for a realist evaluation of remote qualitative data collection in health, Wellcome Trust for the development of future research agendas on the socioethical impact of genomics through interdisciplinary live data-sharing and Diabetes UK for the ELSA study. Felicity Boardman is a member of the UK NSC Foetal, Maternal and Child Health Reference Group, the Bloodspot Task group and the Spinal Muscular Atrophy In Service Evaluation Partnership Board. Yemisi Takwoingi is a member of the ESP Evidence Synthesis Groups Funding Committee 2024. Sian Taylor-Phillips is a member of the UK NSC AI task group, chair of the UK NSC Research Methodology Group and a funded co-applicant of the NIHR WM ESG.

Disclaimer

Publication

Freeman K, Taylor D, Dinnes J, Clark CCA, Kander I, Scandrett K, *et al*. Challenges in evaluating whole genome sequencing for newborn screening: series of systematic reviews and roadmap for evidence generation for policy advisers. *BMJ Medicine* 2025;**4**:e001726. https://doi.org/10.1136/bmjmed-2025-001726

References

1. Department of Health and Social Care. *UK Rare Diseases Framework*. London: Department of Health and Social Care; 2021. URL: www.gov.uk/government/publications/uk-rare-diseases-framework (accessed 16 November 2023).
2. Wright CF, FitzPatrick DR, Firth HV. Paediatric genomics: diagnosing rare disease in children. *Nat Rev Genet* 2018;**19**:25368. https://doi.org/10.1038/nrg.2017.116
3. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Disease gene identification strategies for exome sequencing. *Eur J Hum Genet* 2012;**20**:490–7. https://doi.org/10.1038/ejhg.2011.258
4. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, *et al*. A global reference for human genetic variation. *Nature* 2015;**526**:68–74. https://doi.org/10.1038/nature15393
5. Genomics Education Programme. *Newborn Screening: Time for a Genomic Approach?* London: NHS England; 2021. URL: www.genomicseducation.hee.nhs.uk/blog/newborn-screening-time-for-a-genomic-approach/ (accessed 14 August 2022).
6. Genomics England. *Newborn Genomes Programme*. London: NHS; 2021. URL: https://files.genomicsengland.co.uk/documents/Newborns-Vision-Final_SEP_2021-11-02-122418_jjne.pdf (accessed 16 March 2024).
7. UK National Screening Committee. *UK NSC: Evidence Review Process*. London: UK National Screening Committee; 2024. URL: www.gov.uk/government/publications/uk-nsc-evidence-review-process/uk-nsc-evidence-review-process (accessed 19 May 2024).
8. Beaumont RN, Wright CF. Estimating diagnostic noise in panel-based genomic analysis. *Genet Med* 2022;**24**:2042–50. https://doi.org/10.1016/j.gim.2022.06.008
9. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, *et al*.; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;**17**:405–24. https://doi.org/10.1038/gim.2015.30
10. Forrest IS, Chaudhary K, Vy HMT, Petrazzini BO, Bafna S, Jordan DM, *et al*. Population-based penetrance of deleterious clinical variants. *JAMA* 2022;**327**:350–9. https://doi.org/10.1001/jama.2021.23686
11. Wright CF, West B, Tuke M, Jones SE, Patel K, Laver TW, *et al*. Assessing the pathogenicity, penetrance, and expressivity of putative disease-causing variants in a population setting. *Am J Hum Genet* 2019;**104**:275–86. https://doi.org/10.1016/j.ajhg.2018.12.015
12. Woerner AC, Gallagher RC, Vockley J, Adhikari AN. The use of whole genome and exome sequencing for newborn screening: challenges and opportunities for population health. *Front Pediatr* 2021;**9**:663752. https://doi.org/10.3389/fped.2021.663752
13. Nicholls SG. Proceduralisation, choice and parental reflections on decisions to accept newborn bloodspot screening. *J Med Ethics* 2012;**38**:299–303. https://doi.org/10.1136/medethics-2011-100040
14. Bendor-Samuel OM, Wishlade T, Willis L, Aley P, Choi E, Craik R, *et al*.; GPPAD Study Group. Successful integration of newborn genetic testing into UK routine screening using prospective consent to determine eligibility for clinical trials. *Arch Dis Child* 2023;**108**:26–30. https://doi.org/10.1136/archdischild-2022-324270
15. Owens K, Sankar P, Asfaha DM. How clinicians conceptualize 'actionability' in genomic screening. *J Pers Med* 2023;**13**:290. https://doi.org/10.3390/jpm13020290
16. UK National Screening Committee. *Criteria for a Population Screening Programme*. London: UK National Screening Committee; 2022. URL: www.gov.uk/government/publications/evidence-review-criteria-national-screening-programmes/

criteria-for-appraising-the-viability-effectiveness-and-appropriateness-of-a-screening-programme (accessed 7 August 2024).

17. Genomics England. *How We Choose Conditions*. London: Genomics England Limited; 2024. URL: www.genomicsengland.co.uk/initiatives/newborns/choosing-conditions (accessed 8 June 2023).
18. Leeds Institute of Health Sciences. *AUHE Information Specialists, University of Leeds Checking for Duplicates Guidance*. Leeds: University of Leeds; 2019. URL: https://information-specialists.leeds.ac.uk/wp-content/uploads/sites/71/2019/03/Duplicate_checking_guidance.pdf (accessed 15 May 2023).
19. Whiting P, Savović J, Higgins JPT, Caldwell DM, Reeves BC, Shea B, *et al*.; ROBIS Group. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;**69**:225–34. https://doi.org/10.1016/j.jclinepi.2015.06.005
20. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;**158**:280–6. https://doi.org/10.7326/0003-4819-158-4-201302190-00009
21. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetc R, *et al*. Chapter 7: systematic reviews of etiology and risk. In Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, *editors. JBI Manual for Evidence Synthesis*. Adelaide, SA: JBI; 2020. pp. 219–69. https://doi.org/10.46658/jbimes-20-08
22. Murad MH, Sultan S, Haffar S, Bazerbachi F. Methodological quality and synthesis of case series and case reports. *BMJ Evid Based Med* 2018;**23**:60–3. https://doi.org/10.1136/bmjebm-2017-110853
23. Hunter JE, Irving SA, Biesecker LG, Buchanan A, Jensen B, Lee K, *et al*. A standardized, evidence-based protocol to assess clinical actionability of genetic disorders associated with genomic variation. *Genet Med* 2016;**18**:1258–68. https://doi.org/10.1038/gim.2016.40
24. ClinGen. *ClinGen Pediatric Actionability Workgroup Protocol: Generation of Summary Reports and Semi-Quantitative Metric Scoring*. Bethesda, MD: National Institutes of Health (NIH); 2020. URL: www.clinicalgenome.org/site/assets/files/5075/peds_combined_evidence_curation_and_scoring_protocol_07292020.pdf (accessed 7 March 2024).
25. Goddard KAB, Whitlock EP, Berg JS, Williams MS, Webber EM, Webster JA, *et al*. Description and pilot results from a novel method for evaluating return of incidental findings from next-generation sequencing technologies. *Genet Med* 2013;**15**:721–8. https://doi.org/10.1038/gim.2013.37
26. Berg JS, Foreman AKM, O'Daniel JM, Booker JK, Boshe L, Carey T, *et al*. A semiquantitative metric for evaluating clinical actionability of incidental or secondary findings from genome-scale sequencing. *Genet Med* 2016;**18**:467–75. https://doi.org/10.1038/gim.2015.104
27. Schwarze K, Buchanan J, Taylor JC, Wordsworth S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet Med* 2018;**20**:1122–30. https://doi.org/10.1038/gim.2017.247
28. Nurchis MC, Riccardi MT, Radio FC, Chillemi G, Bertini ES, Tartaglia M, *et al*. Incremental net benefit of whole genome sequencing for newborns and children with suspected genetic disorders: systematic review and meta-analysis of cost-effectiveness evidence. *Health Policy* 2022;**126**:337–45. https://doi.org/10.1016/j.healthpol.2022.03.001
29. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan – a web and mobile app for systematic reviews. *Syst Rev* 2016;**5**:210. https://doi.org/10.1186/s13643-016-0384-4
30. Boonsimma P, Ittiwut C, Kamolvisit W, Ittiwut R, Chetruengchai W, Phokaew C, *et al*. Exome sequencing as first-tier genetic testing in infantile-onset pharmacoresistant epilepsy: diagnostic yield and treatment impact. *Eur J Hum Genet* 2023;**31**:179–87. https://doi.org/10.1038/s41431-022-01202-x
31. Koul R, Alfutaisi A, Abdelrahim R, Altihilli K. Pyridoxine responsive seizures: beyond aldehyde dehydrogenase 7A1. *J Neurosci Rural Pract* 2019;**10**:613–6. https://doi.org/10.1055/s-0039-1697775

32. Coughlin CR II, Swanson MA, Spector E, Meeks NJL, Kronquist KE, Aslamy M, *et al*. The genotypic spectrum of ALDH7A1 mutations resulting in pyridoxine dependent epilepsy: a common epileptic encephalopathy. *J Inherit Metab Dis* 2019;**42**:353–61. https://doi.org/10.1002/jimd.12045

33. Jiao X, Xue J, Gong P, Wu Y, Zhang Y, Jiang Y, Yang Z. Clinical and genetic features in pyridoxine-dependent epilepsy: a Chinese cohort study. *Dev Med Child Neurol* 2020;**62**:315–21. https://doi.org/10.1111/dmcn.14385

34. Jiao X, Gong P, Niu Y, Zhang Y, Yang Z. A rare presentation characterized by epileptic spasms in ALDH7A1, pyridox(am)ine-5'-phosphate oxidase, and PLPBP deficiency. *Front Genet* 2022;**13**:804461. https://doi.org/10.3389/fgene.2022.804461

35. Mills PB, Footitt EJ, Mills KA, Tuschl K, Aylett S, Varadkar S, *et al*. Genotypic and phenotypic spectrum of pyridoxine-dependent epilepsy (ALDH7A1 deficiency). *Brain* 2010;**133**:2148–59. https://doi.org/10.1093/brain/awq143

36. Plecko B, Mills P. *PNPO Deficiency*. Seattle, WA: GeneReviews®, University of Washington; 2022. URL: www.ncbi.nlm.nih.gov/books/NBK581452/ (accessed 18 March 2024).

37. Amsterdam UMC. *PDE H2M*. 2023. URL: https://pdeonline.org/index.html (accessed 6 October 2023).

38. Basura GJ, Hagland SP, Wiltse AM, Gospe SM Jr. Clinical features and the management of pyridoxine-dependent and pyridoxine-responsive seizures: review of 63 North American cases submitted to a patient registry. *Eur J Pediatr* 2009;**168**:697–704. https://doi.org/10.1007/s00431-008-0823-x

39. Salviat F, Gauthier-Villars M, Carton M, Cassoux N, Lumbroso-Le Rouic L, Dehainault C, *et al*. Association between genotype and phenotype in consecutive unrelated individuals with retinoblastoma. *JAMA Ophthalmol* 2020;**138**:843–50. https://doi.org/10.1001/jamaophthalmol.2020.2100

40. Temming P, Viehmann A, Biewald E, Lohmann DR. Sporadic unilateral retinoblastoma or first sign of bilateral disease? *Br J Ophthalmol* 2013;**97**:475–80. https://doi.org/10.1136/bjophthalmol-2012-302666

41. Hulsenbeck I, Frank M, Biewald E, Kanber D, Lohmann DR, Ketteler P. Introduction of a variant classification system for analysis of genotype-phenotype relationships in heritable retinoblastoma. *Cancers* 2021;**13**:1605. https://doi.org/10.3390/cancers13071605

42. NHS England. *Whole Genome Sequencing*. NHS England; 2022. URL: www.genomicseducation.hee.nhs.uk/genotes/knowledge-hub/whole-genome-sequencing/ (accessed 18 August 2023).

43. Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat Rev Genet* 2018;**19**:269–85. https://doi.org/10.1038/nrg.2017.117

44. NHS England. *Mosaicism*. London: NHS England; 2022. URL: www.genomicseducation.hee.nhs.uk/genotes/knowledge-hub/mosaicism/ (accessed 15 April 2024).

45. Dehainault C, Garancher A, Castéra L, Cassoux N, Aerts I, Doz F, *et al*. The survival gene MED4 explains low penetrance retinoblastoma in patients with large RB1 deletion. *Hum Mol Genet* 2014;**23**:5243–50. https://doi.org/10.1093/hmg/ddu245

46. Jacob P, Bhavani GS, Udupa P, Wang Z, Hariharan SV, Delampady K, *et al*. Exome sequencing in monogenic forms of rickets. *Indian J Pediatr* 2023;**90**:1182–90. https://doi.org/10.1007/s12098-022-04393-9

47. Marik B, Bagga A, Sinha A, Khandelwal P, Hari P, Sharma A. Genetic and clinical profile of patients with hypophosphatemic rickets. *Eur J Med Genet* 2022;**65**:104540. https://doi.org/10.1016/j.ejmg.2022.104540

48. Gaucher C, Walrant-Debray O, Nguyen TM, Esterle L, Garabedian M, Jehan F. PHEX analysis in 118 pedigrees reveals new genetic clues in hypophosphatemic rickets. *Hum Genet* 2009;**125**:401–11. https://doi.org/10.1007/s00439-009-0631-z

49. Del Pino M, Viterbo GL, Arenas MA, Perez Garrido N, Ramirez P, Marino R, *et al*. Growth in height and body proportion from birth to adulthood in hereditary hypophosphatemic rickets: a retrospective cohort study. *J Endocrinol Invest* 2022;**45**:1349–58. https://doi.org/10.1007/s40618-022-01768-9

50. Rafaelsen S, Johansson S, Raeder H, Bjerknes R. Hereditary hypophosphatemia in Norway: a retrospective population-based study of genotypes, phenotypes, and treatment complications. *Eur J Endocrinol* 2016;**174**:125–36. https://doi.org/10.1530/EJE-15-0515

51. Ariceta G, Beck-Nielsen SS, Boot AM, Brandi ML, Briot K, de Lucas Collantes C, *et al*. The International X-Linked Hypophosphatemia (XLH) Registry: first interim analysis of baseline demographic, genetic and clinical data. *Orphanet J Rare Dis* 2023;**18**:304. https://doi.org/10.1186/s13023-023-02882-4

52. Acar S, Demir K, Shi Y. Genetic causes of rickets. *J Clin Res Pediatr Endocrinol* 2017;**9**:88–105. https://doi.org/10.4274/jcrpe.2017.S008

53. Cetica V, Sieni E, Pende D, Danesino C, De Fusco C, Locatelli F, *et al*. Genetic predisposition to hemophagocytic lymphohistiocytosis: report on 500 patients from the Italian registry. *J Allergy Clin Immunol* 2016;**137**:188–96. e4. https://doi.org/10.1016/j.jaci.2015.06.048

54. Shabrish S, Kelkar M, Yadav RM, Bargir UA, Gupta M, Dalvi A, *et al*. The spectrum of clinical, immunological, and molecular findings in familial hemophagocytic lymphohistiocytosis: experience from India. *Front Immunol* 2021;**12**:612583. https://doi.org/10.3389/fimmu.2021.612583

55. Amirifar P, Ranjouri MR, Abolhassani H, Moeini Shad T, Almasi-Hashiani A, Azizi G, *et al*. Clinical, immunological and genetic findings in patients with UNC13D deficiency (FHL3): a systematic review. *Pediatr Allergy Immunol* 2021;**32**:186–97. https://doi.org/10.1111/pai.13323

56. Pagel J, Beutel K, Lehmberg K, Koch F, Maul-Pavicic A, Rohlfs AK, *et al*. Distinct mutations in STXBP2 are associated with variable clinical presentations in patients with familial hemophagocytic lymphohistiocytosis type 5 (FHL5). *Blood* 2012;**119**:6016–24. https://doi.org/10.1182/blood-2011-12-398958

57. Trizzino A, zur Stadt U, Ueda I, Risma K, Janka G, Ishii E, *et al*.; Histiocyte Society HLH Study Group. Genotype-phenotype study of familial haemophagocytic lymphohistiocytosis due to perforin mutations. *J Med Genet* 2008;**45**:15–21. https://doi.org/10.1136/jmg.2007.052670

58. Henter JI, Horne A, Aricó M, Egeler RM, Filipovich AH, Imashuku S, *et al*. HLH-2004: diagnostic and therapeutic guidelines for hemophagocytic lymphohistiocytosis. *Pediatr Blood Cancer* 2007;**48**:124–31. https://doi.org/10.1002/pbc.21039

59. Martin-Rivada A, Palomino Perez L, Ruiz-Sala P, Navarrete R, Cambra Conejero A, Quijada Fraile P, *et al*. Diagnosis of inborn errors of metabolism within the expanded newborn screening in the Madrid region. *JIMD Rep* 2022;**63**:146–61. https://doi.org/10.1002/jmd2.12265

60. Maguolo A, Rodella G, Dianin A, Nurti R, Monge I, Rigotti E, *et al*. Diagnosis, genetic characterization and clinical follow up of mitochondrial fatty acid oxidation disorders in the new era of expanded newborn screening: a single centre experience. *Mol Genet Metab Rep* 2020;**24**:100632. https://doi.org/10.1016/j.ymgmr.2020.100632

61. Wang B, Zhang Q, Gao A, Wang Q, Ma J, Li H, Wang T. New ratios for performance improvement for identifying acyl-CoA dehydrogenase deficiencies in expanded newborn screening: a retrospective study. *Front Genet* 2019;**10**:811. https://doi.org/10.3389/fgene.2019.00811

62. Nichols MJ, Saavedra-Matiz CA, Pass KA, Caggana M. Novel mutations causing medium chain acyl-CoA dehydrogenase deficiency: under-representation of the common c.985 A > G mutation in the New York state population. *Am J Med Genet A* 2008;**146A**:610–9. https://doi.org/10.1002/ajmg.a.32192

63. Mesbah Z, Sing Ho K, Fitzsimons P, Monavari AA, Mayne PD, Crushell E. Medium chain acyl-CoA dehydrogenase deficiency (MCADD) in the Irish paediatric population. *Ir Med J* 2019;**112**:1016.

64. Touw CML, Smit GPA, de Vries M, de Klerk JBC, Bosch AM, Visser G, *et al*. Risk stratification by residual enzyme activity after newborn screening for medium-chain acyl-CoA dehyrogenase deficiency: data from a cohort study. *Orphanet J Rare Dis* 2012;**7**:30. https://doi.org/10.1186/1750-1172-7-30

65. Balakrishnan U. Inborn errors of metabolism-approach to diagnosis and management in neonates. *Indian J Pediatr* 2021;**88**:679–89. https://doi.org/10.1007/s12098-021-03759-9

66. Ikegawa K, Hasegawa Y. Presentation and diagnosis of pediatric X-linked hypophosphatemia. *Endocrines* 2023;**4**:128–37. URL: www.mdpi.com/2673-396X/4/1/12 (accessed 12 June 2025).

67. Grosse SD, Khoury MJ, Greene CL, Crider KS, Pollitt RJ. The epidemiology of medium chain acyl-CoA dehydrogenase deficiency: an update. *Genet Med* 2006;**8**:205–12. https://doi.org/10.1097/01.gim.0000204472.25153.8d

68. Coughlin CR, Tseng LA, Bok LA, Hartmann H, Footitt E, Striano P, *et al.*; International PDE Consortium. Association between lysine reduction therapies and cognitive outcomes in patients with pyridoxine-dependent epilepsy. *Neurology* 2022;**99**:e2627–36. https://doi.org/10.1212/WNL.0000000000201222

69. Drugs.com. *Pyridoxine*. Drugs.com; 2024. URL: www.drugs.com/mtm/pyridoxine.html (accessed 15 March 2024).

70. van Karnebeek CDM, Stockler-Ipsiroglu S, Jaggumantri S, Assmann B, Baxter P, Buhas D, *et al*. Lysine-restricted diet as adjunct therapy for pyridoxine-dependent epilepsy: The PDE Consortium Consensus Recommendations. In Zschocke J, Gibson KM, Brown G, Morava E, Peters V, editors. *JIMD Reports*, vol. 15. Berlin; Heidelberg: Springer; 2015. pp. 47–57. https://doi.org/10.1007/8904_2014_296

71. Berry JL, Xu L, Polski A, Jubran R, Kuhn P, Kim JW, Hicks J. Aqueous humor is superior to blood as a liquid biopsy for retinoblastoma. *Ophthalmology* 2020;**127**:552–4. https://.doi.org/10.1016/j.ophtha.2019.10.026

72. Hamel P, Heon E, Gallie BL, Budning AS. Focal therapy in the management of retinoblastoma: when to start and when to stop. *J Am Assoc Pediatr Ophthalmol Strab* 2000;**4**:334–7. https://doi.org/10.1067/mpa.2000.107902

73. Moncrieff MW. Early biochemical findings in familial hypophosphataemic, hyperphosphaturic rickets and response to treatment. *Arch Dis Child* 1982;**57**:70–2.

74. Poon KS, Sng AA, Ho CW, Koay ESC, Loke KY. Genetic testing confirmed the early diagnosis of X-linked hypophosphatemic rickets in a 7-month-old infant. *J Invest Med High Impact Case Rep* 2015;**3**:598167. https://doi.org/10.1177/2324709615598167

75. Lucchini G, Marsh R, Gilmour K, Worth A, Nademi Z, Rao A, *et al*. Treatment dilemmas in asymptomatic children with primary hemophagocytic lymphohistiocytosis. *Blood* 2018;**132**:2088–96. https://doi.org/10.1182/blood-2018-01-827485

76. Bok LA, Been JV, Struys EA, Jakobs C, Rijper EAM, Willemsen MA. Antenatal treatment in two Dutch families with pyridoxine-dependent seizures. *Eur J Pediatr* 2010;**169**:297–303. https://doi.org/10.1007/s00431-009-1020-2

77. Jiao X, Gong P, Wu Y, Zhang Y, Yang Z. Analysis of the phenotypic variability as well as impact of early diagnosis and treatment in six affected families with ALDH7A1 deficiency. *Front Genet* 2021;**12**:644447. https://doi.org/10.3389/fgene.2021.644447

78. Strijker M, Tseng LA, van Avezaath LK, Oude Luttikhuis MAM, Ketelaar T, Coughlin CR II, *et al*. Cognitive and neurological outcome of patients in the Dutch pyridoxine-dependent epilepsy (PDE-ALDH7A1) cohort, a cross-sectional study. *Eur J Paediatr Neurol* 2021;**33**:112–20. https://doi.org/10.1016/j.ejpn.2021.06.001

79. Tseng LA, Abdenur JE, Andrews A, Aziz VG, Bok LA, Boyer M, *et al*. Timing of therapy and neurodevelopmental outcomes in 18 families with pyridoxine-dependent epilepsy. *Mol Genet Metab* 2022;**135**:350–6. https://doi.org/10.1016/j.ymgme.2022.02.005

80. Abramson DH, Beaverson K, Sangani P, Vora RA, Lee TC, Hochberg HM, *et al*. Screening for retinoblastoma: presenting signs as prognosticators of patient and ocular survival. *Pediatrics* 2003;**112**:1248–55. https://doi.org/10.1542/peds.112.6.1248

81. Chantada GL, Dunkel IJ, Qaddoumi I, Antoneli CBG, Totah A, Canturk S, *et al*. Familial retinoblastoma in developing countries. *Pediatr Blood Cancer* 2009;**53**:338–42. https://doi.org/10.1002/pbc.21970

82. Moll AC, Imhof SM, Meeteren AY, Boers M. At what age could screening for familial retinoblastoma be stopped? A register based study 1945–98. *Br J Ophthalmol* 2000;**84**:1170–2. https://doi.org/10.1136/bjo.84.10.1170

83. Rothschild PR, Levy D, Savignoni A, Lumbroso-Le Rouic L, Aerts I, Gauthier-Villars M, *et al*. Familial retinoblastoma: fundus screening schedule impact and guideline proposal: a retrospective study. *Eye* 2011;**25**:1555–61. https://doi.org/10.1038/eye.2011.198

84. Soliman SE, ElManhaly M, Dimaras H. Knowledge of genetics in familial retinoblastoma. *Ophthalmic Genet* 2017;**38**:226–32. https://doi.org/10.1080/13816810.2016.1195846

85. Imhof SM, Moll AC, Schouten-van Meeteren AYN. Stage of presentation and visual outcome of patients screened for familial retinoblastoma: nationwide registration in the Netherlands. *Br J Ophthalmol* 2006;**90**:875–8. https://doi.org/10.1136/bjo.2005.089375

86. Soliman SE, Dimaras H, Khetan V, Gardiner JA, Chan HSL, Heon E, Gallie BL. Prenatal versus postnatal screening for familial retinoblastoma. *Ophthalmology* 2016;**123**:2610–7. https://doi.org/10.1016/j.ophtha.2016.08.027

87. Kaliki S, Maniar A, Patel A, Palkonda VAR, Mohamed A. Clinical presentation and outcome of retinoblastoma based on age at presentation: a review of 1450 children. *Int Ophthalmol* 2020;**40**:99–107. https://doi.org/10.1007/s10792-019-01155-z

88. Makitie O, Doria A, Kooh SW, Cole WG, Daneman A, Sochett E. Early treatment improves growth and biochemical and radiographic outcome in X-linked hypophosphatemic rickets. *J Clin Endocrinol Metab* 2003;**88**:3591–7. https://doi.org/10.1210/jc.2003-030036

89. Quinlan C, Guegan K, Offiah A, Neill RO, Hiorns MP, Ellard S, *et al*. Growth in PHEX-associated X-linked hypophosphatemic rickets: the importance of early treatment. *Pediatr Nephrol* 2012;**27**:581–8. https://doi.org/10.1007/s00467-011-2046-z

90. Kruse K, Hinkel GK, Griefahn B. Calcium metabolism and growth during early treatment of children with X-linked hypophosphataemic rickets. *Eur J Pediatr* 1998;**157**:894–900. https://doi.org/10.1007/s004310050962

91. Roza M, Miguel MA, Galbe M, Mejido L, Mencia C. Early treatment of familial hypophosphataemic rickets. *Arch Dis Child* 1983;**58**:1020–2. https://doi.org/10.1136/adc.58.12.1020

92. Abdenur JE, Chamoles NA, Specola N, Schenone AB, Jorge L, Guinle A, *et al*. MCAD deficiency: acylcarnitines (AC) by tandem mass spectrometry (MS-MS) are useful to monitor dietary treatment. *Adv Exp Med Biol* 1999;**466**:353–63.

93. Alcaide P, Ferrer-Lopez I, Gutierrez L, Leal F, Martin-Hernandez E, Quijada-Fraile P, *et al*. Lymphocyte medium-chain acyl-CoA dehydrogenase activity and its potential as a diagnostic confirmation tool in newborn screening cases. *J Clin Med* 2022;**11**:2933. https://doi.org/10.3390/jcm11102933

94. Anderson DR, Viau K, Botto LD, Pasquali M, Longo N. Clinical and biochemical outcomes of patients with medium-chain acyl-CoA dehydrogenase deficiency. *Mol Genet Metab* 2020;**129**:13–9. https://doi.org/10.1016/j.ymgme.2019.11.006

95. Gong Z, Liang L, Qiu W, Zhang H, Ye J, Wang Y, *et al*. Clinical, biochemical, and molecular analyses of medium-chain acyl-CoA dehydrogenase deficiency in Chinese patients. *Front Genet* 2021;**12**:577046. https://doi.org/10.3389/fgene.2021.577046

96. Haas M, Chaplin M, Joy P, Wiley V, Black C, Wilcken B. Healthcare use and costs of medium-chain acyl-CoA dehydrogenase deficiency in Australia: screening versus no screening. *J Pediatr* 2007;**151**:121–6, 126.e1. https://doi.org/10.1016/j.jpeds.2007.03.011

97. Li H, Benson LA, Henderson LA, Solomon IH, Kennedy AL, Soldatos A, *et al*. Central nervous system-restricted familial hemophagocytic lymphohistiocytosis responds to hematopoietic cell transplantation. *Blood Adv* 2019;**3**:503–7. https://doi.org/10.1182/bloodadvances.2018027417

98. Wilcken B, Haas M, Joy P, Wiley V, Chaplin M, Black C, *et al*. Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency in Australia: a cohort study. *Lancet* 2007;**369**:37–42. https://doi.org/10.1016/S0140-6736(07)60029-4

99. Wilcken B, Haas M, Joy P, Wiley V, Bowling F, Carpenter K, *et al*. Expanded newborn screening: outcome in screened and unscreened patients at age 6 years. *Pediatrics* 2009;**124**:e241–8. https://doi.org/10.1542/peds.2008-0586

100. Wilson CJ, Champion MP, Collins JE, Clayton PT, Leonard JV. Outcome of medium chain acyl-CoA dehydrogenase deficiency after diagnosis. *Arch Dis Child* 1999;**80**:459–62. https://doi.org/10.1136/adc.80.5.459

101. Busiello R, Adriani M, Locatelli F, Galgani M, Fimiani G, Clementi R, *et al*. Atypical features of familial hemophagocytic lymphohistiocytosis. *Blood* 2004;**103**:4610–2. https://doi.org/10.1182/blood-2003-10-3551

102. Li Y, Zhu R, Liu Y, Song J, Xu J, Yang Y. Medium-chain acyl-coenzyme A dehydrogenase deficiency: six cases in the Chinese population. *Pediatr Int* 2019;**61**:551–7. https://doi.org/10.1111/ped.13872

103. Howard C, Gorman I, Crushell E, Knerr I, Hughes J, Boruah R, *et al*. Medium chain acyl-CoA dehydrogenase deficiency: 3 years of newborn screening. *Ir Med J* 2023;**116**:743.

