



PHILIPPINE GUIDELINES on PERIODIC HEALTH EXAMINATION



Screening for Congenital and Developmental Disorders



PERIODIC HEALTH EXAMINATION TASK FORCE 2021



DISCLAIMER

This guideline is intended to be used by specialists, general practitioners, allied health professionals who are primary care providers. Although adherence to this guideline is encouraged, it should not restrict the healthcare providers in using their sound clinical judgment in handling individual cases.

Payors and policymakers, including hospital administrators and employers, can also utilize this CPG, but this document should not be the sole basis for evaluating insurance claims. Recommendations from this guideline should not also be treated as strict rules on which to base legal action.

TABLE OF CONTENTS

DISCLAIMER	2
TABLE OF CONTENTS	3
ABBREVIATIONS AND ACRONYMS.....	5
ACKNOWLEDGMENT	7
EXECUTIVE SUMMARY	7
SUMMARY OF RECOMMENDATIONS.....	9
1. INTRODUCTION	13
2. SCOPE & PURPOSE.....	15
3. GUIDELINE DEVELOPMENT METHODOLOGY.....	15
3.1 Organization of the Process	15
3.2 Creation of the Evidence Summaries.....	15
3.3 Composition of the CPG Panel.....	16
3.4 Formulation of the Recommendations.....	17
3.5 Managing Conflicts of Interest.....	18
3.6 External Review Process	18
3.7 Planning for Dissemination and Implementation	18
4.RECOMMENDATIONS AND PANEL DISCUSSION	19
4.1 Pulse Oximetry Screening for Critical Congenital Heart Disease.....	19
4.2 Newborn Screening for Cystic Fibrosis.....	27
4.3 Newborn Screening for Sickle Cell Disease.....	35
4.4 Newborn Screening for Thalassemia	44
4.5 Newborn Screening for G6PD Deficiency	51
4.6 Newborn Screening for Homocystinuria and Methionine Adenosyltransferase Deficiency	55
4.7 Newborn Screening for Tyrosinemia	62
4.8 Newborn Screening for Fatty Acid Oxidation Disorders.....	68
4.9 Newborn Screening for Biotinidase Deficiency	77
4.10 Newborn Screening for Beta-Ketothiolase Deficiency.....	84
4.11 Newborn Screening for Holocarboxylase Synthetase Deficiency	89
4.12 Newborn Screening for Isovaleric Acidemia	93
4.13 Screening for Developmental Delay.....	102

4.14 Screening for Autism Spectrum Disorder	114
4.15 Screening for Learning Disorders	121
5. RESEARCH IMPLICATIONS.....	132
6. DISSEMINATION AND IMPLEMENTATION	134
7. APPLICABILITY ISSUES	134
8. UPDATING OF THE GUIDELINES	134
9. APPENDICES	135
1. Pulse Oximetry Screening for Critical Congenital Heart Disease.....	136
2. Newborn Screening for Cystic Fibrosis	170
3. Newborn Screening for Sickle Cell Disease	180
4. Newborn Screening for Thalassemia	199
5. Newborn Screening for G6PD Deficiency	208
6. Newborn Screening for Homocystinuria and Methionine Adenosyltransferase Deficiency	213
7. Newborn Screening for Tyrosinemia	231
8. Newborn Screening for Fatty Acid Oxidation Disorders	253
9. Newborn Screening for Biotidinase Deficiency	267
10. Newborn Screening for Beta-Ketothiolase Deficiency	279
11. Newborn Screening for Holocarboxylase Synthetase Deficiency.....	284
12. Newborn Screening for Isovaleric Acidemia	292
13. Screening for Developmental Delay	302
14. Screening for Autism Spectrum Disorder	337
15. Screening for Learning Disorders.....	349
PERIODIC HEALTH EXAMINATION TASK FORCE ON CONGENITAL AND DEVELOPMENTAL DISORDERS 2021.....	358
PERIODIC HEALTH EXAMINATION PHASE 2 TASK FORCE 2021	360
CONFLICT OF INTEREST DECLARATION	361

ABBREVIATIONS AND ACRONYMS

Acronyms	Acronyms full name
AAFP	American Academy of Family Physicians
AAP	American Academy of Pediatrics
ASD	Autism Spectrum Disorder
ASQ	Ages and Stages Questionnaire
BD	Biotidinase deficiency
BKTD	Beta-ketothiolase deficiency
BSID	Bayley Scales of Infant and Toddler Development
CHD	Congenital heart disease
CCHD	Critical congenital heart disease
CE	Capillary electrophoresis
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator
CPT1D	Carnitine palmitoyl transferase type 1 deficiency
CPT2D	Carnitine palmitoyl transferase type 2 deficiency
CSBS-ITC	Communication and Symbolic Behavior Scales Infant Toddler Checklist
CTFPHC	Canadian Task Force on Preventive Health Care
DBS	Dried blood spots
DOH	Department of Health
EAG	Expert Advisory Group
ECFS	European Cystic Fibrosis Society
EIBI	Early intensive behavioral intervention
E-IMD	European Registry and Network for Intoxication Type Metabolic Diseases
ELISA	Enzyme-linked immunosorbent assay
ENBS	Expanded newborn screening
EO	Early onset
ERE	Evidence review experts
ESAT	Early Screening of Autistic Traits
EtD	Evidence-to-Decision
FAOD	Fatty acid oxidation disorders
FPR	False positive rate
FYI	First-Year Inventory
GA2	Glutaric aciduria type 2
HCLS	Holocarboxylase synthetase deficiency
HCU	Homocystinuria
HHS	Department of Health and Human Services
HOL	Hour of life
HP	Heel prick
HPLC	High performance liquid chromatography
HRQOL	Health-related quality of life
HSCT	Hemopoietic stem cell transplantation
HTA	Health technology assessment
ICER	Incremental cost effectiveness ratio
IEF	Isoelectric focusing
IEM	Inborn errors of metabolism
IMD	Immunodeficiency
IRT	Immunoreactive trypsinogen testing
ITC	Infant Toddler Checklist
IVA	Isovaleric acidemia
IVD	Isovaleryl CoA dehydrogenase

LCHADD	Long chain 3-hydroxy acyl CoA dehydrogenase deficiency
LO	Late onset
LY	Life years
MAT	Methionine adenosyltransferase
M-CHAT-F	Modified Checklist for Autism in Toddlers
M-CHAT R/F	Modified Checklist for Autism in Toddlers Revised with Follow-Up
MCT	Medium-chain triglycerides
MS	Mass spectroscopy
MTPD	Mitochondrial trifunctional protein deficiency
NBS	Newborn screening
NICE	National Institute for Health and Care Excellence
NIH-ICE	National Institutes of Health – Institute of Clinical Epidemiology
NPV	Negative predictive value
PAP	Pancreatitis-associated protein
PBD	Profound biotidinase deficiency
PDDST	Pervasive Developmental Disorders Screening Test
PDQ	Psychological Development Questionnaire
PEDS	Parents' Evaluation of Developmental Status
PHEX	Periodic Healthy Examination
PICO	Population, Intervention, Comparator and Outcome
PLS	Preschool Language Scale
POS	Pulse oximetry screening
PPV	Positive predictive value
PSI	Parenting Stress Index
PWD	Persons with disability
QALY	Quality-adjusted life-years
RCT	Randomized controlled trial
RUSP	Recommended Uniform Screening Panel
SACS	Social Attention and Communication Surveillance
SCD	Sickle cell disease
SCQ	Social Communication Questionnaire
SLD	Specific learning disability
STE	Social and system demographics, Technology effects and effectiveness & Economic analysis evidence assessment
SUAC	Succinylacetone
tHcy	Total homocysteine
Tyr	Tyrosinemia
USPSTF	US Preventive Services Task Force
UK NSC	UK National Screening Committee
VAB-S	Vineland Adaptive Behavior Scales
WPPSI	Wechsler Preschool and Primary Scale of Intelligence
YACHT-18	Young Autism and other developmental disorders Checkup Tool
YLD	Years lived with disabilities

ACKNOWLEDGMENT

This CPG on PHEX 2021 was prepared by the National Institutes of Health - Institute of Clinical Epidemiology (NIH-ICE).

The project would not have been possible without the initiative and financial support from the DOH. Additional funding for the PHEX 2: Guidelines on Screening for Congenital and Developmental Disorders was provided by the Newborn Screening Reference Center (NSRC). The DOH and NSRC neither imposed any condition nor exerted any influence on the operations and the final output formulation.

The NIH-ICE undertook extensive technical work in (1) searching and synthesizing the evidence while ensuring objectivity in each stage of the process, (2) presenting the evidence in the panel discussion and, documenting and writing the final report. They were also indispensable in carrying out the legwork, coordinating among various individuals, groups, and committees, and facilitating the *en banc* meeting. The CPG Central Steering Committee and the Task Forces Steering Committee were responsible for overall organization and management and is accountable for the quality of the CPG.

Lastly, this guideline is invaluable because of the contribution and participation of panelists from different sectors of healthcare who committed their time and effort to share their knowledge, experience, and expertise in analyzing the scientific evidence and their values and preferences in formulating the recommendations with consideration of patients and the current healthcare system in the country.

The content of this CPG is an intellectual property of the Department of Health (DOH). Kindly provide the proper citations when using any part of this document in lectures, research papers, and any other format presented to the public. The electronic version of this material can be accessed online on the DOH website.

Queries, suggestions, and other concerns regarding this CPG may be directed to the DOH National Practice Guidelines Program office by email (egmd@doh.gov.ph) or to DOH-HPDPB and UP-NIH.

EXECUTIVE SUMMARY

This Clinical Practice Guideline for the Periodic Health Examination (Congenital and Developmental Disorders) is an output from the joint undertaking of the Department of Health and National Institutes of Health- Institute of Clinical Epidemiology.

This clinical practice guideline is a systematic synthesis of evidence to address screening for congenital and developmental disorders among apparently healthy neonates and children. The CPG provides twenty (20) recommendations on fifteen (15) prioritized questions in the screening for certain congenital and developmental disorders.

Recommendations are based on the appraisal of the best available evidence on each of the fifteen identified clinical questions. The CPG is intended to be used by general practitioners and specialists in the primary care setting, policy makers, employers and administrators, allied health practitioners and even patients. The guideline development process followed the widely accepted Grading of Recommendations, Assessment, Development, and Evaluation or the GRADE approach including GRADE Adolopment¹, a systematic process of adapting evidence summaries and the GRADE Evidence to Decision or EtD framework². It included 1) identification of critical questions and critical outcomes, 2) retrieval of current evidence, 3) assessment and synthesis of the evidence base for these critical questions, 4) formulation of draft recommendations, 5) convening of a multi-sectoral stakeholder panel to discuss values and preferences and assess the strength of the recommendations, and 6) planning for dissemination, implementation, impact evaluation and updating.

The recommendations in this CPG shall hold and will be updated after 3 years or when new evidence arise.

¹ Schunemann H, Wiercioch W, Brozek J, Etxeandia-Ikobaltzeta I, Mustafa R, Manja V. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLPMENT. *J Clin Epidemiol.* 2017;81:101-10.

² Schunemann HJ, Mustafa R, Brozek J, Santesso N, Alonso-Coello P, Guyatt G, et al. GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. *J Clin Epidemiol.* 2016;76:89-98.

SUMMARY OF RECOMMENDATIONS

Recommendation	Certainty of Evidence	Strength of Panel Recommendation
Question 1: Should asymptomatic, apparently healthy newborns be screened for critical congenital heart disease?		
Among asymptomatic, apparently healthy newborns, we recommend for the screening of critical congenital heart disease using pulse oximetry.	Moderate	Strong
Question 2: Should asymptomatic, apparently healthy newborns be screened for Cystic Fibrosis?		
Among asymptomatic, apparently healthy newborns, we recommend against the screening of cystic fibrosis.	Very low	Strong
Question 3: Should asymptomatic, apparently healthy newborns be screened for Sickle Cell Disease?		
Among asymptomatic, apparently healthy newborns, we recommend against the screening of sickle cell disease.	Very low	Strong
Question 4: Should asymptomatic, apparently healthy newborns be screened for Thalassemia?		
Among asymptomatic, apparently healthy newborns, we recommend for the screening of thalassemia using HPLC (BIORAD KIT).	Very low	Strong
Question 5: Should asymptomatic, apparently healthy newborns be screened for G6PD deficiency?		
Among asymptomatic, apparently healthy newborns, we recommend for the screening of G6PD deficiency using fluorescence assay (PE neonatal kit).	Very low	Strong

Question 6: Should asymptomatic, apparently healthy newborns be screened for Homocystinuria and Methionine Adenosyltransferase Deficiency?

Among asymptomatic apparently healthy newborns, we recommend against the screening of homocystinuria. Very low Strong

Among asymptomatic, apparently healthy newborns, we recommend against the screening of methionine adenosyltransferase deficiency. Very low Strong

Question 7: Should asymptomatic, apparently healthy newborns be screened for Tyrosinemia?

Among asymptomatic, apparently healthy newborns, we recommend against the screening of tyrosinemia I/II. Very low Strong

Question 8: Should asymptomatic, apparently healthy newborns be screened for fatty acid oxidation disorders?

Among asymptomatic, apparently healthy newborns, we recommend against the screening of long chain 3-hydroxy acyl CoA dehydrogenase deficiency (LCHADD) and mitochondrial trifunctional protein deficiency (MTPD). Very low Strong

Among asymptomatic, apparently healthy newborns, we recommend against the screening of carnitine palmitoyl transferase types 1 and 2 (CPT1, CPT2) and glutaric aciduria type 2 (GA2). Very low Strong

Question 9: Should asymptomatic, apparently healthy newborns be screened for Biotidinase deficiency?

Among asymptomatic, apparently healthy newborns, we recommend against the screening of biotidinase deficiency. Very low Strong

Question 10: Should asymptomatic, apparently healthy newborns be screened for Beta-ketothiolase deficiency?

Among asymptomatic, apparently healthy newborns, we recommend against the screening of beta-ketothiolase deficiency.

Very low

Strong

Question 11: Should asymptomatic, apparently healthy newborns be screened for Holocarboxylase Synthetase deficiency?

Among asymptomatic, apparently healthy newborns, we recommend against the screening of holocarboxylase synthetase deficiency.

Very low

Strong

Question 12: Should asymptomatic, apparently healthy newborns be screened for Isovaleric Acidemia?

Among asymptomatic, apparently healthy newborns, we recommend against the screening of isovaleric acidemia.

Very low

Strong

Question 13: Should asymptomatic, apparently healthy children with neonatal risk factors be screened for developmental delay?

Among asymptomatic, apparently healthy children born preterm, we recommend for the screening of developmental delay at 3-5 months, 12 months, 24 months corrected age and at 36-48 months of age.

Low

Strong

Among asymptomatic, apparently healthy children who have any of the following risk factors: maternal alcohol use during pregnancy, gestational diabetes, gestational hypertension or maternal obesity, we recommend for the screening of developmental delay at 9-, 18- and 24-30 months.

Low

Strong

Among asymptomatic, apparently healthy children who were exposed to maternal cigarette smoking during pregnancy, there is insufficient evidence to recommend for or against the screening of developmental delay.

Low

N/A

Among asymptomatic, apparently healthy children whose mothers were anemic, there is insufficient evidence to

Low

N/A

recommend for or against the screening of developmental delay.

Question 14: Should symptomatic, apparently healthy children be screened for Autism Spectrum Disorder?

Among asymptomatic, apparently healthy children, we recommend for the screening of autism spectrum disorder between the ages 18 to 24 months using the M-CHAT R/F.

Moderate

Strong

Question 15: Should asymptomatic, apparently healthy children be screened for specific learning disorders (reading disability)?

Among asymptomatic, apparently healthy children, there is insufficient evidence to recommend for or against the screening of specific learning disorders (reading disability) in the primary health care setting.

Very low

N/A

1. INTRODUCTION

The Philippine Guidelines on Periodic Health Examination (PHEX) was first published in 2004.(1) It was a comprehensive appraisal and synthesis of evidence on screening interventions committed to providing early prevention services among apparently healthy Filipinos. It was a long-awaited publication and the first to offer evidence-based recommendations for screening tests made possible through the concerted effort of various medical and paramedical organizations composed of more than a hundred experts, researchers, and stakeholders.(1) It was inspired by the Canadian and the US Preventive Services Task Forces, but it was tailored to the Philippine setting.

Due to the evolving technology, scientific evidence, and health policies, there is a pressing need to update this guideline. This 2021 Philippine Guidelines will support the objectives stated in the Universal Health Care Act(2) that all Filipinos are given access to quality and affordable medical services, including primary care benefits.

In the guideline development, evidence-based recommendations for the prioritized health screening were formulated using the GRADE Evidence-to-Decision (EtD) framework.(4, 5) The EtD framework aims to facilitate the adaptation of recommendations and decisions of experts and stakeholders based on specific contexts, essential health outcomes, benefits, and harms while looking through the equity, applicability, and feasibility lenses.

The evidence collated to answer the research questions on screening tests are used in formulating the recommendations. They can be classified into two: (1) screening for a risk factor and (2) screening for early disease. Screening for the former is directed towards determining the effective management of the condition as a risk factor, and screening for the latter is focused on the performance of the tests that will be used to detect and subsequently treat that early disease and prevent it from progressing.

Health screening also carries potential harm, for example, mislabeling the person as being ill. It can pose a threat to the psychological, social, or physical well-being and even to the individual's financial stability. Because of these probable adverse effects of screening, criteria are set to determine if screening for a particular condition can be beneficial and pragmatic. The voting panel members used these criteria (5) aligned with the EtD framework: (1) burden of illness must be high, (2) screening tests must be accurate enough, (3) early treatment must be more effective than late treatment, (4) confirmatory tests and early management must be safe and available, and (5) costs of screening must be proportional with the potential benefit.

Aside from the regulatory agencies and policymakers in the national government, the target users of this guideline on screening strategies include primary care providers, general physicians, specialists, training institutions, payors, patients, the general public, and industry partners.

References

1. Dans A, Morales D. Philippine Guidelines on Periodic Health Examination (PHEX): Effective Screening for Diseases among Apparently Healthy Filipinos. Manila: The Publications Program; 2004.
2. Duque F, Villaverde M, Morales R. Implementing Rules and Regulations of the Universal Health Care Act (Republic Act No. 11223). In: Health Do, editor.: PhilHealth.
3. PhilHealth. Circular No. 2020-0022: Implementing Guidelines for the PhilHealth Konsultasyong Sulit at Tama (PhilHealth Konsulta) Package. In: Corporation PHI, editor. Pasig City: PhilHealth; 2020.
4. Alonso-Coello P, Schünemann H, Moberg J, Brignardello-Petersen R, Davoli M, Tweek S, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ. 2016;353:i2016.
5. Schunemann H, Wiercioch W, Brozek J, Etcheandia-Ikobaltzeta I, Mustafa R, Manja V. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. J Clin Epidemiol. 2017;81:101-10.

2. SCOPE & PURPOSE

The Screening for Congenital and Developmental Disorders is composed of key questions that revolve around screening asymptomatic, apparently healthy newborns and children for **some metabolic disorders included in the Expanded Newborn Screening panel**, critical congenital heart diseases and developmental disorders such as Autism Spectrum Disorder, Specific Learning Disorder, and developmental delay.

3. GUIDELINE DEVELOPMENT METHODOLOGY

3.1 Organization of the Process

Following the international standards, the DOH (1) outlined the guideline development process into four phases: 1) preparation and prioritization, 2) CPG generation, 3) CPG appraisal, and 4) implementation in the Manual for CPG Development [1].

In the preparation and prioritization phase, the Steering Committee set the CPG objectives, scope, target audience, and clinical questions. They consulted different stakeholders in prioritizing and developing the guideline questions. They identified and formed the working groups involved in creating the evidence base and finalizing the recommendations for each clinical question included.

The evidence review experts (ERE) or the technical working group were tasked to review existing CPGs, appraise and summarize the evidence, and draft the initial recommendations. The evidence summaries were then presented to the consensus panel members to finalize the recommendations.

The consensus panel comprised of multisectoral representatives tasked to review the evidence summaries and develop recommendations during the *en banc* meeting. In the meeting, they prioritized critical and important outcomes; discussed necessary considerations revolving around the recommendations and voted on each recommendation and its strength. They participated in a modified Delphi activity to decide on recommendations that were not resolved during the *en banc* meeting.

3.2 Creation of the Evidence Summaries

The clinical questions were developed using the PICO (population, intervention, comparator, and outcome) format. The ERE searched and appraised international practice guidelines related to periodic health screening, including but not limited to those of the Canadian Task Force on Preventive Health Care, U.S. Preventive Services Task Force, National Institute for Health and Care Excellence. If the CPG were of good quality and done within 5 years (2016-2021), the evidence summaries of the CPG were adopted.

The results of the appraisal of existing CPGs and their evidence summaries determined the need for a systematic search in electronic databases (MEDLINE via PubMed, CENTRAL, Google Scholar) for the need to do de-novo systematic reviews and meta-analysis for each question. All searches were done from May to Nov. of 2021. Details on the time periods were discussed under the specific questions. Please see evidence summaries in Appendices. Relevant local databases and websites of medical societies were also utilized in the search. Keywords were based on PICO (MeSH and free text) set for each question. The ERE also contacted authors of related articles to verify details and identify other research studies for appraisal, if needed.

At least two reviewers worked on each PICO question. The search strategy and inclusion criteria were based on the PICO question and are included in their respective evidence summaries. Evidence reviewers appraised the directness, methodological validity, results, and applicability of each relevant article included. RevMan, STATA, and GRADEPro were used for the quantitative synthesis of important clinical outcomes for each question. The ERE generated evidence summaries for each of the sixteen (16) questions. Each evidence summary included evidence on the burden of the problem, and diagnostic performance, benefits, harm, and social and economic impact of the screening test/intervention. Evidence/information that will facilitate in the decision (i.e. cost of screening test, cost-effectiveness studies, qualitative studies) were also included in the evidence summaries. The Quality of Evidence was assessed using the GRADE approach. [2] See table 1.

Table 1. Basis for Assessing the Quality of the Evidence using GRADE Approach

Certainty of Evidence	Interpretation
High	We are very confident that the true effect lies close to that of the estimate of the effect
Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very Low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect
Factors that lower quality of the evidence are:	
<ul style="list-style-type: none"> • Risk of bias • Important inconsistency of results • Some uncertainty about directness • High probability of reporting bias • Sparse data/Imprecision • Publication bias 	
Additional factors that may increase quality are:	
<ul style="list-style-type: none"> • All plausible residual confounding, if present, would reduce the observed effect • Evidence of a dose-response gradient • Large effect 	

3.3 Composition of the CPG Panel

The Steering Committee convened the Consensus Panel (CP), considering possible conflicts of interests of each panel member. To ensure fairness and transparency, the

composition was guided by the DOH manual (1). Content experts and other key stakeholders were invited to join the CP. The key stakeholders included policymakers, patient advocates, and physicians from different settings (e.g., public primary care settings, private practice, occupational health settings). In the choice of CP, the task force made sure that all stakeholders were part of the target population for the CPGs (See PERIODIC HEALTH EXAMINATION TASK FORCE ON CONGENITAL & DEVELOPMENTAL DISORDERS 2021).

3.4 Formulation of the Recommendations

Draft recommendations were formulated based on the quality of evidence, trade-offs between benefit and harm, cost-effectiveness, applicability, feasibility, equity, resources and uncertainty due to research gaps. Prior to the series of online consensus panel meetings, the consensus panel received the draft recommendations together with evidence summaries based on the EtD framework shown in Table 2. These recommendations, together with the evidence summaries, were presented during the *en banc* meeting.

Table 2.Detailed considerations based on the EtD framework (3)

- | |
|--|
| <ol style="list-style-type: none">1. Is the problem a priority?2. How accurate is the test?3. How substantial are the desirable anticipated effects?4. How substantial are the undesirable anticipated effects?5. What is the overall certainty of the evidence?6. Is there important uncertainty about or variability in how much people value the main outcomes, including adverse effects and burden of the test and downstream outcomes of clinical management guided by the test results?7. Does the balance between desirable and undesirable effects favor the test or the comparison?8. How large are the resource requirements (costs)?9. What is the certainty of the evidence of resource requirements (costs)?10. Does the cost-effectiveness of the test favor the test or the comparison?11. What would be the impact on health equity?12. Is the test acceptable to key stakeholders?13. Is the test feasible to implement? |
|--|

The strength of each recommendation (i.e. strong or weak) was determined by the panel considering all the factors mentioned above. Strong recommendation means that the panel is “confident that the desirable effects of adherence to a recommendation outweigh the undesirable effects” while weak recommendation means that the “desirable effects of adherence to a recommendation probably outweigh the undesirable effect but is not confident.” (4)

The recommendation for each question and its strength was determined through voting. A consensus decision was reached if 75% of all CP members agreed. (2) If consensus was not reached in the first voting, questions, and discussions were encouraged. Two further rounds of voting on an issue were conducted. Evidence-based draft recommendations were also revised based on input arrived at by consensus in the *en banc* discussions.

3.5 Managing Conflicts of Interest

The Central Executive Committee convened an Oversight Committee (OC) whose task was to thoroughly review the declaration of conflict of interest (DCOI) of each of the Task Force members particularly the Consensus Panelists (CP) and make recommendations on how to manage the COI. For TF members with potential significant COIs, the member of OC conducted additional investigations with due diligence to ensure the integrity of the CPG process and the final recommendations.

All task force members submitted a DCOI and their curriculum vitae (CV) prior to the initiation of guideline development process. The disclosure included a 4-year period of personal potential intellectual and/or financial conflicts of interest (COI).

Management of the COI of the Consensus Panel, Technical Coordinators, and Task Force Steering Committees were deliberated and decided by the OC, using the pre-agreed criteria. A full description of the methods can be found in the [Final Technical Report](#).

Those with significant potential COI were either not allowed to join or vote depending on the COI. See [Conflict of Interest Declaration](#) at the end of the document.

3.6 External Review Process

The CPGs were reviewed by independent stakeholders, who were not members of the Task Force. They were also presented in conferences and to relevant societies for their comments and suggestions.

3.7 Planning for Dissemination and Implementation

The SC discussed with relevant stakeholders such as DOH and PhilHealth to prepare a dissemination plan that will actively promote the adoption of this guideline with strategies for copyrights. Suggestions ranged from making guidelines available on websites, press conferences, social media sites, professional society conventions, and journal publications.

References

1. DOH, PHIC. Manual for Clinical Practice Guideline Development 2018.
2. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011 Apr;64(4):401-6. doi: 10.1016/j.jclinepi.2010.07.015. Epub 2011 Jan 5. PMID: 21208779.
3. Schünemann H, Wiercioch W, Brozek J, Etxeandia-Ikobaltzeta I, Mustafa R, Manja V. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. *J Clin Epidemiol*. 2017;81:101-10.
4. Guyatt, G. H., Oxman, A. D., Kunz, R., Falck-Ytter, Y., Vist, G. E., Liberati, A., Schünemann, H. J., & GRADE Working Group (2008). Going from evidence to recommendations. *BMJ (Clinical research ed.)*, 336(7652), 1049–1051. <https://doi.org/10.1136/bmj.39493.646875.AE>

4.RECOMMENDATIONS AND PANEL DISCUSSION

4.1 Pulse Oximetry Screening for Critical Congenital Heart Disease

RECOMMENDATION

Among asymptomatic, apparently healthy newborns, we recommend for the screening of critical congenital heart disease using pulse oximetry. (**Strong recommendation, moderate certainty of evidence.**)

Considerations

The consensus panel considered the following when formulating this recommendation:

- The screening tool (use of pulse oximetry) is cheap and effective with moderate certainty of evidence.
- If detected early, there is a great benefit on the management of the patients that screen positive.
- The panel agreed that there could be inequity in management of the critical congenital heart disease due to the lack of hospitals that can perform these corrective surgeries.

4.1.1 Burden of disease

Congenital Heart Disease (CHD) is a large rapidly emerging problem in global health. Last 2017 there were an estimated 261,247 deaths globally because of CHD (1). In the Philippines, it is estimated that CHD accounts for 0.79% of all deaths, amounting to around 4.49 in 100,000 deaths across all age groups, and in less than 5 years old, 30.7 in 100,000 deaths (2), as well as one of the top ten leading causes of mortality and morbidity (3). Critical Congenital Heart Disease (CCHD) is a subset of CHDs that require early intervention within the first 28 days of life. It is usually diagnosed by echocardiography, either antenatally or postnatally. Prenatal screening with antenatal 2D Echo has been utilized by high income countries like the Netherlands, New Zealand, UK, and Denmark, but there are still missed cases especially for less devastating conditions (4). In the United States, it occurs in 17.3 live births out of 10,000 (5), while another study has determined a mean prevalence of 19.1 births out of 10,000 in North and South America, Europe, and Asia (6). There is no published data in the Philippines regarding the prevalence of CCHD as of this writing, but there is an ongoing pilot study since 2018 involving multiple centers noting 14 cases of CCHD in 32,105 screened newborns (4 in 10,000) with 99.29% completion of data collection (7).

Infants with CCHDs are usually asymptomatic during the first few days of life (8). If undiagnosed, babies with CCHD rapidly deteriorate without intervention due to the closure of the ductus arteriosus after birth. As the duct closes in these so called ductal dependent lesions, patients would present with signs of hemodynamic and circulatory collapse like hypoxemia, shock, acidosis, and eventually mortality if without treatment

(9). Intervention is usually by administering urgent prostaglandin infusion (to keep the Patent Ductus Arteriosus open), circulatory and ventilatory support then eventually surgery. It is argued that patients who have surgery after delayed diagnosis of CCHD may have worse outcomes for survival (9). Patients who have complex CHD also have increased risk of impaired developmental outcomes, especially for single ventricle pathologies (10).

4.1.2 Benefits and Harms of Screening Tests

There were no recent direct RCT studies found on screening for CCHD vs no screening among apparently healthy and/or asymptomatic newborns on mortality and other patient important outcomes like early intervention and referral. All studies used a prospective or retrospective cohort model, where all newborns who fit the study criteria were screened within the predefined study periods. However, a large 11-year retrospective cohort of screened and unscreened infants completed in the UK last 2018 (11) noted that there were 5 mortalities after 1 year in the screened cohort ($n=76,232$) vs unscreened cohort ($n = 61,944$ live births). There is paucity of available data as well on survival outcomes of patients with CCHD after the perioperative period (5).

Focus of evidence for benefit and harm of screening will then be on the diagnostic accuracy of pulse oximetry screening, and availability of evidence on effective treatment. However, there were no RCTs available for the management of CCHD, (early: prostaglandin infusion for ductal dependent lesions, late: surgery). A study analyzing temporal trends in US cases of CCHD found out that diagnosis of CCHD after 24 hours of life is associated with improved survival after 1 year compared to a diagnosis of CCHD before 24 HOL (71.7% vs 82.5% $p <.001$), despite surgical complications and long-term morbidity (5). A study done last 2017 has found out that there was a noted decrease in early infant (between 24 hours to 6 months of age) mortality from CCHD since implementation of pulse oximetry screening in the United States. Compared to states without mandatory pulse oximetry screening (POS), states with mandatory POS experienced a significant 33.4% decrease in mortality (95% CI 10.6% - 50.3%) after implementation of POS from CCHD, with an absolute decrease of 3.9 deaths (95% CI 3.6-4.1) per 100,000 births. No significant difference was found from states who had non mandatory screening policies (12). Taking this into consideration, there is a 1 in 10,000 risk of mortality until 1 year of age from CCHD if no screening was informed with a moderate certainty of evidence. Despite concern for false positive rates, a recent study has emphasized that POS false positives are still clinically important secondary outcomes (sepsis, pneumonia, persistent pulmonary hypertension of the newborn), and recommends tracking these data as well (13). Low and low middle-income countries may benefit from POS as an effective tool to detect early onset sepsis (14).

The UK National Screening Committee, based on their pilot study done last 2016 (15) did not recommend POS as a population-based screening test (on top of antenatal screening and newborn physical exam), due to harms associated with further investigations following a positive screen like delayed discharge and parental anxiety (16). Supplementing this conclusion was the UK 2018 cohort (10), noting that there was no statistically significant difference in the post discharge diagnosis rate between screened (7/100,000) and unscreened cohorts (13/100,000) RR = 0.52 (95% CI 0.2-1.42).

4.1.3 Diagnostic Performance of Screening Tests

A Cochrane systematic review published last 2017 (21 studies total, N = 436,758 participants, search until March 2017) has determined that pulse oximetry screening to detect CCHD had an overall sensitivity of 76.3% (95% confidence interval [CI] 69.5 to 82.0) with a low certainty of evidence for true positives and false negatives, and a specificity of 99.9% (95% CI 99.5 to 99.9) with a high certainty of evidence for false positives and true negatives (18). After a systematic search for studies done beyond March 2017 until September 2021, 11 studies were further added (13, 19-28). Sensitivity and specificity were pooled from studies completed beyond January 2015 until March 2017 from the Cochrane Review (29-34) as well as the additional 11 studies, totaling 17 studies with N = 418,219. POS had an overall sensitivity of 71 % (95% CI 53 – 85) with a low certainty of evidence with a specificity of 100% (95% CI 100 – 100) with a moderate certainty of evidence. There was substantial overall heterogeneity for sensitivity ($I^2 = 59.45$). Bias issues especially in the conduct of the reference standard and in flow and timing of analyses for some of the studies pooled affected the certainty of evidence.

Subgroup analysis was done to investigate possible sources of heterogeneity (PIs see Appendix for overall table).

There were 13 studies that screened preterm newborns and term newborns (13, 19, 20, 22-28, 30, 31, 34) and 4 studies (21, 29, 32, 33) that only included term newborns. Summary estimates of sensitivity were 77% (95% CI 64-86) for the studies that included preterm newborns and 58% (95% CI 17-90) for studies only included term newborns. Studies that included antenatal screening (23, 26, 27) had a pooled sensitivity of 84% (95% CI 73-94). Studies that excluded antenatally screened newborns for CCHD (13, 19-22, 24, 25, 29-34) had a pooled sensitivity of 72% (95% CI 49 – 87).

There were 13 studies that did CCHD screening after the 24th hour of life (13, 19, 21, 22, 24-28, 31-34) which had a summary sensitivity estimate of 68% (95% CI 48-73). 14 studies which administered CCHD screening in two limbs (pre and post ductal) (13, 19, 21-28, 30-33) instead of 1 limb (post ductal) only (20, 29, 34) had a summary sensitivity estimate of 72% (95% CI 53-85).

4.1.4 Cost Implication

There is a lack of cost effectiveness studies in the Philippine setting regarding POS, although POS has been practiced at some private and public institutions. Presented (Table 1) is the cost of POS and 2D echo as it would be itemized in a hospital bill if they would be paid on an out-of-pocket basis. Cost of surgical treatment for CCHD in the Philippines differ between healthcare settings as well, but it is noted that PhilHealth has provided a Z package coverage for Tetralogy of Fallot Surgery only (PHP 320,000).

Parameter	Screening intervention	Confirmatory Test
-----------	------------------------	-------------------

	<i>Newborn Pulse Oximetry Screening</i>	<i>2D Echo</i>
Unit cost of procedure: Public (PGH)	Free	Php 600-750
Unit Cost of Procedure: Private	Php 1,800 per single use probe	Php 7000 before professional fees

Table 1: Cost of Pulse Oximetry Screening and 2D echo in a Public and Private Hospital in the Philippines

Cost effectiveness studies from different countries (Canada, UK, Colombia) showed that pulse oximetry screening for CCHD was cost effective compared with no screening (36-38). One of these studies (37) has estimated the potential reduction in treatment cost with timely CCHD diagnosis, where late CCHD was associated with 52% more admissions, 18% more hospital days, and 35% higher inpatient costs during infancy compared to patients with CCHD who were detected earlier. Overall, cost per life year gained could be as low as 12,000 USD (39).

The UK National Screening Center included other clinically significant diagnoses with a positive screen (persistent pulmonary hypertension of the newborn, meconium aspiration syndrome, respiratory distress syndrome, congenital pneumonia, sepsis, transient tachypnea of the newborn, lung malformation) (40). They hesitantly conclude that it may be cost effective due to the lack of comparator data from the initial pilot study (15).

	The Canadian Public Health Association (Canadian Dollar) (36) 1 CAD = Php 40.82	Colombia (Converted to USD from Colombian Peso) (38) 1 Colombian Peso = Php 0.013	UK National Screening Committee POS on CCHD + other relevant outcomes (£ GBP) (40) 1 UK Pound = Php 69.25
Conclusion	Cost-effective	Cost effective	Possibly cost effective
Cost of POS per individual	\$27.17	1 st week of life: \$ 7 1 st year of life: \$ 39	Analysis 1: £48.87 Analysis 2: £157.59
Incremental Cost Effective Ratio (ICER)	\$1110.79 per Quality Adjusted Life Months (QALM)	\$ 100 for a correctly detected case within 1 st week of life,	Analysis 1: £ 27,502 per timely diagnosis

		\$39,050 for the 1 st year of life	Analysis 2: £ 9,996 per timely diagnosis
--	--	---	--

Table 2: Summary Table for More Recent Cost Effectivity Studies

4.1.5 Equity, Acceptability, and Feasibility

Parental anxiety for a positive screening result could be an issue for the acceptability of POS. A positive screen can contribute to parental anxiety that the child has a serious health condition, that there is a possibility for misinterpretation that CCHD has been totally ruled out if there is a negative screen and that the need to obtain consent may be a barrier to the implementation of POS (41). Mothers were predominantly satisfied with POS and that anxiety for false positive results is not significantly different from mothers who were given true negative results (42).

POS may be more advantageous if universally implemented, because not all mothers would have optimal prenatal care and antenatal screening, and that POS would catch these babies of disadvantaged mothers more (43). A review has determined that POS may be of more benefit towards low to low-middle income countries due to most having low antenatal detection rates (14).

In the US, challenges identified in implementation of screening were lack of uniform legislative mandates for the screening programs, lack of funding/resources, screening algorithm method and interpretation, limited availability of echocardiography especially in rural areas, and lack of integration of screening data among programs (44).

In India, factors like high incidence of home births, limited equipment in government hospitals and personnel, limited facilities for confirmatory diagnosis and transport systems after a positive screen have been identified challenges, with proposed solutions such as combining physical exam with POS, training, awareness and education of caregivers, task shifting of echocardiography to selected pediatricians in district hospitals and strengthening existing infrastructure (45). There is also the need for quick and timely access to these services at the secondary and tertiary level for successful diagnosis and surgical treatment of detected cases (46).

4.1.6 Recommendations from Other Groups

Group	Recommendation	Strength of recommendation and certainty of evidence
American Academy of Pediatrics (8)	Recommends screening	
Canadian Pediatric Society (47)	Recommends screening	
UK National Screening Committee (49)	Needs further research	Unclear certainty of evidence

Pulse Oximetry Screening then has been endorsed by the AAP (8) and has been mandatorily implemented since 2011 in the United States. As of 2019, 15 countries have national screening recommendations for POS in CCHD, with 6 other countries have more than 90% of births screened (48). The UK NSC has expressed hesitation for the national mandatory implementation due to issues with harms regarding false positive diagnosis, and that there is no documented benefit so far over the current practice of antenatal screening and newborn physical exam above (16).