104. Li L, Li H, Zhang J, Gan H, Liu R, Hu X, *et al*. Five novel RB1 gene mutations and genotype-phenotype correlations in Chinese children with retinoblastoma. *Int Ophthalmol* 2022;**42**:3421–30. https://doi.org/10.1007/s10792-022-02341-2

105. Chong SC, Law LK, Hui J, Lai CY, Leung TY, Yuen YP. Expanded newborn metabolic screening programme in Hong Kong: a three-year journey. *Hong Kong Med J* 2017;**23**:489–96. https://doi.org/10.12809/hkmj176274

106. Gramer G, Haege G, Fang-Hoffmann J, Hoffmann GF, Bartram CR, Hinderhofer K, *et al*. Medium-chain acyl-CoA dehydrogenase deficiency: evaluation of genotype-phenotype correlation in patients detected by newborn screening. *JIMD Rep* 2015;**23**:101–12. https://doi.org/10.1007/8904_2015_439

107. Yusupov R, Finegold DN, Naylor EW, Sahai I, Waisbren S, Levy HL. Sudden death in medium chain acyl-coenzyme a dehydrogenase deficiency (MCADD) despite newborn screening. *Mol Genet Metab* 2010;**101**:33–9. https://doi.org/10.1016/j.ymgme.2010.05.007

108. Joy P, Black C, Rocca A, Haas M, Wilcken B. Neuropsychological functioning in children with medium chain acyl coenzyme a dehydrogenase deficiency (MCADD): the impact of early diagnosis and screening on outcome. *Child Neuropsychol* 2009;**15**:8–20. https://doi.org/10.1080/09297040701864570

109. Hsu HW, Zytkovicz TH, Comeau AM, Strauss AW, Marsden D, Shih VE, *et al*. Spectrum of medium-chain acyl-CoA dehydrogenase deficiency detected by newborn screening. *Pediatrics* 2008;**121**:e1108–14. https://doi.org/10.1542/peds.2007-1993

110. Gadoury-Levesque V, Dong L, Su R, Chen J, Zhang K, Risma KA, *et al*. Frequency and spectrum of disease-causing variants in 1892 patients with suspected genetic HLH disorders. *Blood Adv* 2020;**4**:2578–94. https://doi.org/10.1182/bloodadvances.2020001605

111. Ahmari AA, Alsmadi O, Sheereen A, Elamin T, Jabr A, El-Baik L, *et al*. Genetic and clinical characteristics of pediatric patients with familial hemophagocytic lymphohistiocytosis. *Blood Res* 2021;**56**:86–101. https://doi.org/10.5045/br.2021.2020308

112. Ruppe MD, Brosnan PG, Au KS, Tran PX, Dominguez BW, Northrup H. Mutational analysis of PHEX, FGF23 and DMP1 in a cohort of patients with hypophosphatemic rickets. *Clin Endocrinol (Oxf)* 2011;**74**:312–8. https://doi.org/10.1111/j.1365-2265.2010.03919.x

113. Adam MP, Feldman J, Mirzaa GM, Pagon RA, Wallace SE, Amemiya A. *GeneReviews®* [Internet]. Seattle, WA: University of Washington; 2024. URL: www.ncbi.nlm.nih.gov/books/NBK1116/ (accessed 8 May 2023).

114. NICE. *Burosumab for Treating X-Linked Hypophosphataemia in Children and Young People*. London: National Institute for Health and Care Excellence; 2018. URL: www.nice.org.uk/guidance/hst8 (accessed 18 March 2024).

115. Coughlin CR II, Tseng LA, Abdenur JE, Ashmore C, Boemer F, Bok LA, *et al*. Consensus guidelines for the diagnosis and management of pyridoxine-dependent epilepsy due to α-aminoadipic semialdehyde dehydrogenase deficiency. *J Inherit Metab Dis* 2021;**44**:178–92. https://doi.org/10.1002/jimd.12332

116. NHS. *Clinical Commissioning Policy: Anakinra for Haemophagocytic Lymphohistiocytosis (HLH) for Adults and Children in All Ages [210701P] (1924)*. London: National Health Service; 2021. URL: www.england.nhs.uk/wp-content/uploads/2021/10/1924-Clinical-commissioning-policy-anakinra-for-haemophagocytic-lymphohistiocytosis-.pdf (accessed 17 March 2024).

117. BIMDG. *British Inherited Metabolic Disease Group*. Cambridge: British Inherited Metabolic Disease Group; 2024. URL: www.bimdg.org.uk/site/index.asp (accessed 18 March 2024).

118. Rothenbuhler A, Schnabel D, Högler W, Linglart A. Diagnosis, treatment-monitoring and follow-up of children and adolescents with X-linked hypophosphatemia (XLH). *Metabolism* 2020;**103S**:153892. https://doi.org/10.1016/j.metabol.2019.03.009

119. Andresen BS, Lund AM, Hougaard DM, Christensen E, Gahrn B, Christensen M, *et al*. MCAD deficiency in Denmark. *Mol Genet Metab* 2012;**106**:175–88. https://doi.org/10.1016/j.ymgme.2012.03.018

120. Al-Nawaiseh I, Ghanem AQ, Yousef YA. Familial retinoblastoma: raised awareness improves early diagnosis and outcome. *J Ophthalmol* 2017;**2017**:5053961. https://doi.org/10.1155/2017/5053961

121. Skalet AH, Gombos DS, Gallie BL, Kim JW, Shields CL, Marr BP, *et al*. Screening children at risk for retinoblastoma. *Ophthalmology* 2018;**125**:453–8. https://doi.org/10.1016/j.ophtha.2017.09.001

122. Green RC, Shah N, Genetti CA, Yu T, Zettler B, Uveges MK, *et al*.; BabySeq Project Team. Actionability of unanticipated monogenic disease risks in newborn genomic screening: findings from the BabySeq project. *Am J Hum Genet* 2023;**110**:1034–45. https://doi.org/10.1016/j.ajhg.2023.05.007

123. Green R, Shah N, Genetti C, Yu T, Zettler B, Schwartz T, *et al*. Medical evaluation of unanticipated monogenic disease risks identified through newborn genomic screening: findings from the BabySeq project. 2022. https://doi.org/10.1101/2022.03.18.22272284

124. Ceyhan-Birsoy O, Murry JB, Machini K, Lebo MS, Yu TW, Fayer S, *et al*. Interpretation of genomic sequencing results in healthy and ill newborns: results from the BabySeq project. *Am J Hum Genet* 2019;**104**:76–93. https://doi.org/10.1016/j.ajhg.2018.11.016

125. Hao C, Guo R, Hu X, Qi Z, Guo Q, Liu X, *et al*. Newborn screening with targeted sequencing: a multicenter investigation and a pilot clinical study in China. *J Genet Genomics* 2022;**49**:13–9. https://doi.org/10.1016/j.jgg.2021.08.008

126. Yang RL, Qian GL, Wu DW, Miao JK, Yang X, Wu BQ, *et al*. A multicenter prospective study of next-generation sequencing-based newborn screening for monogenic genetic diseases in China. *World J Pediatr* 2023;**19**:663–73. https://doi.org/10.1007/s12519-022-00670-x

127. Bodian DL, Klein E, Iyer RK, Wong WSW, Kothiyal P, Stauffer D, *et al*. Utility of whole-genome sequencing for detection of newborn screening disorders in a population cohort of 1,696 neonates. *Genet Med* 2016;**18**:221–30. https://doi.org/10.1038/gim.2015.111

128. Chen T, Fan C, Huang Y, Feng J, Zhang Y, Miao J, *et al*. Genomic sequencing as a first-tier screening test and outcomes of newborn screening. *JAMA Netw Open* 2023;**6**:e2331162. https://doi.org/10.1001/jamanetworkopen.2023.31162

129. Huang X, Wu D, Zhu L, Wang W, Yang R, Yang J, *et al*. Application of a next-generation sequencing (NGS) panel in newborn screening efficiently identifies inborn disorders of neonates. *Orphanet J Rare Dis* 2022;**17**:66. https://doi.org/10.1186/s13023-022-02231-x

130. Jian M, Wang X, Sui Y, Fang M, Feng C, Huang Y, *et al*. A pilot study of assessing whole genome sequencing in newborn screening in unselected children in China. *Clin Transl Med* 2022;**12**:e843. https://doi.org/10.1002/ctm2.843

131. Pavey AR, Bodian DL, Vilboux T, Khromykh A, Hauser NS, Huddleston K, *et al*. Utilization of genomic sequencing for population screening of immunodeficiencies in the newborn. *Genet Med* 2017;**19**:1367–75. https://doi.org/10.1038/gim.2017.57

132. Balciuniene J, Liu R, Bean L, Guo F, Nallamilli BRR, Guruju N, *et al*. At-risk genomic findings for pediatric-onset disorders from genome sequencing vs medically actionable gene panel in proactive screening of newborns and children. *JAMA Netw Open* 2023;**6**:e2326445. https://doi.org/10.1001/jamanetworkopen.2023.26445

133. Roman TS, Crowley SB, Roche MI, Foreman AKM, O'Daniel JM, Seifert BA, *et al*. Genomic sequencing for newborn screening: results of the NC NEXUS project. *Am J Hum Genet* 2020;**107**:596–611. https://doi.org/10.1016/j.ajhg.2020.08.001

134. Cai L, Liu Y, Xu Y, Yang H, Lv L, Li Y, *et al*. Multi-center in-depth screening of neonatal deafness genes: Zhejiang, China. *Front Genet* 2021;**12**:637096. https://doi.org/10.3389/fgene.2021.637096

135. Ma Z, Huang W, Xu J, Qiu J, Liu Y, Ye M, Fan S. Analysis of deafness susceptibility gene of neonates in northern Guangdong, China. *Sci Rep* 2024;**14**:362. https://doi.org/10.1038/s41598-023-49530-2

136. Ye L, Yin Y, Chen M, Gong N, Peng Y, Liu H, Miao J. Combined genetic screening and traditional newborn screening to improve the screening efficiency of congenital hypothyroidism. *Front Pediatr* 2023;**11**:1185802. https://doi.org/10.3389/fped.2023.1185802

137. Luo H, Yang Y, Wang X, Xu F, Huang C, Liu D, *et al*. Concurrent newborn hearing and genetic screening of common hearing loss variants with bloodspot-based targeted next generation sequencing in Jiangxi province. *Front Pediatr* 2022;**10**:1020519. https://doi.org/10.3389/fped.2022.1020519

138. Gold NB, Harrison SM, Rowe JH, Gold J, Furutani E, Biffi A, *et al*. Low frequency of treatable pediatric disease alleles in gnomAD: an opportunity for future genomic screening of newborns. *HGG Adv* 2022;**3**:100059. https://doi.org/10.1016/j.xhgg.2021.100059

139. Breilyn MS, Kenny EE, Abul-Husn NS. Diverse and unselected adults with clinically relevant ACADS variants lack evidence of metabolic disease. *Mol Genet Metab* 2023;**138**:106971. https://doi.org/10.1016/j.ymgme.2022.106971

140. Van Nimwegen K, Vissers L, Willemsen M, Schieving J, Veltman J, Van Der Wilt G, Grutters JP. The cost-effectiveness of whole-exome sequencing in complex paediatric neurology. *Value Health* 2016;**19**:A695. https://doi.org/10.1016/j.jval.2016.09.1998

141. Buchanan-Hughes AM, Agirrezabal I, Eddowes LA, Luheshi LM, Torok ME, Sagoo GS. A system dynamics model for the cost-effectiveness evaluation of bacterial whole-genome sequencing for monitoring outbreaks of clostridium difficile. *Value Health* 2015;**18**:A44. https://doi.org/10.1016/j.jval.2015.03.261

142. Schieving JH. PP05.5–3064: the role of exome sequencing in daily pediatric neurology practice. *Eur J Paediatr Neurol* 2015;**19**:S47. https://doi.org/10.1016/S1090-3798(15)30154-9

143. Towne MC, Beggs AH, Agrawal PB. *Efficiency of Whole Exome/Genome Sequencing for Achieving a Diagnosis in Rare Presentations*. Annual Meeting of the American Society of Human Genetics, 22–26 October 2013, Boston, MA, abstract no. 1379.

144. Pankhurst LJ, Del Ojo Elias C, Votintseva AA, Walker TM, Cole K, Davies J, *et al*.; COMPASS-TB Study Group. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. *Lancet Resp Med* 2016;**4**:49–58. https://doi.org/10.1016/S2213-2600(15)00466-X

145. McDonell LM, Warman Chardon J, Schwartzentruber J, Foster D, Beaulieu CL, Majewski J, *et al*.; FORGE Canada Consortium. The utility of exome sequencing for genetic diagnosis in a familial microcephaly epilepsy syndrome. *BMC Neurol* 2014;**14**:22. https://doi.org/10.1186/1471-2377-14-22

146. Neveling K, Feenstra I, Gilissen C, Hoefsloot LH, Kamsteeg EJ, Mensenkamp AR, *et al*. A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. *Hum Mutat* 2013;**34**:1721–6. https://doi.org/10.1002/humu.22450

147. Sawyer SL, Schwartzentruber J, Beaulieu CL, Dyment D, Smith A, Chardon JW, *et al*. Exome sequencing as a diagnostic tool for pediatric-onset ataxia. *Hum Mutat* 2014;**35**:45–9. https://doi.org/10.1002/humu.22451

148. Dewey FE, Grove ME, Pan C, Goldstein BA, Bernstein JA, Chaib H, *et al*. Clinical interpretation and implications of whole-genome sequencing. *JAMA* 2014;**311**:1035–45. https://doi.org/10.1001/jama.2014.1717

149. Ghaoui R, Cooper ST, Lek M, Jones K, Corbett A, Reddel SW, *et al*. Use of whole-exome sequencing for diagnosis of limb-girdle muscular dystrophy: outcomes and lessons learned. *JAMA Neurol* 2015;**72**:1424–32. https://doi.org/10.1001/jamaneurol.2015.2274

150. Bonnefond A, Philippe J, Durand E, Muller J, Saeed S, Arslan M, *et al*. Highly sensitive diagnosis of 43 monogenic forms of diabetes or obesity through one-step PCR-based enrichment in combination with next-generation sequencing. *Diabet Care*. 2014;**37**:460–7. https://doi.org/10.2337/dc13-0698

151. Lee EJ, Dykas DJ, Leavitt AD, Camire RM, Ebberink E, García de Frutos P, *et al*. Whole-exome sequencing in evaluation of patients with venous thromboembolism. *Blood Adv* 2017;**1**:1224–37. https://doi.org/10.1182/bloodadvances.2017005249

152. Bennette CS, Gallego CJ, Burke W, Jarvik GP, Veenstra DL. The cost-effectiveness of returning incidental findings from next-generation genomic sequencing. *Genet Med* 2015;**17**:587–95. https://doi.org/10.1038/gim.2014.156

153. Chrystoja CC, Diamandis EP. Whole genome sequencing as a diagnostic test: challenges and opportunities. *Clin Chem* 2014;**60**:724–33. https://doi.org/10.1373/clinchem.2013.209213

154. Crawford S, Gong C, Randolph LM, Yieh L, Hay JW. PIH27 diagnosing newborns with suspected severe mitochondrial disorders: a cost-effectiveness study comparing early whole exome sequencing to standard of care. *Value Health* 2020;**23**:S156. https://doi.org/10.1016/j.jval.2020.04.422

155. Yuen T, Carter MT, Szatmari P, Ungar WJ. Cost-effectiveness of genome and exome sequencing in children diagnosed with autism spectrum disorder. *Appl Health Econ Health Policy* 2018;**16**:481–93. https://doi.org/10.1007/s40258-018-0390-x

156. Mfumbilwa ZA, Simons M, Ramaekers B, Retèl VP, Mankor JM, Groen HJM, *et al*. Exploring the cost effectiveness of a whole-genome sequencing-based biomarker for treatment selection in patients with advanced lung cancer ineligible for targeted therapy. *PharmacoEconomics* 2024;**42**:419–34. https://doi.org/10.1007/s40273-023-01344-w

157. Simons MJHG, Retel VP, Ramaekers BLT, Butter R, Mankor JM, Paats MS, *et al*. Early cost effectiveness of whole-genome sequencing as a clinical diagnostic test for patients with inoperable stage IIIB,C/IV non-squamous non-small-cell lung cancer. *PharmacoEconomics* 2021;**39**:1429–42. https://doi.org/10.1007/s40273-021-01073-y

158. Simons MJHG, Uyl-de Groot CA, Retel VP, Mankor JM, Ramaekers BLT, Joore MA, van Harten WH. Cost-effectiveness and budget impact of future developments with whole-genome sequencing for patients with lung cancer. *Value Health* 2023;**26**:71–80. https://doi.org/10.1016/j.jval.2022.07.006

159. Avram CM, Caughey AB, Norton ME, Sparks TN. Cost-effectiveness of exome sequencing versus targeted gene panels for prenatal diagnosis of fetal effusions and non-immune hydrops fetalis. *Am J Obstetr Gynecol* 2022;**4**:100724. https://doi.org/10.1016/j.ajogmf.2022.100724

160. Wu Y, Jayasinghe K, Stark Z, Quinlan C, Patel C, McCarthy H, *et al*.; KidGen Collaborative Investigators. Genomic testing for suspected monogenic kidney disease in children and adults: a health economic evaluation. *Genet Med* 2023;**25**:100942. https://doi.org/10.1016/j.gim.2023.100942

161. Sanford Kobayashi E, Waldman B, Engorn BM, Perofsky K, Allred E, Briggs B, *et al*. Cost efficacy of rapid whole genome sequencing in the pediatric intensive care unit. *Front Pediatr* 2021;**9**:809536. https://doi.org/10.3389/fped.2021.809536

162. Stark Z, Schofield D, Martyn M, Rynehart L, Shrestha R, Alam K, *et al*. Does genomic sequencing early in the diagnostic trajectory make a difference? A follow-up study of clinical outcomes and cost-effectiveness. *Genet Med* 2019;**21**:173–80. https://doi.org/10.1038/s41436-018-0006-8

163. Stark Z, Schofield D, Alam K, Wilson W, Mupfeki N, Macciocca I, *et al*. Prospective comparison of the cost-effectiveness of clinical whole-exome sequencing with that of usual care overwhelmingly supports early use and reimbursement. *Genet Med* 2017;**19**:867–74. https://doi.org/10.1038/gim.2016.221

164. Schofield D, Rynehart L, Shresthra R, White SM, Stark Z. Long-term economic impacts of exome sequencing for suspected monogenic disorders: diagnosis, management, and reproductive outcomes. *Genet Med* 2019; **21**:2586–93. https://doi.org/10.1038/s41436-019-0534-x

165. Lavelle TA, Feng X, Keisler M, Cohen JT, Neumann PJ, Prichard D, *et al*. Cost-effectiveness of exome and genome sequencing for children with rare and undiagnosed conditions. *Genet Med* 2022;**24**:1349–61. https://doi.org/10.1016/j.gim.2022.03.005

166. Nurchis MC, Radio FC, Salmasi L, Heidar Alizadeh A, Raspolini GM, Altamura G, *et al*. Bayesian cost-effectiveness analysis of whole genome sequencing versus whole exome sequencing in a pediatric population with suspected genetic disorders. *Eur J Health Econ* 2023;**25**:999–1011. https://doi.org/10.1007/s10198-023-01644-0

167. Nurchis MC, Radio FC, Salmasi L, Heidar Alizadeh A, Raspolini GM, Altamura G, *et al*. Cost-effectiveness of whole-genome vs whole-exome sequencing among children with suspected genetic disorders. *JAMA Netw Open* 2024;**7**:e2353514. https://doi.org/10.1001/jamanetworkopen.2023.53514

168. Oerton J, Khalid JM, Besley G, Dalton RN, Downing M, Green A, *et al*. Newborn screening for medium chain acyl-CoA dehydrogenase deficiency in England: prevalence, predictive value and test validity based on 1.5 million screened babies. *J Med Screen* 2011;**18**:173–81. https://doi.org/10.1258/jms.2011.011086

169. International Consortium on Newborn Sequencing (ICoNS). *We Are ICoNS the International Consortium on Newborn Sequencing*. London: ICoNS; 2024. URL: www.iconseq.org/ (accessed 14 May 2024).

170. Downie L, Bouffler SE, Amor DJ, Christodoulou J, Yeung A, Horton AE, *et al*. Gene selection for genomic newborn screening: moving toward consensus? *Genet Med* 2024;**26**:101077. https://doi.org/10.1016/j.gim.2024.101077

171. Horton R, Lucassen A. Ethical issues raised by new genomic technologies: the case study of newborn genome screening. *Camb Prisms Precis Med* 2022;**1**:1–16. https://doi.org/10.1017/pcm.2022.2

172. Turnbull C, Firth HV, Wilkie AOM, Newman W, Raymond FL, Tomlinson I, *et al*. Population screening requires robust evidence – genomics is no exception. *Lancet* 2024;**403**:583–6. https://doi.org/10.1016/S0140-6736(23)02295-X

173. Jackson L, Weedon MN, Green HD, Mallabar-Rimmer B, Harrison JW, Wood AR, *et al*. Influence of family history on penetrance of hereditary cancers in a population setting. *EClinicalMedicine* 2023;**64**:102159. https://doi.org/10.1016/j.eclinm.2023.102159

174. Horton R, Wright CF, Firth HV, Turnbull C, Lachmann R, Houlston RS, Lucassen A. Challenges of using whole genome sequencing in population newborn screening. *BMJ* 2024;**384**:e077060. https://doi.org/10.1136/bmj-2023-077060

175. Kinsella S, Hopkins H, Cooper L, Bonham JR. A public dialogue to inform the use of wider genomic testing when used as part of newborn screening to identify cystic fibrosis. *Int J Neonatal Screen* 2022;**8**:32. https://doi.org/10.3390/ijns8020032

176. Clark CCA, Boardman FK. Expanding the notion of 'benefit': comparing public, parent, and professional attitudes towards whole genome sequencing in newborns. *New Genet Soc* 2022;**41**:96–115. https://doi.org/10.1080/14636778.2022.2091533

177. Chudleigh J, Holder P, Clark C, Moody L, Cowlard J, Allen L, *et al*. Parents' and childrens' views of wider genomic testing when used as part of newborn screening to identify cystic fibrosis. *SSM Qual Res Health* 2024;**6**:100455. https://doi.org/10.1016/j.ssmqr.2024.100455

178. Pearce A, Mitchell LA, Best S, Young MA, Terrill B. Publics' knowledge of, attitude to and motivation towards health-related genomics: a scoping review. *Eur J Hum Genet* 2024;**32**:747–58. https://doi.org/10.1038/s41431-024-01547-5

179. Hopkins H, Kinsella S, Evans G. *Implications of Whole Genome Sequencing for Newborn Screening – A Public Dialogue: A Findings Report Hopkins Van Mil July 2021*. London: UK National Screening Committee; 2021. URL: www.gov.uk/government/publications/implications-of-whole-genome-sequencing-for-newborn-screening (accessed 15 February 2024).

180. Etchegary H, Winsor M, Power A, Simmonds C. Public engagement with genomic medicine: a summary of town hall discussions. *J Commun Genet* 2021;**12**:27–35. https://doi.org/10.1007/s12687-020-00485-1

181. Werner KM. The gender gap in caring for children with medical complexity. *J Perinatol* 2023;**43**:835–6. https://doi.org/10.1038/s41372-023-01652-1

182. Adhikari AN, Gallagher RC, Wang Y, Currier RJ, Amatuni G, Bassaganyas L, *et al*. The role of exome sequencing in newborn screening for inborn errors of metabolism. *Nat Med* 2020;**26**:1392–7. https://doi.org/10.1038/s41591-020-0966-5

183. Schwartz O, Vill K, Pfaffenlehner M, Behrens M, Weiß C, Johannsen J, *et al*. Clinical effectiveness of newborn screening for spinal muscular atrophy: a nonrandomized controlled trial. *JAMA Pediatr* 2024;**178**:540–7. https://doi.org/10.1001/jamapediatrics.2024.0492

184. University of York Centre for Reviews and Dissemination. *DARE and NHS EED Archives Secure on CRD Website until At Least the End of March 2024*. 2024. URL: www.crd.york.ac.uk/CRDWeb/ (accessed 30 May 2024).

185. CADTH. *CADTH Search Filters Database*. Ottawa, ON: Canadian Agency for Drugs and Technologies in Health; 2024. URL: https://searchfilters.cadth.ca/link/16 (accessed 17 May 2024).

186. Zhang Y, Wei WB, Zhao J, Xu X, Wang F. Spectrum and tissue distribution of RB1 pathogenic alleles in mosaic retinoblastoma patients. *Ophthalmic Genet* 2022;**43**:795–805. https://doi.org/10.1080/13816810.2022.2098985

187. Coughlin CR II, van Karnebeek CDM, Al-Hertani W, Shuen AY, Jaggumantri S, Jack RM, *et al*. Triple therapy with pyridoxine, arginine supplementation and dietary lysine restriction in pyridoxine-dependent epilepsy: neurodevelopmental outcome. *Mol Genet Metab* 2015;**116**:35–43. https://doi.org/10.1016/j.ymgme.2015.05.011

188. Alfadhel M, Sirrs S, Waters PJ, Szeitz A, Struys E, Coulter-Mackie M, Stockler-Ipsiroglu S. Variability of phenotype in two sisters with pyridoxine dependent epilepsy. *Can J Neurol Sci* 2012;**39**:516–9. https://doi.org/10.1017/s0317167100014050

189. Marguet F, Barakizou H, Tebani A, Abily-Donval L, Torre S, Bayoudh F, *et al*. Pyridoxine-dependent epilepsy: report on three families with neuropathology. *Metab Brain Dis* 2016;**31**:1435–43. https://doi.org/10.1007/s11011-016-9869-z

190. Rankin PM, Harrison S, Chong WK, Boyd S, Aylett SE. Pyridoxine-dependent seizures: a family phenotype that leads to severe cognitive deficits, regardless of treatment regime. *Dev Med Child Neurol* 2007;**49**:300–5. https://doi.org/10.1111/j.1469-8749.2007.00300.x

191. Yeghiazaryan NS, Striano P, Spaccini L, Pezzella M, Cassandrini D, Zara F, Mastrangelo M. Long-term follow-up in two siblings with pyridoxine-dependent seizures associated with a novel ALDH7A1 mutation. *Eur J Paediatr Neurol* 2011;**15**:547–50. https://doi.org/10.1016/j.ejpn.2011.05.011

192. Ulvi H, Mungen B, Yakinci C, Yoldas T. Pyridoxine-dependent seizures: long-term follow-up of two cases with clinical and MRI findings, and pyridoxine treatment. *J Trop Pediatr* 2002;**48**:303–6. https://doi.org/10.1093/tropej/48.5.303

193. Tseng LA, Hoytema van Konijnenburg EMM, Longo N, Andrews A, van Wegberg A, Coene KLM, *et al*. Clinical reasoning: pediatric seizures of unknown cause. *Neurology* 2022;**98**:1023–8. https://doi.org/10.1212/wnl.0000000000200711

194. Busiello R, Galgani M, De Fusco C, Poggi V, Adriani M, Racioppi L, *et al*. Role of A91V mutation in perforin gene in hemophagocytic lymphohistiocytosis. *Blood* 2004;**104**:1910.

Appendix 1 Search development methods

Overview of the search development methods for the review of five conditions, the review of genomic studies of newborn cohorts reporting penetrance for pathogenic variants and the review of cost-effectiveness evaluations of WGS and WES.

Review of five conditions

Search strategies were developed for each condition by an Information Specialist. The searches were developed in a test database [MEDLINE (Ovid)] and were informed and refined through a series of scoping searches, checks of a proportion of results from these searches and iterative discussions between the Information Specialist (ND), project lead (KF) and members of the reviewing team (IK, JD and SC).

Exploratory scoping searches were carried out for each condition to gain familiarity with the condition and the genetic cause. The scoping searches revealed that searching for the five conditions in one single search would not be a feasible approach due to the complexity and differences across each condition and their genetic causalities. In our scoping searches, we tested searching for the specific genes associated with each condition and the condition, for example, rb1 and RB. Combining the search terms for the condition and the gene with the Boolean operator 'Or' yielded an unmanageable volume of irrelevant results of the specific gene related to other conditions. For example, variants of the RB gene are associated with a large amount of other cancer types. Combining the terms for the gene and the condition using the Boolean operator 'And' would produce an overly specific search, which would have resulted in potentially relevant results being missed. Therefore, an iterative approach was adopted. A standardised search strand for terms related to hereditary/inherited conditions or genetics was utilised for the conditions that can also present for reasons that are not due to the specific genetic variants. These included: fHLH, RB and XLHR. The searches for MCADD and PDE did not need to include the search terms for genetics, as the numbers retrieved were manageable without the addition of any other concepts and the draft search results for these conditions did not yield such a high proportion of irrelevant results. The process of developing and running the searches took a period of 6 weeks, which is longer than average for our Information Specialist team. This was largely due to the complexity of the topic.

Database-specific subject headings and free text words were identified for use in the search concepts by analysing the free text and indexing terms of the results from the scoping searches, text analysis software, including medical subject heading (MeSH) on Demand, Anne O'Tate and PubMed ReMiner by the Information Specialist. Further terms were identified and tested from known relevant papers and resources, including ClinGen and GeneReviews. The searches were peer-reviewed by a Senior Information Specialist and the project lead.

The following databases were searched from inception to November 2023 (see *Appendix 2* for exact dates and full search details): MEDLINE (via Ovid), EMBASE (via Ovid), SCI (via Clarivate) and the Cochrane Library (via Wiley). No date, language or study-type filters were applied. Search results were managed using EndNote 20 and systematically deduplicated using the University of Leeds method.[18]

Review of genomic studies of newborn cohorts reporting penetrance for pathogenic variants

Searches were developed iteratively in a single database (MEDLINE via Ovid) by an experienced information specialist (RC) and the project lead (KF), with input from members of the reviewing team. Developing and running the searches took approximately 5 weeks. The development process took longer than usual because there appears to be little evidence in the newborn screening setting. Despite running scoping searches, we did not have enough examples of published literature to confidently base an initial search on. Therefore, we began with a narrow search combining the concepts of WGS and newborn screening, checking samples of records for potentially relevant literature. Simultaneously, we undertook targeted searches to check for outputs from known large genomic studies, such as the BabySec project. It became evident that most of these ongoing studies had not published results at the time. We question whether a future review of just large genomic studies would be helpful to answer the question around penetrance because (1) conditions are so rare that there is limited scope for evidence on penetrance from 1 or 2 years' worth of data and (2) newborns positive on WGS will be treated and not followed up to symptoms. Due to not finding

much of relevance in this first iteration of the search and our other concerns, we broadened it by adding search terms for related concepts, such as genetic testing and sequencing, ran test searches with and without certain concepts and checked samples for potentially relevant literature. We also considered other approaches, such as searching for specific terms within the main body of an article using a database of full-text publications. The final search combines the concept of newborn screening with either WGS, WES, penetrance, actionability, sequencing or allele frequency. This search retrieved a large, but manageable number of records in MEDLINE and found all five known studies we had identified up to this point.[122,126,128,129,138]

Search strategy

The following databases were searched from inception to January 2024 (see *Appendix 2* for exact dates and full search details): MEDLINE (via Ovid), EMBASE (via Ovid), SCI (via Clarivate) and the Cochrane Library (via Wiley). No date, language or study-type filters were applied. Records were exported to EndNote and systematically deduplicated using a process based on the University of Leeds method.[18]

Review of cost-effectiveness evaluations of whole genome sequencing and whole exome sequencing

The searches for the review of cost-effectiveness studies were developed by an Information Specialist with input from the lead economist (BS) and peer-reviewed by the Senior Information Specialist (RC). Comprehensive database searches were undertaken to identify evidence relating to cost effectiveness and WGS. The search strategies are reported in *Appendix 2*.

We considered updating the search strategy carried out for a systematic review by Schwarze *et al.*;[27] however, as their review was carried out in July 2016, they were able to significantly make use of the NHS Economic Evaluation Database (NHS EED), produced by the Centre for Reviews and Dissemination. NHS EED ceased adding new records in March 2015 and the searches to identify them stopped at the end of 2014;[184] therefore, it was agreed that we would need to carry out broader searches on MEDLINE, EMBASE and the SCI, utilising an economic search filter. Supplementary searches was also carried out on the CEA registry and the review by Scwartz *et al.*[27] was used as a source by cross-checking their included studies.

The search terms were derived from analysing the free text and indexing terms of relevant known studies and analysing the results from initial scoping searches. The terms for the cost-effectiveness search strand was developed from an economic search filter.[185] The MEDLINE (Ovid), EMBASE (Ovid), SCI (WoS – Clarivate), EconLit (EBSCO), the CEA registry, the HTA International database and Google were searched from inception on the 13 February 2023 for economic evaluations, HTA reports and economic models. The search results were limited to English-language studies. The search results were stored and deduplicated in EndNote 20 using the University of Leeds method.[18]

Appendix 2 Search strategy

Search details, including databases, date of search, terms and breakdown of number of results for the traditional review of five conditions, the review of penetrance or actionability of gene variants of rare genetic childhood-onset diseases identified in newborn screening populations using WGS (alternative review) and the review of cost-effectiveness evaluations of WGS and WES.

Review of five conditions

Pyridoxine-dependent epilepsy

Search summary

Database (platform)	Date searched	Concepts	Hits	Notes
MEDLINE (Ovid)	29 September 2023	PDE	545	Removed animal studies English-language limit
EMBASE (Ovid)	29 September 2023	PDE	706	Removed animal studies No Emtree terms focused English-language limit
SCI (WoS)	29 September 2023	PDE	584	
Cochrane Library	29 September 2023	PDE	7	

Total from database searches: 1835.

Total after systematically removing duplicates in EndNote (using University of Leeds method):[18] 994.

Ovid MEDLINE(R) ALL 1946–18 September 2023

PDE	1	(exp vitamin b 6/or exp pyridoxal/or exp pyridoxamine/or exp pyridoxine/) and (Epilepsy/or Seizures/or seizures, febrile/or exp status epilepticus/) and (dependen* or dependan*).ti,ab,kf,rx.	247
	2	((pyridoxine or pyridoxin or pyridoxamine or vitamin b6 or "vitamin b 6") and (dependen* or dependan*) and (epilep* or seizure* or convuls* or spasm*)).ti,ab,kf,rx.	479
	3	Aldehyde Dehydrogenase/df [Deficiency]	168
	4	PDE-ALDH7A1.ti,ab,kf.	23
	5	((AASA or "α-AASA" or alpha aminoadipic semialdehyde) and dehydrogenase deficien*).ti,ab,kf.	16
	6	((Antiquitin or ATQ or ASADH) and deficien*).ti,ab,kf.	58
	7	or/1-6 [pyridoxamine dependent epilepsy]	653
Removing animal studies	8	(exp Animals/or Models, Animal/or Disease Models, Animal/) not Humans/	51,55,276
	9	7 not 8	593
Limit to English language	10	limit 9 to english language	545

EMBASE Classic + EMBASE 1947–19 September 2023

1 pyridoxine-dependent epilepsy/164
2 (pyridoxal/or pyridoxamine/or exp pyridoxamine/or pyridoxine/) and (*Epilepsy/or *febrile convulsion/or *epileptic state/) and (dependen* or dependan*).ti,ab,kf. 285
3 ((pyridoxine or pyridoxin or pyridoxamine or vitamin b6 or "vitamin b 6") and (dependen* or dependan*) and (epilep* or seizure* or convuls* or spasm*)).ti,ab,kf. 756
4 PDE-ALDH7A1.ti,ab,kf.28
5 ((AASA or "α-AASA" or alpha aminoadipic semialdehyde) and dehydrogenase deficien*).ti,ab,kf. 23
6 ((Antiquitin or ATQ or ASADH) and deficien*).ti,ab,kf. 110
7 or/1-6 [pyridoxamine dependent epilepsy]828
8 (exp animal/or juvenile animal/or adult animal/or nonhuman/or animal experiment/or animal model/) not Human/8380712
9 7 not 8 779
10 limit 9 to english language 706

Web of Science – Science Citation Index Expanded (SCI-EXPANDED)–1970–present, Social Sciences Citation Index (SSCI)–1900–present

1 TS=((pyridoxine or pyridoxin or pyridoxamine or "vitamin b6" or "vitamin b 6") and (dependen* or dependan*) and (epilep* or seizure* or convuls* or spasm*))571
2 TS = PDE-ALDH7A1 17
3 TS=((AASA or "α-AASA" or "alpha aminoadipic semialdehyde") and "dehydrogenase deficien*")17
4 TS=((Antiquitin or ATQ or ASADH) and deficien*)118
5 #4 OR #3 OR #2 OR #1 584

Cochrane Library

#1 MeSH descriptor: [] explode all trees0
#2 MeSH descriptor: [Pyridoxal] this term only87
#3 MeSH descriptor: [Pyridoxamine] this term only16
#4 MeSH descriptor: [Pyridoxine] this term only507
#5 #2 OR #3 OR #4 534
#6 MeSH descriptor: [Epilepsy] this term only2204
#7 MeSH descriptor: [Seizures, Febrile] this term only102
#8 MeSH descriptor: [] explode all trees0
#9 #6 OR #7 OR #8 2297
#10 (dependen* or dependan*):TI,AB,KW 93205
#11 #10 AND #5 AND #92
#12 ((pyridoxine or pyridoxin or pyridoxamine or vitamin b6 or "vitamin b 6") and (dependen* or dependan*) and (epilep* or seizure* or convuls* or spasm*)):TI,AB,KW7
#13 PDE-ALDH7A1:TI,AB,KW0
#14 ((AASA or "α-AASA" or alpha aminoadipic semialdehyde) and dehydrogenase deficien*):TI,AB,KW0
#15 ((Antiquitin or ATQ or ASADH) and deficien*):TI,AB,KW0
#16 #1 OR #11 OR #12 OR #13 OR #14 OR #157

Cochrane database of systematic reviews: 1

CENTRAL: 6

Heritable retinoblastoma

Search summary

Database (platform)	Date searched	Concepts	Hits	Notes
MEDLINE (Ovid)	4 October 2023	Retinoblastoma AND Genetics/hereditary	4521	Removed animal studies English-language limit
EMBASE (Ovid)	4 October 2023	Retinoblastoma AND Genetics/hereditary	3345	Removed animal studies English-language limit
SCI (WoS)	4 October 2023	Retinoblastoma AND Genetics/hereditary	471	Removed animal studies English-language limit
Cochrane Library	4 October 2023	Retinoblastoma AND Genetics/hereditary	18	

Total from database searches: 8355.

Total after systematically removing duplicates in EndNote (using University of Leeds method):[18] 5797.

Ovid MEDLINE(R) ALL 1946–2 October 2023

Retinoblastoma	1	Retinoblastoma/	8207
	2	Retinal Neoplasms/ge	919
	3	Genes, Retinoblastoma/	1772
	4	(Retina* adj3 (cancer* or tumor* or tumour* or neoplasm* or glioblastoma* or glioma* or neuroblastoma*)).ti,ab,kf.	1582
	5	or/1-4 [retinoblastoma]	10,873
Hereditary/genetics	6	Genetics/	12,921
	7	Genetics.fs.	4,015,866
	8	Genetic disorder/	14,547
	9	exp genetic predisposition to disease/	157,705
	10	Genetic Diseases, Inborn/	14,547
	11	(Genetic* or gene or genes or family or families or familial or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*).ti,ab,kf.	5,102,985
	12	or/6-11 [Hereditary/genetics]	6,533,487
Retinoblastoma and genetics/hereditary	13	5 and 12 [Retinoblastoma and genetics/hereditary]	5401
Removed animal studies	14	(exp Animals/or exp Models, Animal/or Disease Models, Animal/) not Humans/	5,161,657
	15	13 not 14	5025
Limit to English language	16	limit 15 to english language	4521

EMBASE Classic + EMBASE 1947–Week 39, 2023

1 hereditary retinoblastoma/35
2 *retinoblastoma/ 9424
3 *retina tumor/1197
4 (Retina* adj3 (cancer* or tumor* or tumour* or neoplasm* or glioblastoma* or glioma* or neuroblastoma*)).ti,ab,kf. 2127
5 or/2-4 [retinoblastoma]11626
6 *genetics/127456
7 *genetic disorder/22358
8 *heredity/9388
9 *genetic predisposition/11661
10 *genetic disorder/22358
11 (Genetic* or gene or genes or family or families or familial or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*).ti,ab,kf. 6421949
12 or/6-11 [Hereditary/ genetics]6455505
13 (5 and 12) or 1 [Retinoblastoma and genetics/ hereditary]4028
14 (exp animal/ or exp juvenile animal/ or adult animal/ or animal cell/ or animal tissue/or nonhuman/or animal experiment/or animal model/) not human/8231136
15 13 not 14 3734
16 limit 15 to english language 3345

Web of Science

1 TS = (Retina* near/3 (cancer* or tumor* or tumour* or neoplasm* or glioblastoma* or glioma* or neuroblastoma*))1327
2 TS = (Genetic* or gene or genes or family or families or familial or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*) 6,625,490
3 #1 AND #2 495
4 LIMIT #3 to English language 471

Cochrane Library

IDSearchHits

#1 MeSH descriptor: [Retinoblastoma] this term only37
#2 MeSH descriptor: [Retinal Neoplasms] this term only17
#3 MeSH descriptor: [Genes, Retinoblastoma] this term only1
#4 (Retina* near/3 (cancer* or tumor* or tumour* or neoplasm* or glioblastoma* or glioma* or neuroblastoma*)):ti,ab,kw 54
#5 #1 or #2 or #3 or #475
#6 MeSH descriptor: [Genetics] this term only103
#7 MeSH descriptor: [Genetic Diseases, Inborn] this term only81
#8 MeSH descriptor: [Genetic Predisposition to Disease] explode all trees1602
#9 (Genetic* or gene or genes or family or families or familial or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal):ti,ab,kw 147555
#10 #6 or #7 or #8 or #9147555
#11 #5 and #10 18

X-linked hypophosphataemic rickets

Search summary

Database (platform)	Date searched	Concepts	Hits	Notes
MEDLINE (Ovid)	5 October 2023	X-linked Hypophosphatemic Rickets AND Genetics/hereditary	2078	Removed animal studies English-language limit
EMBASE (Ovid)	5 October 2023	X-linked Hypophosphatemic Rickets AND Genetics/hereditary	3474	Removed animal studies English-language limit
SCI (WoS)	5 October 2023	X-linked Hypophosphatemic Rickets AND Genetics/hereditary	2088	English-language limit
Cochrane Library	5 October 2023	X-linked Hypophosphatemic Rickets AND Genetics/hereditary	278	

Total from database searches: 7918.

Total after systematically removing duplicates in EndNote (using University of Leeds method):[18] 4787.

Ovid MEDLINE(R) ALL 1946–18 September 2023

Hypophosphatemic rickets	1	exp Rickets, Hypophosphatemic/	952
	2	Hypophosphat?emic.ti,ab,kf,rx.	2433
	3	rickets/or (rickets or rachitides or rachitis).ti,ab,kf,rx.	10,233
	4	2 and 3	1644
	5	1 or 4 [Hypophosphatemic rickets]	2080
Hypophosphatemia	6	Hypophosphatemia/	1908
	7	Hypophosphat?emia*.ti,ab,kf,rx.	5397
	8	or/6-7 [Hypophosphatemia]	5864
Vitamin D-resistant rickets	9	(Vitamin D resistant and (rickets or rickets or rachitides or rachitis)).ti,ab,kf,rx.	608
	10	VDRR.ti,ab,kf.	30
	11	or/9-10 [Vitamin D resistant rickets]	612
Hypophosphatemia or Hypophosphatemic rickets or vitamin D-resistant rickets	12	5 or 8 or 11 [Hypophosphatemia or Hypophosphatemic rickets or vitamin D resistant rickets]	7318
Genetics/hereditary	13	Genetics/	12,922
	14	Genetics.fs.	4,016,108
	15	Genetic disorder/	14,548
	16	exp genetic predisposition to disease/	157,711
	17	Genetic Diseases, Inborn/	14,548
	18	(Genetic* or gene or genes or family or families or familial or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*).ti,ab,kf.	5,103,147
	19	or/13-18 [Genetics/hereditary]	6,533,601

Hypophosphatemia or Hypophosphatemic rickets and genetics/ hereditary	20	12 and 19 [Hypophosphatemia or Hypophosphatemic rickets and genetics/hereditary]	2327
Phrase searches for X-linked hypophosphatemic hypophosphataemic rickets	21	((XLHR or XLH or X Linked) adj3 (hypophosphatemia* or hypophosphataemia*)).ti,ab,kf,rx.	659
	22	(((XLHR or XLH or X Linked) adj3 (Hypophosphatemic or hypophosphataemic)) and (rickets or rachitides or rachitis)).ti,ab,kf,rx.	580
	23	(((XLHR or XLH or X Linked) adj3 vitamin d resistant) and (rickets or rachitides or rachitis)).ti,ab,kf,rx.	41
	24	or/20-23	2656
Removal of animal studies	25	(exp Animals/or exp Models, Animal/or Disease Models, Animal/) not Humans/	5,162,107
	26	24 not 25	2286
Limit to English language	27	limit 26 to english language	2078

EMBASE Classic + EMBASE 1947–4 October 2023

1 *hypophosphatemic rickets/496
2 Hypophosphat?emic.ti,ab,kf. 3461
3 *rickets/or (rickets or rachitides or rachitis).ti,ab,kf.13047
4 (2 and 3) or 12517
5 *hypophosphatemia/3187
6 Hypophosphat?emia*.ti,ab,kf. 9143
7 or/5-6 [Hypophosphatemia]10109
8 *vitamin D resistant rickets/964
9 (Vitamin d resistant and (rickets or rickets or rachitides or rachitis)).ti,ab,kf. 838
10 VDRR.ti,ab,kf.49
11 or/8-10 1558
12 4 or 7 or 1111821
13 *genetics/127456
14 *genetic disorder/22364
15 *heredity/ 9395
16 *genetic predisposition/11661
17 *genetic disorder/22364
18 (Genetic* or gene or genes or family or families or familial or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*).ti,ab,kf. 6423367
19 or/13-18 [Genetics/hereditary]6456923
20 12 and 19 3272
21 *X linked hypophosphatemic rickets/760
22 *vitamin D resistant rickets/964
23 *familial hypophosphatemic rickets/160
24 ((XLHR or XLH or X Linked) adj3 (hypophosphatemia* or hypophosphataemia*)).ti,ab,kf. 1064
25 (((XLHR or XLH or X Linked) adj3 (Hypophosphatemic or hypophosphataemic)) and (rickets or rachitides or rachitis)).ti,ab,kf. 800
26 (((XLHR or XLH or X Linked) adj3 vitamin d resistant) and (rickets or rachitides or rachitis)).ti,ab,kf. 54
27 or/20-26 4421
28 (exp animal/ or exp juvenile animal/ or adult animal/ or animal cell/or animal tissue/or nonhuman/or animal experiment/or animal model/) not human/8232138
29 27 not 28 3871
30 limit 29 to english language 3474

Web of Science – Science Citation Index Expanded (SCI-EXPANDED)–1970–present, Social Sciences Citation Index (SSCI)–1900–present

1 ts=((Hypophosphatemic or hypophosphataemic) and (rickets or rachitides or rachitis)) 2259
2 ts=((Hypophosphatemic or hypophosphataemic) and (rickets or rachitides or rachitis)) 5403
3 TS=("Vitamin d resistant" and (rickets or rickets or rachitides or rachitis)) 424
4 TS = VDRR 15
5 **#4 OR #3 OR #2 OR #1 7149**
6 TS = (Genetic* or gene or genes or family or families or familial or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*) 6,609,856
7 #6 and #5 2161
8 #6 AND #5 and English (Languages) 2088

Cochrane Library

#1 MeSH descriptor: [Rickets, Hypophosphatemic] explode all trees35
#2 Hypophosphat?emic:ti,ab,kw136
#3 MeSH descriptor: [Rickets] this term only118
#4 #2 and #38
#5 #1 or #438
#6 MeSH descriptor: [Hypophosphatemia] this term only163
#7 Hypophosphat?emia*:ti,ab,kw796
#8 #6 or #7 796
#9 ("Vitamin d resistant" and (rickets or rickets or rachitides or rachitis)):ti,ab,kw8
#10 VDRR:ti,ab,kw0
#11 #9 or #108
#12 #5 or #8 or #11808
#13 MeSH descriptor: [Genetics] this term only103
#14 MeSH descriptor: [Genetic Diseases, Inborn] this term only81
#15 MeSH descriptor: [Genetic Predisposition to Disease] explode all trees1602
#16 (Genetic* or gene or genes or family or families or familial or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*):ti,ab,kw 147556
#17 #13 or #14 or #15 or #16147556
#18 #12 and 17 186
#19 ((XLHR or XLH or X Linked) near/3 (hypophosphatemia* or hypophosphataemia*)):ti,ab,kw 96
#20 (((XLHR or XLH or X Linked) near/3 (Hypophosphatemic or hypophosphataemic)) and (rickets or rachitides or rachitis)):ti,ab,kw 91
#21 (((XLHR or XLH or X Linked) near/3 vitamin d resistant) and (rickets or rachitides or rachitis)):ti,ab,kw2
#22 #18 or #19 or #20 or #21278

Cochrane Reviews: 4

CENTRAL: 274

Familial haemophagocytic lymphohistiocytosis

Search summary

Database (platform)	Date searched	Concepts	Hits	Notes
MEDLINE (Ovid)	6 October 2023	Hemophagocytic lymphohistiocytosis AND Genetics/hereditary OR Phrase searches for familial hemophagocytic lymphohistiocytosis	2395	Removed animal studies English-language limit

Database (platform)	Date searched	Concepts	Hits	Notes
EMBASE (Ovid)	6 October 2023	Hemophagocytic lymphohistiocytosis AND Genetics/hereditary OR Phrase searches for familial hemophagocytic lymphohistiocytosis	3841	Removed animal studies No Emtree terms focused English-language limit
SCI (WoS)	6 October 2023	Hemophagocytic lymphohistiocytosis AND Genetics/hereditary OR Phrase searches for familial hemophagocytic lymphohistiocytosis	2773	
Cochrane Library	6 October 2023	Hemophagocytic lymphohistiocytosis AND Genetics/hereditary OR Phrase searches for familial hemophagocytic lymphohistiocytosis	37	

Total from database searches: 9046.

Total after systematically removing duplicates in EndNote (using University of Leeds method):[18] 5151.

Ovid MEDLINE(R) ALL 1946–5 October 2023

	1	Lymphohistiocytosis, Hemophagocytic/	3949
Hemophagocytic lymphohistiocytosis	2	((Hemophagocytic or haemophagocytic or erythrophagocytic) adj3 (lymphohistiocytos* or lymphocytos* or histiocytos* or reticulos* or hymphohistiocytos* or syndrome*)).ti,ab,kf,rx.	7098
	3	or/1-2 [Hemophagocytic Lymphohistiocytosis]	7556
Genetics/hereditary	4	Genetics/	12,922
	5	Genetics.fs.	4,016,776
	6	Genetic disorder/	14,549
	7	exp genetic predisposition to disease/	157,723
	8	Genetic Diseases, Inborn/	14,549
	9	(Genetic* or gene or genes or familial or family or families or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*).ti,ab,kf.	5,104,100
	10	or/4-9 [Genetics/hereditary]	6,534,784
Hemophagocytic lymphohistiocytosis and genetics/hereditary	11	3 and 10	2618
Familial hemophagocytic lymphohistiocytosis	12	((Primary adj4 (hemophagocytic or haemophagocytic or erythrophagocytic)) and (lymphohistiocytosis or lymphocytos* or histiocytos* or reticulos* or hymphohistiocytos* or syndrome*)).ti,ab,kf,rx.	210
	13	(FHLH or PHLH).ti,ab,kf.	99
	14	or/11-13	2709
Excluding animal studies	15	(exp Animals/or exp Models, Animal/or Disease Models, Animal/) not Humans/	5,162,502
	16	14 not 15	2657
English language	17	limit 16 to english language	2395

EMBASE Classic + EMBASE 1947–5 October 2023

1 *hemophagocytic syndrome/ 5628
2 ((Hemophagocytic or haemophagocytic or erythrophagocytic) adj3 (lymphohistiocytos* or lymphocytos* or histiocytos* or reticulos* or hymphohistiocytos* or syndrome*)).ti,ab,kf. 11039
3 or/1-211284
4 *genetics/ 127456
5 *genetic disorder/ 22371
6 *heredity/ 9400
7 *genetic predisposition/ 11661
8 *genetic disorder/ 22371
9 (Genetic* or gene or genes or family or families or familial or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*).ti,ab,kf. 6425199
10 or/4-96458755
11 3 and 10 4095
12 ((Primary adj4 (hemophagocytic or haemophagocytic or erythrophagocytic)) and (lymphohistiocytosis or lymphocytos* or histiocytos* or reticulos* or hymphohistiocytos* or syndrome*)).ti,ab,kf. 400
13 (FHLH or PHLH).ti,ab,kf.232
14 or/11-13 4285
15 (exp animal/or exp juvenile animal/or adult animal/or animal cell/or animal tissue/or nonhuman/or animal experiment/or animal model/) not human/8233442
16 14 not 15 4153
17 limit 16 to english language 3841

Web of Science Science Citation Index Expanded (SCI-EXPANDED)--1970–present, Social Sciences Citation Index (SSCI)–1900–present

1 **TS=((Hemophagocytic or haemophagocytic or erythrophagocytic) NEAR/3 (lymphohistiocytos* or lymphocytos* or histiocytos* or reticulos* or hymphohistiocytos* or syndrome*))8455**
2 TS = (Genetic* or gene or genes or familial or family or families or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*) 6,609,856
3 **#2 AND #1 2725**
4 TS=((Primary NEAR/4 (hemophagocytic or haemophagocytic or erythrophagocytic)) and (lymphohistiocytosis or lymphocytos* or histiocytos* or reticulos* or hymphohistiocytos* or syndrome*))255
5 TS = (FHLH or PHLH) 82
6 #5 OR #4 OR #3 2866
7 #5 OR #4 OR #3 and English (Languages) 2773

Cochrane Library

#1 MeSH descriptor: [Lymphohistiocytosis, Hemophagocytic] this term only14
#2 ((Hemophagocytic or haemophagocytic or erythrophagocytic) near/3 (lymphohistiocytos* or lymphocytos* or histiocytos* or reticulos* or hymphohistiocytos* or syndrome*)):ti,ab,kw 94
#3 #1 or #294
#4 MeSH descriptor: [Genetics] this term only103
#5 MeSH descriptor: [Genetic Predisposition to Disease] explode all trees1602
#6 MeSH descriptor: [Genetic Diseases, Inborn] this term only81
#7 (Genetic* or gene or genes or familial or family or families or DNA or hereditary or heritable or heredodegenerative* or inherit* or congenital* or germline or germinal*):ti,ab,kw 147556
#8 #4 or #5 or #6 or #7147556
#9 #3 and #8 33
#10 ((Primary near/4 (hemophagocytic or haemophagocytic or erythrophagocytic)) and (lymphohistiocytosis or lymphocytos* or histiocytos* or reticulos* or hymphohistiocytos* or syndrome*)):ti,ab,kw6

#11 (FHLH or PHLH):ti,ab,kw3
#12 #9 or #10 or #1137

CENTRAL: 37

Medium-chain acyl-CoA dehydrogenase deficiency

Search summary

Database (platform)	Date searched	Concepts	Hits	Notes
MEDLINE (Ovid)	20 September 2023	MCADD	931	Removed animal studies English-language limit
EMBASE (Ovid)	20 September 2023	MCADD	1750	Removed animal studies No Emtree terms focused English-language limit
SCI (WoS)	20 September 2023	MCADD	1813	No date or language limits.
Cochrane Library	20 September 2023	MCADD	17	No date or language limits.

Total pre-duplication: 4511.

Total after systematically removing duplicates in EndNote (using University of Leeds method):[18] 2962.

Ovid MEDLINE(R) ALL 1946–18 September 2023

MCADD	1	Acyl-CoA Dehydrogenase/and deficien*.ti,ab,kf,rx.	658
	2	Acyl-CoA Dehydrogenase/df, ge, me	550
	3	MCADD.ti,ab,kf.	140
	4	((MCAD or MCADH or MCACA) and deficien*).ti,ab,kf.	368
	5	(medium-chain acyl-CoA dehydrogenase adj2 deficien*).ti,ab,kf,rx.	568
	6	(medium-chain acyl-coenzyme A dehydrogenase adj2 deficien*).ti,ab,kf.	87
	7	(medium chain acyl dehydrogenase adj2 deficien*).ti,ab,kf.	0
	8	MCACA dehydrogenase deficien*.tw,kw.	0
	9	Octanoyl-CoA dehydrogenase deficien*.ti,ab,kf.	0
	10	Octanoyl-coenzyme A dehydrogenase deficien*.ti,ab,kf.	0
	11	or/1-10 [MCADD]	1147
Removing animal studies	12	(exp Animals/or exp Models, Animal/or Disease Models, Animal/) not Human/	5,158,092
	13	11 not 12	986
English language	14	limit 13 to english language	931

EMBASE Classic + EMBASE 1947–19 September 2023

1 acyl coenzyme A dehydrogenase/and deficien*.ti,ab,kf. 799
2 medium chain acyl coenzyme A dehydrogenase deficiency/or medium chain acyl coenzyme A dehydrogenase/ 1395
3 MCADD.ti,ab,kf.267

4 ((MCAD or MCADH or MCACA) and deficien*).ti,ab,kf. 605
5 (medium-chain acyl-CoA dehydrogenase adj2 deficien*).ti,ab,kf. 697
6 (medium-chain acyl-coenzyme A dehydrogenase adj2 deficien*).ti,ab,kf. 99
7 (medium chain acyl dehydrogenase adj2 deficien*).ti,ab,kf.0
8 MCACA dehydrogenase deficien*.tw,kw.0
9 Octanoyl-CoA dehydrogenase deficien*.ti,ab,kf.0
10 Octanoyl-coenzyme A dehydrogenase deficien*.ti,ab,kf.0
11 or/1-10 [MCADD]2372
12 (exp animal/or exp juvenile animal/or adult animal/or animal cell/or animal tissue/or nonhuman/or animal experiment/or animal model/) not human/8221162
13 11 not 12 1841
14 limit 13 to english language 1750

Web of Science – Science Citation Index Expanded (SCI-EXPANDED)–1970–present, Social Sciences Citation Index (SSCI)–1900–present

1 TS=("Acyl-CoA Dehydrogenase" and deficien*) 1652
2 TS = (MCADD) 139
3 TS=((MCAD or MCADH or MCACA) and deficien*) 493
4 TS=("medium-chain acyl-CoA dehydrogenase" near/2 deficien*) 517
5 TS=("medium-chain acyl-coenzyme A dehydrogenase" NEAR/2 deficien*) 79
6 TS=("medium chain acyl dehydrogenase" NEAR/2 deficien*) 0
7 TS="MCACA dehydrogenase deficien*" 0
8 TS="Octanoyl-CoA dehydrogenase deficien*" 0
9 TS="Octanoyl-coenzyme A dehydrogenase deficien*" 0
10 #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 1849
11 Limit to English language 1813

Cochrane Library

#1 MeSH descriptor: [Acyl-CoA Dehydrogenases] explode all trees14
#2 deficien*:ti,ab,kw 33019
#3 #1 and #28
#4 MCADD:ti,ab,kw2
#5 ((MCAD or MCADH or MCACA) and deficien*):ti,ab,kw2
#6 (medium-chain acyl-CoA dehydrogenase near/2 deficien*):ti,ab,kw1
#7 (medium-chain acyl-coenzyme A dehydrogenase near/2 deficien*):ti,ab,kw4
#8 (medium chain acyl dehydrogenase near/2 deficien*):ti,ab,kw 13
#9 MCACA dehydrogenase deficien*:ti,ab,kw0
#10 Octanoyl-CoA dehydrogenase deficien*:ti,ab,kw0
#11 Octanoyl-coenzyme A dehydrogenase deficien*:ti,ab,kw0
#12 #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #1117

CDSR: 0

CENTRAL: 17

Review of genomic studies of newborn cohorts reporting penetrance for pathogenic variants

Search summary

Database (platform)	Date searched	Concepts	Hits	Notes
MEDLINE (Ovid)	18 January 2024	WGS/WES/Penetrance/Actionability/Gene frequency/Sequencing/Allele frequency AND Newborn screening	3019	No date or language limits. Found 5 known studies
EMBASE (Ovid)	19 January 2024	WGS/WES/Penetrance/Actionability/Gene frequency/Sequencing/Allele frequency AND Newborn screening	3764	Removed conference abstracts. No date or language limits. No EMTREE headings focused
SCI (WoS)	19 January 2024	WGS/WES/Penetrance/Actionability/Gene frequency/Sequencing/Allele frequency AND Newborn screening	2076	No date or language limits
Other? Cochrane Library	19 January 2024	WGS/WES/Penetrance/Actionability/Gene frequency/Sequencing/Allele frequency AND Newborn screening	133	No date or language limits

Total from database searches: 8992.