References

1. GBD 2017 Congenital Heart Disease Colaborators. Global, regional, and national burden of congenital heart disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Child Adolesc Health.* 2020 Mar; 4(3): 185-200.
2. Institute for Health Metrics and Evaluation (IHME). GBD Compare. Seattle, WA: IHME, University of Washington, 2015. Available from <http://vizhub.healthdata.org/gbd-compare>. (Accessed [August 28, 2021]).
3. Reganion, JG. Transcatheter Intervention of Congenital Heart Diseases: 40-Years Philippine Heart Center Experience. *Philippine Heart Center Journal.* 2015, 20(1):1-9.
4. Ailes EC, Gilboa SM, Honein MA, Oster ME. Estimated Number of Infants Detected and Missed by Critical Congenital Heart Defect Screening. *Pediatrics.* 2015, 135 (6): 1000-1008. doi: 10.1542/peds.2014-3662
5. Oster ME, Lee KA, Honein MA, Riehle-Colarusso T, Shin M, Correa A. Temporal trends in survival among infants with critical congenital heart defects. *Pediatrics.* 2013;131(5):e1502-8. doi:10.1542/peds.2012-3435.
6. Bakker MK, Bergman JEH, Krikov S, et al. Prenatal diagnosis and prevalence of critical congenital heart defects: an international retrospective cohort study. *BMJ Open* 2019;9:e028139. doi:10.1136/bmjopen-2018-028139
7. Del Rosario, JJ. Padilla CD, Alcausin ML. The Philippine Multicenter Pulse Oximetry Screening for Critical Congenital Heart Disease: A Pilot Study. 2018 -Ongoing. HERDIN Registry ID: PHRR201103-003119. Preliminary Findings presented at the Newborn Screening Society of the Philippines 19th National Newborn Screening Convention, October 7, 2021.
8. Wren C, Reinhardt Z, Khawaja K. Twenty-year trends in diagnosis of life-threatening neonatal cardiovascular malformations. *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2008;93(1):F33-5. doi: 10.1136/adc.2007.119032
9. Mahle WT, Newburger JW, Matherne GP, Smith FC, Hoke TR, Koppel R, et al. Role of pulse oximetry in examining newborns for congenital heart disease: a scientific statement from the AHA and AAP. *Pediatrics.* 2009;124(2):823-36. doi:10.1542/peds.2009-1397
10. Huisenga D, La Bastide-Van Gemert S, Van Bergen A, Sweeney J, Hadders-Algra M. Developmental outcomes after early surgery for complex congenital heart disease: a systematic review and meta-analysis. *Dev Med Child Neurol.* 2021;63(1):29–46. doi: 10.1111/dmcn.14512
11. Banait N, Ward-Platt M, Abu-Harb M, Wyllie J, Miller N, Harigopal S. Pulse oximetry screening for critical congenital heart disease: a comparative study of cohorts over 11 years. *J Matern Fetal Neonatal Med.* 2020;33(12):2064–8. doi: org/10.1080/14767058.2018.1538348
12. Abouk R, Grosse SD, Ailes EC, Oster ME. Association of US state implementation of newborn screening policies for critical congenital heart disease with early infant cardiac deaths. *JAMA.* 2017;318(21):2111-8. doi:10.1001/jama.2017.17627
13. Schwartz BN, Hom LA, Von Kohorn I, Becker J, Cuzzi SS, Clarke SEG, et al. Newborn pulse oximetry screening at a community hospital: An 8-year experience. *Pediatrics.* 2021;148(3): e2020049847. doi:10.1542/peds.2020-049847
14. Abbas A, Ewer AK. Newborn pulse oximetry screening: A global perspective. *Early Hum Dev.* 2021;(105457):105457. doi: 10.1016/j.earlhumdev.2021.105457
15. Public Health England. Newborn Pulse Oximetry Screening Pilot End Project Report. May, 2016 <https://legacyscreening.phe.org.uk/documents/pulse-oximetry/NPOSP%20End%20Project%20Report.pdf> (accessed sept 5, 2021)
16. Public Health England. Recommendation to the UK National Screening Committee (UK NSC) for Population Screening for Critical Congenital Heart Disease and Significant Non-Cardiac Conditions Using Pulse Oximetry Screening in Addition to Current Routine Screening. <https://legacyscreening.phe.org.uk/documents/pulse-oximetry/PO%20Research%20Review.pdf> (accessed Sept 5, 2021)

17. Public Health England. UK NSC consultation: pulse oximetry as an additional test in the newborn and infant physical exam. 2019. <https://legacyscreening.phe.org.uk/pulse-oximetry> (accessed September 5, 2021)
18. Plana MN, Zamora J, Suresh G, Fernandez-Pineda L, Thangaratinam S, Ewer AK. Pulse oximetry screening for critical congenital heart defects. *Cochrane Libr [Internet]*. 2018;2018(3). Available from: <http://dx.doi.org/10.1002/14651858.cd011912.pub2>.
19. Almawazini AM, Hanafi HK, Madkhali HA, Majrashi NB. Effectiveness of the critical congenital heart disease screening program for early diagnosis of cardiac abnormalities in newborn infants. *Saudi Med J*. 2017;38(10):1019-24. doi: 10.15537/smj.2017.10.20295
20. Cloete E, Gentles TL, Webster DR, Davidkova S, Dixon LA, Alsweiler JM, et al. Pulse oximetry screening in a midwifery-led maternity setting with high antenatal detection of congenital heart disease. *Acta Paediatr*. 2020;109(1):100-8. doi:10.1111/apa.14934
21. Diller CL, Kelleman MS, Kupke KG, Quary SC, Kochilas LK, Oster ME. A modified algorithm for critical congenital heart disease screening using pulse oximetry. *Pediatrics*. 2018;141(5):e20174065. doi:10.1542/peds.2017-4065
22. Gopalakrishnan S, Karmani S, Pandey A, Singh N, Ratheesh Kumar J, Praveen R, et al. Pulse oximetry screening to detect critical congenital heart diseases in asymptomatic neonates. *Med J Armed Forces India*. 2021;77(2):214–9. doi: 10.1016/j.mjafi.2020.09.004
23. Hamilcikan S, Can E. Critical congenital heart disease screening with a pulse oximetry in neonates. *J Perinat Med*. 2018;46(2):203-7. doi:10.1515/jpm-2017-0006
24. Hu X-J, Ma X-J, Zhao Q-M, Yan W-L, Ge X-L, Jia B, et al. Pulse oximetry and auscultation for congenital heart disease detection. *Pediatrics*. 2017;140(4):e20171154. doi:10.1542/peds.2017-1154
25. Narayen IC, Blom NA, van Geloven N, Blankman EIM, van den Broek AJM, Bruijn M, et al. Accuracy of pulse oximetry screening for critical congenital heart defects after home birth and early postnatal discharge. *J Pediatr*. 2018;197:29-35.e1. doi:10.1016/j.jpeds.2018.01.039
26. Nuntnarumit P, Thanomsingh P, Limrungskul A, Wanitkun S, Sirisopikun T, Ausayapao P. Pulse oximetry screening for critical congenital heart diseases at two different hospital settings in Thailand. *J Perinatol*. 2018;38(2):181-4. doi:10.1038/jp.2017.168
27. Paranka MS, Brown JM, White RD, Park MV, Kelleher AS, Clark RH. The impact of altitude on screening for critical congenital heart disease. *J Perinatol*. 2018;38(5):530-6. doi:10.1038/s41372-018-0043-9
28. El Idrissi Slitine N, Benhaoui F, Sable CA, Martin GR, Hom LA, Fadel A, et al. Pulse oximetry and congenital heart disease screening: Results of the first pilot study in Morocco. *Int J Neonatal Screen*. 2020;6(3):53. doi:10.3390/ijns6030053
29. Gomez-Rodriguez G, Quezada-Herrera A, Amador-Licona N, Carballo-Magdaleno D, Rodriguez-Mejia EJ, Guizar-Mendoza JM. Pulse oximetry as a screening test for critical congenital heart disease in term newborns. *Rev Invest Clin*. 2015;67(2):130-4. PMID: 25938847
30. Jones AJ, Howarth C, Nicholl R, Mat-Ali E, Knowles R. The impact and efficacy of routine pulse oximetry screening for CHD in a local hospital. *Cardiol Young*. 2016;26(7):1397-405. doi:10.1017/S1047951115002784
31. Klausner R, Shapiro ED, Elder RW, Colson E, Loyal J. Evaluation of a screening program to detect critical congenital heart defects in newborns. *Hosp Pediatr*. 2017;7(4):214-8. doi:10.1542/hpeds.2016-0176
32. Ozalkaya E, Akdag A, Sen I, Comert E, Melek Yaren H. Early screening for critical congenital heart defects in asymptomatic newborns in Bursa province. *J Matern Fetal Neonatal Med*. 2016;29(7):1105-7. doi:10.3109/14767058.2015.1035642
33. Van Niekerk AM, Cullis RM, Linley LL, Zuhlk L. Feasibility of Pulse Oximetry Pre-discharge Screening Implementation for detecting Critical Congenital heart Lesions in newborns in a secondary level maternity hospital in the Western Cape, South Africa: The "POPSICLE"• study. *S Afr Med J*. 2016;106(8):817-21. doi:10.7196/SAMJ.2016.v106i8.10071
34. Zuppa AA, Riccardi R, Catenazzi P, D'Andrea V, Cavani M, D'Antuono A, et al. Clinical examination and pulse oximetry as screening for congenital heart disease in low-risk newborn. *J Matern Fetal Neonatal Med*. 2015;28(1):7-11. doi:10.3109/14767058.2014.899573
35. Hoffman JIE. False negative diagnoses of critical congenital heart disease with screening neonatal pulse oximetry. *J Neonatal Perinatal Med*. 2020;13(1):5–9. doi:10.3233/NPM-190297

36. Mukerji, A., Shafey, A., Jain, A. et al. Pulse oximetry screening for critical congenital heart defects in Ontario, Canada: a cost-effectiveness analysis. *Can J Public Health* 111, 804–811 (2020). doi:10.17269/s41997-019-00280-7
37. Peterson C, Dawson A, Grosse SD, Riehle-Colarusso T, Olney RS, Tanner JP, et al. Hospitalizations, costs, and mortality among infants with critical congenital heart disease: how important is timely detection? *Birth Defects Res A Clin Mol Teratol.* 2013;97(10):664-72. doi: 10.1002/bdra.23165
38. Trujillo LD, Reyes SNF, Restrepo TA. et al. Cost-effectiveness analysis of newborn pulse oximetry screening to detect critical congenital heart disease in Colombia. *Cost Eff Resour Alloc* 2019; 17, 11. doi: 10.1186/s12962-019-0179-2
39. Grosse SD, Peterson C, Abouk R, et al. Cost and Cost-Effectiveness Assessments of Newborn Screening for Critical Congenital Heart Disease Using Pulse Oximetry: A Review. *Int J Neonatal Screen* 2017; 3:34.
40. Public Health England. Roberts T, Pickering K, Barton P, Ewer A . Pulse oximetry as a screening test for critical congenital heart defects and other significant diagnoses in newborn infants : a cost effectiveness analysis 2019. <https://legacyscreening.phe.org.uk/documents/pulse-oximetry/PO%20Health%20Economics%20Report.pdf> (accessed September 7, 2021)
41. Kluckow, M. Barriers to the implementation of newborn pulse oximetry screening: A different perspective. *International Journal of Neonatal Screening.* 2018; 4, 4. doi:10.3390/ijns4010004
42. Powell R, Pattison HM, Bhoyar A, Furmston AT, Middleton LJ, Daniels JP, et al. Pulse oximetry screening for congenital heart defects in newborn infants: an evaluation of acceptability to mothers. *Arch Dis Child Fetal Neonatal Ed.* 2013;98(1):F59-63. doi:10.1136/fetalneonatal-2011-301225
43. Gentles T, Cloete E, Mellander M. Comment on Kluckow M. Barriers to the Implementation of Newborn Pulse Oximetry Screening: A Different Perspective. *Int. J. Neonatal Screen.* 2018, 4(1), 4. *International Journal of Neonatal Screening.* 2018;4(2):13.
44. McClain MR, Hokanson JS, Grazel R, Van Naarden Braun K, Garg LF, Morris MR, et al. Critical congenital heart disease newborn screening implementation: Lessons learned. *Matern Child Health J.* 2017;21(6):1240-9. doi:10.1007/s10995-017-2273-4.
45. Kumar, R.K. Screening for congenital heart disease in India: Rationale, practical challenges, and pragmatic strategies. *Ann. Pediatr. Cardiol.* 2016, 9, 111-114.doi:10.4103/0974-2069.181499.
46. Zheleva B, Nair SM, Dobrzycka A, Saarinen A. Considerations for newborn screening for critical congenital heart disease in low- and middle-income countries. *Int J Neonatal Screen.* 2020;6(2):49. doi: 10.3390/ijns6020049
47. Narvey M, Wong KK, Fournier A. Pulse oximetry screening in newborns to enhance detection of critical congenital heart disease. *Paediatr Child Health.* 2017;22(8):494-503. 10.1093/pch/pxx136
48. Martin GR, Cross RR, Hom LA, Klugman D. Updates in congenital heart disease: Can outcomes improve? *Am J Perinatol.* 2019;36(S 02):S22-8. doi: 10.1055/s-0039-1691773.
49. UK National Screening Committee. (2021, July 8). Pulse Oximetry in Newborn Physical Screening. Gov.U; GOV.UK. <https://www.gov.uk/government/consultations/pulse-oximetry-in-newborn-physical-screening>. Accessed October 6, 2021.

4.2 Newborn Screening for Cystic Fibrosis

RECOMMENDATIONS

Among asymptomatic, apparently healthy newborns, we recommend against the screening for cystic fibrosis. (**Strong recommendation, very low certainty of evidence**).

Considerations

The consensus panel considered the following when formulating this recommendation:

- The very low burden of disease among Filipinos coupled with the high cost of confirmatory testing and treatment of the disease led the consensus panel to decide against the screening.
- The panel also agreed that there needs to be stronger evidence to recommend for the screening of cystic fibrosis among the study population.

4.2.1 Burden of disease

Cystic fibrosis (CF) is the most common life-threatening hereditary disease among Caucasian populations. Its estimated global incidence is 1/3,000-1/6,000 livebirths among those of European descent. In Asia and in those of Asian descent, CF is found to be much rarer. While the Middle East has comparable incidence rates (1/2,560 livebirths, Jordan) with western countries, East Asian (1/350,000 livebirths, Japan) and South Asian (1/10,000-1/100,000, India) countries show much lower rates. (1) In the Philippines, there remains a dearth of information regarding CF. In a study done in California, there were 5 newborns of Filipino descent who were diagnosed with CF through a four-step screening process. (2) In the Philippines, however, there is yet to be a detected case of CF. (3)

Patients with CF develop a myriad of complications, most notably, CF-related diabetes, spontaneous pneumothorax, pulmonary hypertension, and recurrent infection. Respiratory sequelae of CF are the leading cause of patient mortality and morbidity (5).

Treatment of CF is centered on patient monitoring and management of complications. Prophylactic antibiotics, laxatives, and nutritional support are given as prevention strategies for avoiding or slowing down the development of complications. While there is still no definite cure for CF, there have been several developments in CFTR modulator therapy in recent years. However, application of these new drugs is limited to certain genotypes and their high costs further inhibit their use (6).

A cost-of-illness study done in the Czech Republic in 2015 reported that the mean total health care cost per CF patient was €14,486 (₱854,689) per year. A majority of these costs went towards medicinal products and devices followed by medical procedures and inpatient care (7).

As CF is a life-long condition demanding rigorous medical management, anxiety and depression for the patient and their family. Parents may experience caregiver fatigue as they spend most of their time attending to their sick child's needs, while siblings may feel neglected and form resentment towards the patient. On a positive note, families who can overcome these challenges tend to be more resilient and are able to have high quality of life. As treatment for CF continues to develop, patients can live longer with more optimistic dispositions (8,10).

4.2.2 Benefits and Harms of Screening Tests

Two observational studies were found which compared the effects of CF NBS with symptomatic/clinical diagnosis regarding the 30-year Survival, Overall Survival (OS), Cumulative Incidence of CF-Related Death, Forced Vital Capacity (FVC).

The first study ($n = 485$) showed a significant increase in 30-year survival (Rate Ratio 1.38; 95% CI 1.10–1.74) regardless of the severity of disease. The same study showed a significant increase in OS in patients with moderate (HR 3.86; 95% CI 1.21–12.38) to severe symptoms (HR 2.23; 95% CI 1.23–4.03), while those with mild symptoms did not see a remarkable increase (HR 0.99; 95% CI 0.44–4.15).

For CF-related deaths, patients with mild (Rate Ratio 0.17; 95% CI 0.05–0.56) to moderate symptoms (Rate Ratio 0.15; 95% CI 0.04–0.55) had a significant decrease, while those presenting with severe symptoms did not (Rate Ratio 0.46; 95% CI 0.20–1.061). (11)

In a study ($n = 2,216$) in patients aged 5-6 years old, there was no significant improvement in FC in those diagnosed via NBS (MD 1.29; 95% CI –0.12–2.70). (12) While most outcomes from these studies showed benefits of CF NBS, the overall certainty of data is very low. This can be attributed to attrition bias, inconsistency, and imprecision issues. See Appendix. (11,12)

Table 1. Effects of Newborn Screening of Cystic Fibrosis on Asymptomatic Newborns

Outcomes	No. of Studies (no. of participants)	Effect (95% CI)	Level of Certainty
30-year Survival	1 Observational Study ($n = 485$)	Rate Ratio 1.38 (1.10–1.74)	Very Low
Overall Survival (Severe Symptoms)	1 Observational Study ($n = 129$)	HR 2.23 (1.23–4.03)	Very Low
Overall Survival (Moderate Symptoms)	1 Observational Study ($n = 120$)	HR 3.86 (1.21–12.38)	Very Low

Outcomes	No. of Studies (no. of participants)	Effect (95% CI)	Level of Certainty
Overall Survival (Mild Symptoms)	1 Observational Study (<i>n</i> = 207)	HR 0.99 (0.44–4.15)	Very Low
Cumulative Incidence of CF-Related Death (Severe Symptoms)	1 Observational Study (<i>n</i> = 129)	Rate Ratio 0.46 (0.20–1.061)	Very Low
Cumulative Incidence of CF-Related Death (Moderate Symptoms)	1 Observational Study (<i>n</i> = 120)	Rate Ratio 0.15 (0.04–0.55)	Very Low
Cumulative Incidence of CF-Related Death (Mild Symptoms)	1 Observational Study (<i>n</i> = 207)	Rate Ratio 0.17 (0.05–0.56)	Very Low
FVC (5-6 years old)	1 Observational Study (<i>n</i> = 2,216)	MD 1.29 (−0.12–2.70)	Very Low

4.2.3 Diagnostic Performance of Screening Tests

There are several distinct tests which can screen for CF: immunoreactive trypsinogen testing (IRT), pancreatitis-associated protein testing (PAP), DNA testing, CFTR gene sequencing. These tests sensitivities that range from 95% to 100% and specificities ranging from 89.99% to 100% (Table 2).(13,15) However, there are certain drawbacks in using single tests to screen for CF such as IRT's high false-positive rates and detecting healthy carriers when using CFTR mutation analysis.(16) In order to optimize screening, these tests are often combined to generate screening strategies (Table 3). Among these, the IRT–DNA–seq strategy appears to be the most effective with a sensitivity of 100% (95% CI 80–100%) and a specificity of 99.99% (95% CI 99.98–99.99%). (16-18)

The six diagnostic cohort studies included in the analysis were found to have very low to low certainty of evidence. This could be attributed to selection bias, attrition, indirectness, inconsistency, and imprecision issues which are described in Appendix C.(13-18)

Table 2. Diagnostic Performance per Screening Test

Procedure	Sensitivity (95% CI)	Specificity (95% CI)
Immunoreactive Trypsinogen Testing	95.60% (90–100%)	98.99% (98.94–99.04%)
Pancreatitis-Associated Protein Testing	95% (90–100%)	89.99% (88.32–91.46%)
DNA Testing	97.90% (96.80–98.70%)	99.40% (98.7–99.9%)
CFTR Gene Sequencing	100% (96.50–100%)	100% (97.20–100%)

Table 3. Diagnostic Performance per Screening Strategy

Procedure	Sensitivity (95% CI)	Specificity (95% CI)
IRT-IRT	100% (96.40–100%)	99.90% (99.80–99.90%)
IRT-PAP	95% (73.10–99.70%)	99.90% (99.88–99.91%)
IRT-PAP-IRT	92.90% (75.70–100%)	100% (99.90–100%)
IRT-DNA	95.80% (95.30–96.30%)	99.90% (99.40–100%)
IRT-DNA-seq	100% (80–100%)	99.99% (99.98–99.99%)
IRT-PAP-DNA-seq	95% (73.10–99.70%)	99.99% (99.99–100%)

4.2.4 Cost Implication

While results show that IRT-PAP is the most economic among the four, it is important to note that PAP is currently not performed in the Philippines. Moreover, CFTR gene sequencing and confirmatory sweat chloride test are also not yet available in local facilities. (13) Should overseas transport cost for specimens be included in the analysis, these values would be expected to change significantly. Table 4 shows approximate cost of CF tests.

Table 4. Cost per Cystic Fibrosis Screening/Confirmative Test

Procedure	Cost
Immunoreactive Trypsinogen Testing	€2.28 (₱134)
Pancreatitis-Associated Protein Testing	€155 (₱9,143)
DNA Testing	\$250 (₱12,666)
CFTR Gene Sequencing	€417 (₱24,598)
Sweat Chloride Test	\$250 (₱12,666)
€1 = ₱58.9; \$1 = ₱50.66	

A study done in the Netherlands in 2015 showed the cost-effectiveness of different screening strategies for CF (Table 5).

Table 5. Cost-Effectiveness of Cystic Fibrosis Screening Strategies

Parameter	Screening Strategy			
	IRT-PAP	IRT-DNA	IRT-DNA-seq	IRT-PAP-DNA-seq
Total Cost of Screening (n = 185,000 children)	€714,000 (₱42,117,119)	€735,000 (₱43,355,857)	€773,000 (₱45,597,385)	€772,000 (₱45,538,397)
Screening Cost per Patient (Total Cost of Screening/185,000)	€3.86 (₱228)	€3.97 (₱234)	€4.18 (₱247)	€4.17 (₱246)
Cost per Life-Year Gained	€23,600 (₱1,392,106)	€ 28,200 (₱1,663,449)	€29,200 (₱1,722,437)	€ 24,300 (₱1,433,398)
Total Program Cost	€30,289,000 (₱1,784,918,938)	€30,349,000 (₱1,788,454,714)	€30,292,000 (₱1,785,095,727)	€30,340,000 (₱1,787,924,348)
Values based on the 2015 cost-effectiveness study conducted in the Netherlands.				€1 = ₱58.9

4.2.5 Equity, Acceptability, and Feasibility

Since the first implementation of CF NBS, there has been discourse regarding its benefits and harm. Some suggest that screening should be continued since early diagnosis leads to better quality of life. While there have been studies showing its benefits to patients'

nutritional status and pulmonary function, there is still inconclusive evidence that screening offers significant improvement. Early diagnosis itself appears to be a controversial issue for families. False positive cases lead to much unnecessary psychosocial stress for the family and parents of CF patients knowing that their next child might also have the disease causes anxiety towards future pregnancies.^[19] On the other hand, some parents prefer early diagnosis as they find that adjusting to their child's needs is easier when started as soon as possible. Another positive outcome from CF NBS programs was the decrease in the incidence of CF as genetic counselling offered to parents resulted in better family planning. (20)

Countries which established CF NBS programs over the last four decades have continued to perform screening for most, if not all, of their newborns. Overall, their decision-making bodies for health policies found the programs worthwhile to keep. It is important to note that these countries are primarily comprised of citizens of Caucasian descent with incidences of CF >1/7000 livebirths. (21) In Asia, while there have been significant incidences of CF in the Middle East, countries in other parts of the continent had very low statistics of CF with no patient from Southeast Asia detected. As such, CF NBS programs are not considered a priority for Asian countries. (21-22)

There have been CF patients of Filipino descent detected in California, USA. These patients were found to have one Caucasian parent which may be concerning as there is a significant number of mixed marriages in the Philippine registry. (2,23) According to 2018 figures, 15,514 Filipinos were married to foreign nationals (3.5% of all marriages). Of these, American, Australian, Canadian, and British citizens were among the top 5 spouses. (23) Given this trend of mixed marriages, CF NBS may become more applicable to future Filipino children. (2)

The implementation of screening programs is often followed by establishing intervention services which add up to the cost to be considered by decision-making bodies. In the Netherlands, the overall cost of CF programs from screening to treatment range from €30,289,000 to €30,349,000 (₱1,780,297,623 to ₱1,783,824,245) (Table 5).(13) These costs are often covered by robust national health insurance policies which is currently limited for cystic fibrosis in the Philippines (PhilHealth: ₱7,300).(7,24)

4.2.6 Recommendations from Other Groups

There are several organizations which recommend CF NBS including the UK National Screening Committee (UK NSC) and the European Cystic Fibrosis Society (ECFS). Both the UK NSC and ECFS guidelines advocate for CF screening for all newborns. (25,26) The ECFS further emphasizes the importance of early multidisciplinary intervention for detected cases. Regarding the establishment of CF NBS programs, the ECFS suggests that countries with a CF incidence of <1/7000 livebirths and with limited health and social services for CF patients might need to reassess its applicability. (26)

References

1. Scotet V, L'Hostis C, Férec C. The changing epidemiology of cystic fibrosis: Incidence, survival and impact of the CFTR gene discovery. *Genes.* 2020;11(6):589.
2. Padilla CD, Aguirre TB. Enhancing Case Detection of Selected Inherited Disorders Through Expanded Newborn Screening in the Philippines. *Acta Medica Philippina.* 2013;47(1).
3. Newborn Screening Reference Center (NSRC). Prevalance as of December 2020 [Internet]. Newborn Screening Reference Center. 2020 [cited 2021Aug31]. Available from: <https://www.newbornscreening.ph/>
4. Keogh RH, Szczesniak R, Taylor-Robinson D, Bilton D. Up-to-date and projected estimates of survival for people with cystic fibrosis using baseline characteristics: A longitudinal study using UK patient registry data. *Journal of Cystic Fibrosis.* 2018;17(2):218–27.
5. Brown SD, White R, Tobin P. Keep them breathing. *Journal of the American Academy of Physician Assistants.* 2017;30(5):23–7.
6. Rafeeq MM, Murad HA. Cystic fibrosis: Current therapeutic targets and future approaches. *Journal of Translational Medicine.* 2017;15(1).
7. Mlčoch T, Klimeš J, Fila L, Vávrová V, Skalická V, Turnovec M, et al. Cost-of-illness analysis and regression modeling in cystic fibrosis: A retrospective prevalence-based study. *The European Journal of Health Economics.* 2017;18(1):73–82.
8. Ernst MM, Johnson MC, Stark LJ. Developmental and Psychosocial issues in cystic fibrosis. *Pediatric Clinics of North America.* 2011;58(4):865–85.
9. Pfeffer PE, Pfeffer JM, Hodson ME. The psychosocial and psychiatric side of cystic fibrosis in adolescents and adults. *Journal of Cystic Fibrosis.* 2003;2(2):61–8.
10. Toth CE. "I am not CF, I have CF": Social Connection and Isolation During Young Adulthood with Cystic Fibrosis [dissertation]. [Philadelphia, Pennsylvania]: Penn Libraries; 2016.
11. Tridello G, Castellani C, Meneghelli I, Tamanini A, Assael BM. Early diagnosis from newborn screening maximises survival in severe cystic fibrosis. *ERJ Open Research.* 2018;4(2):00109–2017.
12. Schlüter DK, Southern KW, Dryden C, Diggle P, Taylor-Robinson D. Impact of newborn screening on outcomes and social inequalities in cystic fibrosis: A UK CF Registry-Based Study. *Thorax.* 2019;75(2):123–31.
13. van der Ploeg CPB, van den Akker-van Marle ME, Vernooij-van Langen AMM, Elvers LH, Gille JJP, Verkerk PH, et al. Cost-effectiveness of newborn screening for cystic fibrosis determined with real-life data. *Journal of Cystic Fibrosis.* 2015;14(2):194–202.
14. Palomaki GE, Bradley LA, Richards CS, Haddow JE. Analytic validity of cystic fibrosis testing: A preliminary estimate. *Genetics in Medicine.* 2003;5(1):15–20.
15. Beauchamp KA, Johansen Taber KA, Grauman PV, Spurka L, Lim-Harashima J, Svenson A, et al. Sequencing as a first-line methodology for cystic fibrosis carrier screening. *Genetics in Medicine.* 2019;21(11):2569–76.
16. Vernooij-van Langen AM, Loeber JG, Elvers B, Triepels RH, Gille JJ, Van der Ploeg CP, et al. Novel strategies in newborn screening for cystic fibrosis: A prospective controlled study. *Thorax.* 2012;67(4):289–95.
17. Sadik I, de Algabe IP, Jiménez R, Benito C, Blasco-Alonso J, Caro P, et al. Initial evaluation of prospective and Parallel assessments of cystic fibrosis newborn screening protocols in Eastern Andalusia: IRT/IRT versus IRT/PAP/IRT. *International Journal of Neonatal Screening.* 2019;5(3):32.
18. Massie RJ, Curnow L, Glazner J, Armstrong DS, Francis I. Lessons learned from 20 years of newborn screening for cystic fibrosis. *Medical Journal of Australia.* 2012;196(1):67–70.
19. Wilfond B, Rothenberg LS. Ethical issues in cystic fibrosis newborn screening: From data to public health policy. *Current Opinion in Pulmonary Medicine.* 2002;8(6):529–34.
20. Dankert-Roelse JE, Vernooij-van Langen A. Newborn screening for cystic fibrosis: Pros and cons. *Breathe.* 2011;24–30.

21. Scotet V, Gutierrez H, Farrell PM. Newborn screening for cf across the globe—where is it worthwhile? *International Journal of Neonatal Screening*. 2020;6(1):18.
22. Singh M, Rebordosa C, Bernholz J, Sharma N. Epidemiology and genetics of cystic fibrosis in Asia: In preparation for the Next-generation treatments. *Respirology*. 2015;20(8):1172–81.
23. Philippine Statistics Authority. Marriages in the Philippines, 2018 [Internet]. Philippine Statistics Authority. 2019 [cited 2021Sep15]. Available from: <https://psa.gov.ph/content/marriages-philippines-2018>
24. Philippine Health Insurance Corporation. Case Rates Search [Internet]. Philippine Health Insurance Corporation. 2012 [cited 2021Sep15]. Available from: <https://crs.philhealth.gov.ph>
25. UK National Screening Committee. Cystic fibrosis (newborn) [Internet]. UK National Screening Committee. 2017 [cited 2021Sep15]. Available from: <https://view-health-screening-recommendations.service.gov.uk/cystic-fibrosis-newborn/>
26. Castellani C, Duff AJA, Bell SC, Heijerman HGM, Munck A, Ratjen F, et al. ECFS best practice guidelines: The 2018 Revision. *Journal of Cystic Fibrosis*. 2018;17(2):153–78.

4.3 Newborn Screening for Sickle Cell Disease

RECOMMENDATIONS

Among asymptomatic, apparently healthy newborns, we recommend against the screening for sickle cell disease. (*Strong recommendation, very low certainty of evidence*).

Considerations

The consensus panel considered the following when formulating this recommendation:

- There is a lack of burden of this disease among Filipinos.
- The very low certainty of evidence did not provide enough basis for the panel to recommend for the screening of sickle cell disease.

4.3.1 Burden of disease

Sickle cell disease (SCD) refers to a group of inherited disorders caused by structural variations in the beta subunit of hemoglobin.(1) The incidence of SCD varies according to different regions and may range from 1 in every 17,271 to 1 in every 5,650 in Canada, 1 in every 6,579 in the United States, 1 in 375 African American newborns, 1 in 2000 in England, 1 in 674 in France and 1 in 54 in Africa.(2) In the Philippines, it has an estimated annual incidence of 20 (Hb SS) to 40 (Hb SC) in every 2 million births.(3) Under the Philippine Pediatric Society (PPS) registry, there are eight registered cases of sickle cell disorders for the past 15 years (2006 to 2021).(4)

SCD can present with acute and chronic complications as a result of vaso-occlusion and hemolysis.(1) Among the most common complications in children include long-term anemia, pain crisis, acute chest syndrome, spleen sequestration, stroke, infections and priapism.(5) Adopting intensive screening procedures and vaccination guidelines for children with SCD has markedly reduced mortality between ages 0 to 4 years.(1) Most recent data also mentioned the following as predictors of mortality: age, reticulocyte count, fetal hemoglobin, tricuspid regurgitant jet velocity (2.5 m/s or more) and log (N-terminal-pro-BNP). Current management includes blood transfusion, disease-modifying drugs (e.g hydroxyurea, voxelotor, crizanlizumab), hematopoietic stem cell transplant and gene therapy.(1)

4.3.2 Benefits and Harms of Screening Tests

Mortality

A moderate quality systematic review published by Runkel *et al.* in 2020 and a HTA report by Blancquaert (6) did not find randomized controlled trials that directly assessed mortality of newborns after screening versus no screening for sickle cell disease.

However, the systematic review yielded one retrospective controlled cohort study published in 2007 which reported that newborn screening for SCD (using umbilical cord blood) followed by early treatment compared to SCD screening with no treatment resulted in a significant reduction of mortality in the first five years of life, with an odds ratio of 0.09

(95% CI 0.04 to 0.22, p < 0.001) for the 5th year of life.(7) This study had high risk of bias due to comparison of non-concurrent groups in different decades, non-blinding of patients and investigators, and unclear control for confounding factors (See Appendices).

Effectiveness of Newborn Screening Programs for SCD

Consistent evidence suggests that early detection and early initiation of penicillin prophylaxis, parental education and follow-up comprehensive care are associated with reduced early mortality and improved survival in children with SCD. (2) (See Appendices).

Evidence from Systematic Reviews/HTAs

In the Social and System Demographics, Technology Effects and Effectiveness & Economic Analysis Evidence Assessment (STE) published in Alberta Canada in 2016, a low-quality HTA showed that there were no randomized controlled trials that compared clinical outcomes of SCD with or without newborn screening suggesting no direct evidence to support its effectiveness in improving outcomes in newborns with SCD.

The HTA found one cohort study by Vichinsky *et al.* (8) which showed that in a program with strong focus on parental education, but without the use of penicillin prophylaxis, the overall mortality rate of children diagnosed with Hb SS in the neonatal period was 1.8%, compared to 8% for children diagnosed at age of 21 months. In another French study (2002) wherein penicillin prophylaxis and parental education were initiated once diagnosis is confirmed, there is significantly fewer splenic sequestration and painful events in the newborn screening group (diagnosed at birth, n = 38) versus the control group (diagnosed at mean age of 24 months, n = 69, p = 0.04). However, only children who survived at least 2 years old were included thus the effectiveness of screening on early mortality cannot be properly assessed.

The HTA by Blancquaert also identified several longitudinal and trend studies which showed the following: (2)

- Yanni *et al.* (9) – In black children with SCD in the US, mortality rate decreased significantly by 68% (95% CI 58 to 75%) for children 3 years and below. Mortality caused by infection decreased from 57 to 23% in children below 4 years old. However, the authors were not able to determine whether expansion of newborn screening for SCD in 1987 contributed to the decrease in mortality rates.
- Frempong and Pearson (10) – Newborn screening with comprehensive follow-up care greatly reduced mortality in children with SCD. Review of death certificates in different time intervals from when newborn screening was not available until it was offered showed the following: no newborn screening (1970-1988) – 13 deaths, limited newborn screening (1988-1990) – 5 deaths (no screening at birth), and universal newborn screening with prophylactic penicillin and vaccination (1990 2002) – no deaths.

- Two studies reported experiences of Jamaica in newborn screening programs providing penicillin prophylaxis, parental education in early diagnosis of acute splenic sequestrations and close monitoring in sickle cell clinics. A study by Lee *et al.* (11) followed 307 of 315 children diagnosed with Hb SS and showed a mortality rate of 14.2% before 2 years old and 23.8% before 10 years old between 1973 to 1975. This study also looked at survival of patients after screening and having penicillin prophylaxis, parental education and close monitoring among three groups over 15 years. Results showed a trend for improved survival from any cause of mortality (Hazard ratio (HR) 1.37, 95% CI 1.00 – 1.88) and improved survival from acute splenic sequestration and pneumococcal septicemia/meningitis after the intervention (HR 2.40 95% CI 1.11 – 5.19). The second study by King *et al.* (12) showed a mortality rate of 0.8% and 8.6% before 2 years and 10 years respectively between 1995 to 2006. The study also mentioned that the mortality caused by splenic sequestration after the parental education program was introduced in 1978 decreased from 28% (3 deaths out 98 episodes) to 0.53% (1 deaths out of 188 episodes).

Evidence from Primary Studies

Several primary studies conducted in Angola, Brazil, India and the United States reported short and long-term outcomes of newborns screening programs for SCD. (6)

In Brazil (13), implementation of SCD screening over 10 years in 1, 217 833 infants using HPLC showed a mortality rate of 3.7% and survival rate of 94% (95% CI 0.918 to 0.914). For the Hb SS genotype, the survival rate was 92.4% (95% CI 0.894 to 0.995). These mortality rates decreased in comparison to the available historical statistical data before the introduction of the newborn screening for SCD.

In the United States (14), they followed up of 940 children with SCD diagnosed through newborn screening. After early detection of SCD and the resulting improved medical care (early initiation of penicillin prophylaxis, initial visit to a sickle cell center and pneumococcal vaccination), 93.9% of children with Hb SS and 98.4% of children with mild SCD lived into adulthood. It was noted that the combination of penicillin and vaccination did not prevent all fatal pneumococcal infections but bacterial sepsis was no longer the leading cause of mortality.

In Angola (15), among 244 patients with SCD followed-up after neonatal screening, there was 3.6% mortality rate and a 6.8% first-year mortality compared to 9.5% national infant mortality in children with SCD. This pilot study provided compelling data warranting expansion of the program to other provinces.

In India, (16), on follow-up of cases with SCD (n = 5) and carriers (n=61) after initial SCD screening, the study showed no death by 2 years of age. Morbidities seen included anemia and splenomegaly at 1.5 to 2.5 years in 2 of the 5 cases.

Harm of Screening for SCD

There were no harms directly related to screening of SCD in newborns.(2,7)

According to the Canada STE, a low-quality HTA (6) did not find any evidence of physical risks related to screening or early treatment, except for a possible anaphylactic reaction from penicillin, which is rare.(2) Error rates for screening techniques (IEF or HPLC) for SCD diagnoses were minimal. Using a different second-line test on a patient's initial blood sample reduced the number of residual false positives before the results were relayed to families. The main concerns of false positive screening results were related to parent's anxiety while waiting for the confirmation of the diagnosis. False negative results may delay the starting of penicillin prophylaxis and parental education, but these are rare.(2)

Five observational studies (2 prospective, 3 retrospective cohorts) did not report harms related to newborn screening in terms of overdiagnosis, false positive screening results and psychosocial aspects.(2)

Based on the UK National Screening Committee (2016), a single relevant study suggests that the SCD screening in the US has not led to widespread harm. There is no sufficient evidence to suggest that the evidence supporting national SCD screening program needs to be reviewed in depth or stopped.(17)

4.3.3 Diagnostic Performance of Screening Tests

Diagnostic Performance of HPLC and IEF in Diagnosing SCD

Based on the STE assessment done in Alberta Canada, both high performance liquid chromatography (HPLC) and isoelectric focusing (IEF) are valid tests for newborn screening of SCD, with estimated sensitivity and specificity above 99%.(2) Reported PPV (99.81 to 100%) and NPV (99.99 to 100) were also high. Tandem mass spectrometry (MS/MS) may have potential for newborn screening for SCD, but it can only detect certain variants of Hb, and has not been fully validated (Appendix D and E).

4.3.4 Cost Implication

There are no locally published economic evaluation studies that estimate the cost-effectiveness/ benefit/ utility of newborn screening for SCD.