Total after systematically removing duplicates in EndNote (using University of Leeds method):[18] 4970.

MEDLINE

Date searched: 18 January 2024

Database segment: Ovid MEDLINE(R) ALL < 1946–17 January 2024 > https://ovidsp.ovid.com/ovidweb.cgi?T=JS&NEWS= N&PAGE=main&SHAREDSEARCHID=62stKMwQjirbNXTZKB33Ey3QPoKWo Peo9IIbQ7b1QNoyW2UjVAZu6ArPkcdDGqva7

WGS or WES	1	WGS/	9616
	2	Exome Sequencing/	8205
	3	(whole genome sequenc* or complete genome sequenc*).ti,ab,kf.	45,870
	4	WGS.ti,ab,kf.	8462
	5	(whole exome sequenc* or complete exome sequenc*).ti,ab,kf.	19,981
	6	WES.ti,ab,kf.	6937
	7	genomic sequenc*.ti,ab,kf.	20,077
	8	Genomics/	68,324
	9	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 (WGS or WES)	153,084
Newborn screening	10	Neonatal Screening/	12,112
	11	((newborn* or neonat* or infan* or baby or babies) adj5 (screen* or test*)).ti,ab,kf.	36,554
	12	Mass Screening/	117,096

	13	exp infant/or infant disease*.ti,ab,kf. or (babies or baby or infan* or neonat* or neo-nat* or newborn* or new-born*).ti,ab,kf. or infan*.jn,jw. or infan*.in.	1,663,069
	14	12 and 13	10,501
	15	10 or 11 or 14 [Newborn screening]	48,363
Penetrance/ actionability	16	9 and 15 [WGS/WES AND Newborn screening]	528
	17	Penetrance/	2542
	18	penetrance.ti,ab,kf.	14,501
	19	(("proportion of carriers" or "proportion of individuals") adj10 phenotyp*).ti,ab,kf.	22
	20	actionability.ti,ab,kf.	894
	21	17 or 18 or 19 or 20 [Penetrance/Actionability]	16,140
	22	15 and 21 [Penetrance/actionability AND Newborn screening]	73
Sequencing	23	Sequence Analysis/or exp High-Throughput Nucleotide Sequencing/or exp Molecular Sequence Annotation/or exp Nanopore Sequencing/or exp Oligonucleotide Array Sequence Analysis/or exp Position-Specific Scoring Matrices/or exp Sequence Analysis, DNA/	499,723
	24	sequenc*.ti,ab,kf.	1,482,656
	25	23 or 24 [Sequencing]	1,706,387
	26	15 and 25 [Sequencing AND Newborn screening]	2656
Gene frequency	27	exp Gene Frequency/	84,475
	28	((allele* or gene*) adj3 frequenc*).ti,ab,kf.	62,822
	29	27 or 28	123,910
	30	15 and 29 [Gene frequency and Newborn screening]	396
	31	16 or 22 or 26	2765
	32	16 or 22 or 26 or 30	3019

EMBASE

Date searched: 19 January 2024

Database segment: EMBASE Classic + EMBASE < 1947–Week 2, 2024 > https://ovidsp.ovid.com/ovidweb. cgi?T=JS&NEWS=N&PAGE=main&SHAREDSEARCHID=4zKpZuDwhGs6myjL7S8Z46JqaFTuk80 EWrwhr8tsaG6QSMVKpxP5critdRCUddTRd

EMBASE Classic + EMBASE < 1947–Week 2, 2024 >

1 exp whole genome sequencing/49249
2 whole exome sequencing/51415
3 (whole genome sequenc* or complete genome sequenc*).ti,ab,kf. 52855
4 WGS.ti,ab,kf.12372
5 (whole exome sequenc* or complete exome sequenc*).ti,ab,kf. 36037
6 WES.ti,ab,kf.14777
7 genomic sequenc*.ti,ab,kf. 23004
8 genomics/ 90596
9 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 [Whole Genome Sequencing (WGS) or Whole Exome Sequencing (WES)]236593
10 newborn screening/ 23602
11 ((newborn* or neonat* or infan* or baby or babies) adj5 (screen* or test*)).ti,ab,kf. 53683
12 mass screening/ 68006

13 exp infant/ or infant disease*.ti,ab,kf. or (babies or baby or infan* or neonat* or neo-nat* or newborn* or new-born*).ti,ab,kf. or infan*.jn,jx. or infan*.in. 1861950
14 12 and 13 5786
15 10 or 11 or 14 [Newborn screening]63916
16 9 and 15 [WGS/WES AND Newborn screening]1268
17 penetrance/ 13345
18 penetrance.ti,ab,kf. 22184
19 (("proportion of carriers" or "proportion of individuals") adj10 phenotyp*).ti,ab,kf. 31
20 actionability.ti,ab,kf. 1512
21 17 or 18 or 19 or 20 [Penetrance/Actionability]26219
22 15 and 21 [Penetrance/actionability AND Newborn screening]147
23 sequence analysis/or exp bisulfite sequencing/or exp high throughput sequencing/or molecular genetics/or exp DNA microarray/or position weight matrix/or DNA sequencing/or Sanger sequencing/or sequence alignment/ 620523
24 sequenc*.ti,ab,kf. 1771003
25 23 or 24 [Sequencing]2040478
26 15 and 25 [Sequencing AND Newborn screening]4448
27 gene frequency/ 227648
28 ((allele* or gene*) adj3 frequenc*).ti,ab,kf. 86658
29 27 or 28257630
30 15 and 29 [Gene frequency and Newborn screening]1382
31 16 or 22 or 26 4794
32 16 or 22 or 26 or 30 5574
33 limit 32 to (conference abstract or "conference review")1810
34 32 not 33 3764

Science Citation Index (Web of Science)

Date searched: 19 January 2024

Note: search reads from bottom to top.

67	#8 OR #13 OR #15 OR #17	2076
17	#7 AND #16	194
16	TS= ((allele* OR gene*) NEAR/2 frequenc*)	82,970
15	#7 AND #14	1901
14	TS = sequenc*	2,047,808
13	#12 AND #7	62
12	#9 OR #10 OR #11	15,124
11	TS = actionability	881
10	TS= (("proportion of carriers" OR "proportion of individuals") NEAR/10 phenotyp*)	22
9	TS = penetrance	14,255
8	#6 AND #7	318
7	TS= ((newborn* OR neonat* OR infan* OR baby OR babies) NEAR/5 (screen* OR test*))	32,897
6	#1 OR #2 OR #3 OR #4 OR #5	90,839
5	TS="genomic sequenc*"	20,189

4	TS = WES	6697
3	TS= (“whole exome sequenc*” OR “complete exome sequenc*”)	19,578
2	TS = WGS	10,344
1	TS=((“whole genome sequenc*” OR “complete genome sequenc*”))	47,997

Cochrane Library (Wiley)

Date searched: 19 January 2024

IDSearchHits
#1 [mh ^“Whole Genome Sequencing”]60
#2 [mh ^“Exome Sequencing”]49
#3 ((“whole genome” NEXT sequenc*) OR (“complete genome” NEXT sequenc*)):ti,ab,kw 336
#4 WGS:ti,ab,kw143
#5 ((“whole exome” NEXT sequenc*) OR (“complete exome” NEXT sequenc*)):ti,ab,kw 348
#6 WES:ti,ab,kw218
#7 (“genomic” NEXT sequenc*):ti,ab,kw 121
#8 [mh ^Genomics]222
#9 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 1137
#10 [mh ^“Neonatal Screening”]172
#11 ((newborn* OR neonat* OR infan* OR baby OR babies) NEAR/5 (screen* OR test*)):ti,ab,kw 3391
#12 [mh ^“Mass Screening”]4576
#13 [mh infant] OR ((“infant” NEXT disease*) OR babies OR baby OR infan* OR neonat* OR neo-nat* OR newborn* OR new-born*):ti,ab,kw OR infan*:so 94730
#14 #12 AND #13 206
#15 #10 OR #11 OR #14 3553
#16 #9 AND #15 30
#17 [mh ^Penetrance]14
#18 penetrance:ti,ab,kw 130
#19 ((“proportion of carriers” OR “proportion of individuals”) NEAR/10 phenotyp*):ti,ab,kw1
#20 actionability:ti,ab,kw 47
#21 #17 OR #18 OR #19 OR #20177
#22 #15 AND #210
#23 [mh ^“Sequence Analysis”] OR [mh “High-Throughput Nucleotide Sequencing”] OR [mh “Molecular Sequence Annotation”] OR [mh “Nanopore Sequencing”] OR [mh “Oligonucleotide Array Sequence Analysis”] OR [mh “Position-Specific Scoring Matrices”] OR [mh “Sequence Analysis, DNA”] OR [mh ^“Sequence Alignment”]1941
#24 sequenc*:ti,ab,kw 35324
#25 #23 OR #2436273
#26 #15 AND #25 124
#27 [mh “Gene Frequency”]930
#28 ((allele* OR gene*) NEAR/3 frequenc*):ti,ab,kw 3781
#29 #27 OR #28 3785
#30 #15 AND #29 11
#31 #16 OR #22 OR #26125
#32 #16 OR #22 OR #26 OR #30133

Cochrane Database of Systematic Reviews (CDSR):

- Reviews:5
- Protocols:0

Trials (CENTRAL): 128

Review of cost-effectiveness evaluations of whole genome sequencing and whole exome sequencing

Search summary

Database (platform)	Date searched	Concepts	Hits	Notes
MEDLINE (Ovid)	13 February 2024	WGS/WES/ AND Economic evaluations/cost-effectiveness/economic models/	1765	Limited to studies published after 2014 English language
EMBASE (Ovid)	13 February 2024	WGS/WES/ AND Economic evaluations/cost-effectiveness/economic models	934	Limited to studies published after 2014 English language
SCI (WoS)	13 February 2024	WGS/WES/ AND Economic evaluations/cost-effectiveness/economic models	520	Limited to studies published after 2014 English language
EconLit	13 February 2024	WGS/WES/	28	Limited to studies published after 2014
CEA Register	13 February 2024	WGS/WES	14	Limited to studies published after 2014

Search summary

Database (platform)	Date searched	Concepts	Hits	Notes
MEDLINE (Ovid)	13 February 2024	WGS/WES/ AND HTAs	7	Limited to studies published after 2014 English language
EMBASE (Ovid)	13 February 2024	WGS/WES/ AND HTAs	51	Limited to studies published after 2014 English language
International HTA Database – INAHTA	13 February 2024	WGS/WES/	9	Limited to studies published after 2014 English language
Google	13 February 2024	WGS/WES/ AND HTAs	4	Limited to studies published after 2014

Ovid MEDLINE(R) ALL 1946–12 February 2024

WGS	1	WGS/	9658
	2	Exome Sequencing/	8254
	3	Sequence Analysis, DNA/	171,477
	4	Chromatin Immunoprecipitation Sequencing/	865
	5	DNA Barcoding, Taxonomic/	4886
	6	DNA Contamination/	356
	7	DNA Mutational Analysis/	63,004
	8	Multilocus Sequence Typing/	8342
	9	Sequence Analysis/	9771
	10	High-Throughput Nucleotide Sequencing/	47,144
	11	Chromatin Immunoprecipitation Sequencing/	865
	12	Ribosome Profiling/	31
	13	RNA-Seq/	6162
	14	Molecular Sequence Annotation/	12,325
	15	Nanopore Sequencing/	709
	16	Oligonucleotide Array Sequence Analysis/	67,182
	17	position-specific scoring matrices/	643
	18	sequence analysis, dna/	171,477
	19	Sequence Alignment/	100,973
	20	(whole genom* sequenc* or complete genom* sequenc*).ti,ab,kf.	47,688
	21	WGS.ti,ab,kf.	8612
	22	(whole exome sequenc* or complete exome sequenc*).ti,ab,kf.	20,242
	23	WES.ti,ab,kf.	7054
	24	genom* sequenc*.ti,ab,kf.	109,978
	25	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24	553,031

CADTH economic evaluations search filter[185]	26	Economics/	27,523
	27	exp "Costs and Cost Analysis"/	268,748
	28	Economics, Medical/	9269
	29	exp Budgets/	14,190
	30	budget*.ti,kf.	9308
	31	exp models, economic/	16,262
	32	(economic adj2 model*).ti,ab,kf.	6283
	33	(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kf.	291,505
	34	(value adj2 (money or monetary)).ti,ab,kf.	3158
	35	(economic* adj2 (evaluat* or cost* or analys* or impact*)).ab,kf.	45,917
	36	(cost minimi* or cost-utilit* or health utilit* or economic evaluation* or economic review* or cost outcome or cost analys?s or economic analys?s or budget* impact analys?s).ti,ab,kf.	43,856
	37	(cost-effective* or pharmacoeconomic* or pharmaco-economic* or cost-benefit or costs).ti,kf.	91,211
	38	(life year or life years or qaly* or cost-benefit analys?s or cost-effectiveness analys?s).ab,kf.	41,630
	39	(cost or economic*).ti,kf. and (costs or cost-effectiveness or markov).ab.	75,793
	40	(resource* adj2 allocation*).ti,ab,kf.	17,442
	41	economics.fs.	443,078
	42	26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41	729,634
WGS and costs search filter	43	25 and 42	3086
Geonomics/WGS and economics indexing terms	44	Genomics/ec [Economics]	459
	45	WGS/ec [Economics]	60
	46	Exome Sequencing/ec [Economics]	39
	47	44 or 45 or 46	551
WGS and costs search filter or Geonomics/WGS and economics indexing terms	48	43 or 47	3368
Limit to 2014	49	(201409* or 201410* or 201411* or 201412* or 2015* or 2016* or 2017* or 2018* or 2019* or 202*).dt,ez,da.	13,437,301
	50	48 and 49	1792
Limit to English language	51	limit 50 to english language	1765

EMBASE Classic + EMBASE 1947–13 February 2024

1 *whole genome sequencing/9006
2 *whole exome sequencing/7167
3 *chromatin immunoprecipitation sequencing/330
4 *DNA sequencing/1034
5 *DNA Barcoding/3289
6 *DNA contamination/168
7 *dna mutational analysis/218
8 *multilocus sequence typing/1914
9 *sequence analysis/11598
10 *High-Throughput Nucleotide Sequencing/13998
11 *chromatin immunoprecipitation sequencing/330
12 *ribosome profiling/124
13 *RNA sequencing/6223
14 *molecular genetics/11252
15 *nanopore sequencing/1157
16 *DNA microarray/13558
17 *sequence analysis/11598
18 *sequence alignment/2638
19 (whole genom* sequenc* or complete genom* sequenc*).ti,ab,kf. 54750
20 WGS.ti,ab,kf.12508
21 (whole exome sequenc* or complete exome sequenc*).ti,ab,kf. 36334
22 WES.ti,ab,kf.14915
23 genom* sequenc*.ti,ab,kf. 117928
24 or/1-23219852
25 *economics/28751
26 *"cost utility analysis"/3255
27 *"cost benefit analysis"/13607
28 *"cost effectiveness analysis"/39789
29 *budget/ 8147
30 budget*.ti,kf.12761
31 economic model/ 3492
32 (economic adj2 model*).ti,ab,kf. 9085
33 (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).ti,kf. 365423
34 (value adj2 (money or monetary)).ti,ab,kf. 4298
35 (economic* adj2 (evaluat* or cost* or analys* or impact*)).ab,kf. 63922
36 (cost minimi* or cost-utilit* or health utilit* or economic evaluation* or economic review* or cost outcome or cost analys?s or economic analys?s or budget* impact analys?s).ti,ab,kf. 67313
37 (cost-effective* or pharmacoeconomic* or pharmaco-economic* or cost-benefit or costs).ti,kf. 134823
38 (life year or life years or qaly* or cost-benefit analys?s or cost-effectiveness analys?s).ab,kf. 64059
39 (cost or economic*).ti,kf. and (costs or cost-effectiveness or markov).ab. 118459
40 (resource* adj2 allocation*).ti,ab,kf. 21677
41 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40474162
42 24 and 41 1126
43 (201409* or 201410* or 201411* or 201412* or 2015* or 2016* or 2017* or 2018* or 2019* or 202*).dc,dd,dp. 17285388
44 42 and 43 943
45 limit 44 to english language 934

Web of Science Search Strategy

Run 14 Feb 2024

Database: Web of Science Core Collection

- WOS.SCI: 1970–2024
- WOS.AHCI: 1975–2024
- WOS.ESCI: 2015–24
- WOS.ISTP: 1990–2024
- WOS.SSCI: 1900–2024
- WOS.ISSHP: 1990–2024

 # Searches:

1 : (“whole genom* sequenc*” or “complete genom* sequenc*”) (Topic)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:13:16 GMT+0000 (Greenwich Mean Time)Results: 49889
2 : WGS OR WES (Topic)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:18:16 GMT+0000 (Greenwich Mean Time)Results: 16807
3 : TS=(“whole exome sequenc*” or “complete exome sequenc*”)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:21:38 GMT+0000 (Greenwich Mean Time)Results: 19790
4 : TS=(“genom* sequenc*”)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:23:23 GMT+0000 (Greenwich Mean Time)Results: 118797
5 : #4 OR #3 OR #2 OR #1 Editions: WOS.SCI Date Run: Wed Feb 14 2024 09:23:32 GMT+0000 (Greenwich Mean Time)Results: 142621
6 : TS=(economic NEAR/2 model*)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:24:03 GMT+0000 (Greenwich Mean Time)Results: 18189
7 : TI=(economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:24:55 GMT+0000 (Greenwich Mean Time)Results: 459913
8 : AK=((economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic* or pharmaco-economic* or expenditure or expenditures or expense or expenses or financial or finance or finances or financed).)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:25:24 GMT+0000 (Greenwich Mean Time)Results: 247357
9 : TS=(value NEAR/2 (money or monetary))Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:25:50 GMT+0000 (Greenwich Mean Time)Results: 4902
10 : TI=(economic* NEAR/2 (evaluat* or cost* or analys* or impact*))Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:28:19 GMT+0000 (Greenwich Mean Time)Results: 31930
11 : AK=((economic* NEAR/2 (evaluat* or cost* or analys* or impact*)))Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:29:30 GMT+0000 (Greenwich Mean Time)Results: 15466
12 : TS=(“cost minimi*” or “cost-utilit*” or “health utilit*” or “economic evaluation*” or “economic review*” or “cost outcome” or “cost analysis” or “costs analyses” or “economic analysis” or “economic analyses” or “budget* impact analysis” or “budget impact analyses”)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:31:01 GMT+0000 (Greenwich Mean Time)Results: 83111
13 : TI=(“cost-effective*” or pharmacoeconomic* or pharmaco-economic* or “cost-benefit” or costs)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:31:32 GMT+0000 (Greenwich Mean Time)Results: 237689
14 : AK=(“cost-effective*” or pharmacoeconomic* or pharmaco-economic* or “cost-benefit” or costs)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:32:56 GMT+0000 (Greenwich Mean Time)Results: 121134
15 : AB=(“life year” or “life years” or qaly* or “cost-benefit analysis” or “cost benefit analyses” or “cost-effectiveness analysis” or “cost effectiveness analyses”)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:34:30 GMT+0000 (Greenwich Mean Time)Results: 36158

16 : AK=("life year" or "life years" or qaly* or "cost-benefit analysis" or "cost benefit analyses" or "cost-effectiveness analysis" or "cost effectiveness analyses")Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:34:40 GMT+0000 (Greenwich Mean Time)Results: 10634
17 : TI=(cost or economic*)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:35:04 GMT+0000 (Greenwich Mean Time)Results: 362312
18 : AK=(cost or economic*)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:35:17 GMT+0000 (Greenwich Mean Time)Results: 190707
19 : #17 OR #18 Editions: WOS.SCI Date Run: Wed Feb 14 2024 09 : 35 : 25 GMT+0000 (Greenwich Mean Time) Results: 454456
20 : AB=((costs or cost-effectiveness or markov))Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:35:44 GMT+0000 (Greenwich Mean Time)Results: 1372570
21 : #19 AND #20 Editions: WOS.SCI Date Run: Wed Feb 14 2024 09:35:52 GMT+0000 (Greenwich Mean Time) Results: 215564
22 : TS=(resource* NEAR/2 allocation*)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:36:39 GMT+0000 (Greenwich Mean Time)Results: 52120
23 : #22 OR #16 OR #21 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #6 OR #7 OR #8 Editions: WOS.SCI Date Run: Wed Feb 14 2024 09:37:14 GMT+0000 (Greenwich Mean Time)Results: 661626
24 : #23 AND #5 Editions: WOS.SCI Date Run: Wed Feb 14 2024 09:37:22 GMT+0000 (Greenwich Mean Time)Results: 606
25 : #23 AND #5 and 2014 or 2015 or 2017 or 2018 or 2016 or 2019 or 2020 or 2021 or 2022 or 2023 or 2024 (Publication Years)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:37:46 GMT+0000 (Greenwich Mean Time) Results: 522
26 : #23 AND #5 and 2014 or 2015 or 2017 or 2018 or 2016 or 2019 or 2020 or 2021 or 2022 or 2023 or 2024 (Publication Years) and English (Languages)Editions: WOS.SCIDate Run: Wed Feb 14 2024 09:37:53 GMT+0000 (Greenwich Mean Time)Results: 520

EconLit

S1	TI (("whole genom* sequenc*" or "complete genom* sequenc*")) OR AB (("whole genom* sequenc*" or "complete genom* sequenc*")) OR SU (("whole genom* sequenc*" or "complete genom* sequenc*"))	10
S2	TI ((WGS OR WES)) OR AB ((WGS OR WES)) OR SU ((WGS OR WES))	58
S3	TI (("whole exome sequenc*" or "complete exome sequenc*")) OR AB (("whole exome sequenc*" or "complete exome sequenc*")) OR SU (("whole exome sequenc*" or "complete exome sequenc*"))	2
S4	TI "genom* sequenc*" OR AB "genom* sequenc*" OR SU "genom* sequenc*"	17
S5	S1 OR S2 OR S3 OR S4	71
S6	S1 OR S2 OR S3 OR S4 Limiters – Publication Date: 20140101-20231231	28

CEA registry

Title, Abstract or Keyword genomic sequencing 3 results
Title, Abstract or Keyword genome 7 results
Title, Abstract or Keyword exome sequencing 3 results
Title, Abstract or Keyword WGS 0 results
Title, Abstract or Keyword WES 1 result

HTA search

Ovid MEDLINE(R) ALL 1946–13 February 2024

1 Whole Genome Sequencing/9662

2 Exome Sequencing/8258
3 Sequence Analysis, DNA/171483
4 Chromatin Immunoprecipitation Sequencing/866
5 DNA Barcoding, Taxonomic/4888
6 DNA Contamination/356
7 DNA Mutational Analysis/63005
8 Multilocus Sequence Typing/8343
9 Sequence Analysis/9772
10 High-Throughput Nucleotide Sequencing/47158
11 Chromatin Immunoprecipitation Sequencing/866
12 Ribosome Profiling/31
13 RNA-Seq/6162
14 Molecular Sequence Annotation/12326
15 Nanopore Sequencing/709
16 Oligonucleotide Array Sequence Analysis/67183
17 position-specific scoring matrices/643
18 sequence analysis, dna/171483
19 Sequence Alignment/100974
20 (whole genom* sequenc* or complete genom* sequenc*).ti,ab,kf. 47719
21 WGS.ti,ab,kf. 8618
22 (whole exome sequenc* or complete exome sequenc*).ti,ab,kf. 20252
23 WES.ti,ab,kf. 7063
24 genom* sequenc*.ti,ab,kf. 110021
25 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24553104
26 exp Technology Assessment, Biomedical/12286
27 (technology assessment* or HTA or HTAs or technology overview* or technology appraisal*).ti,ab,kf. 11825
28 (biomedical technology assessment* or bio-medical technology assessment*).mp,hw. 27
29 (health adj2 technology assessment).jw. 4924
30 (cochrane or (health adj2 technology assessment) or evidence report).jw. 21864
31 26 or 27 or 28 or 29 or 3038796
32 25 and 31 115
33 (2022* or 2023* or 2024*).dt,ez,da. 3788244
34 32 and 337

EMBASE Classic + EMBASE 1947–13 February 2024

1 whole genome sequencing/48989
2 whole exome sequencing/52032
3 chromatin immunoprecipitation sequencing/5462
4 DNA sequencing/18044
5 DNA Barcoding/8198
6 DNA contamination/1086
7 dna mutational analysis/2196
8 multilocus sequence typing/17016
9 sequence analysis/190474
10 High-Throughput Nucleotide Sequencing/97146
11 chromatin immunoprecipitation sequencing/5462
12 ribosome profiling/345
13 RNA sequencing/85069
14 molecular genetics/159939
15 nanopore sequencing/3027
16 DNA microarray/67752

17 sequence analysis/190474
18 sequence alignment/73757
19 (whole genom* sequenc* or complete genom* sequenc*).ti,ab,kf. 54750
20 WGS.ti,ab,kf.12508
21 (whole exome sequenc* or complete exome sequenc*).ti,ab,kf. 36334
22 WES.ti,ab,kf.14915
23 genom* sequenc*.ti,ab,kf. 117928
24 or/1-23801244
25 biomedical technology assessment/17793
26 (technology assessment* or HTA or HTAs or technology overview* or technology appraisal*).ti,ab,kf. 19704
27 (biomedical technology assessment* or bio-medical technology assessment*).mp,hw. 17819
28 (health adj2 technology assessment).jw. 6580
29 (cochrane or (health adj2 technology assessment) or evidence report).jw. 31781
30 25 or 26 or 27 or 28 or 2959598
31 (2022* or 2023* or 2024*).dc,dd,dp. 4656130
32 24 and 30 and 31 52
33 limit 32 to english language 51

International HTA Database – INAHTA

("whole genom* sequenc*" or "complete genom* sequenc*" OR WGS OR WES OR "whole exome sequenc*" or "complete exome sequenc*" or "genom* sequenc*") 9 results – limited to 2022–2024 0 results

Google

"Health technology assessment" "whole genome sequencing" 2 results

Appendix 3 Quality appraisal

Quality appraisal for the review of five conditions, including details of tailoring of Murad et al.'s[22] quality appraisal tool for the assessment of case series and case reports for Q2 (studies exploring the prevalence of different genetic variants) and Q4 (studies exploring the impact of earlier vs. later treatment), the ROBIS-219 tool used to appraise one systematic review exploring the prevalence of different genetic variants in fHLH and assessment of bias and applicability of included studies according to these tools.

Gene/variant frequency in patients with the condition(s) of interest

Subset of items from Murad *et al.*'s[22] quality appraisal tool for the assessment of case series and case reports used for Q2 (studies exploring the prevalence of different genetic variants)

Quality appraisal tables

The majority of studies were appraised using the tailored Murad *et al.*[22] tool; one[55] was a systematic review and was assessed with the ROBIS-2 tool[19] and is presented below in a separate table.

Evidence on early versus late treatment

Subset of items from Murad *et al.*'s [22] quality appraisal tool for the assessment of case series and case reports tailored for Q4 (studies exploring the impact of earlier vs. later treatment).

TABLE 16 Tailoring of Murad's quality appraisal tool for Q2 (studies exploring the prevalence of genetic variants in children with disease)

Domain	Questions	Rating
Selection	1a. Does the patient(s) represent(s) the whole experience of the investigator (centre) or is the selection method unclear to the extent that other patients with similar presentation may not have been reported?	Yes – all subjects/consecutive subjects/random subjects over a certain time period who meet inclusion criteria are included (registry studies should include all patients in the respective country/countries, for 'yes', information on adequate completeness should be provided) No – no information/unclear selection of subjects included in the study/only proportion of eligible patients had sequence data
	1a (comment). Comment if needed (e.g. justification for 'No' response)	N/A
	1b. Were the criteria for selection for sequencing clear and appropriate (e.g. not based on disease severity)?	Yes – all eligible patients (refer to question 1a) were sequenced, or selection of eligible patients who were sequenced was random/not selective No – not all eligible patients were sequenced and patients who were sequenced were selected inappropriately Unclear – not all eligible patients were sequenced and it is not reported how patients who were sequenced were selected (appropriateness unclear)
	1b (comment). Comment if needed (e.g. justification for 'No' response)	NA
Ascertainment	2a. Were cases (i.e. presence of clinical features and/or biochemical features of the relevant condition) adequately ascertained?	Yes – clinical or biochemical diagnosis of condition ascertained from clinical records/clinical assessment No – self-reporting/parent-reporting Unclear – not reported
	2b. Were cases (i.e. presence of clinical features and/or biochemical features of the relevant condition) adequately defined?	Yes – unambiguous label of condition/group of conditions with clear clinical and/or biochemical definitions No – some ambiguity in label of condition/group of conditions and/or no clear clinical and/or biochemical definitions (Ines to think about whether there is inconsistency in using terms in fHLH, inability to clearly categorise conditions based on symptoms/biochemical test is actually a risk of bias or just reporting)
	3a. Was the presence of variants (genetic make-up) adequately ascertained?	Yes – genetic make-up ascertained from clinical records (retrospective) or genetic testing (prospective) No – outcomes are self-reported/parent-reported Unclear – not reported
Reporting	4. Is the case(s) described with sufficient details to allow other investigators to replicate the research or to allow practitioners make inferences related to their own practice?	Yes – definition of disease (clinical and/or biochemical)/eligibility criteria/extent of sequencing (genes/variants/number of exons …) specified/definition of genetic disease (annotation/calling or prioritisation rules) (all present)
Applicability	5. Is the reported genetic spectrum based only on sequencing techniques or were additional tests needed to fully identify variants [i.e. are the results applicable to the review question (WGS)]?	Yes – genetic make-up reported only based on sequencing based techniques No – sequencing was followed by additional methods to identify, for instance, deletions/insertions Unclear – type of genetic test not reported (e.g. registry studies) or multiple types of genetic test were used, and it was unclear which variants were identified by sequencing

TABLE 17 Quality appraisal of studies exploring the gene/variant frequency in patients with the condition(s) of interest using the tailored Murad's tool

		Selection				Ascertainment			Reporting	Applicability
Condition	Study	1a	1a (comment)	1b	1b (comment)	2a	2b	3a	4	5
PDE	Coughlin 2019[32]	No	Time period not reported; registry study	No	No information about tests was reported	Yes	Yes	Yes	No	Unclear
PDE	Jiao 2020[33]	No	No information about participant selection criteria	No	Basis for selection for testing was not reported	Yes	Yes	Yes	No	Yes
PDE	Koul 2019[31]	Yes		No	'Target mutations testing in *ALDH7A1* was done in all the cases', but no further detail. Also, it is unclear if all received this testing: 'four were confirmed on genetic testing; eight were their siblings'	Yes	Yes	Yes	No	Unclear
PDE	Boonsimma 2023[30]	Yes		Yes		Yes	Yes	Yes	Yes	Yes
hRB	Salviat 2020[39]	Yes		Yes		Yes	Yes	No	No	Unclear
hRB	Hulsenbeck 2021[41]	Yes		Yes		Yes	Yes	No	No	Unclear
hRB	Temming 2013[40]	Yes		Yes		Yes	Yes	No	No	Unclear
hRB	Zhang 2022[186]	No	Unclear how many patients/ families are actually included and how they have been selected	No	Criteria for selection not reported (five chosen for further mosaicism testing with no explanation other than that they agreed to participate in mosaicism research)	Yes	Yes	Yes	No	No
XLHR	Marik 2022[47]	Yes		Yes		Yes	Yes	Yes	Yes	Yes
XLHR	Ariceta 2023[51]	Yes		No	282/360 children had genetic test results available	Yes	Yes	Yes	No	Unclear
XLHR	Gaucher 2009[48]	No	No clear indication of the resemblance of all eligible patients	Yes		Yes	Yes	Yes	No	Yes
XLHR	Del Pino 2022[49]	Yes		No	42/96 underwent molecular testing	Yes	Yes	Yes	No	No
XLHR	Rafaelsen 2016[50]	Yes		Yes		Unclear	No	Yes	No	No
XLHR	Jacob 2023[46]	No	Unclear whether included participants resembled all eligible patients	Yes		Yes	Yes	Yes	No	Yes

continued

TABLE 17 Quality appraisal of studies exploring the gene/variant frequency in patients with the condition(s) of interest using the tailored Murad's tool (*continued*)

		Selection				Ascertainment			Reporting	Applicability
Condition	Study	1a	1a (comment)	1b	1b (comment)	2a	2b	3a	4	5
fHLH	Pagel 2012[56]	No	Patients from national HLH reference centre in Germany + some cases sent in from various other countries requested from members of histiocyte society – did not report the response rate	Yes		Yes	Yes	Yes	Yes	Yes
fHLH	Shabrish 2021[54]	No	Patients have been included from 20 centres and no clear details on whether anyone has been excluded	Yes		Yes	Yes	Yes	Yes	Yes
fHLH	Trizzino 2008[57]	No	Members of histiocyte society invited to contribute their patients	Yes		Yes	Yes	Yes	Yes	Yes
fHLH	Cetica 2016[53]	No	Exact time period not reported	Yes		Yes	Yes	Yes	Yes	Yes
MCADD	Martin-Rivada 2022[59]	Yes		Yes		Yes	Yes	Yes	Yes	Unclear
MCADD	Wang 2019[61]	Yes		Yes		Yes	Yes	Yes	Yes	Yes
MCADD	Touw 2012[64]	Yes		No	68/84 patients had the sequencing information	Yes	Yes	Yes	Yes	Yes
MCADD	Maguolo 2020[60]	Yes		Yes		Yes	Yes	Yes	Yes	Yes
MCADD	Mesbah 2019[63]	Yes		Yes		Yes	Yes	Yes	No	Unclear
MCADD	Nichols 2008[62]	Yes		Yes		Yes	Yes	Yes	Yes	Yes

TABLE 18 Quality appraisal of one systematic review exploring the gene/variant frequency in patients with the condition(s) of interest using the ROBIS-2 tool

Condition: fHLH **Study: Amirifar *et al.*[55]**		
PHASE 1: ASSESSING RELEVANCE (optional) **Not applicable in this review**		
PHASE 2: IDENTIFYING CONCERNS WITH THE REVIEW PROCESS		
DOMAIN 1: STUDY ELIGIBILITY CRITERIA		
1.1 Did the review adhere to pre-defined objectives and eligibility criteria?	Probably yes	No protocol to check
1.2 Were the eligibility criteria appropriate for the review question?	Yes	
1.3 Were eligibility criteria unambiguous?	Yes	
1.4 Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality and outcomes measured)?	Yes	Selection criterion on study design is very broad, so no risk of bias by excluding certain study types
1.5 Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language and availability of data)?	Yes	
Concerns regarding specification of study eligibility criteria	**Low**	
DOMAIN 2: IDENTIFICATION AND SELECTION OF STUDIES		
2.1 Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Yes	
2.2 Were methods additional to database searching used to identify relevant reports?	Yes	
2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Probably yes	
2.4 Were restrictions based on date, publication format or appropriate language?	Probably yes	
2.5 Were efforts made to minimise error in selection of studies?	No information	No details on how many people were sifted and how discrepancies were discussed
Concerns regarding methods used to identify and/or select studies	**Low**	
DOMAIN 3: DATA COLLECTION AND STUDY APPRAISAL		
3.1 Were efforts made to minimise error in data collection?	Probably yes	No information on how data extraction form was developed and whether it was piloted
3.2 Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Probably yes	
3.3 Were all relevant study results collected for use in the synthesis?	Yes	
3.4 Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	No information	
3.5 Were efforts made to minimise error in risk of bias assessment?	No information	

continued

TABLE 18 Quality appraisal of one systematic review exploring the gene/variant frequency in patients with the condition(s) of interest using the ROBIS-2 tool (*continued*)

Condition: fHLH **Study: Amirifar *et al.*[55]**		
Concerns regarding methods used to collect data and appraise studies	**Unclear**	**Rationale: no information on the risk of bias assessment**
DOMAIN 4: SYNTHESIS AND FINDINGS		
4.1 Did the synthesis include all studies that it should?	Probably yes	
4.2 Were all pre-defined analyses reported or departures explained?	No information	No protocol to check
4.3 Was the synthesis appropriate given the nature and similarity in the RQs, study designs and outcomes across included studies?	No information	No information on individual studies and no justification on methods of synthesis provided: 'Data were combined without being weighted, and the analysis was done as if the data were derived from a single sample' but probably ok for the simple analysis of frequencies of types of variants
4.4 Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	No information	No information on individual included studies, so it is impossible to judge the level of heterogeneity and the type of adjustment that would have been appropriate
4.5 Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	No information	Probably not appropriate for this study
4.6 Were biases in primary studies minimal or addressed in the synthesis?	No information	
Concerns regarding the synthesis and findings	**Unclear**	**Rationale: no information on heterogeneity of studies and the synthesis method is not justified, no information provided on included study characteristics, list of included studies or a protocol**
PHASE 3: JUDGING RISK OF BIAS		
1. Concerns regarding specification of study eligibility criteria	Low	Rationale: N/A
2. Concerns regarding methods used to identify and/or select studies	Low	Rationale: N/A
3. Concerns regarding methods used to collect data and appraise studies	Unclear	Rationale: no information about the risk of bias assessment
4. Concerns regarding the synthesis and findings	Unclear	Rationale: no information on heterogeneity of studies and the synthesis method is not justified, no information provided on included study characteristics, list of included studies or a protocol
RISK OF BIAS IN THE REVIEW		
A. Did the interpretation of findings address all of the concerns identified in domains 1–4?	No	
B. Was the relevance of identified studies to the review's RQ appropriately considered?	Probably yes	
C. Did the reviewers avoid emphasizing results on the basis of their statistical significance?	Probably no	
Risk of bias in the review	**Unclear**	**Rationale: the study did not provide sufficient detail on the conduct of the review, the include studies and the synthesis of the studies to make an informed assessment. No protocol is appended or referenced**

TABLE 19 Tailoring of Murad's quality appraisal tool for Q4 (studies exploring the benefit of earlier vs. later treatment)

Domain	Questions	Rating
Selection	1. **Does the patient(s) represent(s) the whole experience of the investigator (centre) or is the selection method unclear to the extent that other patients with similar presentation may not have been reported?**	Yes – all subjects/consecutive subjects/random subjects over a certain time period who meet inclusive criteria are included No – no information/unclear selection of subjects included in the study
Ascertainment	2. **Was the exposure (Rx delivery) adequately ascertained?**	Yes – treatment/management received ascertained from clinical records, no problems with compliance No – self-reporting/parent reporting/poor compliance Unclear – not reported
	3. **Was the outcome adequately ascertained?**	Yes – outcome ascertained from clinical records (retrospective) No – outcomes are self-reported/parent reported Unclear – not reported
Causality	4. **Was follow-up long enough for outcomes to occur?**	Separately for each outcome For instance, for yes fHLH – relapse after (time) HSCT fHLH – mortality after (time) HSCT XLHR (Pi/D) – growth => 6 months to assess height velocity XLHR (Pi/D) – biochemical findings time => 3–4 months XLHR (Pi/D) – radiological findings time (activity score) => 6 months XLHR (Pi/D) – nephrocalcinosis => 1 year hRB – detection of RB; FU at least 2 year (mean age at onset is 15 months) PDE – sufficient time for developmental outcomes to become apparent; FU of at least a few years MCADD (dietary advice) – decompensation event/hospitalisation MCADD (dietary advice) – mortality time? otherwise no
Reporting	5. **Is the case(s) described with sufficient details to allow other investigators to replicate the research or to allow practitioners make inferences related to their own practice?**	Yes – definition of disease/eligibility criteria/definition of treatment/definition of early vs. late/outcome measures and time points (all present) No – NR

FU, follow-up.

Quality appraisal table

TABLE 20 Quality appraisal of studies exploring early vs. late treatment in patients with the condition(s) of interest using the tailored Murad *et al.* tool

		Selection		Ascertainment		Causality		Reporting
Condition	Study	1	1 (comment)	2	3	4	4 (comment)	5
PDE	Tseng 2022[79]	No	Does not state 'all' sibling pairs were identified; no dates were reported	Yes	Yes	Yes	No specific FU period stated, but registry is updated annually and age at last evaluation was 13.23 years (± 7.53) (for early-treated sibling)	Yes
PDE	Coughlin 2022[68]	Yes		Yes	Yes	Yes	No specific FU period stated, but registry is updated annually and mean age at evaluation was 7.68 years	Yes
PDE	Jiao 2021[77]	No	No details	Yes	Yes	No	Mostly ok, but one 'early' patient aged only 6 months at last FU	No
PDE	Strijker 2021[78]	No	Timing not clearly reported	Yes	Yes	No	Timing NR	Yes
PDE	Bok 2010[76]	No	Only reports two families of three in the Dutch PDE cohort	Yes	Yes	Yes	Family A evaluated at 4 years and family B evaluated at 12 years	Yes
hRB	Abramson 2003[80]	Yes		Yes	Yes	Yes	Mean FU 109 months (9 years)	No
hRB	Chantada 2009[81]	Yes		Yes	Yes	Yes	Median FU 79 months	No
hRB	Moll 2000[82]	Yes		Yes	Yes	No	FU length NS	No
hRB	Rothschild 2011[83]	Yes		Yes	Yes	Yes	Median 107 months	Yes
hRB	Soliman 2017[84]	Yes		Yes	Yes	Yes	Median 4.8 years	No
XLHR	Makitie 2003[88]	Yes		Yes	Yes	Yes	Height: at the end of the first treatment year, pre-puberty (9 years) and predicted adult height S-Pi, S-ALP and severity of rickets: at the end of the first treatment year and at pre-puberty (median 10.8, NR and 10.4 years, respectively)	Yes
XLHR	Quinlan 2012[89]	No	Children with XLHR followed at Great Ormond Street Hospital between 1990 and 2006; 46/61 HR patients were PHEX-sequenced, and 6/23 with PHEX variant had unavailable growth data because treatment commenced in different hospitals	Yes	Yes	Yes	For most recent data available (not adult height): at medium treatment years of 8.5 vs. 11.9 (p = 0.557) (IQR: 4.0–15.2 vs. 6.2–14.3) Biochemical markers: throughout treatment Nephrocalcinosis: higher in those treated for longer, not reported comparatively for groups	Yes

TABLE 20 Quality appraisal of studies exploring early vs. late treatment in patients with the condition(s) of interest using the tailored Murad *et al.* tool (*continued*)

		Selection		Ascertainment		Causality		Reporting
Condition	Study	1	1 (comment)	2	3	4	4 (comment)	5
XLHR	Rafaelsen 2016[50]	Yes		Yes	Yes	Unclear	At last registered consultation, *Figure 1* suggests that final adult height was not reached by most, number of patients were only followed-up for very short term, there was no minimum treatment duration as inclusion criterion, no median (range, IQR) was reported for the last registered consultation	Yes
fHLH	Luccinic 2018[75]	No	Survey was sent to key physician in each of the paediatric HSCT centres and they were asked to collect and report data on index cases and subsequent asymptomatic family member(s) with primary HLH. Not clear if all patients in each centre would be reported (and no way to ascertain this)	Yes	Yes	Yes	Median 41 months for patients symptomatic at diagnosis, 41.5 for patients asymptomatic at diagnosis. Expert suggested minimum is 24 months	Yes
MCADD	Abdenur 1999[92]	No	Two cases only, no information about how they were selected	Yes	Yes	Yes	Follow-up time 18–19 months	Yes
MCADD	Alcaide 2022[93]	No	Unclear if all positive cases detected by NBS test are included in this analysis. Unclear whether all sibling-detected cases are included (there is only one reported here)	Yes	Yes	Unclear	Unknown follow-up time	Yes
MCADD	Gong 2021[95]	No	Not clear whether these represent all the MCADD patients from the study centre	Yes	Yes	Yes	Variable follow-up length, follow-up until study close (December 2019) Unclear when the post-treatment data were collected	No
MCADD	Haas 2007[96]	Yes		No	Yes	Yes	Evaluated service use in the first 4 years of life	No
MCADD	Li 2019[102]	No	Unclear if all MCADD cases from that centre are included	Yes	Yes	Yes	Follow-up time 3–11 years	No
MCADD	Wilcken 2007[98]	Yes		No	Yes	Yes	To age 6 years	No
MCADD	Wilcken 2009[99]	Yes		No	Yes	Yes	Between 2 and 4 years	Yes
MCADD	Wilson 1999[100]	Yes		Yes	Yes	Yes	10 months–14 years, median 6 years	Yes

FU, follow-up; IQR, interquartile range; NS, not specified; S-ALP, serum alkaline phosphate; S-Pi, serum phosphate.

Appendix 4 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow charts

Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow charts illustrating the selection process of studies for the review of five conditions, the review of genomic studies of newborn cohorts reporting penetrance for pathogenic variants and the review of cost-effectiveness evaluations of WGS and WES.

Pyridoxine-dependent epilepsy

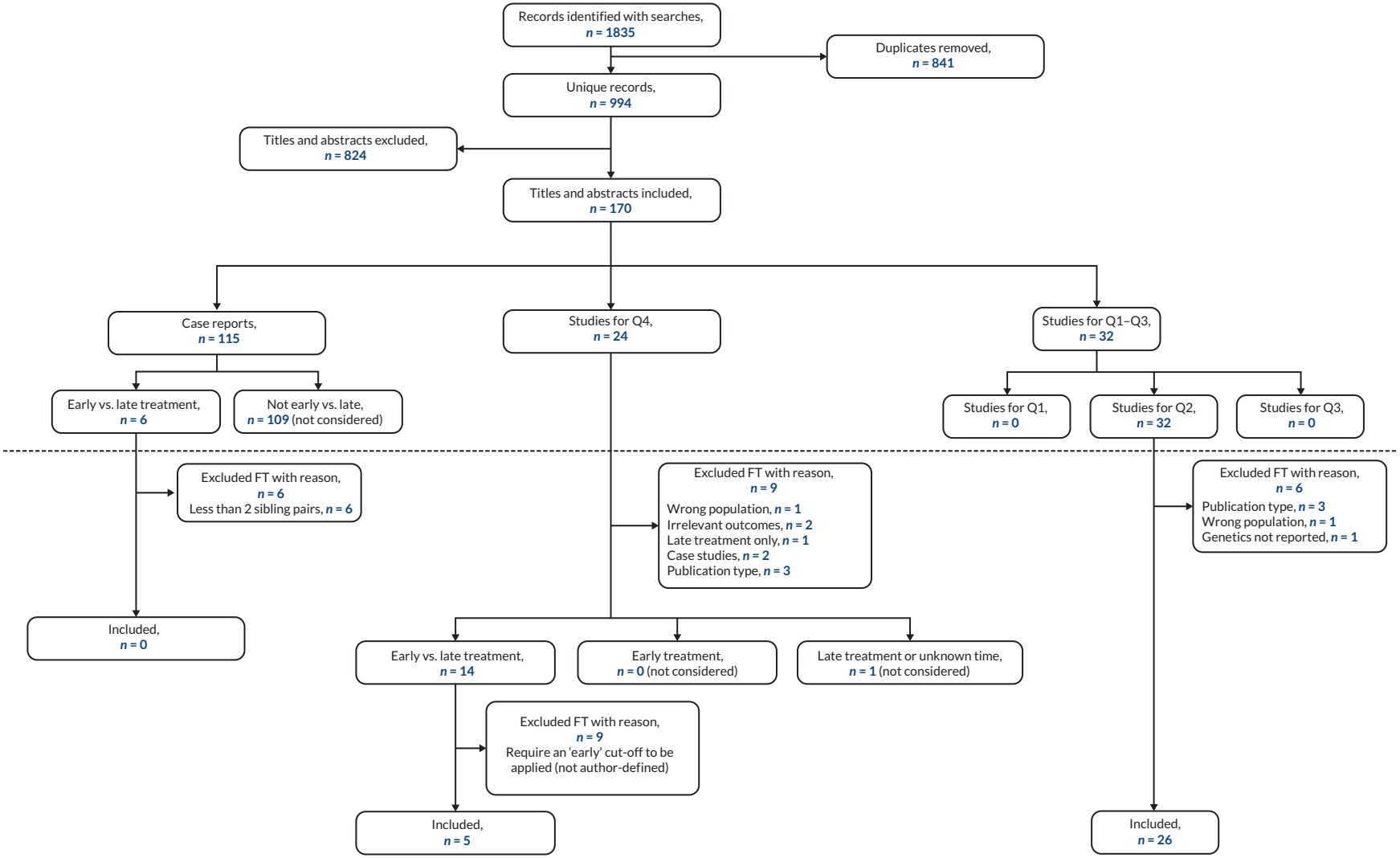

FIGURE 2 The PRIMA flow chart for the review of PDE. FT, full texts.

DOI: 10.3310/DJRF1124

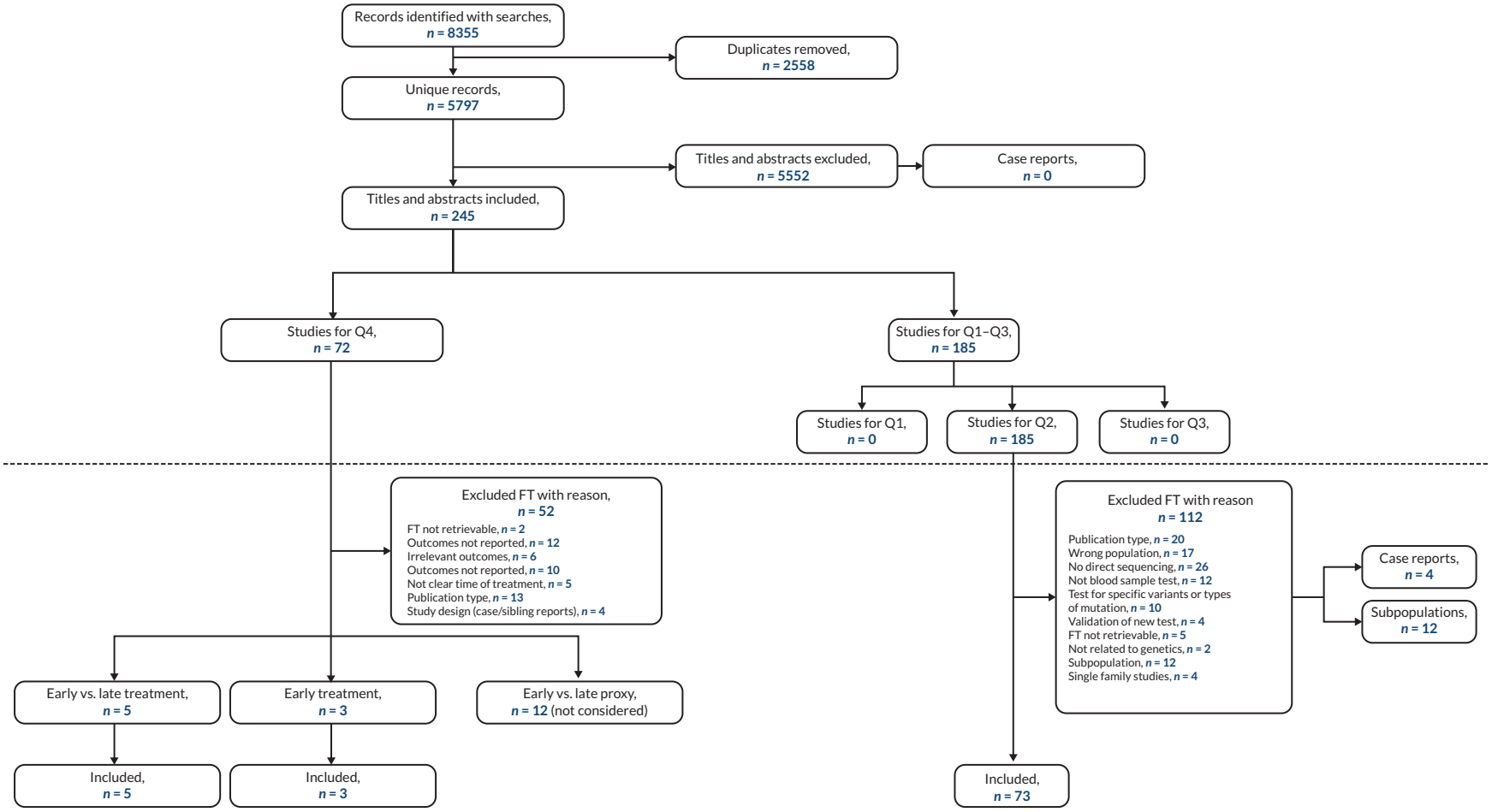

FIGURE 3 The PRISMA flow chart for the review of hRB.

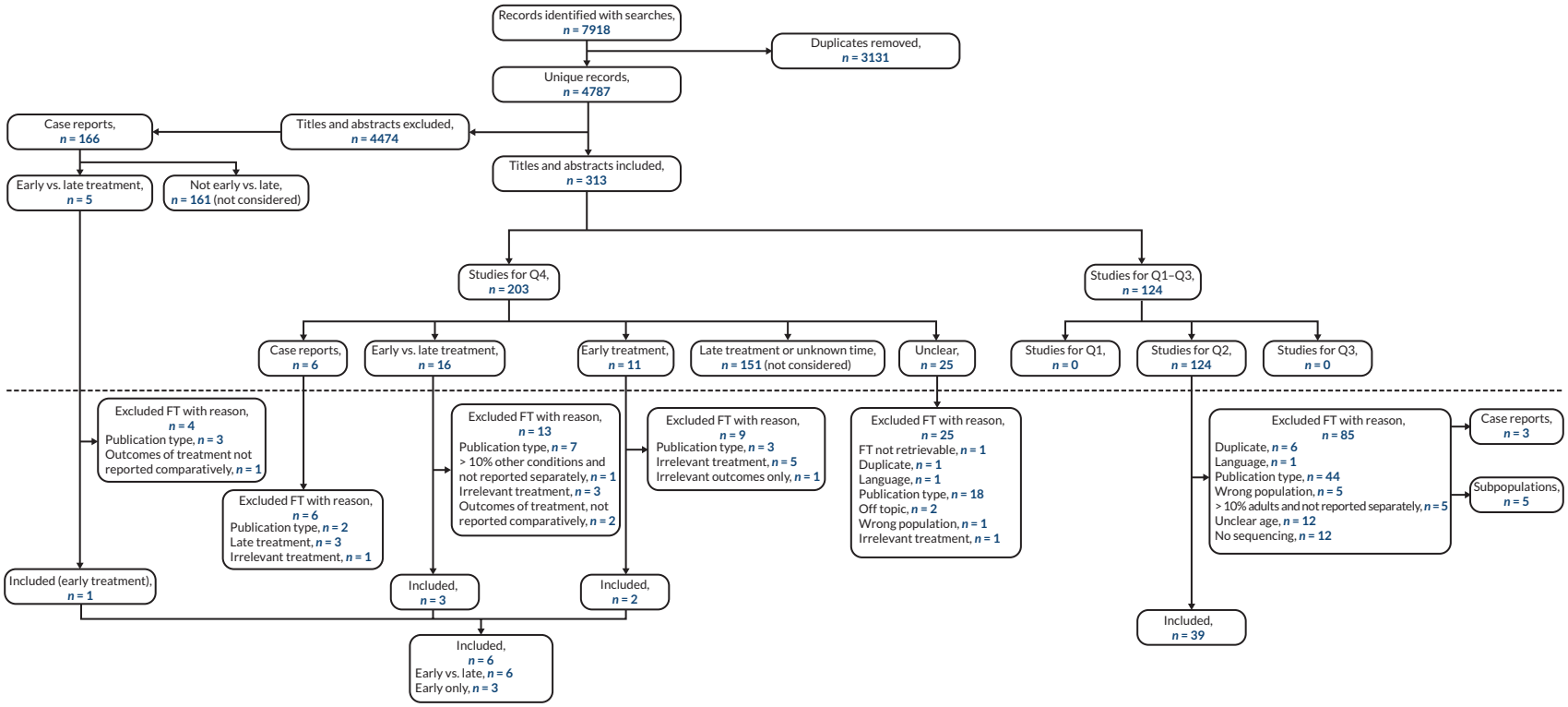

FIGURE 4 The PRISMA flow chart for the review of XLHR.

DOI: 10.3310/DJRF1124

Familial haemophagocytic lymphohistiocytosis

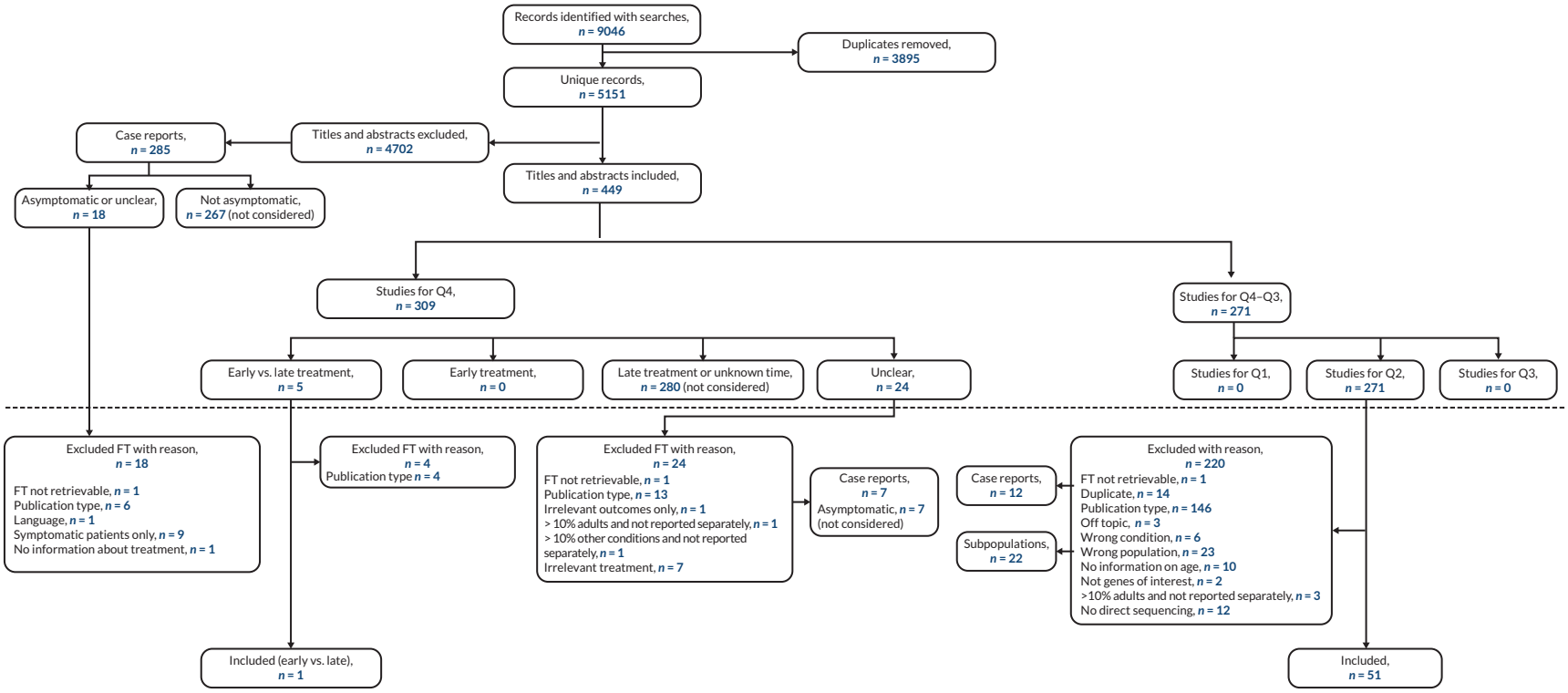

FIGURE 5 The PRISMA flow chart for the review of fHLH.