The table below shows the current estimated cost for the screening of SCD through HPLC using dried blood samples and the cost of its confirmatory test (Hb S Targeted Sanger Sequencing) in the Philippines. The total cost for both screening and confirmatory test for every newborn with positive SCD screening is approximately Php 6,550.

Table 1. Estimated cost of screening and confirmatory test for SCD for every newborn

Parameter	Screening intervention	Confirmatory Test ^a
	HPLC (from ENBS dried blood spot sample)	Hb S Targeted Sanger Sequencing
Unit cost of screening intervention (PHP)	Php 1,750.00 (Covered by PhilHealth)	Php 2,300.00
Additional costs	-	Php 2,500 (DNA extraction from blood sample)
Total	Php 1,750	Php 4,800

^a From the Molecular Genetics Laboratory, National Institutes of Health, Manila Philippines

Cost-effectiveness of screening versus no screening for SCD (2)

Based on an STE done in Alberta Canada in 2016, the estimated added/incremental cost of screening for SCD is CAN\$ 12.21 (~Php 457.00) compared to no screening which costs CAN\$ 2.21 (~Php 87.72). The incremental cost-effectiveness ratio (ICER) is CAN\$ 2,620.73 (~Php 98,094) (Table 2).

Table 2. Cost-effectiveness of screening versus no screening for SCD*

	Lifelong cost per infant screened (Canadian \$)	Life year	Incremental cost (Canadian \$)	Life year saved	ICER
No SCD Screening	\$2.21 (Php 78.90)**	79.04959			
SCD Screening	\$14.42 (Php 514.80)	79.05425	\$12.21 (Php 435.90)**	0.00466	\$2,620.73 (Php 93,560)**

*The model was intended to capture the incremental cost associated with adding SCD to an existing program. The cost inputs reflect incremental costs and not absolute costs.

** Estimated costs in Philippine peso are based on exchange rate conversions at the time of publication of the Canada STE (March 2016): 1 CAN\$ = Php 35.7

Appendix F shows the additional cost inputs and their sources for the confirmatory tests, treatment and subsequent consultations for patients with SCD after screening.(2)

4.3.5 Equity, Acceptability, and Feasibility

In the Philippines, screening for SCD is currently being done under the expanded newborn screening program that started in 2014, however, studies on patient preferences, acceptability and equity are still scarce.

In a cross-sectional multi-site study done in three hospitals in Saint Lucia (2017), the knowledge and attitudes toward heel prick (HP) testing for SCD of healthcare workers (n=70) and mothers (n=132) were assessed using surveys and focused group discussions (FGD). Around 63.6% of mothers reported knowing the reason why the newborns would be tested. In addition, 83.3% indicated that the test will be beneficial for the baby and 89.4% stated that early detection may help if a genetic disease is diagnosed. However, 88.6% said that more information on the heel-prick test was necessary. In the FGDs, the mothers mentioned that they were generally not concerned about the pain of the heel prick test might bring about to the baby (18).

In a study done in Nigeria (2018), attitudes and acceptability of newborn screening for SCD among different socio-demographic groups including undergraduate students, health professionals, parents of children with SCD and patients with SCD were assessed. Among 1301 participants, 86% (n=1119) supported newborn screening but there was a statistically significant relationship between support for NBS and age (21 to 30 years supports NBS versus 41- 60 years old, p = 0.003), educational status (secondary level, p = 0.00) and religion (p = 0.00). The study suggests that the main barriers to the use of NBS are likely to be financial and practical more than social or cultural (19).

A study in the UK (2016) explored parents experiences on receiving the initial positive newborn screening result for sickle cell disease. Semi-structured, qualitative interviews were done among 22 parents whose children had been diagnosed with SCD or cystic fibrosis through NBS and were less than a year of age. The study revealed that most parents through that initial positive results should be delivered by health professionals with knowledge on the specific condition, preferring that both parents are present. It was also mentioned that genetic counseling should include the impact of the NBS results on parental relationships. There must also be consideration on the needs to be given to strategies that will support parents of babies who have positive screening results in terms of psychological health and assistance in sharing the diagnosis (20).

4.3.6 Recommendations from Other Groups

Group	Recommendation	Strength of recommendation and certainty of evidence
USPSTF / AAFP (2008)	Early detection of sickle cell anemia followed by prophylactic oral penicillin substantially reduces risk of serious infections during the first few years of life. Additional benefits results from pneumococcal conjugate vaccination and parental education about early warning signs of infection. Lastly detection of sickle cell disease permits counseling for family members about disease management and future reproductive decisions.(21)	High Certainty
AAP (2012)	High invasive of pneumococcal disease (IPD) in SCD justifies newborn screening, daily prophylactic penicillin and immunization with pneumococcal conjugate and polysaccharide vaccines.(22)	Not Mentioned
United Kingdom Clinical Practice Guidelines (2010)	All newborns should be screened for SCD. Screening should be extended to infants under 1 year of age newly arrived in the UK. Newborn screening and, when necessary, follow-up testing and referral, should be carried according to the guidelines of the NHS Sickle Cell and Thalassemia Screening Program.(2)	Not mentioned

References

1. Mangla A. Sickle cell anemia [Internet]. StatPearls [Internet]. U.S. National Library of Medicine; 2021 [cited 2021Aug25]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK482164/&lang=en/>
2. Institute of Health Economics. Newborn blood spot screening for galactosemia, tyrosinemia type I, homocystinuria, sickle cell anemia, sickle cell/beta-

- thalassemia, sickle cell/hemoglobin C disease and severe combined immunodeficiency. Edmonton: Institute of Health Economics (IHE) (2016).
- 3. Padilla CD, Therrell BL, Alcausin MM, de Castro RC, Gepte MB, Reyes ME, et al. Successful implementation of newborn screening for Hemoglobin disorders in the Philippines. International Journal of Neonatal Screening. 2021;7(2):30.
 - 4. Pediatrics disease registry program - [cited 2021Aug25]. Available from: <https://pps.ivan.com/search.do>
 - 5. Stanford children's health [Internet]. Stanford Children's Health - Lucile Packard Children's Hospital Stanford. [cited 2021Aug25]. Available from: <https://www.stanfordchildrens.org/en/topic/default?id=sickle-cell-disease-in-children-90-P02327&sid=>
 - 6. Blancqueart IR. *Avis scientifique sur le dépistage néonatal de l'anémie falciforme: état des connaissances et enjeux pour le Québec*. Institut National de Santé Publique du Québec; 2010. Available from: http://www.inspq.qc.ca/pdf/publications/1171_AnemieFalciforme.pdf.
 - 7. Runkel B, Klüppelholz B, Rummer A, Sieben W, Lampert U, Bollig C, et al. Screening for sickle cell disease in newborns: A systematic review. Systematic Reviews. 2020;9(1).
 - 8. Vichinsky E, Hurst D, Earles A, Kleman K, Lubin B. Newborn screening for sickle cell disease: effect on mortality. Pediatrics 1988;81(6):749-755.
 - 9. Yanni E, Grosse SD, Yang Q, Olney RS. Trends in pediatric sickle cell disease-related mortality in the United States, 1983-2002. Journal of Pediatrics 2009;154(4):541-5.
 - 10. Frempong T, Pearson HA. Newborn screening coupled with comprehensive follow-up reduced early mortality of sickle cell disease in Connecticut. Connecticut Medicine 2007;71(1):9- 12.
 - 11. Lee A, Thomas P, Cupidore L, Serjeant B, Serjeant G. Improved survival in homozygous sickle cell disease: lessons from a cohort study. BMJ 1995;311(7020):1600-1602.
 - 12. King L, Fraser R, Forbes M, Grindley M, Ali S, Reid M. Newborn sickle cell disease screening: the Jamaican experience (1995-2006). Journal of Medical Screening 2007;14(3):117- 122.
 - 13. Lobo CLC, Ballas SK, Domingos ACB, Moura PG, do Nascimento EM, Cardoso GP, et al. Newborn screening program for hemoglobinopathies in Rio de Janeiro, Brazil. Pediatric Blood and Cancer 2014;61(1):34-9.
 - 14. Quinn CT, Rogers ZR, McCavit TL, Buchanan GR. Improved survival of children and adolescents with sickle cell disease. Blood 2010;115(17):3447-3452.
 - 15. McGann PT, Ferris MG, Ramamurthy U, Santos B, De O, V, Bernardino L, et al. A prospective newborn screening and treatment program for sickle cell anemia in Luanda, Angola. American Journal of Hematology 2013;88(12):984-989.
 - 16. Panigrahi S, Patra PK, Khodiar PK. Neonatal screening of sickle cell anemia: A preliminary report. Indian Journal of Pediatrics 2012;79(6):747-750.
 - 17. UK National Screening Committee [Internet]. [cited 2021Aug25]. Available from: https://legacyscreening.phe.org.uk/policydb_download.php?doc=647 (March 2015)

18. Alexander S, Belmar-George S, Eugene A, Elias V. Knowledge of and attitudes toward heel prick screening for sickle cell disease in Saint Lucia. *Revista Panamericana de Salud Pública*. 2017;41:1.
19. Nnodu OE, Adegoke SA, Ezenwosu OU, Emadi II, Ugwu N, Ohiaeri C, et al. A multi-centre survey of acceptability of newborn screening for sickle cell disease in Nigeria. *Cureus*. 2018;10(3):e2354.
20. Chudleigh J, Buckingham S, Dignan J, O'Driscoll S, Johnson K, Rees D, et al. Parents' experiences of receiving the initial positive newborn screening (NBS) result for cystic fibrosis and sickle cell disease. *Journal of Genetic Counseling*. 2016;25(6):1215–26.
21. US Preventive Services Task Force. Screening for sickle cell disease in newborns: Recommendation statement [Internet]. American family physician. U.S. National Library of Medicine; 2008 [cited 2021Sept15]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/18540496>.
22. McCavit TL. Sickle cell disease [Internet]. American Academy of Pediatrics. American Academy of Pediatrics; 2012 [cited 2021Sept15]. Available from: <https://pedsinreview.aappublications.org/content/33/5/195>

4.4 Newborn Screening for Thalassemia

RECOMMENDATIONS

Among asymptomatic, apparently healthy newborns, we recommend for the screening of thalassemia using HPLC (BIORAD KIT). (*Strong recommendation, very low certainty of evidence*).

Considerations

The consensus panel considered the following when formulating this recommendation:

- There is a high burden of illness among Filipinos.
- Despite the very low certainty of evidence, the panel agreed that the benefit of early recognition and treatment for those who test positive outweighs the risks.

4.4.1 Burden of disease

Thalassemia involves variations in the structural characteristics in the alpha (α)-globin and beta (β)-globin genes or quantitative differences in α - or β -globin chain production, which can range from asymptomatic to severe phenotypic presentation. Thalassemia is the most common single gene disorder, and has high prevalence in Mediterranean, Middle East, and Asian countries.(1-3) High prevalence of thalassemia is seen in consanguineous marriage and in malaria-endemic regions.(1)

In a study of the Filipino population that underwent newborn screening of hemoglobinopathies, the estimated prevalence of hemoglobinopathies is 1 in 1020 births.(4) α -thalassemias, with their varying phenotypes, have a higher overall frequency while β -thalassemias are clinically more significant. α -thalassemia gene carrier frequency was noted to be 6.8% and 9.1% in Filipinos in a limited population study in Hawaii and Taiwan, respectively.(5-7)

Treatment for thalassemia depends on the phenotype and ranges from no treatment for asymptomatic phenotypes or carriers to transfusion for severe phenotypes. Allogeneic hemopoietic stem cell transplantation (HSCT) is the only method currently available to cure transfusion-dependent thalassemia major; it affords a median 2-year overall survival and thalassemia-free survival were 88±1% and 81±1%, respectively.(8)

A quality of life study reported that quality of life of children with β -thalassemia major is significantly lower than their normal siblings.(9) Another study showed that the quality of life among children with β -thalassemia major had similar results and that the chronic illness would negatively affect parents on their physical, emotional and cognitive levels.(10) A study by Zolaly et al. on thalassemia patients and depression noted that depression symptoms were detected in 60 % of patients, anxiety symptoms were detected in half of the studied group, and stress symptoms were detected in 38.7% of patients.(11) The chronicity of the illness not only affects the patients but also the caregivers. A cross-sectional qualitative study indicated that there is a negative correlation between the caregivers and their psychological well-being. Their

responsibilities as caregiver had a crucial role in increasing their anxiety and stress levels which further affected their psychological and emotional health.(12)

4.4.2 Benefits and Harms of Screening Tests

Despite a comprehensive literature search, there were no recent clinical practice guidelines, systematic reviews, meta-analysis, randomized clinical trials, or cohort studies found to directly address the benefits and harms of newborn screening for thalassemias.

The common strategy for thalassemia screening by measuring red cell indices in blood counts, and then screened for thalassemia phenotype, and if necessary, genetic diagnosis. Early recognition of the disease may play a pivotal role in the outcome of the disease.(13) Mean life expectancy is better in those afflicted that had early recognition and optimal care of thalassemia patients (41.5 years) versus those with no care (3 years), with noted difference between mean life expectancy of 38.5 years gained.(14)

4.4.3 Diagnostic Performance of Screening Tests

Screening for hemoglobinopathies not only focus on clinically significant genotypes, but also on the screening of carrier status. Multiple diagnostic examinations such as isoelectric focusing (IEF), high-performance liquid chromatography (HPLC), capillary electrophoresis (CE), mass spectroscopy (MS) have been used to define abnormal hemoglobin. The newborn screening program in the Philippines uses high-performance liquid chromatography to screen neonates. (4) As of press time, there were no systematic reviews comparing different modalities of newborn screening for thalassemias.

In a observational study of by Allaf et al, 343,036 newborns underwent HPLC screening and they that given an optimum threshold for β -thalassemia screening showed that $HbA \leq 0.16$ multiple of the median had a sensitivity of 100% and a specificity of 95.3%. (15) Other studies performed newborn screening using other modalities such as MS and CE. An observational study by Yu et al involving 1,081 neonates showed using MS to determine alpha (α) versus beta (β) globin level ratio cut-offs ($\alpha:\beta < 0.90$; $\beta+, 1.05-1.50$; and $\beta0, \alpha:\beta \geq 1.50$) to differentiate between normal individuals and patients with thalassemia had a sensitivity 90%. (16) Yang et al did an observational study that screened 1193 neonates for beta-thalassemia using CE; they determined that the cut-off ratio of the α -globin over the fetal globin is >1.4 , the sensitivity is 91.38% and specificity 91.89%. (17) In a study by Wu et al using neonatal cord blood to determine alpha thalassemia status in 1169 newborns, they used an automatic CE system to determine the Hb Bart's levels in cord blood but with a sensitivity of 58% and specificity 100% hence concluding that it is not an ideal screening test. (18) On the other hand, a study in Thailand by Uaprasert et al. of neonatal cord blood to determine alpha-thalassemia, comparing IEF and HPLC; IEF yielded sensitivity and specificity of 85.4% and 72.4%, respectively, while HPLC yielded sensitivity and specificity of 76.4% and 89.5%, respectively. (19)

4.4.4 Cost Implication

Currently, there are cost-effective analysis study of newborn screening that probes into thalassemia patients its impact on clinical outcomes and cost of the disease. The table below summarizes the costing parameters based on the Philhealth package for newborn

screening and subsequent confirmatory examination based on the existing rate (as of August 2021) of the Institute of Human Genetics for biochemical and molecular genetic testing. The initial screening is HPLC of dried blood on filter paper, hemoglobin variants detected will need further confirmatory tests as follows:

Table 1. Estimated cost of screening and confirmatory test

	Test	Cost	Turnaround time
Screening test	Expanded newborn screening	1,500	
Confirmatory tests	Alpha-thalassemia		
	Alpha Globin Multiplex PCR-based mutation detection	5,200	4 weeks
	Alpha Globin StripAssay	15,900	4 weeks
	Beta-thalassemia		
	Beta Globin StripAssay	14,800	4 weeks

Molecular Genetics Laboratory - National Institutes of Health, Manila

Thalassemia can range from no cost of treatment to asymptomatic and carrier states to high cost of transfusion dependent thalassemia that is a big burden to the patients and caretakers. A study by Esmaeilzadeh et al in 2016 based in Iranian thalassemia patients; the average annual cost per patient was estimated \$8, 321.80 (₱424, 628.17 as per exchange rate September 25, 2021 of \$1=₱51.30) in which \$7, 286.80 (₱371, 816.26) was related to direct medical costs, \$461.40 (₱23, 543.40) to direct non-medical costs, \$573.50 (₱ 29, 263.41) to indirect costs.(20)

4.4.5 Equity, Acceptability, and Feasibility

A study by Boardman et al in the UK using mixed-methods (quantitative and qualitative) integrative analysis in UK families shows that most have a strong support for screening both newborn and preconception.(21) Factors that served as hurdles to screening included cultural, social, and religious sectors that see thalassemia as a burden with an associated social stigma.(21)

Patients and caregivers that must deal with chronic illness experience a great deal of physical and emotional disruption. People struggle to understand their options if there is lack of adequate information and right support from frontline health professionals.(22)

Early screening of hemoglobinopathies lead to early detections and prompt management. A study done in India on public health perspective on screening for hemoglobinopathies revealed that parents have no reservations in sharing information of their affected children with their relatives and most of the relatives have accepted the risk of being a carrier.(24) In addition, the communication needs to be improved for all the families to accept the risk of having a child with thalassemia and that there is a need to make the screening more

readily available and to motivate high-risk groups through awareness-raising program.(23)

4.4.6 Recommendations from Other Groups

For other European countries, the thrust of neonatal screening for hemoglobinopathy has primarily focused on sickle cell disease (SCD) and not on thalassemia; countries that have included SCD screening are France, Malta, Spain and UK.(24) UK Thalassemia Society recommends that the diagnosis should be anticipated from antenatal screening, and established by prenatal diagnosis where requested or by neonatal testing.(25) If not, affected infants may be identified through the newborn screening program. The UK National Screening Committee includes the screening of thalassemia as a consequence of screening of SCD that also detects other hemoglobinopathies.(26) The baby and parents should be seen for testing as soon as possible and preferably within two weeks. If the diagnosis has not been made at these stages, and the presentation is a clinical one, then assessment and treatment may be urgent, within one to two days.

In the United States of America, occurrence of thalassemia disorders is increasing. However, there is no national newborn screening for thalassemia in USA but certain hemoglobinopathies such as sickle beta thalassemia can be detected.(27) The state of California screens newborns for hemoglobin H disease utilizing high performance liquid chromatography to measure hemoglobin Barts. For thalassemia patients, complete DNA testing prior to commencement of treatment is required to determine prognosis, appropriate therapy, and family counseling.(28)

In India, hemoglobinopathies are the first among genetic disorders for which a national policy for prevention and control has been framed. The elements of the policy are guided by WHO directives and guidelines on hemoglobinopathies including thalassemia and sickle cell disorders.(29) However, this newborn screening program for thalassemia is optional and does not cover all neonates.(30)

References

1. Goonasekera HW, Paththinige CS, Dissanayake VHW. Population Screening for Hemoglobinopathies. *Annu Rev Genomics Hum Genet.* 2018 Aug 31;19:355-380. doi: 10.1146/annurev-genom-091416-035451. Epub 2018 May 11. PMID: 29751732.
2. Wen W, Guo M, Peng HB, Ma L. Optimization and application of a dried blood spot-based genetic screening method for thalassemia in Shenzhen newborns. *World J Pediatr.* 2019 Dec;15(6):610-614. doi: 10.1007/s12519-018-00222-2. Epub 2019 Apr 27. PMID: 31030386.
3. Zhong YH, Ye LX, Cai XJ, Xie CL, Chen JG. [Gene characteristics and changing trend of neonatal thalassemia in Dongguan, China]. *Zhongguo Dang Dai Er Ke Za Zhi.* 2020 May;22(5):454-459. Chinese. doi: 10.7499/j.issn.1008-8830.1910112. PMID: 32434640; PMCID: PMC7389399.
4. Padilla CD, Therrell BL Jr, Alcausin MMLB, de Castro RC Jr, Gepte MBP, Reyes MEL, Jomento CM, Suarez RCN, Maceda EBG, Abarquez CG, Posecion JEWC, Andal AP, Elizaga ALG, Halili-Mendoza BC, Otayza MPVK, Hoppe CC. Successful Implementation of Newborn Screening for Hemoglobin Disorders in the Philippines. *Int J Neonatal Screen.* 2021 Jun 17;7(2):30. doi: 10.3390/ijns7020030. PMID: 34204320; PMCID: PMC8293152.