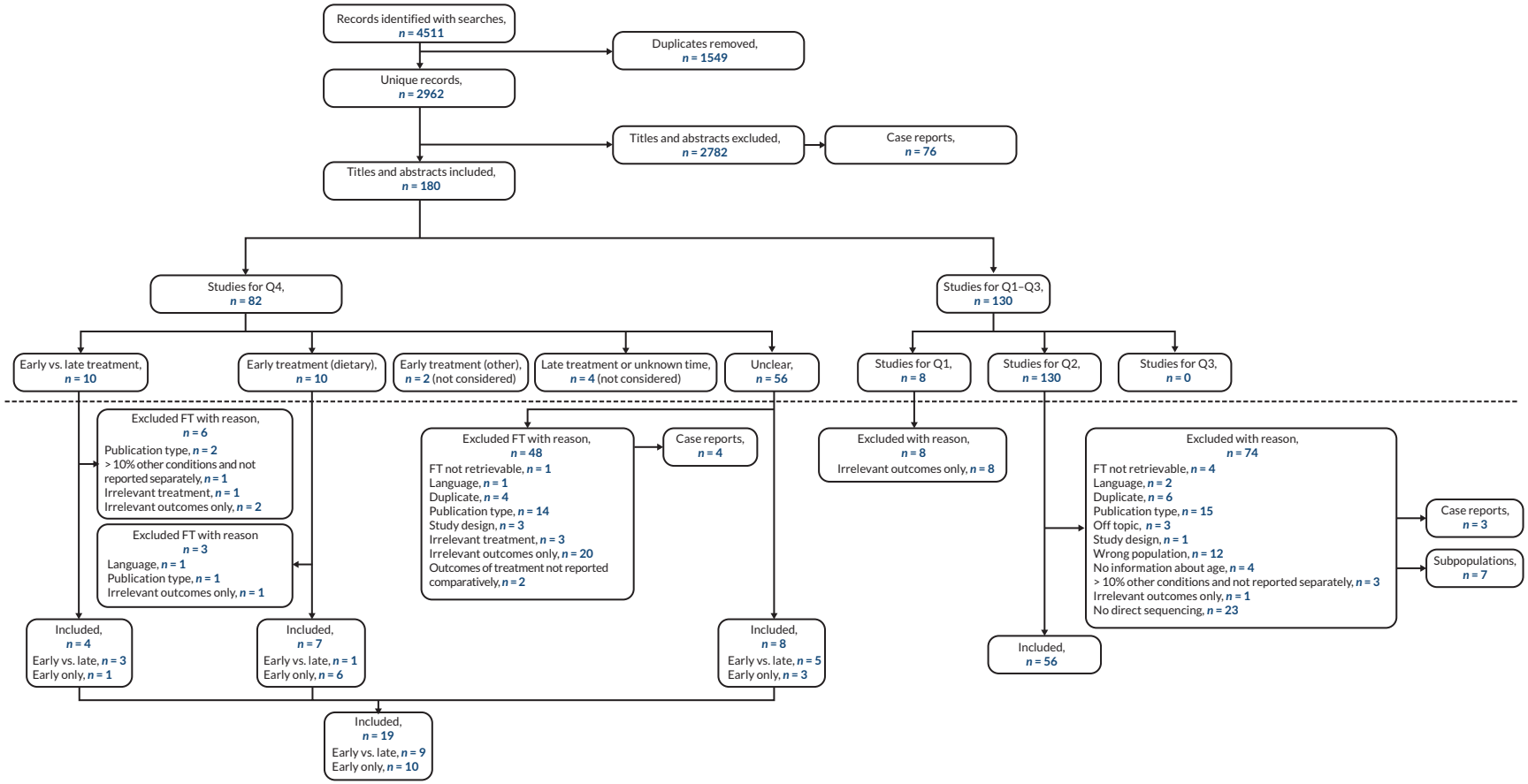

FIGURE 6 The PRISMA flow chart for the review of MCADD.

DOI: 10.3310/DJRF1124

Review of genomic studies of newborn cohorts reporting penetrance for pathogenic variants

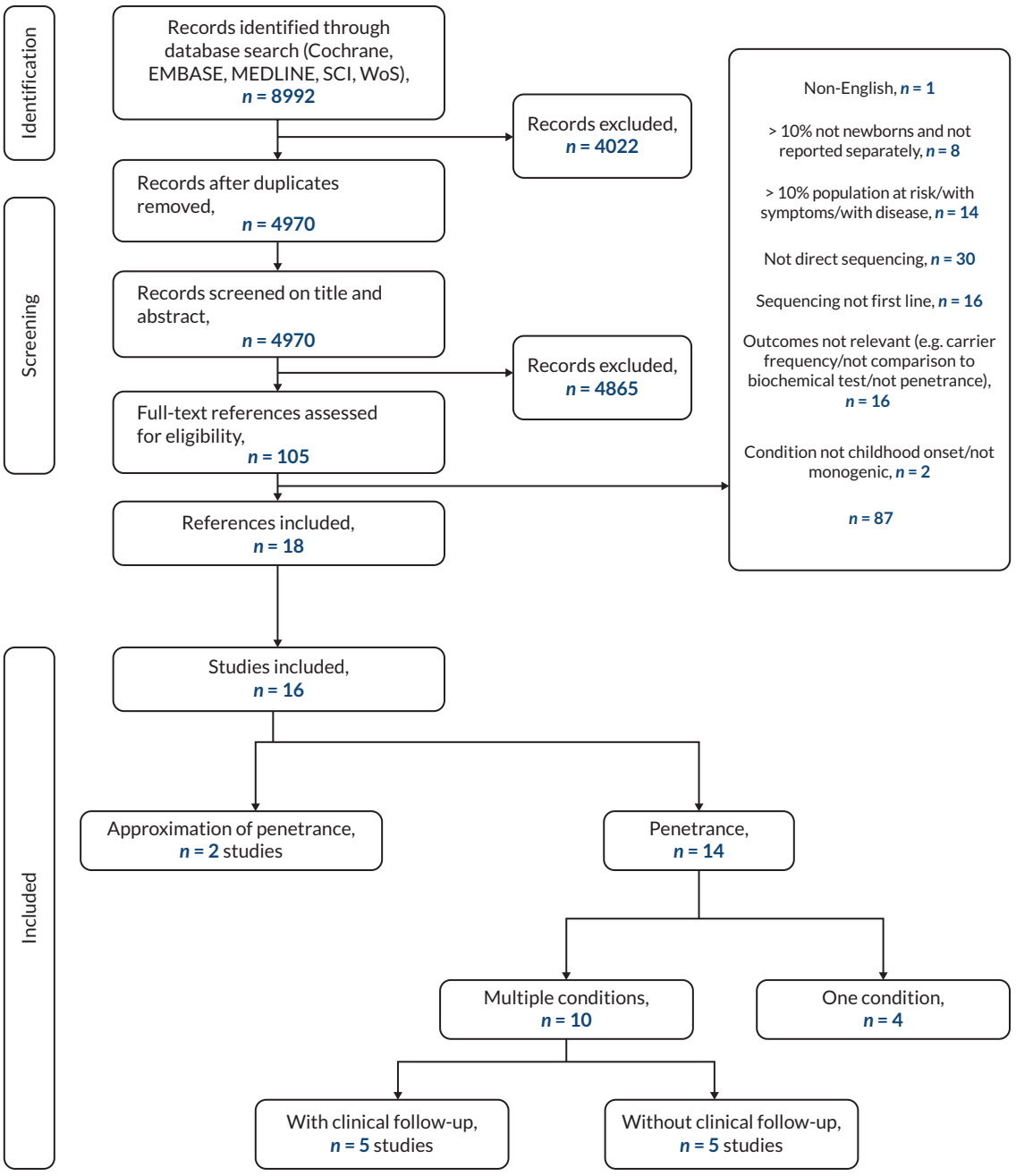

FIGURE 7 The PRISMA flow chart for the review of genomic studies of newborn cohorts reporting penetrance for pathogenic variants.

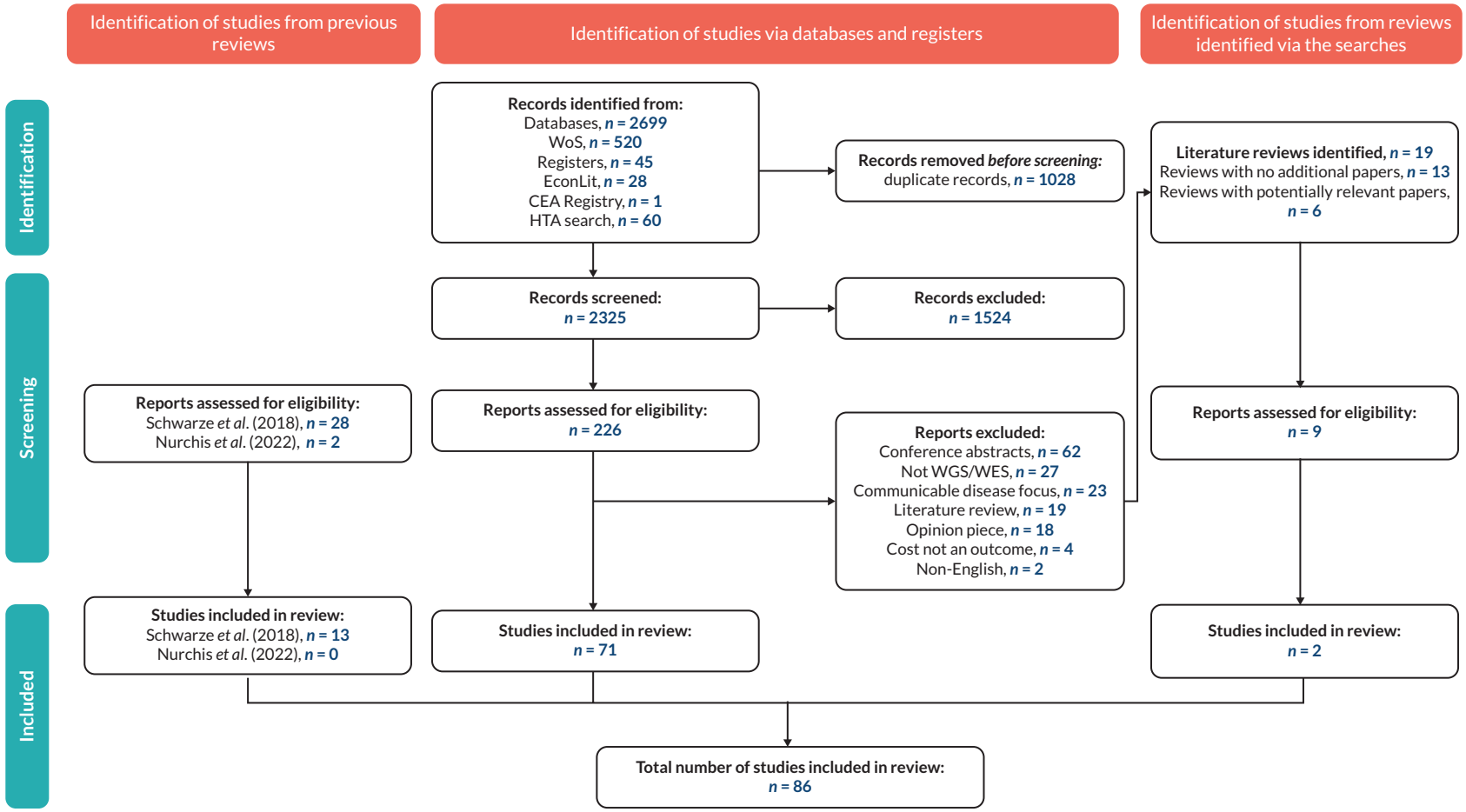

FIGURE 8 The PRISMA flow chart for the review of cost-effectiveness evaluations of WGS and WES.

DOI: 10.3310/DJRF1124

Appendix 5 Summary tables for the review of the five conditions

Data extraction tables for the traditional review, including studies exploring gene/variant frequency in patients with the condition(s) of interest (including study design and recruitment dates, number and definition of cases, test description, genes/variants considered, gene frequency in cases and number of negative tests, variant frequency and expressivity), and studies presenting evidence on early versus late treatment, including study design, number and definition of cases, definitions of early and late treatment, outcome measure and time point and results.

Gene/variant frequency in patients with condition(s) of interest

Pyridoxine-dependent epilepsy

TABLE 21 Summary tables for the studies exploring gene/variant frequency in patients with PDE

Study reference; country (ethnicity)	Study design; recruitment dates	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases test negatives	Variant frequency	Expressivity
Childhood-onset pharmacoresistant seizures							
Boonsimma 2023;[30] Thailand	Prospective, single-arm, single centre Recruitment: June 2016–December 2020	Patients with infantile-onset (age ≤ 12 months) pharmacoresistant epilepsy seen or referred for genetic testing at a tertiary care centre; 103 included, all unrelated	WES (trio) Short- and long-read GS	First step: 728 'epilepsy' genes id from GEL Second step: candidate pathogenic variants selected according to pre-specified criteria	6/103, 5.8% with biallelic *ALDH7A1* variants 6 (100%) compound heterozygous 4/6 id on WES 2/6 id after second tier GS • 1 short-read GS • 1 long-read GS Other genes id: 55 more patients with genetic variants id; 2 'treatable' (*PNPO or BTD* variant), and 36 variants that could 'inform' treatment decisions [*SCN1A* (*n* = 13), *SCN2A* (*n* = 3), *SCN8A* (*n* = 4), *ATP1A3* (*n* = 3), *KCNA2* (*n* = 1), *KCNT1* (*n* = 3), *KCNQ2* (*n* = 8), and *PDHA1* (*n* = 1)]	8 ALDH7A1 variants identified; 5 recurrent and 3 novel (incl. 2 CNVs) Of 5 recurrent variants, 2 occurred in more than 1 patient in the study sample, suggesting possible founder effect [c.1061 A > G (*p*. Tyr354Cys) (4 patients); c.1547 A > G (*p*.Tyr516Cys) (2 patients)]	NR (4/6 reported to have normal development at 8 months to 4 years FU, including 2 of the 4 patients with possible founder variants)

TABLE 21 Summary tables for the studies exploring gene/variant frequency in patients with PDE (*continued*)

Study reference; country (ethnicity)	Study design; recruitment dates	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases test negatives	Variant frequency	Expressivity
Clinically or biochemically defined PDE							
Koul 2019,[31] NR	Retrospective, single-arm Recruitment: 1998–2018	All children with refractory neonatal or infantile seizures who failed to respond to antiepileptic drugs but later responded to PN; 35 included	Targeted variant testing in *ALDH7A1* (no further detail) followed by WES if negative on initial test	*ALDH7A1*; if no variant found using initial test, exome sequencing conducted on multiple genes (including *PLPBP*, *PRRT2* and *ALDH7A1*)	4/35, 11% *ALDH7A1* variant Other genes id: 31/35 (89%); 12/31 (39%) PLPHP variant and 2/31 (6%) PRRT2 variant detected by WES	NR	**Age at onset of seizures (range)** *ALDH7AI + (n = 4)*: 30 minutes to 1 day *PLPBP + (n = 12)*: 1 hour to 10 days *ALDH7A1- and PLPBP- (n = 19)*: 4 hour to 29 months **Delayed development [*n* (%)]** *ALDH7AI + (n = 4)*: 3 (75%) *PLPBP + (n = 12)*: 2 (17%) *ALDH7A1- and PLPBP- (n = 19)*: 7 (37%)
ALDH7A1-positive PDE							
Coughlin 2015;[187] International	Retrospective, single-arm, multicentre registry study Recruitment: NR	Any subject with a clinical suspicion of PDE and at least one pathogenic variant identified in ALDH7A1; 185 included	NR; registry study	NR	182/185, 98% biallelic *ALDH7A1* variants (i.e. 367 alleles with variants) 73 (39.5%) homozygous 109 (58.9%) compound heterozygous Other genes id: 3/185 (1.6%) with only one variant allele identified	Denominator 367 alleles: 209 (57%) missense 66 (18%) splicing errors 29 (8%) inDel 29 (8%) SNV terminations 18 (5%) SNV synonymous 15 (4%) CNVs Recurrent variants: 4 variants accounted for 38% of all variants identified (140/367) c.1279G > C (94, 25.6%) c.834G > A (20, 5.4%) c.1513G > C (13, 3.5%) c.1566–1G > C (13, 3.5%) 49 missense variants occurred for a total of 209 times; 32/49 were only identified in a single individual, 17/49 were recurrent	NR

continued

TABLE 21 Summary tables for the studies exploring gene/variant frequency in patients with PDE (*continued*)

Study reference; country (ethnicity)	Study design; recruitment dates	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases test negatives	Variant frequency	Expressivity
Jiao 2020;[33] China	Retrospective, single-arm, single centre Recruitment: NR	Patients with genetically confirmed PDE (*ALDH7A1 or PLPBP*) at a single centre; 33 included	PCR panel test Sequencing	Each exon (1–18) and exon-intron boundary of the *ALDH7A1* gene	31/33, 94% biallelic *ALDH7A1* variants 5 (16%) homozygous 26 (84%) compound heterozygous Other genes id:2/33 (6%) with PNPO variant	26 *ALDH7A1* variant types: 17 (65%) missense, 2 (8%) nonsense, 3 (12%) splicing sites, 4 (15%) deletions 9 variants recurrent within the study sample; 7 nonsense variants and 2 splicing site	NR

Heritable retinoblastoma

TABLE 22 Summary tables for the studies exploring gene/variant frequency in patients with hRB

Study reference; country	Study design; recruitment	Number and definition of cases	Test	Genes/ variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
Any RB							
Salviat 2020;[39] France	Retrospective, single-arm, single centre; 1 January 2000–30 September 2017	Unrelated RB index cases consecutively referred for genetic consultation; 1371 patients out of 1404 completed genetic counselling	Multiple methods dependent on year of testing (including sequencing and sequential testing); first-line screening done on tumour when available, followed by germline screening	*RB1*; whole spectrum of pathogenic variants	606/1371 (44.2%) with germline *RB1* variant (including 28/606 (4.6%) with germline mosaicism); 45/606 (7.4%) identified through first-line tumour screening; 561/606 (92.6%) identified through germline screening 765/1371 (54.5%) with no germline RB1 variant; 517/765 (67.6%) without tumour screening; 248/765 (32.4%) with somatic pathogenic variant	Nonsense[a]: 222 (36.6%) Frameshift[a]: 140 (23.0%) Out-of-frame splice variants[a]: 110 (18.2%) Large rearrangements[a]: 65 (10.8%) In-frame splice variants: 35 (5.8%) Missense: 25 (4.1%) Promoter sequence pathogenic variants: 5 (0.8%) In-frame deletion: 4 (0.7%) Novel low-penetrance region identified in exon 24	**Percentage of bilateral RB** *Variants with loss of RB protein (n = 537)*: 84.2% *Variants with no loss of RB protein (n = 69)*: 65.2% *Nonsense variants (n = 222)*: 89.2% **Mean (SD) age at diagnosis** *Variants with loss of RB protein (n = 537)*: 12.3 m (11.3) *Variants with no loss of RB protein (n = 69)*: 16.3 m (13.2) **Percentage of stage D/E RB** *Variants with loss of RB protein (n = 537)*: 63.5% *Variants with no loss of RB protein (n = 69)*: 49.3%

TABLE 22 Summary tables for the studies exploring gene/variant frequency in patients with hRB (*continued*)

Study reference; country	Study design; recruitment	Number and definition of cases	Test	Genes/ variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
Sporadic RB							
Temming 2013;[40] Germany	Retrospective, single-arm, single centre; 1961–2006	Patients treated and undergoing ophthalmological follow-up (up to age 5 years) for unilateral RB with genetic data available and no FH of RB; 195 patients selected out of 868 diagnosed at the centre	Multiple methods dependent on year of testing; including sequencing and sequential testing	*RB1*; no further detail	40/195 (20.5%) with germline RB1 variant [including 11/40 (27.5%] with germline mosaicism) 155/195 (79.5%) with no germline RB1 variant	*For the 29 patients with a heterozygous RB1 variant* Whole gene deletion: 10 (34%) Mild variant: 6 (21%) Premature termination: 13 (45%)	9/195 (4.6%) developed metachronous bilateral RB; 8/9 had a heterozygous variant and 1/9 had germline mosaicism
Germline/hereditary RB							
Hulsenbeck 2021;[41] Germany (includes national and international patients)	Retrospective, single-arm, single centre; 1992–2011	Patients with a heterozygous pathogenic constitutional RB1 variant (somatic mosaicism excluded) with complete phenotype data who did not undergo screening for RB; 287 patients selected out of 821 diagnosed at the centre	Multiple methods; including sequencing and sequential testing	RB1; no further detail	287/287 (100.0%) with germline RB1 variant None with no germline variant *3/821 without RB1 variants characterized by high MYCN amplification*	Nonsense or frameshift variant (REC-I): 199 (69.3%) Whole RB1 gene deletions (REC-II): 39 (13.6%) Missense or in-frame SNV (REC-III): 45 (15.7%) SNV or indels 3′ end (REC-IV): 1 (0.3%) SNV in promotor region (REC-V): 3 (1.0%)	**Percentage of bilateral RB** *REC-I (n = 199)*: 93.5% *REC-II (n = 39)*: 76.9% *REC-III (n = 45)*: 80.0% **Median (range) age at diagnosis** *REC-I (n = 199)*: 7.3 m (0.2-48.0) *REC-II (n = 39)*: 10.3 m (0.4-40.9) *REC-III (n = 45)*: 11.6 m (0.9–45.1)

m, months.
a Associated with absence of RB protein.

X-linked hypophosphataemic rickets

TABLE 23 Summary tables for the studies exploring gene/variant frequency in patients with XLHR

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/ variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
Hereditary rickets							
Jacob 2023;[46] India	Prospective, single-arm, single centre	Unrelated individuals with suspected hereditary rickets	Exome sequencing	Not specified	3/10 (33%) PHEX	3 recurrent (previously reported) truncating PHEX variants identified	Phenotype of c1586_1586 + 1del comparatively more severe than c1482 + 5G > C and c.58C > T
	Recruitment: 2018	10 probands with childhood-onset disease, 2/10 with diagnosis in early adulthood			Test 'negatives': CYP27B1 (n = 3)	c1482 + 5G > C, c1586_1586 + 1del and c.58C > T. The c1586_1586 + 1del	
					CYP2R1 (n = 1)		
					VDR (n = 1)		
					SLC34A3 (n = 1)		
					SLC2A2 (n = 1)		
Hypophosphataemic rickets							
Gaucher 2009;[48]	Prospective, single-arm, single centre	Hypophosphataemic rickets associated with tubular phosphate wasting (low serum phosphate and TmP/GFR), along with bone deformities and radiological evidence of rickets	Sequencing	PHEX	93/118 (78%) PHEX (44 sporadic and 49 familial cases)	78 different variants identified. 60/78 (77%) were novel	Some uncertainty about pathogenicity of novel variants, e.g. c.1206A > G (single nucleotide replacement): unclear relationship between variant and disease (patient also displays a missense variant)
France (European, North African, Caribbean, Asian)	Recruitment: NR	118 probands (62 sporadic; 56 familial)/209 patients of 118 families with hypophosphataemic rickets 'pre-screened' to exclude patients with tubulopathy, hypercalciuria or hyperparathyroidism, all indicative of other HR types		22 exons plus adjacent intronic sequences and the untranslated region at the 3-prime end	PHEX-negative patients: 25/118 Further analysis revealed: 1 case large deletion, 1 case different variant (FGF23), 1 case tumour-induced osteomalacia. Rest remains unexplained (possible reasons: large deletions, mosaicism, other causal genes)	22/78 (28%) nonsense 23/78 (30%) frameshift 18/78 (23%) splice sites 15/78 (19%) missense	2x c.505G > A variant (single nucleotide replacement): both patients carry additional variants (1x insertion leading to frameshift and 1x deletion leading to a frameshift) c.849 + 6insT may be responsible for late-onset disease (10–12 years)

TABLE 23 Summary tables for the studies exploring gene/variant frequency in patients with XLHR (*continued*)

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/ variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
Marik 2022;[47] India	Prospective, single-arm, single centre (consecutive patients) Recruitment: May 2015–July 2019	Refractory hypophosphataemic rickets (lack of radiological healing despite treatment); blood levels of phosphate below normal values for age 66 unrelated patients (42 familial cases); mean age at onset of symptoms 22.5 ± 14.3 months	WES	PHEX, FGF23, DMP1, ENPP1, CLCN5, CTNS, SLC2A2, GATM, SLC34A1, EHHADH, SLC4A1, ATP6V1B1	24/66 (36.4%) PHEX 40/66 (60.6%) other affected genes: ENPP1 (n = 2), CLCN5 (n = 3), CTNS (n = 3), SLC2A2 (n = 6), SLC4A1 (n = 7), ATP6V1B1 (n = 4), CYP27B1 (n = 5), VDR (n = 4); n = 1 each for FGF23, DMP1, SLC34A1, EHHADH, GATM, ATP6V0A4, FGFR1 2/66 (3.0%) VUS; 1 PHEX and 1 HRAS variant considered negative	24 different PHEX variants 13/24 novel variants 11/24 known variants 19 pathogenic, 4 likely pathogenic and 1 VUS	One patient was carrying two PHEX variations c.2048T > A, c.2071-1G > C, while her mother who was clinically mildly affected had only one PHEX variation c.2048T > A; p.(Leu683His) which could be the reason behind the variability in disease severity Patients carrying PHEX deletions/insertions/non-sense for a premature stop codon and truncated PHEX protein showed a severe phenotype Patients with: PHEX insertion c.985dup showed an additional feature of dolichocephaly, PHEX deletion c.1202del had multiple pseudofractures, PHEX deletion c.1965del showed loss of permanent teeth, sensorineural hearing loss in left ear and severe enthesopathy
Hereditary hypophosphataemic rickets							
del Pino 2022;[49] Argentina	Retrospective, single-arm, single centre Recruitment: 1992–2019	Patients diagnosed with hereditary hypophosphataemic rickets (following clinical and biochemical selection of patients, diagnosis was based on clinical examination, laboratory tests and X-ray) 42/96 included children underwent molecular testing	Sequencing + MLPA (to detect gene deletions and duplications)	*PHEX* Entire coding region (exons 1–22) and splice sites in flanking intronic regions	36/42 (85.7%) PHEX No confirmed variant 6/42 (14.3%)	36/36 deleterious alterations or large deletions	NR

continued

TABLE 23 Summary tables for the studies exploring gene/variant frequency in patients with XLHR (*continued*)

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/ variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
Rafaelsen 2016;[50] Norway	Retrospective, single-arm, multicentre Recruitment: 2009–14	Patients with hereditary hypophosphataemia; defined by serum phosphate below age-dependent reference range in repeated samples combined with TmP/GFR not due to primary or secondary hyperparathyroidism, Fanconi syndrome or other tubulopathy, vitamin D-dependent rickets, vitamin D deficiency or hypophosphatemia secondary to acute metabolic derangements. ± FH or genetic diagnosis 19 probands/28 patients (22 familial and 6 sporadic cases) from 19 families Median age at diagnosis was 2.1 years (range 0.1–15.5 years)	Sanger sequencing + MLPA (if a disease-causing variant was not found, the inheritance pattern suggested a sporadic case or X-linked dominant disease to look for mid-size deletions and insertions)	*PHEX* and *FGF23*, *DMP1*, *ENPP1*, *KL*, *FAM20C* in successive order if *PHEX* negative All exons and intron–exon boundaries of selected genes	13/19 (68.4%) *PHEX* Other genes: 1/19 (5.3%) *FAM20C* 1/19 (5.3%) *SLC34A3* 4/19 (21.1%) No confirmed variant (possible reason: variants in other genes)	13 different *PHEX* variants, 9/13 novel *PHEX* variants 9 novel variants: 1 large duplication, 2 frameshift and premature stop codons, 2 triplet deletions, 2 missense, 1 nonsense, 1 splice site	There were no differences in growth, dental involvement, persistent bowing, or development of nephrocalcinosis in a comparison of non-sense PHEX variants with missense PHEX variants
X-linked hypophosphataemic rickets							
Ariceta 2023;[51] Multinational	Retrospective, single-arm registry study Recruitment: August 2017–November 2020	People of all ages diagnosed with XLHR based on clinical judgement of an XLH-treating expert physician (FH, clinical, radiological and biochemical findings), or by genetic testing (positive for *PHEX* variant), small number of non-*PHEX* variants (*n* = 15; 2.6%) 282/579 children with genetic test results included	Genetic testing	NR	253/282 (89.7%) *PHEX* Other genes: 4/282 (1.4%) *FGF23*, 1/282 (0.4%) *SLC34A3*, 7/282 (2.5%) Other 17/282 (6.0%) No confirmed variant	NR	NR

Familial haemophagocytic lymphohistiocytosis

TABLE 24 Summary tables for the studies exploring gene/variant frequency in patients with fHLH