5. Hall, J.G.; Yuen, J.; Hunt, J.A.; Chu, B.M.; Hsia, Y.E. Gene frequencies for the alpha-globin variants among Chinese, Laotians and Filipinos. *Am. J. Hum. Genet.* **1991**, *49*, 470, Abstract 2679.
6. Ko, T.M.; Hwa, H.L.; Liu, C.W.; Li, S.F.; Chu, J.Y.; Cheung, Y.P. Prevalence study and molecular characterization of alpha-thalassemia in Filipinos. *Ann. Hematol.* **1999**, *78*, 355–357.
7. Ko, T.M.; Caviles, A.P., Jr.; Hwa, H.L.; Liu, C.W.; Hsu, P.M.; Chung, Y.P. Prevalence and molecular characterization of beta-thalassemia in Filipinos. *Ann. Hematol.* **1998**, *77*, 257–260.
8. Baronciani D, Angelucci E, Potschger U, Gaziev J, Yesilipek A, Zecca M, Orofino MG, Giardini C, Al-Ahmari A, Marktel S, de la Fuente J, Ghavamzadeh A, Hussein AA, Targhetta C, Pilo F, Locatelli F, Dini G, Bader P, Peters C. Hemopoietic stem cell transplantation in thalassemia: a report from the European Society for Blood and Bone Marrow Transplantation Hemoglobinopathy Registry, 2000-2010. *Bone Marrow Transplant.* 2016 Apr;51(4):536-41. doi: 10.1038/bmt.2015.293. Epub 2016 Jan 11. PMID: 26752139 Modell B, Khan M, Darlison M. Survival in beta-thalassaemia major in the UK: data from the UK Thalassaemia Register. *Lancet* (London, England). 2000 Jun;355(9220):2051-2052. DOI: 10.1016/s0140-6736(00)02357-6. PMID: 10885361.
9. Wahyuni, M., M. Ali, N. Rosdiana, and B. Lubis. "Quality of Life Assessment of Children With Thalassemia". *Paediatrica Indonesiana*, Vol. 51, no. 3, June 2011, pp. 163-9, doi:10.14238/pi51.3.2011.163-9.
10. Ayoub MD, Radi SA, Azab AM, Abulaban AA, Balkhoyor AH, Bedair SW, Aljaouni SK, Kari JA. Quality of life among children with beta-thalassemia major treated in Western Saudi Arabia. *Saudi Med J.* 2013 Dec;34(12):1281-6. PMID: 24343469.
11. Zolaly MA, Zolaly FM, Al Belowi L, Shuqdar R, Al Belowi MA Sr, Alwasaidi TA, Albadrani M. Depression, Anxiety, and Stress Symptoms in Patients With Beta Thalassemia Major in Almadinah Almunawwarah, Saudi Arabia. *Cureus.* 2020 Nov 7;12(11):e11367. doi: 10.7759/cureus.11367. PMID: 33304700; PMCID: PMC7721071.
12. Anum, J., & Dasti, R. (2015). *Caregiver Burden, Spirituality, and Psychological Well-Being of Parents Having Children with Thalassemia*. *Journal of Religion and Health*, *55*(3), 941–955. doi:10.1007/s10943-015-0127-1
13. Moorthie S, Blencowe H, W Darlison M, Lawn JE, Mastroiacovo P, Morris JK, Modell B; Congenital Disorders Expert Group. An overview of concepts and approaches used in estimating the burden of congenital disorders globally. *J Community Genet.* 2018 Oct;9(4):347-362. doi: 10.1007/s12687-017-0335-3. Epub 2017 Oct 11. PMID: 29022220; PMCID: PMC6167265.
14. Modell B, Darlison MW, Moorthie S, Blencowe H, Petrou M, Lawn J (2016) Epidemiological methods in community genetics and the Modell Global Database of Congenital Disorders (MGDb) downloaded from UCL Discovery. <http://discovery.ucl.ac.uk/1532179/>
15. Allaf B, Pondarre C, Allali S, De Montalembert M, Arnaud C, Barrey C, Benkerrou M, Benhaim P, Bensaid P, Brousse V, Dollfus C, Eyssette-Guerreau S, Galacteros F, Gajdos V, Garrec N, Guillaumat C, Guitton C, Monfort-Gouraud M, Gouraud F, Holvoet L, Ithier G, Kamdem A, Koehl B, Malric A, Missud F, Monier B, Odièvre MH, Joly P, Renoux C, Patin F, Pissard S, Couque N. Appropriate thresholds for accurate screening for β-thalassemias in the newborn period: results from a French center for newborn screening. *Clin Chem Lab Med.* 2020 Aug 19;59(1):209-216. doi: 10.1515/cclm-2020-0803. PMID: 32813673.
16. Yu C, Huang S, Wang M, Zhang J, Liu H, Yuan Z, Wang X, He X, Wang J, Zou L. A novel tandem mass spectrometry method for first-line screening of mainly beta-thalassemia from dried blood spots. *J Proteomics.* 2017 Feb 10;154:78-84. doi: 10.1016/j.jprot.2016.12.008. Epub 2016 Dec 20. PMID: 28007617.

17. Yang JL, Cai R, Chen DY, Tan JQ, Huang LH. [Screening indices and their cut-off values for full-term neonates carrying β-thalassemia gene]. *Zhongguo Dang Dai Er Ke Za Zhi*. 2018 Dec;20(12):990-993. Chinese. doi: 10.7499/j.issn.1008-8830.2018.12.003. PMID: 30572986; PMCID: PMC7389490.
18. Wu MY, Xie XM, Li J, Li DZ. Neonatal screening for α-thalassemia by cord hemoglobin Bart's: how effective is it? *Int J Lab Hematol*. 2015 Oct;37(5):649-53. doi: 10.1111/ijlh.12376. Epub 2015 May 8. PMID: 25955662.
19. Uaprasert N, Settapiboon R, Amornsiriwat S, Sarnthammakul P, Thanapat T, Rojnuckarin P, Sutcharitchan P. Diagnostic utility of isoelectric focusing and high performance liquid chromatography in neonatal cord blood screening for thalassemia and non-sickling hemoglobinopathies. *Clin Chim Acta*. 2014 Jan 1;427:23-6. doi: 10.1016/j.cca.2013.09.041. Epub 2013 Oct 2. PMID: 24095765. Esmaeilzadeh F, Azarkeivan A, Emamgholipour S, Akbari Sari A, Yaseri M, Ahmadi B, Ghaffari M. Economic Burden of Thalassemia Major in Iran, 2015. *J Res Health Sci*. 2016 Summer;16(3):111-115. PMID: 27840337; PMCID: PMC7191027.
20. Esmaeilzadeh F, Azarkeivan A, Emamgholipour S, Akbari Sari A, Yaseri M, Ahmadi B, Ghaffari M. Economic Burden of Thalassemia Major in Iran, 2015. *J Res Health Sci*. 2016 Summer;16(3):111-115. PMID: 27840337; PMCID: PMC7191027.
21. Boardman FK, Clark C, Jungkurth E, Young PJ. Social and cultural influences on genetic screening programme acceptability: A mixed-methods study of the views of adults, carriers, and family members living with thalassemia in the UK. *J Genet Couns*. 2020 Dec;29(6):1026-1040. doi: 10.1002/jgc4.1231. Epub 2020 Mar 1. PMID: 32114710; PMCID: PMC7754126.
22. Streetly A. Equipping nurses to support genuine patient choice. *Br J Nurs*. 2006 Jun 8-21;15(11):594. doi: 10.12968/bjon.2006.15.11.21225. PMID: 16835526.
23. Thiagarajan A, Bhattacharya S, Sharma N, Srivastava A, Dhar DK. Need for a universal thalassemia screening programme in India? A public health perspective. *J Family Med Prim Care*. 2019 May;8(5):1528-1532. doi: 10.4103/jfmpc.jfmpc_90_19. PMID: 31198708; PMCID: PMC6559078.
24. Loeber JG, Platis D, Zetterström RH, Almashanu S, Boemer F, Bonham JR, Borde P, Brincat I, Cheillan D, Dekkers E, Dimitrov D, Fingerhut R, Franzson L, Groselj U, Hougaard D, Knapkova M, Kocova M, Kotori V, Kozich V, Kremezna A, Kurkijärvi R, La Marca G, Mikelsaar R, Milenkovic T, Mitkin V, Moldovanu F, Ceglarek U, O'Grady L, Oltarzewski M, Pettersen RD, Ramadza D, Salimbayeva D, Samardzic M, Shamsiddinova M, Songailiené J, Szatmari I, Tabatadze N, Tezel B, Toromanovic A, Tovmasyan I, Usurelu N, Vevere P, Vilarinho L, Vogazianos M, Yahyaoui R, Zeyda M, Schielen PCJ. Neonatal Screening in Europe Revisited: An ISNS Perspective on the Current State and Developments Since 2010. *Int J Neonatal Screen*. 2021 Mar 5;7(1):15. doi: 10.3390/ijns7010015. PMID: 33808002; PMCID: PMC8006225.
25. Standards for the Clinical Care of Children and Adults with Thalassaemia in the UK, 3rd Edition (2016). [cited 20 Aug 2021]
26. Lobitz S, Telfer P, Cela E, Allaf B, Angastiniotis M, Backman Johansson C, Badens C, Bento C, Bouva MJ, Canatan D, Charlton M, Coppinger C, Daniel Y, de Montalembert M, Ducoroy P, Dulin E, Fingerhut R, Frömmel C, García-Morin M, Gulbis B, Holtkamp U, Inusa B, James J, Kleanthous M, Klein J, Kunz JB, Langabeer L, Lapouméroulie C, Marcao A, Marín Soria JL, McMahon C, Ohene-Frempong K, Périmi JM, Piel FB, Russo G, Sainati L, Schmugge M, Streetly A, Tshilolo L, Turner C, Venturelli D, Vilarinho L, Yahyaoui R, Elion J, Colombatti R; with the endorsement of EuroBloodNet, the European Reference Network in Rare Haematological Diseases. Newborn screening for sickle cell disease in Europe: recommendations from a Pan-

- European Consensus Conference. Br J Haematol. 2018 Nov;183(4):648-660. doi: 10.1111/bjh.15600. Epub 2018 Oct 18. PMID: 30334577.
27. State-Based Monitoring for Selected Hemoglobinopathies [Internet]. Centers for Disease Control and Prevention; c2021 [updated 2020 December 16; cited 2021 September 24]. Available from: <https://www.cdc.gov/ncbddd/hemoglobinopathies/features/keyfinding-state-based.html>
28. Standards of Care Guidelines for Thalassemia (2012). Comprehensive Thalassemia Center at Children's Hospital Oakland [cited 22 Aug 2021]. Available from: <https://thalassemia.com/documents/SOCGuidelines2012.pdf>
29. Prevention and Control of Hemoglobinopathies in India (2016). Ministry of Health and Family Welfare. Available from: https://nhm.gov.in/images/pdf/in-focus/NHM_Guidelines_on_Hemoglobinopathies_in_India.pdf
30. Thiagarajan A, Bhattacharya S, Sharma N, Srivastava A, Dhar DK. Need for a universal thalassemia screening programme in India? A public health perspective. J Family Med Prim Care. 2019 May;8(5):1528-1532. doi: 10.4103/jfmpc.jfmpc_90_19. PMID: 31198708; PMCID: PMC6559078.

4.5 Newborn Screening for G6PD Deficiency

RECOMMENDATIONS

Among asymptomatic, apparently healthy newborns, we recommend for the screening of G6PD deficiency using fluorescence assay (PE neonatal kit). (**Strong recommendation, very low certainty of evidence.**)

Considerations

The consensus panel considered the following when formulating this recommendation:

- There is a high prevalence of this disease.
- The panel agreed that the management is cost-effective and preventative against the development of kernicterus, which would greatly benefit those patients that screen positive.

4.5.1 Burden of disease

G6PD deficiency is the most common X-linked enzyme deficiency and is estimated to affect 400 million people worldwide with about 11 million infants born with this disorder yearly.(1) Global incidence has recorded highest rates in Africa, Southeast Asia and the Middle East with more than 33% of the global health burden of hyperbilirubinemia is a result of G6PD deficiency.(2)

The American Academy of Pediatrics (AAP) clinical practice guideline on the management of neonatal hyperbilirubinemia recognizes G6PD deficiency as a major etiologic risk factor for the development of severe hyperbilirubinemia, with kernicterus being one of the most severe and life-threatening clinical consequences. Once an infant is diagnosed with G6PD deficiency, the relative risk for severe hyperbilirubinemia is said to be 3.5 fold higher than those who are not deficient and is strongly associated with mortality and long term neurodevelopmental impairment.(3,4)

The USA Pilot Kernicterus Registry has also listed G6PD deficiency as a likely cause for kernicterus in 20.8% of cases with those with African American and South East Asian descent to be the most affected. Acute manifestations are triggered by drug-sensitive or oxidant stress-mediated hemolysis, sepsis, ingestion or exposure to certain foods, dyes, and chemicals. Chronic sequelae include hemolytic anemia, cholelithiasis, splenomegaly and protection against malaria.(5)

In 2020, there were 248,285 confirmed new cases in the Philippines, with a prevalence rate of 1:60.(6) The Philippine Pediatric Society has reported 524 cases of hemolytic anemia due to G6PD Deficiency since 2006.(7) Patients identified with G6PD deficiency do not require specific treatment. Timely counseling on strict avoidance of food and environmental triggers should be done to prevent a hemolytic crisis.

4.5.2 Benefits and Harms of Screening Tests

Hospitalization Rate due to Hemolytic Crisis (3 Observational Studies, Very Low Certainty of Evidence)

Three observational studies were identified that displayed the benefit of screening for G6PD deficiency in decreasing the number of hospitalized infants due to a hemolytic crisis, with some requiring exchange transfusion.

Mallouh et al compared outcomes of infants who were screened for G6PD deficiency and those who were not screened over a 5-year period.(8) Among the 33,943 infants screened, 6,246 (18.4%) were diagnosed with G6PD deficiency. These infants were identified prior to discharge from the nursery, their serum bilirubin levels were checked if clinically jaundiced and subsequently managed with phototherapy or exchange transfusion and the mothers were counselled as to how to avoid oxidizing agents. Among those who screened positive for G6PD deficiency, 42 (0.7%) required hospitalization and exchange transfusion. A total of 10,005 infants were not screened during the same time period and among these, 7 infants were hospitalized due to kernicterus (clinically assessed to have no other etiology apart from possible G6PD deficiency).

On post hoc analysis of this study, screening for G6PD deficiency showed an odds ratio of 1.79 (95% CI 0.81 to 3.99 p = 0.15) indicating a non-significant risk with a trend towards harm in decreasing hospitalization due to a hemolytic crisis. The certainty of evidence was downgraded due to a serious risk of bias as the infants in the control group did not undergo a confirmatory test for G6PD deficiency, as well as serious indirectness and imprecision.

Two other cross-sectional studies done by Cohan et al and Meloni et al evaluated the efficacy of their national screening program for G6PD deficiency.(9,10) A total of 1502 newborns were included in the studies and divided into 2 groups. 1061 patients with a hemolytic crisis due to G6PD deficiency who were admitted before the G6PD screening program and 441 patients who were admitted due to G6PD deficiency after G6PD screening. This indicated that among the hospitalizations due to hemolytic crisis within the 3-10 year follow up period, 70.6% were from the period wherein a screening program has not yet been implemented. The odds ratio however is not estimable due to lack of data given in both studies.

Prevention of Kernicterus (1 Observational Study, Very low Certainty of Evidence)

The study mentioned above done by Mallouh et al also looked at cases of kernicterus among those that were admitted at the medical facility. There were no cases of kernicterus among the infants who were screened and 7 cases were noted among those who were not screened. On post hoc analysis, screening for G6PD deficiency had an odds ratio of 0.0197 (95% CI 0.001 to 0.34 p=0.007) showing a significant benefit of G6PD screening in the prevention of kernicterus.

4.5.3 Diagnostic Performance of Screening Tests

Diagnostic Performance of the Fluoroimmunoassay (Perkin Elmer)

(1 Observational Study, Very Low Certainty of Evidence)

Verma et al evaluated the performance of newborn testing of G6PD Deficiency (along with Congenital hypothyroidism and Congenital Adrenal Hyperplasia).(11) A total of 13,376 newborns were screened using time resolved fluoroimmunoassay. Those who were screened positive were recalled for confirmatory testing using quantitative analysis using an enzyme-linked immunosorbent assay (ELISA). Among the newborns screened, 138 screened positive for G6PD Deficiency. 20% (30/138) of those recalled was reported to be false positive. The study reported a sensitivity of 1.00 (95% CI) and specificity of 0.99 (95% CI), however true and false negative results were not indicated.

4.5.4 Cost Implication

A cost effectiveness study on newborn screening in the Philippines done in 2009 demonstrated that the benefits of a newborn screening program far outweighs the societal cost of no screening or a “do nothing” alternative.(12) Potential costs incurred in the treatment, management and productivity loss for late diagnosed cases displayed a 15.44 net benefit (cost shown in USD) of G6PD screening as compared to no screening. The table below shows the current unit cost of screening through the expanded newborn screening program of the Philippine Health Insurance Corporation (Philhealth).(13)

Parameter	Cost of ENBS (in Php)	Cost of Confirmatory Screening (in Php)
Unit cost of screening	1,750 (Expanded NBS)	400

The cost of hospitalization due to complications of G6PD deficiency are covered by Philhealth with the following case rates for Level 3 admission:

Diagnosis/Procedure	Cost (in Php)
Hemolytic Anemia due to G6PD Deficiency	10,000
Kernicterus	7,400
Neonatal Jaundice due to Hemolytic Anemia	7,400
Exchange Blood Transfusion	5,680

4.5.5 Equity, Acceptability, and Feasibility

One study was done through an electronic survey in the United States among 472 pediatric providers that stated that screening for G6PD has been helpful and influential in their management which included more counseling on jaundice and follow up and avoidance of hemolytic crisis triggers. The survey also showed their perception of how parents deal with a diagnosis of G6PD deficiency with 74% understanding that this is a risk for neonatal hyperbilirubinemia but unlikely to cause significant problems in the future and 13% of providers felt that the diagnosis seemed to increase parental worry.(14)

4.5.6 Recommendations from Other Groups

There are no existing clinical practice guidelines or recommendations from other groups regarding screening for G6PD deficiency among asymptomatic newborns. However, the World Health Organization recommends screening all newborns in populations with a prevalence of 3-5% or more alongside vigilant monitoring for the development of jaundice and institution of therapy as early as possible. The WHO has also emphasized the value of parental and health worker education to prevent complications of G6PD Deficiency.(15) Countries that have implemented universal newborn screening for G6PD deficiency such as Singapore, Greece, Saudi Arabia, Taiwan, Hong Kong and the Philippines have shown decreased incidence of severe hyperbilirubinemia and kernicterus.(16,17)

References

1. Silao C, Padilla C, Uy E, Delos Reyes E. Glucose-6-Phosphate Dehydrogenase Deficiency in Filipino Neonates with Jaundice. *Acta Medica Philippina*. 2009; 43 (2): 22-25
2. Del Favero J, Jnah A, Newberry D. Glucose-6-Phosphate Dehydrogenase Deficiency and the Benefits of Early Screening. *Neonatal Network*. 2020; 39 (5): 270-282
3. Kaplan M, Hammerman C, Bhutani V. Parental education and the WHO neonatal program: A quarter century later. *Journal of Perinatology*. 2015; 35:779-784
4. Arain Y, Bhutani V. Prevention of Kernicterus in South Asia: Role of Neonatal G6PD deficiency and its identification. *Indian Journal of Pediatrics* 2014; 81 (6): 599-607
5. Bhutani, V. Jaundice due to Glucose-6-Phosphate Dehydrogenase Deficiency. *NeoReviews*. 2012; 13(6):e166-177
6. Newborn Screening Reference Center. Prevalence of Hemoglobinopathies as of December 2020. Available from: <https://newbornscreening.ph/images/stories/ResourcesPrevalence/prevalence-2020.pdf> [Accessed August 5, 2021]
7. Philippine Pediatric Society. ICD 10 Registry. Available from: <https://pps.org.ph/icd-10-registry/> [Accessed August 26, 2021]
8. Mallouh A, Imseeh G, Abu-Osba Y, Hamdan J. Screening for glucose-6-phosphate dehydrogenase deficiency can prevent severe neonatal jaundice. 1992; 12:391-395
9. Cohan N, Karimi M et al. The efficacy of a neonatal screening programme in decreasing the hospitalization rate of patients with G6PD deficiency in Southern Iran. *Journal of Medical Screening*. 2010; 17 (2) 66-67
10. Meloni T, Forteleoni G. Marked Decline of Favism after Neonatal Glucose -6-Phosphate Dehydrogenase Screening and Health Education: The Northern Sardinian Experience. *Acta Haematologica*,1992; 87(1-2): 29-31
11. Verma J, Papai R, Thomas D. Newborn Screening for Congenital Hypothyroidism, Congenital Adrenal Hyperplasia and Glucose 6 Phosphate Dehydrogenase Deficiency for Improving Health Care in India. *Journal of Pediatric Intensive Care* 2020; 9:40-44
12. Padilla C, Dans L, Tamondong, M et al. Cost Benefit Analysis of the Newborn Screening Program of the Philippines. *Acta Medica Philippina*. 2009; 43 (2): 46-52
13. Philhealth. Philhealth Circular No. 2018-00214 Enhancement of Newborn Care Package. Available from: https://www.philhealth.gov.ph/circulars/2018/TS_circ2018-0021.pdf [Accessed August 26, 2021]
14. Bernardo J, Nock M. Pediatric Provider Insight into Newborn Screening for G6PD Deficiency. *Clinical Pediatrics*. 2015;54(6):575-578.
15. WHO Working Group. Glucose-6-phosphate dehydrogenase deficiency. *Bulletin of the World Health Organization*. 1989; 67(6): 601-611
16. Watchko J, Kaplan M, Stark A, Stevenson D, Bhutani V. Should we screen newborns for glucose-6-phosphate dehydrogenase deficiency in the United States? *Journal of Perinatology*. 2013;33(7):499-504.
17. Padilla C, Therrel B. Newborn Screening in the Asia Pacific Region. *Journal of Inherited Metabolic Disease*. 2007; 30:490-506

4.6 Newborn Screening for Homocystinuria and Methionine Adenosyltransferase Deficiency

RECOMMENDATIONS

1. Among asymptomatic, apparently healthy newborns, we recommend against the screening of homocystinuria. (**Strong recommendation, very low certainty of evidence**).
2. Among asymptomatic, apparently healthy newborns, we recommend against the screening of methionine adenosyltransferase deficiency. (**Strong recommendation, very low certainty of evidence**).

Considerations

The consensus panel considered the following when formulating this recommendation:

- The very low burden of disease among Filipinos coupled with the high cost of confirmatory testing and treatment of the disease led the consensus panel to decide against the screening.

4.6.1 Burden of disease

Homocystinuria (HCU)

The worldwide true incidence of homocystinuria is unknown but frequency has been based on existing screening programs. Data varies from country to country with the highest incidence reported in Qatar at 1:1,800, and Europe at 1:6,400 to 1:65,000. North American and Japanese frequencies are estimated at 1:200,000-300,000 and 1:900,000 respectively.(1-4) Locally, as of June 30, 2021 the Philippines has detected 2 cases of homocystinuria since 1996: 1 case upon introduction of the 5-test Newborn Screening (NBS) Program (HCY, PKU, CH, CAH, Gal), and another tested because of presenting symptoms (non-NBS case). No new case has been reported since 2019 after the implementation of compulsory expanded NBS program. Incidence has been recorded at 1:1,604,500.(5,6)

The spectrum of the disease will depend on the affected gene and the severity of mutations. Patients may present with neurodevelopmental and ocular symptoms, connective tissue involvement, or thromboembolism. Milder forms, especially the pyridoxine-responsive variants, may remain asymptomatic or manifest later in adulthood. Significant morbidity may be expected if diagnosis is delayed.(2,7) Treatment involves a combination of low methionine/low protein diet and betaine which provides methyl groups for homocysteine remethylation.(2)

Methionine Adenosyltransferase Deficiency (MAT)

There has been difficulty in ascertaining the true incidence of MAT deficiency since most patients are asymptomatic, although some cases may develop neurological symptoms later in life.(8,9) Reported incidence in Spain based on data from three centers is 1:22,874.(8) In 2015, Chien et al consolidated data from relevant publications of known patients (n=64, 24/64 detected via NBS, 18/24 screened developed neurocognitive abnormalities later in life) with MAT deficiency found to be either homozygotes or heterozygotes for MAT1A mutation.(10) Philippine newborn screening data as of December 2020, reported only 1 case of hypermethioninemia (no genetic testing yet to confirm MAT deficiency) detected out of 3,209,001 screened babies.(6)

Heterozygous individuals generally have a benign course even without treatment, but homozygotes and compound heterozygotes can present with neurocognitive symptoms.(2) Due to limited data, no consensus with a high level of evidence exists yet, and optimal management remains to be adequately defined and there has been no definitive answer to the question of the best management for severe cases.(8,10,11) Two therapeutic options have been described for MAT deficiency: methionine diet restriction and oral administration of AdoMet. A methionine restricted diet has been suggested in asymptomatic patients to maintain levels of plasma methionine below 600 umol/L, and management is augmented with AdoMet supplementation if symptoms develop despite diet restriction.(9,10)

4.6.2 Benefits and Harms of Screening Tests

No randomized controlled trials (RCTs) comparing outcomes of patients with treatment versus no treatment were available both for homocystinuria and MAT deficiency. However, based on a summary by the Institute of Health Economics of Canada, no harm nor any adverse event and safety concern has been reported related to the screening for homocystinuria. Few studies, and low-level evidence (case reviews) suggest that treatment may reduce morbidity, but screening has not shown potential to prevent mortality. Some evidence indicate that early treatment may decrease neurocognitive and ocular issues compared to those who are non-responsive to treatment, late-detected or not-treated (Table 1).(3)

Table 1. Effectiveness of screening and early detection on the development of disorder-sequelae(3)

	With screening and early detection	Without screening and delayed detection
Developmental delay	9.00% - 18.00%	89.00% - 99.00%
Spinal osteoporosis	5.00% - 10.00%	50.00% - 67.00%
Lens dislocation	5.00% - 10.00%	100.00%
Mortality rate	10.00% - 20.00%	14.00% - 24.00%

A case review of 18 adult patients diagnosed with homocystinuria in Japan concluded that HCU detected by NBS had more favorable outcomes compared with those clinically diagnosed. However, symptoms might not be preventable and long-term outcomes may worsen even in patients detected by NBS.(12)

Correspondingly, newborn screening for MAT deficiency carries the same lack in reportable harm for the procedure. Screening has been postulated to be justified since most cases are detected by NBS although most are mild and asymptomatic. More clinical data is required to support any recommendation. While the most common cause for hypermethioninemia is the dominant form due to the heterozygosity for mutation in MAT1A which is considered benign, some cases of MAT I/III deficiency can develop neurological symptoms later in life.(9) No randomized control trial was available to compare outcomes for patients screened vs. not screened, and treatment vs. no treatment. Chien et al (2015) reviewed instead the course and management of diagnosed cases and based on the outcomes it was stated that the usefulness of methionine restriction is limited. Due to the limited data, there has been no defined optimum strategy for the management of severe MAT I/III deficiency and benefit from screening has not been fully established.(10)

4.6.3 Diagnostic Performance of Screening Tests

Pooling of estimates could not be done for the accuracy of screening tests for homocystinuria and MAT deficiency due to a very low number of studies, incomplete follow up of patients, and differences in reference standard used. Available evidence showing the diagnostic accuracy of the screening tests for each is presented below.

Homocystinuria

Homocystinuria may be screened by measuring either methionine or homocysteine levels on dried blood spot. The disease is then confirmed via molecular genetic testing and mutation analysis.(4,13) Historically, neonatal screening for homocystinuria was done by measurement of methionine concentration in dried blood spots (DBS). However, the test accuracy has been low and approximately 20% of cases are missed.(14) Thereafter, total homocysteine was also measured, either as first or second tier in dried blood spots. Due to the limitations of using elevated methionine as the first-tier screening test, quantitative analysis of total homocysteine (tHcy) was developed.(15)

Two articles with very low certainty of evidence (Appendices F, G, H) demonstrate the sensitivity of using measurements of methionine and tHcy on DBS. In 1998, Yap and Naughten retrospectively evaluated Ireland's 25 years of experience in screening for Homocystinuria using measurement of methionine levels in DBS. From 1,580,000 newborns screened, 21 cases were identified and four were missed, giving a sensitivity of 84%. (16) Gan-Schreier et al estimated the sensitivity of methionine levels on DBS at 50%, and tHcy on DBS at 100%, based on 46,046 newborns screened in 2006-2009 in Qatar. No follow-up of babies who screened negative was done, hence true negatives were not identified.(13)

MAT Deficiency

One retrospective study with very low certainty of evidence (Appendix I) reviewed the use of methionine levels in DBS in 68,624 newborns in three centers in Spain to screen for MAT deficiency. Eighteen cases were detected, and sensitivity was measured at 100%. True negatives were also not identified due to incomplete follow-up of screen negative newborns.(8)

4.6.4 Cost Implication

At least three journals report on the economic and cost-utility analysis of newborn screening, which includes HCU but not MAT I/III deficiency.(3,17,18) Two out of three reports deemed that it was cost-effective to screen for HCU along with other rare metabolic disorders. NSC Health Economics estimates approximately 30% more cost in the lifetime management for HCU patients not screened (Table 2).

Expense for diagnostic testing for screen positives in the Philippines is detailed in Table 3. Consequently, management outlay is provided in Table 4 for confirmed cases and this will include dietary restriction and concomitant Betaine and B6 administration (for B6 responsive cases) depending on physician prescription. Under no screening, it is assumed that diagnosis costs occur in the year of symptomatic presentation but are the same as for screen diagnosis. Similarly, it is assumed management costs once diagnosed are comparable. Dietary management is expected to be lifelong.(17,19)

Table 2. Summary of comparison of different economic analyses for screening for HCU

	NSC Health Economics Report(17) (Pounds)	Children's Health Services Research, Indiana (USD)(18)	Institute of Health Economics, Canada (Canadian dollar)(3)
Key consideration	Cost-effective	Cost-effective	Not cost-effective: minimal impact due to rarity of disease
Confirmation cost in false positives/Cost of false-positive result	£ 475 (£ 428-£ 521)	\$ 2600	Not mentioned
Discounted lifetime cost of management without screening	£ 172,197 (£ 127,031-£ 229,408)	Not mentioned	Not mentioned
Discounted lifetime cost of management with screening	£ 235,730 (£ 174,305-£ 311,839)	Not mentioned	Not mentioned

Table 3. Summary of Costs for Screening and Diagnosis for HCU/MAT I/III

Screening Intervention	Confirmatory test				
	Urine Metabolic Screen	Comprehensive Urine Metabolic Profile	Urine Organic Acid Analysis	Plasma Quantitative Amino Acid Analysis	Genetic testing
Cost Phi 1750 (as part of the expanded newborn screening)	Phi 5,900 ^a Phi 4,250 ^b	Phi 14,000 ^a Phi 10,300 ^b	Phi 12,500 ^a Phi 9,100 ^b	Phi 17,500 ^a Phi 12,300 ^b	~Phi 12500 (USD 250)

*Prices obtained from the Biochemical Genetics Laboratory of the Institute of Human Genetics (tel no. 02-5310-1780 loc 103)

^aRegular price

^bIndigent/Charity price

Table 4. Cost of available medications/medical food for patients with HCU/MAT I/III

Medication/Medical Food	Disease	Unit Cost
Betaine 500mg in 1 ml, 100 ml, powder for oral solution	HCU	Phi 19,500/bottle
Betaine (98%) anhydrous powder, 100mg/bottle	HCU	Phi 16,412.37/bottle
Vitamin B6: Pyridoxine oral, 100mg/tab, 100 tabs per bottle	HCU	Phi 800/bottle
HCU Medical Food: Amino acid-based methionine-free powdered formula	HCU/MAT I/III	Phi 9,338/can

4.6.5 Equity, Acceptability, and Feasibility

No qualitative study was found describing the patients' values and preferences with regards to screening for HCU and MAT deficiency. To assess the patient's perspective and acceptability of the screening program, the Foundation for Genomics and Population Health in England solicited the opinion of "Climb", the umbrella group for inherited metabolic conditions. The group was strongly in favor of an expansion of newborn screening as it has the potential to prevent death and substantial disability, with improvement of quality of life. Earlier diagnosis also empowers families and despite the rarity of these conditions, they emphasized that attention to these disorders should be similar to common diseases.(19)

In the Philippines, the expanded newborn screening has been introduced and operational since 2014 and mandated to be mandatory for all newborns in 2019.

4.6.6 Recommendations from Other Groups

Guidelines available are detailed in Table. These guidelines were written as part of the European network and registry for homocystinuria and methylation defects (E-HOD) which aims to provide and formulate recommendations for the screening, diagnosis, and treatment for this rare group of disorders. Most of the data available came from case

reports, case series, or physician experience hence, the recommendations had low levels of evidence.(2,4,9)

Table 5. Summary of recommendations from different clinical practice guidelines

Guidelines	Year Published	Recommendation for Screening	Grade of Recommendation
Newborn Screening for Homocystinuria and Methylation Disorders: Systematic Review and Proposed Guidelines(2)	2015	<p>NBS for CBS deficiency can be performed by detecting elevated methionine, methionine-to-phenylalanine ratio and/or hyperhomocysteinemia in DBS. However, total homocysteine has only exceptionally been used as a primary marker.</p> <p>Individuals with MAT I/III deficiency have predominantly detected due to increased methionine concentrations. There is no primary marker to differentiate between MATI/III and CBS deficiency.</p>	C (level 3 evidence: non-analytical studies such as case reports and case series)
Guidelines for the diagnosis and management of Cystathione Beta-Synthase Deficiency (CBS)(4)	2016	Newborn screening for homocystinuria can be detected elevated Met, Met-to-phenylalanine ratio and/or tHcy. Specificity of Met as a primary marker may be substantially increased by analyzing tHcy as a second-tier marker.	C (level 3 evidence: non-analytical studies such as case reports and case series)
Consensus recommendations for the diagnosis, treatment, and follow-up of inherited methylation disorders(9)	2015	Most patients with MAT I/III deficiency have been detected by newborn screening. More clinical and laboratory data are required to make a firm recommendation to initiate newborn screening.	D (level 4 evidence – expert opinion)

References

1. Accinni R, Campolo J, Parolini M, De Maria R, Caruso R, Maiorana A et al. Newborn screening of homocystinuria: quantitative analysis of total homocyst(e)ine on dried blood spot by liquid chromatography with fluorimetric detection. *Journal of Chromatography B*. 2003;785(2):219-226.
2. Huemer M, Kožich V, Rinaldo P, Baumgartner M, Merinero B, Pasquini E et al. Newborn screening for homocystinurias and methylation disorders: systematic review and proposed guidelines. *Journal of Inherited Metabolic Disease*. 2015;38(6):1007-1019.9

3. Institute of Health Economics | [Internet]. Ihe.ca. 2016 [cited 5 August 2021]. Available from: <https://www.ihe.ca/advanced-search/newborn-blood-spot-screening-for-galactosemia-tyrosinemia-type-i-homocystinuria-sickle-cell-anemia-sickle-cell-beta-thalassemia-sickle-cell-hemoglobin-c-disease-and-severe-combined-immunodeficiency>
4. Morris A, Kožich V, Santra S, Andria G, Ben-Omran T, Chakrapani A et al. Guidelines for the diagnosis and management of cystathione beta-synthase deficiency. *Journal of Inherited Metabolic Disease.* 2016;40(1):49-74.
5. Institute of Human Genetics. Metabolic Registry. Manila: National Institutes of Health, University of the Philippines Manila; 2021.
6. Prevalence as of December 2020 [Internet]. Newbornscreening.ph. 2021 [cited 22 August 2021]. Available from: <https://www.newbornscreening.ph/images/stories/ResourcesPrevalence/prevalence-2020.pdf>
7. Keller R, Chrastina P, Pavlíková M, Gouveia S, Ribes A, Kölker S et al. Newborn screening for homocystinurias: Recent recommendations versus current practice. *Journal of Inherited Metabolic Disease.* 2019;42(1):128-139.
8. Couce M, Bóveda M, García-Jiménez C, Balmaseda E, Vives I, Castiñeiras D et al. Clinical and metabolic findings in patients with methionine adenosyltransferase I/III deficiency detected by newborn screening. 2013.
9. Barić I, Staufner C, Augoustides-Savvopoulou P, Chien Y, Dobbelaere D, Grünert S et al. Consensus recommendations for the diagnosis, treatment and follow-up of inherited methylation disorders. *Journal of Inherited Metabolic Disease.* 2016;40(1):5-20.
10. Chien Y, Abdenur J, Baronio F, Bannick A, Corrales F, Couce M et al. Mudd's disease (MAT I/III deficiency): a survey of data for MAT1A homozygotes and compound heterozygotes. 2015.
11. Nashabat M, Al-Khenaizan S, Alfadhel M. Methionine adenosyltransferase I/III deficiency: beyond the central nervous system manifestations. *Therapeutics and Clinical Risk Management.* 2018;Volume 14:225-229.
12. Yamada K, Yokoyama K, Aoki K, Taketani T, Yamaguchi S. Long-Term Outcomes of Adult Patients with Homocystinuria before and after Newborn Screening. *International Journal of Neonatal Screening.* 2020;6(3):60.
13. Gan-Schreier H, Kebbewar M, Fang-Hoffmann J, Wilrich J, Abdoh G, Ben-Omran T et al. Newborn Population Screening for Classic Homocystinuria by Determination of Total Homocysteine from Guthrie Cards. *The Journal of Pediatrics.* 2010;156(3):427-432.
14. Bowron A, Barton A, Scott J, Stansbie D. Blood Spot Homocysteine: A Feasibility and Stability Study. *Clinical Chemistry.* 2005;51(1):257-258.
15. Alodaib A, Carpenter K, Wiley V, Wotton T, Christodoulou J, Wilcken B. Homocysteine Measurement in Dried Blood Spot for Neonatal Detection of Homocystinurias. *JIMD Reports.* 2011;:1-6.
16. Yap S, Naughten E. Homocystinuria due to cystathione β -synthase deficiency in Ireland: 25 years' experience of a newborn screened and treated population with reference to clinical outcome and biochemical control. *Journal of Inherited Metabolic Disease.* 1998;21(7):738-747.
17. Expanded newborn screening for inborn errors of the metabolism health economics [Internet]. Legacyscreening.phe.org.uk. 2013 [cited 5 August 2021]. Available from: https://legacyscreening.phe.org.uk/policydb_download.php?doc=411
18. Carroll A, Downs S. Comprehensive Cost-Utility Analysis of Newborn Screening Strategies. *Pediatrics.* 2006;117(Supplement 3):S287-S295.
19. Expanded newborn screening [Internet]. Phgfoundation.org. 2010 [cited 31 August 2021]. Available from: <https://www.phgfoundation.org/report/expanded-newborn-screening>

4.7 Newborn Screening for Tyrosinemia

RECOMMENDATIONS

Among asymptomatic, apparently healthy newborns, we recommend against the screening of tyrosinemia I/II. (*Strong recommendation, very low certainty of evidence*).