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
HLH							
Cetica 2016;[53]	Retrospective, single-arm multicentre	Patients with a clinical diagnosis of HLH as defined by the diagnostic criteria recommended by the Histiocyte Society that was subsequently confirmed	NR; registry study	Testing strategy based on results from immunologic assays:	171/426 (40.1%) with biallelic pathogenic variants	*PRF1*: 28/69 homozygous	26 families with 2 siblings: 16 pairs with similar age at which the disease manifested, 9 pairs with age difference of between 46 and 207 months, and 1 pair sibling 1 with disease onset at 6.7 years, and sibling 2 unaffected at the age of 25 years
Italy (Southern European; Eastern European; African; Asian; Hispanic; others)	Recruitment: time reported to the Italian registry: 1989–2014	426 sequenced/500 (33/500 no material, 41/500 normal on immunological testing no suspicion of genetic disease, i.e. no consanguinity, familial recurrence, pigment deficiency, or disease reactivation)		• *PRF1* in patients with perforin expression deficit	69/426 (16.2%) *PRF1*	34 different variants	50.8% of the 171 patients with biallelic variants died
		For 500 patients: mostly children, 44/500 (8.8%) adults/ Median age (range) at diagnosis: 2.2 years (0–60 years)		• *UNC13D*, then *STX11*, then *STXBP2* in patients with degranulation defect	62/426 (14.6%) *UNC13D*	• Missense (*n* = 20)	
				• *RAB27a*, then *LYST* in patients with pigment deficiency		• Nonsense (*n* = 6)	
				• *SH2D1A* in male patients with defective *SAP* expression and/or inhibitory instead of activating 2B4 receptor function		• Deletions/insertions (*n* = 8)	

continued

TABLE 24 Summary tables for the studies exploring gene/variant frequency in patients with fHLH (*continued*)

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
				• *XIAP* in male patients	1/426 (0.2%) STX11	Most frequent variants: c.1122G > A (*n* = 19 patients); c.272C > T (*n* = 11); c.657C > A (*n* = 10); c.695G > A (*n* = 9)	
				Coding exons and exon-intron boundaries were sequenced	9/426 (2.1%) STXBP2	*UNC13D*: 22/62 homozygous	
					Other HLH genes: *RAB27*a (*n* = 10)	37 different variants	
					LYST (*n* = 3)	• Missense (*n* = 11)	
					XIAP (*n* = 4)	• Nonsense (*n* = 8)	
					SD2D1A (*n* = 13)	• Deletions/insertions (*n* = 12)	
					Test negative: 255/426 (59.9%)	• Splicing (*n* = 6)	
					Monoallelic variants: 43/426 (10.1%) in 1 (*n* = 41) or 2 (*n* = 2) HLH-related genes[a]	Most frequent variants: c.75311G > T (*n* = 19); c.2346_2349delGGAG (*n* = 11); c.1847A > G (*n* = 8)	
					25/426 *PRF1*, 10/426 *UNC13D*, 2/426 *STX11*, 6/426 *STXBP2*	*STX11*: 1 variant (deletion/insertion)	
					Other HLH genes	*STXBP2*: 7 different variants	
					RAB27A (*n* = 2)	• Missense (*n* = 4)	
					No variants found (secondary HLH): 197/426 (46.2%)	• Deletions/insertions (*n* = 3)	
					No variants found (assumed missed genetic cases): 15/426 (3.5%) (potentially underlying genes not yet discovered at time of testing)		

TABLE 24 Summary tables for the studies exploring gene/variant frequency in patients with fHLH (*continued*)

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
fHLH (4 genes only)							
Shabrish 2021;[54] India	Retrospective, single-arm, multicentre	Patients fitting into HLH criteria (including perforin expression and degranulation assay on natural killer cells) of the Histiocyte Society referred to ICMR-National Institute of Immunohematology or collaborating FPID centres or other tertiary care centres in India	Sanger sequencing (for perforin-deficient patients)	*PRF1*, *UNC13D*, *STX11*, *STXBP2*	86/98 (87.8%) with biallelic disease	Details of variant spectrum available for *n* = 72/98	Across all 4 genes, patients with homozygous variants had earlier disease onset (median 10 months) than patients with compound heterozygous variants (median 3 years)
	Recruitment: 2010–20	Diagnosis confirmed by molecular analysis	NGS or WES (for patients with degranulation abnormalities)		40/98 (40.8%) *PRF1*	*PRF1* (*n* = 34)	All but 2 patients (1 homozygous, 1 compound heterozygous) with PRF1 had < 10% perforin expression
		98/101 under 18 years of age			30/98 (30.6%) *UNC13D*	• 25 different variants	Patients with compound heterozygous variants in *UNC13D* (*n* = 7), *STXBP2* (*n* = 1) had a natural killer cell degranulation < 10% apart from 2 (*UNC13D*)
					7/98 (7.1%) *STX11*	• Missense (19/25)	3 patients with atypical presentations (2 with *PRF1* variants, 1 with *STXBP2* variant)
					9/98 (9.2%) *STXBP2*	• Nonsense (4/25)	
					Test negatives: 12/98 (12.2%)	• Frameshift (2/25)	
					Individuals with monoallelic disease: 8/98 (8.2%) *PRF1*	*UNC13D* (*n* = 23)	

continued

TABLE 24 Summary tables for the studies exploring gene/variant frequency in patients with fHLH (*continued*)

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
					2/98 (2.0%) *UNC13D*	• 28 different variants • Missense (10/28) • Nonsense (1/28) • Frameshift (11/28) • Splice site and intronic (6/28) *STX11* (*n* = 6) • 5 different variants • 2/5 novel • Missense (3/5) • Nonsense (1/5) • Frameshift (1/5) *STXBP2* (*n* = 9) • 8 different variants • 2/8 novel • Missense (5/8) • Splice site and intronic (3/8)	

TABLE 24 Summary tables for the studies exploring gene/variant frequency in patients with fHLH (*continued*)

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
Single gene/variant							
Trizzino 2008;[57]	Retrospective, single-arm, multicentre	Patients with FHL2 with documented biallelic PRF1 variants	Direct sequencing	*PRF1* (exons 2 and 3) and exon–intron boundaries	124/124 with biallelic *PRF1* variants	63 different variants (inc. 15 novel):	Younger age of onset for patients with two disruptive variants compared to patients with missense variants only ($p < 0.001$)
Multinational (USA, Italy, Germany, Japan, Sweden, Czech Republic) (Caucasian, Turkish, African American, Japanese, Hispanic, Arabic, Asian, African, Moroccan)	Recruitment: 1983–2006	124 patients/median (range) age at presentation: 3 months (15 days to 26.3 years)				• 11 nonsense • 10 frameshift • 38 missense • 4 in-frame deletion Missense variants only ($n = 34$) 1 missense + 1 disruptive (nonsense/frameshift) variant ($n = 28$) Disruptive 5 Italian 50delT variant in 21 patients (16 homozygous, 5 compound heterozygous), mainly African or African American 1090–91delCT variants only ($n = 62$)	Nonsense variants significantly associated with absent natural killer cell activity ($p = 0.008$)

continued

TABLE 24 Summary tables for the studies exploring gene/variant frequency in patients with fHLH (*continued*)

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
						Most common variants: 1122G > A in 32 patients (22 homozygous, 10 compound heterozygous), 17 Turkish, variant in 7 patients (2 homozygous, 5 compound heterozygous), all Japanese	
Amirifar 2021;[55] multinational	Systematic review	Patients with *UNC13D* variant meeting ≥ 5 (severe) or ≤ 4 (mild) HLH 2004 criteria (confirmed immunologic and genetic evaluation)	NR	NR	269/322 (83.5%) with biallelic disease	117/269 homozygous 152/269 compound heterozygous	60% of severe feature and 30% of mild feature patients carried homozygous *UNC13D* variants (p = 0.001)
		322 patients with *UNC13D* variant/median age of onset 6 months			Test negatives: Monoallelic disease	For complete sample (n = 322):	Missense variants (20.5%) were the most prevalent variation in severe feature patients
					50/322 (15.5%)	• Missense 131 (20.47%)	Splice errors (35%) are the most prevalent alteration in mild form of the disease
					Not reported: 3/322 (0.9%)	• Nonsense 69 (10.87%)	Frequency of frameshift and nonsense variants was 14% in severe feature patients and 11% in mild feature patients
						• Deletion 4 (0.63%)	Severe feature patients had exonic variants 1.6-fold higher than intronic variants. The frequency of exonic and intronic variants was equal in mild feature patients
						• Frameshift 87 (13.59%)	
						• Splice error 224 (35.0%)	

TABLE 24 Summary tables for the studies exploring gene/variant frequency in patients with fHLH (*continued*)

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
						• Inversion 38 (5.94%) • Unknown 85 (13.28%) • 53.4% of variants in exons • 46.6% of variants in introns	
Pagel 2012;[56] multinational (Germany, Turkey) (White, Turkish, Arab, Asian)		Patients with confirmed biallelic *STXBP2* variants	Direct sequencing	*STXBP2* (exons 1–19) and adjacent intronic sequences	37/37 with biallelic *STXBP2* variants	25/37 homozygous	24 patients with biallelic missense variants, in-frame or frameshift deletions and splice-site variants other than exon 15: classic course of disease with early onset and rapidly fatal course if HSCT could not be performed, variable ethnic backgrounds
		37 (28 families) patients/ median (range) age at diagnosis 2 months (3 days to 19 years); 1/37 (2.7%) > 18 years				12/37 compound heterozygous	13 patients with either homozygous or compound heterozygous variants with 1 allele carrying an exon 15 splice-site variant: mild and atypical course. Most developed chronic recurrent course with long episodes of absent HLH symptoms and recurrent reactivations with spontaneous remission or response to steroids-only treatment. Mainly central European ethnic background

continued

TABLE 24 Summary tables for the studies exploring gene/variant frequency in patients with fHLH (*continued*)

Study reference; country (ethnicity)	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
						9 novel variants 9 different missense variants 4 different splice-site variants Small deletions or insertions Frequent variants: c.1430C > T (detected in 5 patients of Arab origin), c.1621G > A (detected in 7 patients of mainly White origin), c.247-1 G > C splice-site variant in exon 15 (detected in 12 mainly German and Turkish patients)	

FPID, foundation of primary immunodeficiency; FHL/fHLH, familial haemophagocytic lymphohistiocytosis.
a That is, 43 patients have a total of 45 monoallelic variants.

TABLE 25 Summary tables for the studies exploring gene/variant frequency in patients with MCADD

Study reference; country	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
IEM/IMD							
Martin-Rivada 2022;[59] Spain	Retrospective, single-arm newborn screening program experience (14 IEM conditions) Recruitment: 2011–9	All patients who received an initial abnormal IEM screening result on NBS test and subsequently referred for biochemical and molecular genetic diagnosis. Genetic testing conducted in all biochemically confirmed cases 224/902 with positive initial screening test had biochemically confirmed disease and underwent genetic testing	NR	***ACADM*** *PAH, DNAJC12, PCBD1, GCDH, MCCC1, MCCC2, PCCB, MAT1A, FAH, HPD, BCKDHB, BCKHBA, DBT, BCAT2, CBS, SLC3A1, OTC, ASS1, HMGCL, ACADVL, HADHA, SLC22A5, CPT2, ETFB, CPT1A, MMACHC, MMADHC, MMUT, MMAB*	*222/224 (99.1%) with biochemically confirmed disease also had a genetically confirmed* ***diagnosis 43/224 (19.2%)*** ***ACADM*** *Other IEM genes:* *PAH (n = 83)* *DNAJC12 (n = 4)* *PCBD1 (n = 2)* *GCDH (n = 12)* *MCCC1 (n = 2)* *MCCC2 (n = 5)* *PCCB (n = 5)* *MAT1A (n = 6)* *FAH (n = 3)* *HPD (n = 1)* *BCKDHB (n = 2)* *BCKHBA (n = 1)* *DBT (n = 1)* *BCAT2 (n = 2)* *CBS (n = 2)* *SLC3A1 (n = 1)* *OTC (n = 2)* *ASS1 (n = 2)* *HMGCL (n = 1)* ***ACADVL (n = 13)*** ***HADHA (n = 2)*** ***SLC22A5 (n = 11)*** ***CPT2 (n = 2)*** ***ETFB (n = 1)*** ***CPT1A (n = 1)*** ***MMACHC (n = 6)*** ***MMADHC (n = 1)*** ***MMUT (n = 2)*** ***MMAB (n = 2)*** **Test negative:** *n* = 2 (patients with biochemical HPA)	14 different genotypes for *ACADM*, 4 of which recurred within the study sample and 10 occurred only in 1 patient each Most common variant 985A > G (59/86 alleles, 66%); 22 cases homozygous 985A > G and 16 cases compound heterozygous, including 985A > G Remaining 5 cases compound heterozygous for other variants	1 MCADD case with clinical symptoms before newborn screening results were available (and very high C8 levels), homozygous for 985A > G

continued

DOI: 10.3310/DJRF1124

TABLE 25 Summary tables for the studies exploring gene/variant frequency in patients with MCADD (*continued*)

Study reference; country	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
FAOD							
Maguolo 2020;[60] Italy	Retrospective, single-arm, newborn screening program experience Recruitment: February 2014–April 2019	Patients diagnosed with FAODs via newborn screening or clinical diagnosis 20/23 newborns with initial positive screening test had biochemically confirmed disease 5/20 biochemical MCADD positive (diagnosed via NBS test) Further 10/11 with clinical dx reported [mean age at onset 29.3 years (SD 26.5 years)] excluded from review	Sequencing (NGS) with custom designed panel; variants confirmed by Sanger sequencing	***ACADM***, *ACADVL*, *CPT1A*, *CPT2*, *ETFA*, *ETFB*, *ETFDH*, *FLAD1*, *HADHA*, *HADHB*, *SLC23A2-*, *SLC25A32*, *SLC52A1*, *SLC52A2*, *SLC52A3*	3/20 (15%) *ACADM* Test negative: 2/20 (10%) *ACADM* monoallelic Other FAOD genes: 5/20 (25%) *ACADVL* (VLCADD) *n* = 3 *ETFDH* (MADD) *n* = 2 Other monoallelic genes: *ACADVL (VLCADD)* *n* = 1 No genetic information 10/20 *SLC22A5* (CUD) (*n* = 3) *ACADS* (SCADD) (*n* = 6) VLCADD (*n* = 1)	3 biallelic variants: 1 homozygous 985A > G, 2 compound heterozygous: c.817_829del/c.388-14A > G (*n* = 1) c.244insT/c.978G > A (*n* = 1) 2 missense 1 nonsense 1splicing 1 frameshift 5/5 previously reported	3/3 enzymatic activity < 5% (activity < 10% correlates with certainly symptomatic, 10–20% insufficient evidence, > 20% might never manifest severe symptoms, i.e. mild biochemical phenotype)
ACAD							
Wang 2019;[61] China	Retrospective, single-arm newborn screening program experience Recruitment: 2014–8	Newborns recalled for further testing following initial positive result for ACAD deficiency on NBS test (HPLC–MS/MS assay). Cases with subsequent positive result on second specimen referred for diagnostic testing and genetic analysis 20/83 newborns with positive screening result had confirmed ACAD on diagnostic testing 4/20 with confirmed MCADD	NSG + Sanger sequencing	306 genes Expanded IEM panel reporting ***ACADM***, *ACADS*, *ACADVL*	3/20 (15%) Biallelic *ACADM* Test negative: 1/20 (5%) monoallelic *ACADM* Other genes: 15/20 (75%) *ACADS* (SCAD) (*n* = 10) *ACADVL* (VLCAD) (*n* = 5) No genetic information *SCAD n* = 1	3/3 cases compound heterozygous 5 different variants 1 occurred twice: c.449_452DelCTGA(Het) Two novel variants: c.589A > G and c.1248T > G, 3 previously reported variants: c.449_452delCTGA, c.970G > A, c.1238G > A	NR

TABLE 25 Summary tables for the studies exploring gene/variant frequency in patients with MCADD (*continued*)

Study reference; country	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
Clinical–biochemical MCADD							
Mesbah 2019;[63] Ireland	Retrospective, single-arm Recruitment: 1 January 1998–30 August 2016	Children < 18 years with clinical dx of MCADD identified via the national centre for IMD and metabolic laboratory at Temple Street Children's University Hospital 17 children; 4 dx by family screening and 2 post-mortem Average age at clinical presentation: 1.48 years (0.005–2.86)	Genetic testing (not further specified)	*ACADM*	11/17 (64.7%) *ACADM* Missed cases (one variant identified only), 3/17 (17.6%) 3/17 clinical MCADD cases had no reported genetic results; likely no genetic testing conducted	8/11 biallelic *ACADM* variant homozygous (985A > G) 3/11 compound heterozygous (985A > G + a second different variant)	NR
Nichols 2008;[62] USA	Retrospective, single-arm newborn screening program experience Recruitment: first 18 months from MCADD inclusion in NYS NBS programme	Newborns with C8 levels ≥ 0.3 µmol/l on NBS test referred for molecular testing 511 newborns with molecular test results	Sequential: 1. PCR/FRET c.985A > G and c.199T > C variant detection If 1 or 2 variants detected OR C8 ≥ 0.4 µmol/l 2. Complete *ACADM* sequence	12 exons of *ACADM*	MS/MS positives (*n* = 511): PCR/FRET (2 ACADM variants): 8/511 (1.6%) biallelic *ACADM* Test negative: 157/511 (30.7%) monoallelic Neither variant 83/511 (16.2%) Follow-up testing positive (*n* = 20): PCR/FRET + *ACADM* sequencing: 17/20 (85%) biallelic *ACADM* Test negative: 3/20 (15%) monoallelic	17/20 with biallelic MCADD 13 different variants 5 novel variants 6/17 homozygote 11/17 compound heterozygote Homozygote: 6/17 985A > G/985A > G Compound heterozygote: 5/17 985A > G compound heterozygote 1/17 c.985A > G/c.199T > C 2/17 c.199T > C compound heterozygote 3/17 compound heterozygote different variants	'Mild' MCADD c.199Y > C/c.134A > G

continued

TABLE 25 Summary tables for the studies exploring gene/variant frequency in patients with MCADD (*continued*)

Study reference; country	Study design; recruitment	Number and definition of cases	Test	Genes/variants considered	Gene frequency in cases, test negatives	Variant frequency	Expressivity
Gene (ACADM) positive MCADD							
Touw 2012;[64] Netherlands	Retrospective, single-arm newborn screening program experience Recruitment: 2007–10	Children from Dutch birth cohort with a clinical follow-up in a metabolic centre after positive result for MCADD on population NBS programme Diagnosis confirmed by *ACADM* gene analysis 68 children with *ACADM* genotypes/84 MCADD confirmed patients; 108 with initial positive NBS test	Sequencing of all exons and adjacent intron regions	*ACADM*	68/68 *ACADM*	10 known (previously reported) genotypes 7 novel genotypes Classic genotypes: 53/68, 77.9%: 985A > G/985A > G (*n* = 42) 985A > G compound heterozygous (*n* = 8) 233T > C/233T > C (*n* = 2) 233T > C compound heterozygous (*n* = 1) Genotypes not previously recognised in clinically confirmed cases: 15/68, 22.1% 985A > G compound heterozygous (*n* = 12) c.233 T > C/c.1066A > T (*n* = 1) c.250 C > T/c.199 T > C c.799 G > A/c.865 G > A	Median residual MCAD enzyme activity: In patients with classic ACADM genotypes: 0% (range 0–8% and 0–5% in leukocytes and lymphocytes, respectively) In patients with clinically unrecognised genotypes: 25% (range 0–63%) 6/17 genotypes with enzyme activity 0%: 985A > G/985A > G c.233 T > C/c.789A > G c.985A > G/c.216 + 1 G > T c.985A > G/c.470 C > T c.233 T > C/c.1066A > T c.985A > G/c.928 G > A

Evidence on early versus late treatment

Pyridoxine-dependent epilepsy

TABLE 26 Summary tables for the studies exploring the impact of early vs. late treatment in patients with PDE

Study reference	Study design	Number and definition of cases	Definition early vs. late	Outcome measure and time point	Results
Bok 2010;[76] Netherlands	Family study	4 (2 families); families in which the mother used PN daily during the second pregnancy from the first trimester onwards (all genetically confirmed)	Antenatal ***PN*** (following birth, asymptomatically or following first seizure); no antenatal or asymptomatic ***PN***	Total IQ Motor performance Outcomes measured at ages 5 and 4 years (family 1) and ages 14 and 12 years (family 2)	**Total IQ** *Family 1* Early-treated sibling: 98 Late-treated sibling: 73 *Family 2* Early-treated sibling: 106 Late-treated sibling: 80 **Motor performance** *Family 1* Early-treated sibling: hypotonic, walking at 27 m Late-treated sibling: hypotonic, walking at 31 m *Family 2* Early-treated sibling: normal Late-treated sibling: normal
Jiao 2021;[77] China	Family study	6 (3 families); patients with PDE (genetically confirmed)	Second born siblings treated with ***PN*** immediately after first seizure; first born siblings treated with ***PN*** at 10 m (onset 8 days), 4 years (onset 3 days), and 7 years (onset 5 months)	Psychomotor development evaluated through clinical judgement and parental questionnaires; timing not specified	**Psychomotor development** *Family 1* Early-treated sibling: normal Late-treated sibling: mild delay *Family 2* Early-treated sibling: normal Late-treated sibling: severe delay *Family 3* Early-treated sibling: normal Late-treated sibling: normal
Tseng 2022;[79] international	Retrospective, single-arm, multicentre registry study	37 (18 families); families with PDE (confirmed by elevation of α-AASA or genetically confirmed), where at least one sibling has confirmed ALDH7A1 and there is a difference in age at initiation of treatment of at least 7 days	Author defined: sibling with the shortest delay in treatment initiation (***PN and/or LRT***); sibling with the longest delay in treatment initiation	Full-scale IQ (two scales used); mean age 12.5 years in early-treated group and 15.3 years in late-treated group Clinical assessments (if standardised assessment not available); timing not specified	PN monotherapy **Mean full-scale IQ** *Available for 1 sibling pair* Early-treated sibling: 106.0 Late-treated sibling: 80.0 **Clinical assessments** *Available for 9 sibling pairs* Early-treated sibling showed a better outcome in 4 pairs (44%), similar outcome in 4 pairs (44%) and a worse outcome in 1 pair (11%) PN and LRT **Mean (SD) full-scale IQ** *Early-treated (n = 3):* 76.0 (21.63) *Late-treated (n = 5):* 77.40 (27.93) **Clinical assessments** *Available for 8 sibling pairs* Early-treated sibling showed a better outcome in 5 pairs (62.5%), similar outcome in 2 pairs (25%), and a worse outcome in 1 pair (12.5%).

continued

TABLE 26 Summary tables for the studies exploring the impact of early vs. late treatment in patients with PDE (*continued*)

Study reference	Study design	Number and definition of cases	Definition early vs. late	Outcome measure and time point	Results
Coughlin 2022;[68] international	Single-arm, multicentre; patients recruited using registry and ambispective clinical data collected	60; individuals (any age) with a confirmed diagnosis of PDE–ALDH7A1 (either elevated α-AASA or genetically confirmed) and at least one developmental assessment	***PN + LRT*** in the first 6m of life; ***PN*** monotherapy or ***PN + LRT*** at > 6 m of age	Standardised developmental assessment; mean (SD) age at developmental assessment 7.68 years (5.49 years)	**Mean standardised developmental assessment** *Early-treated (n = 14 assessments for 8 patients):* 87.3 (95% CI: 79.5 to 95.0) *Late-treated (n = 98 assessments from 46 patients):* 73.8 (95% CI: 73.8 to 77.5) **Difference in mean standardised developmental assessment** 21.9 (95% CI: 1.7 to 42.0) *After adjustment for confounders*
Strijker 2021;[78] Netherlands	Cross-sectional, single-arm; retrospective data obtained from international registry	12; patients with PDE (genetically confirmed)	*LRT* started at < 3 years; *LRT* started at ≥ 3 years, ***no LRT*** given	IQ (several scales used in different patients); timing not specified Neurological outcome (2 standardised, age-specific neurological assessments); timing not specified	**IQ** *Early-treated (n = 3):* All had IQ> 70 *Late-treated (n = 4):* 1 had IQ> 70 and 3 had IQ≤ 70 *No LRT treatment (n = 5):* 2 had IQ> 70 and 3 had IQ≤ 70 **Neurological outcome** *Early-treated (n = 3):* 2 neurologically normal, 1 with complex MND *Late-treated (n = 4):* 3 neurologically normal, 1 with complex MND *No LRT treatment (n = 5):* 2 neurologically normal, 3 with complex MND

m, months; MND, minor neurological dysfunction; PN, pyridoxine.

TABLE 27 Summary tables for the studies exploring the impact of early vs. late treatment in patients with hRB

Study reference	Study design	Number and definition of cases	Definition of early vs. late	Outcome measure and time point	Results
Abramson 2003;[80] USA	Retrospective, single-arm, single centre	264 from sample of 1831; all RB patients who were seen at a single centre between 1914 and 2000	SFH; no screening, clinical presentation of RB with (FH+) or without FH (FH−)	Age at diagnosis Ocular survival; 1 and 5 years Survival; 1 and 5 years	**Mean age at dx, months (SD)** SFH (*n* = 86): 7.7 (1.5) all FH + (*n* = 264): 10.8 (0.8); NR separately for FH + with no screening **Ocular survival rate (Kaplan–Meier)** 1 year *SFH (n = 86):* 83.2% *FH + (n = 178):* 47.5% 5 years *SFH (n = 86):* 67.7% *FH + (n = 178):* 58.2% **Survival rate (Kaplan–Meier)** 1 year *SFH (n = 86):* 100.0% *FH + (n = 178): 97.3%* 5 years *SFH (n = 86):* 93.2% *FH + (n = 178):* 87.4%
Moll 2000;[82] Netherlands	Retrospective, single-arm, multicentre registry study	75; patients with a positive FH for RB identified on the Dutch RB register	RB detected by screening from birth; no screening (hereditary aspect of RB was not known to proband's parents), clinical presentation of RB	Age at diagnosis	**Mean (median [range]) age at diagnosis** *Screened (n = 60):* 4.9 m [1.9 m (1 day–48 m)] *Not screened (n = 25):* 17.2 m [10.0 m (1.5–63.0 m)]
Rothschild 2011;[83] France	Retrospective, single-arm, single centre	59; all RB patients at a national referral centre with familial RB between 1995 and 2004	RB detected by screening due to a FH of RB, either 'IS' or 'S' if the screening did not meet recommendations; fundus examination due to clinical presentation of RB (NS)	Age at diagnosis; Enucleation; timing not specified Mortality; timing not specified	**Age at diagnosis; $p < 0.001$** IS (*n* = 16): median 0 m (range 0–7) S (*n* = 23): median 4 m (range 0–35) NS (*n* = 20): median 9 m (range 2–57) **Enucleation of at least one eye [*n* (%)]** *IS (n = 16):* 0 (0%) *S (n = 23):* 2 (8.7%) *NS (n = 20):* 13 (65.0%) **Mortality [*n* (%)]** *IS (n = 16):* 0 (0%) *S (n = 23):* 0 (0%) *NS (n = 20):* 1 (5.0%)

continued

DOI: 10.3310/DJRF1124

TABLE 27 Summary tables for the studies exploring the impact of early vs. late treatment in patients with hRB (*continued*)

Study reference	Study design	Number and definition of cases	Definition of early vs. late	Outcome measure and time point	Results
Soliman 2017;[84] Egypt	Retrospective, single-arm, single centre	13 (10 families); RB cases with a FH at presentation at hospital	RB detected by screening; no screening, clinical presentation of RB	Age at diagnosis Enucleation; evaluated at last follow-up [median (range) follow-up: 4.8 year (1.2–9.1)] Mortality; timing not specified	**Age at diagnosis (mean)** *Screened* 4.7 m *Not screened* 16.7 m **Enucleation of at least one eye [*n* (%)]** *Screened (n = 5):* 0 (0%) *Not screened (n = 8):* 6 (75%) **Mortality** No deaths
Chandata 2009;[81] USA, Argentina, Brazil, Turkey, Jordan, Venezuela	Retrospective, single-arm, multicentre	92; patients with fRB, diagnosed at one of five referral centres	Developed country (USA) with more screening for fRB; DC with limited screening for fRB	Enucleation; timing not specified Probability of event-free survival, where an event is defined as extraocular relapse, second malignancies, and death; 5 years	**Enucleation of at least one eye [*n* (%)]** *USA (n = 32):* 8 (25.0%) *DC (n = 60):* 43 (71.7%) **Probability (SE) of event-free survival** *USA (n = 32):* 0.92 (0.05) *DC (n = 60):* 0.81 (0.07)

DC, developing countries; fRB, familial RB; IS, intensively screened; m, months; S, screened; SE, standard error; SFH, RB detected by routine screening due to a FH of RB.

X-linked hypophosphataemic rickets

TABLE 28 Summary tables for the studies exploring the impact of early vs. late treatment in patients with XLHR

Study reference	Study design	Number and definition of cases	Definition of early vs. late	Outcome measure and time point	Results
Makitie 2003;[88] Canada	Retrospective, single-arm, single centre	19; all patients with XLHR (confirmed by medical history, physical examination, radiography, biochemistry tests and FH)	Before development of clinically manifest rickets based on positive FH and biochemistry; treatment onset age < 1.0 year vs. after development of clinical signs of rickets; onset age ≥ 1.0 year	Height, predicted adult height, rickets activity End of first treatment year	Group 1 (early): *n* = 8 Group 2 (late): *n* = 11 **Height**: median *z*-score: Group 1: −0.7 SDS vs. Group 2: −1.8 SDS; *p* = 0.009, at pre-puberty (9.0 years): −1.3 SDS vs. −2.0 SDS; *p* = 0.054 **Predicted adult height**: *z*-score group 1: −0.2 SDS vs. group 2: −1.2 SDS; *p* = 0.06 **S-Pi** (mean change from diagnosis to end of first treatment year): *z*-score Group 1: + 2.2 SDS vs. Group 2 + 0.8 SDS; *p* = 0.005. Median at end of the first treatment year: Group 1: −1.5 SDS vs. Group 2: −3.0 SDS, median at pre-puberty (median age 10.8 years): Group 1: −2.6 SDS vs. Group2: −2.5 SDS **S-ALP** improved in both groups during the first treatment year. Median at end of the first treatment year: group 1: −0.8 SDS vs. Group 2: + 2.6 SDS, median at pre-puberty (median age 10.8 year): group 1: −0.7 SDS vs. group2: + 2.0 SDS **Rickets activity** (median score at the end of the first treatment): no change + 2.0 group 1 vs. improved to 4.0 ± 0.4 group 2; *p* = 0.052; (median score at pre-puberty, median age 10.4 years): 4.0 ± 0.5 group 1 vs. 5.0 ± 0.7 group 2; *p* = 0.27
Rafaelsen 2016;[50] Norway	Retrospective, single-arm, multicentre	19; children who had serum phosphate below the age-dependent reference range in repeated samples combined with tubular maximum reabsorption rate of phosphate per glomerular filtration rate	Treatment start age < 1.0 years vs. treatment start age > 1.0 years	Height, biochemical parameters, nephrocalcinosis Skeletal X-ray examinations At each clinic visit from time of diagnosis to the time of study inclusion	Group 1 (early): *n* = 10 Group 2 (late): *n* = 9 At last registered consultation: **Height:** *z*-score: group 1: −1.4 (−2.6 to 0.8); group 2: −2 (−6.3 to 0.3) Delta *z*-score: group 1: −0.4 (−3.1 to 2.0); group 2 : 0 (−1.1 to 1.3) **Persistent bowing: group 1 :** 5/10; group 2 : 7/9 **Dental involvement:** group 1 2/10; group 2 : 7/9
Quinlan 2012;[89] UK	Retrospective, single-arm, single centre	23; children with PHEX variant, treatment duration > 2 years, > 2 clinic visits per year	Treatment start age < 1.0 years vs. treatment start age > 1.0 years	Height, biochemical parameters, rickets activity, renal ultrasound	Group 1 (early): *n* = 10 Group 2 (late): *n* = 13 **Height** [score (SD)]: G1: −0.7 (−1.5 to 0.3) vs..G2: −2.0 (−2.3 to −1.0); *p* = 0.009 at medium treatment years of 8.5 vs. 11.9 (*p* = 0.557) **Rickets severity score** (available for 20 patients, 11 in G1 and 9 in G2): median= 1 in both groups **s-Pi:** median SD score: G1: −0.54 (−0.45 to −0.60) vs. G2: −0.52 (−0.49 to −0.59); *p* = 0.92 **s-ALP** median: G1: 0.9 (0.8 to 1.2) vs. G2: 0.8 (0.7 to 0.9); *p* = 0.13 throughout treatment for both biochemical markers

Familial haemophagocytic lymphohistiocytosis

TABLE 29 Summary tables for the studies exploring the impact of early vs. late treatment in patients with fHLH

Study reference (extraction approach)	Study design	Number and definition of cases	Definition of early vs. late	Outcome measure and time point	Results
Lucchini 2018;[75] international (per-protocol analysis)	Retrospective, multicentre	51/66; fHLH (sibling pairs/ triplets) with variant on either PRF1, UNC13D, STX11 or STXBP2, either symptomatic or asymptomatic at diagnosis (and diagnosed following sibling diagnosis) One asymptomatic case who did not receive any treatment excluded	**Early** Cases that were asymptomatic at diagnosis and did not activate before treatment initiation *Group 1:* Asymptomatic cases without activation prior to HSCT +/− prophylactic treatment *Group 2:* Asymptomatic cases with activation after the start of prophylactic treatment **Late** Cases that were symptomatic at diagnosis (index cases) or asymptomatic cases that activated before treatment initiation *Group 3:* index cases who were symptomatic at diagnosis and treated with active disease protocol +/− HSCT *Group 4:* asymptomatic cases with activation before treatment initiation and treated with active disease protocol + HSCT	Mortality, cause of death, those in complete remission at end of follow-up; variable follow-up	**Mortality [*n* (%)]; cause of death** *Group 1 (n = 15):* 1 (7%); sudden death while in complete remission *Group 2 (n = 3):* 1 (33%); disease progression *Group 3 (n = 26):* 10 (38%); 4 transplant-related mortalities, 5 disease progressions, 1 died as a result of infection *Group 4 (n = 7):* 2 (29%); 2 transplant-related mortalities **Complete remission [*n* (%)]; median [range] length of follow-up (months)** *Group 1 (n = 15):* 14 (93%); 37 (17–89) *Group 2 (n = 3):* 1 (33%); 48 *Group 3 (n = 26):* 10 (38%); 41.5 (12–144) *Group 4 (n = 7):* 5 (7%); 48 (12–72) *Note that one patient in group 2 was lost to follow-up*
Lucchini 2018;[75] international (intention to treat)	Retrospective, multicentre	52/66; fHLH (sibling pairs/ triplets) with variant on either *PRF1*, *UNC13D*, *STX11 or STXBP2*, either symptomatic or asymptomatic at diagnosis (and diagnosed following sibling diagnosis)	**Early** Cases that were asymptomatic at diagnosis **Late** Cases that were symptomatic at diagnosis	Mortality, cause of death, those in complete remission at end of follow-up; variable follow-up	**Mortality [*n* (%)]; cause of death** *Early (n = 26):* 4 (15%); 2 transplant-related mortalities, 1 HLH progression, 1 sudden death in complete remission *Late (n = 26):* 10 (38%); 4 transplant-related mortalities, 5 HLH progressions, 1 'died as a result of infection' **Complete remission [*n* (%)]; median (range) length of follow-up (months)** *Early (n = 26):* 21 (81%); 42 (12–144) *Late (n = 26):* 16 (62%); 41.5 (12–144) *Note that one patient in the early-treated group was lost to follow-up*

Medium-chain acyl-CoA dehydrogenase deficiency

TABLE 30 Summary tables for the studies exploring the impact of early vs. late treatment in patients with MCADD

Study reference	Study design	Number and definition of cases	Definition of early vs. late	Outcome measure and time point	Results
Abdenur 1999;[92] Argentina	Retrospective, 2 patient case series	2; MCADD confirmed by biochemical testing and variant analysis	**Early** Management with high carbohydrate diet from 5-m age following positive NBS test or sibling detected (asymptomatic) **Late** Management with high carbohydrate diet following symptomatic clinical presentation (Reye-like syndrome)	Description of the clinical status of the patients; assessed at age 34 m for early-treated case and 24 m for late-treated case	**Clinical status description** *Early (n = 1):* no decompensation, height, weight and head circumference at 25th–50th centile, physically normal *Late (n = 1):* severe seizure disorder and cerebral palsy, nasogastric feeding
Alcaide 2022;[93] Spain	Retrospective, single-arm, multicentre	33; neonates with genetically confirmed or strong clinical and biochemical features of MCADD	**Early** Management with avoidance of fasting and carnitine supplementation following positive NBS test (includes symptomatic patients), or detected through family screening (asymptomatic) **Late** Management with avoidance of fasting and/or carnitine supplementation following clinical presentation *Note that one patient in the early-treated group had no management and management was poorly followed in another patient*	Description of the clinical status of the symptomatic patients; no standardised follow-up time	**Clinical status description of symptomatic patients** *Early (n = 31):* 9/31 patients symptomatic Conduct disorder (age 10 years) *n* = 1 Autistic behaviour, language retardation and macrosomy *n* = 1 Hospital admissions for vomiting and food intolerance *n* = 1 Fasting hypoglycaemia *n* = 1 Autistic behaviour and hospitalisation for decompensation (age 2 years) *n* = 1 Hypotonia and food intolerance *n* = 1 Creatine kinase elevation *n* = 1 Jaundice *n* = 1 Liver crisis and decompensation (age 2 years) *n* = 1 *Late (n = 2):* both patients symptomatic Convulsive status epilepticus, fasting hypoglycaemias and conduct disorder (age 16 m) *n* = 1 Cyclical vomiting, abdominal pain, temperature episode (age 10 years) *n* = 1
Anderson 2020;[94] USA	Retrospective, single-arm, single centre	90; diagnosed with MCADD following NBS screening, FH or clinical presentation	**Early** Management with avoidance of fasting, carnitine supplementation and diet following positive NBS test (asymptomatic) **Late** Management with avoidance of fasting, carnitine supplementation and diet following symptomatic clinical presentation or detected through family screening	Number of hypoglycaemia-related hospital admissions and emergency room visits; timing not reported	**Mean (95% CI) number of hypoglycaemia related hospital days and ER visits per patient** *Early (n = 76):* 0.62 (0.21 to 1.04) *Late (n = 16):* 2.15 (0.67 to 3.58) **Mean (95% CI) number of hypoglycaemia related hospital days and ER visits per patient-year** *Early (n = 76):* 0.09 (0.03 to 0.15) *Late (n = 16):* 0.11 (0.04 to 0.19)

continued

TABLE 30 Summary tables for the studies exploring the impact of early vs. late treatment in patients with MCADD (*continued*)

Study reference	Study design	Number and definition of cases	Definition of early vs. late	Outcome measure and time point	Results
Gong 2021;[95] China	Retrospective, single-arm, single centre	24; genetically confirmed MCADD patients recruited between 2009 and 2019	**Early** Management with carnitine supplementation and diet following positive NBS test (asymptomatic) **Late** Management with carnitine supplementation and diet following symptomatic clinical presentation	Biomarkers; measured at presentation (aged between 0.5 m and 36 m) Mortality, description of the clinical status of the patients; at study close (range of follow-up times depending on age at start of study)	**Mean C6** *Early (n = 18):* 0.52 *Late (n = 4):* 0.16 **Mean C8** *Early (n = 18):* 2.09 *Late (n = 6):* 0.60 **Mean C10** *Early (n = 16):* 0.27 *Late (n = 5):* 0.10 **Mean C8: C2 ratio** *Early (n = 15):* 0.25 *Late (n = 4):* 0.07 **Mean C8: C10 ratio** *Early (n = 18):* 9.23 *Late (n = 5):* 7.94 **Mortality [*n* (%)]** *Early (n = 15):* 0 (%) *Late (n = 5):* 0 (0%) **Clinical status description** *Early (n = 15):* all healthy and asymptomatic *Late (n = 5):* 1 patient with intermittent fasting hypoglycaemia, 1 patient with hemiplegia due to disease episode and 3 healthy and asymptomatic

TABLE 30 Summary tables for the studies exploring the impact of early vs. late treatment in patients with MCADD (*continued*)

Study reference	Study design	Number and definition of cases	Definition of early vs. late	Outcome measure and time point	Results
Haas 2007;[96] Australia[a]	Retrospective, single-arm, multicentre	59; children with MCADD diagnosed clinically or by NBS screening between April 1994 and March 2002	**Early** Management following positive NBS test (asymptomatic) **Late** Management following symptomatic clinical presentation	Mortality, healthcare use (inpatient stays, emergency department visits, outpatient services); within the first 4 years of life	**Mortality [*n* (%)]** *Early (n = 24):* 1 (4%) *Late (n = 35):* 6 (17%) Inpatient stays **Patients with inpatient stays (*n* (%))** *Early (n = 24):* 10 (42%) *Late (n = 35):* 25 (71%) **Mean (SD) number of admissions per year in children admitted at least once** *Early (n = 10):* 1.1 (1.4) *Late (n = 25):* 0.9 (0.9) **Mean (SD) length of stay** *Early (n = 10):* 2.5 days (2.6) *Late (n = 25):* 2.6 days (3.1) Emergency department visits **Patients with emergency department visits [*n* (%)]** *Early (n = 24):* 5 (21%) *Late (n = 35):* 12 (34%) **Mean (SD) number of visits per year in children who visited the ED at least once** *Early (n = 5):* 0.5 (0.2) *Late (n = 12):* 1.3 (1.3) Outpatient services **Patients who used outpatient services [*n* (%)]** *Early (n = 24):* 15 (63%) *Late (n = 35):* 21 (60%) **Mean (SD) number of uses per annum in children who used outpatient services at least once** *Early (n = 15):* 1.5 (1.0) *Late (n = 21):* 1.6 (1.4)
Li 2019;[97] China	Retrospective, single-arm	6; children with MCADD diagnosed clinically or by NBS screening between January 2007 and June 2017	**Early** Management with diet following positive NBS test (asymptomatic) or management with carnitine supplementation, avoidance of fasting and diet following positive NBS test (symptomatic) **Late** Management with diet ± carnitine supplementation following symptomatic clinical presentation	Description of the clinical status of the patients; variable follow-up (from 3 years to 11 years)	**Clinical status description** *Early (n = 4):* all normal *Late (n = 2):* all normal

continued

TABLE 30 Summary tables for the studies exploring the impact of early vs. late treatment in patients with MCADD (*continued*)

Study reference	Study design	Number and definition of cases	Definition of early vs. late	Outcome measure and time point	Results
Wilcken, 2007;[98] Australia[a]	Retrospective, single-arm, multicentre	81; children with MCADD diagnosed clinically or by NBS screening born between 1 April 1994 and 31 March 2004 **Cohort 1** includes those diagnosed between **1994–2002** (*n* = 59); **cohort 2** also includes patients diagnosed between **2002 and 2004** (*n* = 81)	**Early** Management with avoidance of fasting, sick-day regimen following positive screening **Late** Management with avoidance of fasting, sick-day regimen following symptomatic clinical presentation	Diagnosed by age 2 years (cohort 2) Diagnosed by age 4 years (cohort 1) Mean age at diagnosis (cohort 1) Mortality, severe episodes; assessed at age 2 years (cohort 2) and age 4 years (cohort 1) Number admitted to hospital, total length of stay; assessed at age 4 years (cohort 1) Neuropsychological outcome (intellectual ability score WJ III); assessed at age > 4 years (cohort) *Note that outcomes assessed at age 4 years are for those in cohort 1 and outcomes assessed at age 2 are for those in cohort 2*	**Diagnosed by age 2 years [*n* (%)]** *Early (n = 41):* 41 (100%) *Late (n = 40):* 28 (70%) **Diagnosed by age 4 years (*n* (%))** *Early (n = 24):* 24 (100%) *Late (n = 35):* 26 (74%) **Mean (range) age at diagnosis (cohort 1)** *Early (n = 24):* 0.5 m (0.1–3) *Late (n = 26):* 16 m (0.1–93) **Mortality [*n* (%)]** At age 2 years *Early (n = 41):* 1 (2%) *Late (n = 28):* 4 (14%) At age 4 years *Early (n = 24):* 1 (4%) *Late (n = 26):* 5 (19%) **Severe episodes [*n* (%)]** At age 2 years *Early (n = 41):* 1 (2%) *Late (n = 28):* 18 (64%) At age 4 years *Early (n = 24):* 2 (8%) *Late (n = 26):* 18 (69%) **Admission to hospital [*n* (%)] (cohort 1)** *Early (n = 24):* 10 (42%) *Late (n = 26):* 22 (85%) **Mean (SD) length of stay per admission days (cohort 1)** *Early (n = 10):* 2.35 (2.6) *Late (n = 22):* 2.95 (3.4) **Mean (SD) WJ III score (cohort 1)** *Early (n = 25):* 103.6 (11.7) *Late (n = 13):* 104.9 (14.8)

TABLE 30 Summary tables for the studies exploring the impact of early vs. late treatment in patients with MCADD (*continued*)

Study reference	Study design	Number and definition of cases	Definition of early vs. late	Outcome measure and time point	Results
Wilcken 2009;[99] Australia[a]	Retrospective, single-arm, multicentre	59; children with MCADD diagnosed clinically or by NBS screening between 1 April 1994 and 31 March 2002 (of 116 diagnosed with IEM)	**Early** Management following positive screening **Late** Management following symptomatic clinical presentation	Mortality, intellectual handicap, school placement (extra help/special class/special school), physical handicap; outcomes measured at 6 years or latest clinical review for those lost to follow-up	**Mortality [*n* (%)]** *Early (n = 24):* 5 (21%) *Late (n = 35):* 6 (17%) **Intellectual handicap** *Early (n = 22):* 0 (0%) *Late (n = 28):* 1 (4%) (mild handicap) **School placement (extra help/special class/special school) [*n* (%)]** *Early (n = 15):* 0 (0%) *Late (n = 26):* 2 (8%) (extra help) **Physical handicap** *Early (n = 22):* 0 (0%) *Late (n = 29):* 0 (0%)
Wilson 1999;[100] UK	Retrospective, single-arm, single centre	41; all children with MCADD seen acutely or as outpatients between September 1993 and September 1997	**Early** Management with avoidance of fasting, sick-day regimen and carnitine supplementation following asymptomatic sibling detection **Late** Management with avoidance of fasting, sick-day regimen and carnitine supplementation following symptomatic clinical presentation	Number (and severity) of episodes requiring hospital admission before diagnosis Mortality, number (and severity) of episodes requiring hospital admission after diagnosis; at latest follow-up variable follow-up median 6 years (10 m–14 years)	**Admissions before diagnosis [*n* (%)]** *Early (n = 8):* 0 (0%) *Late (n = 33):* 17 (52%) (8 admitted for a coma, 9 admitted for lethargy/hypoglycaemia) **Mortality [*n* (%)]** *Early (n = 8):* 0 (0%) *Late (n = 33):* 0 (0%) **Admissions after diagnosis [*n* (%)]** *Early (n = 8):* 2 (25%) (admitted for lethargy/hypoglycaemia) *Late (n = 33):* 12 (36%) (2 admitted for coma, 10 admitted for lethargy/hypoglycaemia)

a Cohort overlap.

Evidence from case reports

TABLE 31 Case studies evaluating early vs. late treatment

Study reference	Condition	Outcome
Alfadhel 2012[188]	PDE	Earlier treatment did not improve clinical outcome
Marguet 2016[189]	PDE	Earlier treatment did not improve clinical outcome
Rankin 2007[190]	PDE	Earlier treatment did not improve clinical outcome
Yeghiazaryan 2011[191]	PDE	Earlier treatment did not improve clinical outcome
Ulvi 2002[192]	PDE	Better outcome in sibling who received early treatment
Tseng 2022[193]	PDE	Better outcome in sibling who received early treatment
Busiello 2004[194]	fHLH	Presented twins with fHLH2, one symptomatic and the other asymptomatic but did not report on treatment for either twin

Appendix 6 Reference lists from the Clinical Genome Resource paediatric reports for X-linked hypophosphataemic rickets, medium-chain acyl-CoA dehydrogenase deficiency, pyridoxine-dependent epilepsy and heritable retinoblastoma

List of references cited in the ClinGen paediatric reports for XLHR, MCADD, PDE and hRB and assessment against our inclusion and exclusion criteria for the review of five conditions.

TABLE 32 List of references from the ClinGen paediatric reports for XLHR, MCADD, PDE and hRB

Condition	Date of search	References underlying ClinGen scores for penetrance and treatment effectiveness and final assertion	Reference in WGS review
XLHR	5 November 2018	MD Ruppe. X-Linked Hypophosphatemia. 2012 Feb 09 [Updated 2017 Apr 13]. In: MP Adam, HH Ardinger, RA Pagon, *et al.*, editors. *GeneReviews®* [Internet]. Seattle, WA: University of Washington, Seattle; 1993–2024. URL: www.ncbi.nlm.nih.gov/books/NBK83985	N/A
		Online Mendelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD. HYPOPHOSPHATEMIC RICKETS, X-LINKED DOMINANT; XLHR. MIM: 307800: 3 March 2017. World Wide Web. URL: http://omim.org	N/A
		Carpenter TO, Imel EA, Holm IA, Jan de Beur SM, Insogna KL. A clinician's guide to X-linked hypophosphatemia. *J Bone Miner Res* 2011;**26**:1381–8	Excluded on T/A (non-systematic review)
		Carpenter TO, Whyte MP, Imel EA, Boot AM, Hogler W, Linglart A, *et al*. Burosumab therapy in children with X-linked hpophosphatemia. *N Engl J Med* 2018;**378**:1987–8	Excluded on FT (children with active rickets – not early vs. late)
		Whyte MP, Carpenter TO, Gottesman GS, Mao M, Skrinar A, San Martin J, Imel EA. Efficacy and safety of burosumab in children aged 1–4 years with X-linked hypophosphataemia: a multicentre, open-label, phase 2 trial. *Lancet Diabetes Endocrinol* 2019	Excluded on T/A (children with conventional therapy – not early vs. late)
PDE	29 May 2019	SM Gospe. Pyridoxine-Dependent Epilepsy. 2001 Dec 07 [Updated 2017 Apr 13]. In: MP Adam, HH Ardinger, RA Pagon, *et al.*, editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993–2024. URL: www.ncbi.nlm.nih.gov/books/NBK1486	N/A
		Stockler S, Plecko B, Gospe SM Jr, Coulter-Mackie M, Connolly M, van Karnebeek C, *et al*. Pyridoxine dependent epilepsy and antiquitin deficiency: clinical and molecular characteristics and recommendations for diagnosis, treatment and follow-up. *Mol Genet Metab* 2011;**104**:48–60	Excluded on T/A (non-systematic review)
		van Karnebeek CD, Stockler-Ipsiroglu S, Jaggumantri S, Assmann B, Baxter P, Buhas D, *et al*. Lysine-restricted diet as adjunct therapy for pyridoxine-dependent epilepsy: the PDE consortium consensus recommendations. *JIMD Rep* 2014;**15**:1–1s1	Excluded on T/A (consortium recommendation)

continued

TABLE 32 List of references from the ClinGen paediatric reports for XLHR, MCADD, PDE and hRB (*continued*)

Condition	Date of search	References underlying ClinGen scores for penetrance and treatment effectiveness and final assertion	Reference in WGS review
MCADD	3 August 2020	Merritt JL, Chang IJ. Medium-Chain Acyl-Coenzyme A Dehydrogenase Deficiency. GeneReviews® (1993)	N/A
		New England Consortium of Metabolic Programs. *Medium Chain Acyl-CoA Dehydrogenase Deficiency (MCADD)*. 2013. URL: www.newenglandconsortium.org/mcadd	N/A
		Frazier DM. Medium chain acyl CoA dehydrogenase deficiency (MCADD). Genetic Metabolic Dietitians International: Nutrition Guidelines. 2008. URL: http://gmdi.org/Resources/Nutrition-Guidelines/MCAD	N/A
		British Inherited Metabolic Diseases Group. *Medium Chain Acyl-CoA Dehydrogenase Deficiency (MCADD) – Acute Illness/Decomposition* (standard version). 2020. URL: www.bimdg.org.uk/store/guidelines/ER-MCADD-v5_232766_05042017.pdf	N/A
		British Inherited Metabolic Diseases Group. *Adult Emergency Management*; *Medium Cain Fat Oxidation Disorders*. 2018. URL: www.bimdg.org.uk/store/guidelines/ADULT_MCAD-rev_2015_566641_09012016.pdf	N/A
		Dixon M Champion M. *British Inherited Metabolic Diseases Group – Dietitians' Group. Medium Chain Acyl-CoA Dehydrogenase Deficiency (MCADD) – Dietary Management Guidelines for Dietitians*. 2007. URL: www.bimdg.org.uk/store/guidelines/ER-MCADD-v5_232766_05042017.pdf	N/A
		Andresen BS, Lund AM, Hougaard DM, Christensen E, Gahrn B, Christensen M, et al. MCAD deficiency in Denmark. *Mol Genet Metab* 2012;**106**:175–88	Included in Q2
hRB	30 April 2018	National Retinoblastoma Strategy Canadian Guidelines for Care: Strategie therapeutique du retinoblastome guide clinique canadien. *Can J Ophthalmol / J Can d'ophtalmol* 2009;44:S1–88	Excluded on T/A (guideline)
		Lohmann D, Gallie B, Dommering C, Gauthier-Villars M. Clinical utility gene card for: retinoblastoma. *Eur J Hum Genet*. 2011;**19**	Excluded on T/A (Gene card)
		Skalet AH, Gombos DS, Gallie BL, Kim JW, Shields CL, Marr BP, *et al*. Screening children at risk for retinoblastoma: consensus report from the American Association of Ophthalmic Oncologists and Pathologists. *Ophthalmology* 2018;**125**:453–8	Excluded on T/A (Consensus report)
		DR Lohmann, BL Gallie. Retinoblastoma. 2000 Jul 18 [Updated 2015 Nov 19]. In: MP Adam, HH Ardinger, RA Pagon, *et al.*, editors. GeneReviews® [Internet]. Seattle, WA: University of Washington, Seattle; 1993–2024. URL: www.ncbi.nlm.nih.gov/books/NBK1452	N/A
		Aerts I, Lumbroso-Le Rouic L, Gauthier-Villars M, Brisse H, Doz F, Desjardins L. Retinoblastoma. *Orphanet J Rare Dis* 2006;**1**:31	Excluded on T/A (non-systematic review)
		Online Medelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD. RETINOBLASTOMA; RB1. MIM: 180200 : 2016 Aug 04. World Wide Web. URL: http://omim.org	N/A

Appendix 7 Clinical Genome Resource information on penetrance and treatment effectiveness for X-linked hypophosphataemic rickets, medium-chain acyl-CoA dehydrogenase deficiency, pyridoxine-dependent epilepsy and heritable retinoblastoma

Information on penetrance and treatment effectiveness from the paediatric actionability reports available on ClinGen for XLHR, MCADD, PDE and hRB.

TABLE 33 The ClinGen information on penetrance and treatment effectiveness for XLHR, MCADD, PDE and hRB

Condition	Gene	Penetrance	Expressivity	Patient management
XLHR	*PHEX*	Penetrance is believed to be 100% by age 1 year (GeneReviews, OMIM, **Tier 3**)	Individuals with XLHR have variable expressivity (OMIM, **Tier 3**) The severity can differ among members of the same family (GeneReviews, **Tier 4**)	Treatment for XLHR focuses on improving pain and correcting bone deformities Treatment from time of diagnosis until growth is complete Treatment: oral phosphate administered three to five times daily and high-dose calcitriol Treatment effectiveness: partially corrects leg deformities, decreases the number of necessary surgeries, improves adult height Many patients can still have suboptimal growth and bone healing with treatment. Early treatment (< 1 year of age) had higher median height *z*-score lower serum alkaline phosphatase, less marked radiographic signs of rickets early treatment does not completely normalize skeletal development. (GeneReviews, OMIM, Carpenter 2011, **Tier 3**) Burosumab has not been incorporated into practice guidelines to date. An open-label study showed improved renal tubular phosphate reabsorption, serum phosphorus levels, standing height and physical function, reduced pain and severity of rickets at week 64 of treatment (Carpenter 2018, Whyte 2019, **Tier 5**)
MCADD	*ACADM*	In affected individuals without established diagnosis of MCADD, at least 18% and up to 25% individuals die during their first metabolic crisis. (GeneReviews, **Tier 3**) This MCADD estimate is four times higher in screened newborns that in clinically presenting cases during the 10-year period prior to initiation of newborn screening [potentially due to (a) a different in variant spectrum – screened population with high numbers of variants associated with a milder biochemical phenotype, and (b) a reduced penetrance of 50% of the common c.985A > G variant] (Andresen 2012, **Tier 5**)	MCADD may remain asymptomatic A 'milder' biochemical phenotype can still develop into life-threatening symptoms (GeneReviews, **Tier 3**)	The standard treatment is lifelong supplementation of pyridoxine. One registry study showed that 75% of patients were seizure-free on monotherapy, 13% were controlled on pyridoxine plus additional anticonvulsants, 13% continued to have seizures (Stockler 2011, **Tier 2**)

continued

TABLE 33 The ClinGen information on penetrance and treatment effectiveness for XLHR, MCADD, PDE and hRB (*continued*)

Condition	Gene	Penetrance	Expressivity	Patient management
PDE	*ALDH7A1*	No penetrance data from unselected populations All individuals with PDE have intractable seizures that respond both clinically and electrographically to large daily supplements of pyridoxine (GeneReviews, **Tier 4**)	Substantial interpatient differences in the required effective dose related to the underlying *ALDH7A1* pathogenic variant Seizure types are variable even in the individual patient Patients treated as early as in the neonatal period may have developmental delay and intellectual disability, whereas other patients with prolonged status epilepticus and later diagnosis may have a normal IQ (Stockler 2011, **Tier 3**) No clear genotype–phenotype correlations are known. Pathogenic missense variants that result in residual enzyme activity may be associated with a more favourable developmental phenotype (GeneReviews, van Karnebeek, **Tier 3**)	Management by avoidance of prolonged fasting to prevent primary manifestations in asymptomatic patients with MCADD Infants and children require frequent feedings, with maximum 'safe fasting times' of 4–12 hours depending on age. Times are based on the few single patient reports of controlled fasting studies and feeding practices in screened populations (Frazier 2008, Dixon 2007, **Tier 2**) Immediate management during an acute crisis is key to preventing sudden death. Treatment is aimed at reversal of catabolism and prevention of hypoglycaemia by giving simple carbohydrates by mouth or intravenous fluids (New England Consortium of Metabolic Programs 2013, Frazier 2008, British Inherited Metabolic Disease Group 2018, Inherited Metabolic Disease Group 2020, Dixon 2007, **Tier 2**) Low-dose L-carnitine supplementation is recommended when carnitine levels are below the normal range (Frazier 2008, **Tier 2**)
hRB	*RB1*	> 90% patients with bilateral RB and 15% of patients with unilateral RB have germinal RB1 variants. (National Retinoblastoma Strategy Canadian Guidelines for Care, Lohmann 2011, GeneReviews, Aerts 2006, **Tier 3**) < 10% of families show a 'low penetrance' phenotype with reduced expressivity and incomplete penetrance (i.e. ≤ 25%) (GeneReviews, **Tier 4**)	RB1 gene has 3 expression patterns: unilateral or bilateral retinoblastoma, retinoma, or no visible retinal pathology except for 'normal degeneration' with age (OMIM, **Tier 3**)	Clinical screening, including examination by an ophthalmologist with experience in RB, from birth. If abnormal frequent examinations under anaesthesia by a paediatric anaesthetist Early diagnosis, when tumours are small, maximizes survival and vision outcomes and reduces the need for chemotherapy, enucleation, and radiotherapy (one retrospective study) (National Retinoblastoma Strategy Canadian Guidelines of care 2009, Skalet 2018, **Tier 2**)

EME
HSDR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library