Considerations

The consensus panel considered the following when formulating this recommendation:

- The very low burden of disease among Filipinos coupled with the high cost of confirmatory testing and treatment of the disease led the consensus panel to decide against the screening.

4.7.1 Burden of disease

Inborn errors of tyrosine metabolism may manifest as hypertyrosinemia and these include Tyrosinemia type I, II, and III.(1) The three different types differ in incidence with type 1 being the most severe and seen more frequently especially in regions like Turkey, Quebec, and India. The frequency of Tyr 1 varies between 1:16,000 and 1:250,000, compared to Tyr II and Tyr III where cases are fewer than 1 in 250,000. Most patients who present with Tyr II and Tyr III are reported as case reports, with Tyr III being the rarest form.(2,3,4,5) In the Philippines, 10 have been diagnosed with Tyr I (1 in 267,416), and no cases of Tyr II and Tyr III have been detected since the introduction of expanded newborn screening in 2014.(6)

Tyr I can either be early onset or late onset. The disease can be fatal before 10 years of age in untreated children as they can present with severe liver disease, renal disease, rickets, and neurologic symptoms.(7) In comparison, Tyr II patients present with oculocutaneous symptoms and due to the rarity of Tyr III, the full clinical spectrum of the disorder is still unknown, with some cases described as mild.(1,4)

Management for Tyr I and II has been described in literature composing mainly of dietary modification (low tyrosine, low phenylalanine). In recent years, the treatment for Tyr I has evolved to include 2-(2-nitro-4-trifluoro-methylbenzoyl)-1,3 cyclohexanedione (NTBC/nitisinone). In combination with dietary restriction, medical therapy was associated with a 4-year survival probability of 88% for children under two years of age.(7,9)

4.7.2 Benefits and Harms of Screening Tests

Studies on the assessment of the balance of the benefits and harms of newborn screening in general were not identified. However, there may be some harm from false positive results and carriers are not identified.(7)

Benefits of screening impact greatly on the improvement in outcomes of patients diagnosed. Based on data from Quebec, prior to population-wide screening, mortality rate

for Tyr I was at 90% for patients under 2 years old due to acute liver failure. Screening for Tyr I is expected to decrease the severity of liver disease as well as kidney damage.(8)

No randomized controlled trial was available for review of evidence of outcomes between screening and non-screening for Tyr. However, a systematic review by Geppert et al reviewed observational studies on the comparison of clinical outcomes of Tyr I patients receiving earlier versus later nitisinone treatment.(10) After assessing the quality of the systematic review using AMSTAR 2 (Appendix D), individual studies were evaluated, and a GRADE table was generated (Appendix F). Only post hoc analysis could be done to these studies due to the high risk of confounding and applicability concerns, which mainly consisted of case reviews of known patients with Tyr I. Pooled estimates could not be derived due to differences in population, intervention, and outcomes. Summary of outcomes and level of evidence is provided below.

Table 1. Summary of outcomes and certainty of evidence based on a systematic review by Geppert et. al.

Outcomes*	No. of Studies (no. of participants)	Certainty of Evidence
<i>Early vs. late treatment</i>	2 ^{11,12} (n = 67)	Very low
Death		
<i>Early vs. late treatment</i>	3 ^{11,12,13} (n = 98)	Very low
Need for liver transplantation		
<i>Early vs. late treatment</i>	2 ^{11,14} (n = 117)	Very low
Patients with neurological crisis		
<i>Screen detection vs symptomatic detection, all with direct nitisinone initiation</i>	2 ^{11,12} (n = 45)	Very low
Death		
<i>Screen detection vs symptomatic detection, all with direct nitisinone initiation</i>	3 ^{11,12,13} (n = 70)	Very low
Need for liver transplantation		
<i>Screen detection vs symptomatic detection, all with direct nitisinone initiation</i>	1 ¹¹ (n = 29)	Very low
Patients with neurological crisis		

*In the post hoc analyses, proportions between two groups were compared using the chi-square test; in cases of expected values smaller than five, a Fisher exact test was used.

*Pooling of results not possible

Results of post hoc analyses suggested an association between earlier treatment and fewer liver transplants (0% vs 25-60%), but no influence or impact on neurologic crisis. No effect of treatment was found in terms of mortality between early and late medical

therapy with nitisinone ($p = 0.49, 0.07$). There is some evidence that earlier treatment may have some benefit, but available studies carry high confounding bias.(10)

4.7.3 Diagnostic Performance of Screening Tests

Historically, tyrosine levels on DBS were used to screen for Tyr I in newborns. This resulted in high false positive rates due to the high incidence of benign transient Tyrosinemia in neonates as well as high false negative rates and misdiagnoses since newborn with Tyr I can have normal tyrosine levels at birth. To overcome this, it was strongly recommended to measure succinylacetone (SUAC) instead in DBS, which is pathognomonic when present in blood or urine of patients with Tyr I. This reduced both false negative and false positive rates thereby optimizing resources and avoiding adverse psychological effects and stress.(15,16)

There were no randomized controlled trials available on SUAC as first tier for Tyr screening in asymptomatic newborns. A systematic review instead by Stinton et al. with moderate quality (Appendix E) was evaluated. It included five observational studies which reported newborn screening experiences and five retrospective case-control studies. Meta-analysis however was not possible due to incomplete follow-up of screen negatives for confirmation of true negative cases. Most of the included studies had high risk of bias due to indirectness in population screened, incomplete data on index tests, SUAC values not pre-specified, and incomplete follow-up. (Appendix L)(17)

Studies that reported on screening experiences had different confirmatory tests and had incomplete follow-up of their screen negatives, not making it possible to pool results. Across all studies, sensitivity of SUAC for Tyr I screening is 100%, with no available data on specificity. Certainty of evidence for these studies was assessed to be very low and factors affecting this is summarized in the GRADE tables. (Appendices G, H)(17)

Five retrospective case-control studies using stored samples from known Tyr I cases were also included in the systematic review which showed 100% for both sensitivity and specificity. Certainty of evidence was also assessed to be very low due to incomplete data on controls and confirmatory tests, indirectness in population screened, and incongruence in SUAC cut-offs. (Appendix I)

4.7.4 Cost Implication

One study on cost-effectiveness of screening for Tyrosinemia was identified. Analysis by Cipriano et al (2007) found Tyrosinemia I to be among the least cost-effective disorders to screen for. The cost-effectiveness however was determined before succinylacetone was used routinely instead of tyrosine.(7,18) Save for certain regions, Tyr I is considered a very rare condition and patients will eventually present clinically, hence only modest change may be expected in the prevalence of the disease as a result of screening, although some evidence may suggest benefit when treatment is initiated early.(8,10) Important to note is that cost-utility is improved when certain combinations of screening is done compared to singular screening for a certain disorder such as Tyr I.(8) No local data was identified evaluating the cost-effectiveness of the combination of diseases screened using the expanded newborn screening program.

Screen positives for Tyr I will undergo confirmatory testing to determine true and false positive cases. Cost of diagnostic tests is itemized in Table 2. Consequently, after confirmation of diagnosis, patients are prescribed with their medical regimen which will include NTBC and diet restriction (costs of medications enumerated in Table 3). Dietary management is assumed to be given for life.(8)

Table 2. Summary of Costs for Screening and Diagnosis for Tyr (in Php)

	Screening Intervention	Confirmatory test				
		Urine Metabolic Screen	Comprehensive Urine Metabolic Profile	Urine Organic Acid Analysis	Plasma Quantitative Amino Acid Analysis	Genetic testing
Cost	Php 1750 (as part of the expanded newborn screening)	Php 5,900 ^a Php 4,250 ^b	Php 14,000 ^a Php 10,300 ^b	Php 12,500 ^a Php 9,100 ^b	Php 17,500 ^a Php 12,300 ^b	~Php 12500

*Prices obtained from the Biochemical Genetics Laboratory of the Institute of Human Genetics (tel no. 02-5310-1780 loc 103)

^aRegular price

^bIndigent/Charity price

Table 3. Cost of available medications/medical food for patients with Tyr I

Medication/Medical Food	Unit Cost
Nitisinone (NTBC) 2mg/tab, 60 tabs/bottle	Php 45,500/bottle
Nitisinone (NTBC) 5mg/tab, 60 tabs/bottle	Php 94,400/bottle
Nitisinone (NTBC) 10mg/tab, 60 tabs/bottle	Php 145,500/bottle
TYROS 1 – amino acid-based phenylalanine- and tyrosine-free powdered formula, 0-12 months, 1-3 years old	Php 3,647/can
TYROS 2 – amino acid-based phenylalanine- and tyrosine-free powdered formula, 1-10 years old	Php 6,375/can

4.7.5 Equity, Acceptability, and Feasibility

No direct evidence or qualitative study was available regarding general acceptability, patient values and preferences in reference to newborn screening for tyrosinemia. The Institute of Health Economics in Alberta relied on first-hand medical experience of their Expert Advisory Group (EAG) to assess the acceptability of the overall screening program, diagnostic testing, and treatment of metabolic diseases (tyrosinemia included) on behalf of their patient population. Based on the opinion of physicians in the EAG, newborn screening appears to be generally accepted by doctors, parents, caregivers, and the general public.(8)

Locally, in terms of feasibility, the expanded newborn screening which includes screening for Tyrosinemia, has been introduced and operational since 2014 and mandated to be compulsory in 2019.

4.7.6 Recommendations from Other Groups

Two practice guidelines for the screening, diagnosis, and management of Tyr I were evaluated using the AGREE II Tool. (Appendices J, K) Recommendations from both guidelines were based on a review of available evidence and consensus of medical professionals and stakeholders. The summary of their recommendations with regards to screening is detailed below. No recommendation could be found for Tyr II and III.

Table 4. Summary of Recommendations from Available Clinical Practice Guidelines

Guidelines	Year Published	Recommendation for Screening	Grade of Recommendation
External Review with Evidence Summary Newborn Blood Spot Screening for Tyrosinemia Type 1 in the UK(7)	2017	Newborn screening for TYR1 could be considered but more information would be required. The evidence on the screening test shows that most of those with positive test results would have TYR1, but there is not enough information available following up babies who have had a negative screening test.	Not mentioned Data obtained from observational studies and case-series
Diagnosis and treatment of tyrosinemia type I: a US and Canadian consensus group review and recommendations(19)	2017	Blood succinylacetone should be used as the primary marker to detect patients with tyrosinemia I through NBS	Evidence quality B
		Tyrosine as a primary marker for NBS of tyrosinemia I is not recommended	Evidence quality D

References

1. Szymanska E, Sredzinska M, Ciara E, Piekutowska-Abramczuk D, Ploski R, Rokicki D et al. Tyrosinemia type III in an asymptomatic girl. 2015.
2. Mayorandan S, Meyer U, Gokcay G, Segarra N, de Baulny H, van Spronsen F et al. Cross-sectional study of 168 patients with hepatorenal tyrosinaemia and implications for clinical practice. 2014.
3. Al-Essa M, Rashed M, Ozand P. Tyrosinemia Type II: Report of the First Four Cases in Saudi Arabia. Annals of Saudi Medicine. 1998;18(5):466-468.
4. Barroso F, Correia J, Bandeira A, Carmona C, Vilarinho L, Almeida M et al. TYROSINEMIA TYPE III: A CASE REPORT OF SIBLINGS AND LITERATURE REVIEW. Revista Paulista de Pediatria. 2020;38.
5. Newborn screening information for tyrosinemia, type III | Baby's First Test | Newborn Screening | Baby Health [Internet]. Babysfirsttest.org. 2021 [cited 29 July 2021]. Available from: <https://www.babysfirsttest.org/newborn-screening/conditions/tyrosinemia-type-iii>

6. Prevalence as of December 2020 [Internet]. Newbornscreening.ph. 2021 [cited 22 August 2021]. Available from:
<https://www.newbornscreening.ph/images/stories/ResourcesPrevalence/prevalence-2020.pdf>
7. Newborn Blood Spot Screening for Tyrosinaemia Type 1 in the UK [Internet]. Legacyscreening.phe.org.uk. 2017 [cited 25 July 2021]. Available from:
https://legacyscreening.phe.org.uk/policydb_download.php?doc=1138
8. Institute of Health Economics | [Internet]. Ihe.ca. 2016 [cited 5 August 2021]. Available from:
<https://www.ihe.ca/advanced-search/newborn-blood-spot-screening-for-galactosemia-tyrosinemia-type-i-homocystinuria-sickle-cell-anemia-sickle-cell-beta-thalassemia-sickle-cell-hemoglobin-c-disease-and-severe-combined-immunodeficiency>
9. Javadi M, Mirdehghan S, Bagheri A, Einollahi B, Dowlati Y. Two cases of tyrosinemia type ii, and its rare occurrence in two brothers [Internet]. Mjiri.iums.ac.ir. 2021. Available from:
http://mjiri.iums.ac.ir/browse.php?a_id=1207&sid=1&slc_lang=en
10. Geppert J, Stinton C, Freeman K, Fraser H, Clarke A, Johnson S et al. Evaluation of pre-symptomatic nitroso treatment on long-term outcomes in Tyrosinemia type 1 patients: a systematic review. *Orphanet Journal of Rare Diseases*. 2017;12(1).
11. Larochelle J, Alvarez F, Bussières J, Chevalier I, Dallaire L, Dubois J et al. Effect of nitroso (NTBC) treatment on the clinical course of hepatorenal tyrosinemia in Québec. *Molecular Genetics and Metabolism*. 2012;107(1-2):49-54.
12. McKiernan P, Preece M, Chakrapani A. Outcome of children with hereditary tyrosinemia following newborn screening. *Archives of Disease in Childhood*. 2015;100(8):738-741.
13. Bartlett D, Lloyd C, McKiernan P, Newsome P. Early nitroso treatment reduces the need for liver transplantation in children with tyrosinemia type 1 and improves post-transplant renal function. *Journal of Inherited Metabolic Disease*. 2014;37(5):745-752.
14. Simoncelli M, Samson J, Bussières J, Lacroix J, Dorais M, Battista R et al. Cost-Consequence Analysis of Nitroso for Treatment of Tyrosinemia Type I. *The Canadian Journal of Hospital Pharmacy*. 2015;68(3).
15. Giguère Y, Berthier M. Newborn Screening for Hereditary Tyrosinemia Type I in Québec: Update. *Advances in Experimental Medicine and Biology*. 2017;:139-146.
16. Malvagia S, Forni G, Ombrone D, la Marca G. Development of Strategies to Decrease False Positive Results in Newborn Screening. *International Journal of Neonatal Screening*. 2020;6(4):84.
17. Stinton C, Geppert J, Freeman K, Clarke A, Johnson S, Fraser H et al. Newborn screening for Tyrosinemia type 1 using succinylacetone – a systematic review of test accuracy. *Orphanet Journal of Rare Diseases*. 2017;12(1).
18. Cipriano L, Rupar C, Zaric G. The Cost-Effectiveness of Expanding Newborn Screening for up to 21 Inherited Metabolic Disorders Using Tandem Mass Spectrometry: Results from a Decision-Analytic Model. *Value in Health*. 2007;10(2):83-97.
19. Chinsky J, Singh R, Ficicioglu C, van Karnebeek C, Grompe M, Mitchell G et al. Diagnosis and treatment of tyrosinemia type I: a US and Canadian consensus group review and recommendations. *Genetics in Medicine*. 2017;19(12):1380-1380

4.8 Newborn Screening for Fatty Acid Oxidation Disorders

RECOMMENDATIONS

1. Among asymptomatic, apparently healthy newborns, we recommend against the screening of LCHADD and MTPD. (*Strong recommendation, very low certainty of evidence*).
2. Among asymptomatic, apparently healthy newborns, we recommend against the screening of CPT1D, CPT2D and GA2. (*Strong recommendation, very low certainty of evidence*).

Considerations

The consensus panel considered the following when formulating this recommendation:

- The very low burden of disease among Filipinos coupled with the high cost of confirmatory testing and treatment of the disease led the consensus panel to decide against the screening of LCHADD, MTPD, CPT1D, CPT2D and GA2.

4.8.1 Burden of disease

Fatty acid oxidation disorders (FAODs) are a group of autosomal recessive inherited conditions wherein enzymes necessary for fatty acid breakdown is either absent or defective.(1) Data from Australia, Germany, and the USA shows a combined incidence of 1:250,000 for Long chain 3-hydroxy acyl CoA dehydrogenase deficiency (LCHADD), 1:750,000 for mitochondrial trifunctional protein deficiency (MTPD) and between 1:750,000 to 1:2,000,000 for carnitine palmitoyl transferase deficiency types 1 and 2 (CPT1D, CPT2D) and glutaric aciduria type 2 (GA2).(2) According to the metabolic registry of the Clinical Genetics and Research Unit – Institute of Human Genetics, 1 case of CPT1D and 3 cases of GA2 were clinically detected and confirmed by mutational analysis. As of June 2021, no cases of CPT1D, CPT2D, LCHADD, MTPD and GA2 were identified by the expanded newborn screening since its implementation.(3)

FAODs are metabolic disorders of clinical significance due to its effect on pediatric morbidity and mortality. LCHADD and MTPD have 3 clinical phenotypes: an early-onset severe form, infant-onset form, and a milder, late-onset form. Approximately 38% of infants die before, or within 3 months of diagnosis.⁴ Long-term complications include fatigue, rhabdomyolysis, hypoglycemia, cardiomyopathy, neuropathy, retinopathy, organ failure and death.(4,5) Early-onset CPT2D have an extremely high neonatal mortality rate while the late-onset form is characterized only by episodic rhabdomyolysis triggered by exercise. CPT1D may often be benign, although early presentation with hypoketotic hypoglycemia certainly occurs.(6)

Patients with FAODs averaged 1.94 hospitalizations annually, with a mean of 17.55 total hospital days per year.(7) The disease has negative impact on health-related quality of life (HRQOL) for both patients and caregivers.(8) There is no cure for FAODs but a number of management strategies are available. The mainstay of treatment are fasting avoidance, a low-fat and high-carbohydrate food plan, and/or taking supplements such

as medium-chain triglycerides (MCT) and carnitine.(4) A number of studies have shown that earlier treatment might result in better long-term outcomes than later treatment and newborn screening is one method by which earlier diagnosis and intervention can be achieved.(5)

4.8.2 Benefits and Harms of Screening Tests

LCHADD / MTPD

A systematic review was conducted to investigate whether pre-symptomatic dietary management following newborn screening provides better outcome than treatment following symptomatic detection. A total of 174 patients with LCHADD, 18 people with MTPD and 12 people with undifferentiated LCHADD/ MTPD were included across all the studies. Patients were then classified to either screened (includes all patients identified via NBS screening and cascade testing) or unscreened (absence of newborn screening or clinically detected following false negative screening results) group and analysis of outcomes were done.(5)

Mortality (5 Cohort study designs, N = 139, Very Low Certainty of Evidence)

Follow up analysis showed that among the 5 studies that reported on mortality, there was a statistically significant fewer deaths among the screened group compared to the unscreened group (OR 0.36, 95% CI 0.13 to 1.02, p=0.05). There is a 64% decrease in the odds of mortality favoring the screened group.(14-18)

Cardiac Pathology (4 Cohort study designs, N = 82, Very Low Certainty of Evidence)

Of the 6 studies that reported on cardiac problem, 4 reported on cardiomyopathy. Pooled analysis showed significantly fewer cases of cardiomyopathy among the screened group (OR 0.28, 95% CI 0.09 to 0.85, p=0.02). Arrhythmia was reported in 1 study however results of analysis showed no statistically significant difference between the two groups.(14,16,18,21)

Liver Pathology (2 Cohort study designs, N = 25, Very Low Certainty of Evidence)

Three (3) studies reported on the incidence of liver problems. Hepatopathy was investigated in 2 studies and pooled analysis showed statistically fewer cases among the screen detected group compared to the unscreened group (OR 0.06, 95% CI 0.01 to 0.60, p=0.02). 1 study reported on the incidence of Reye syndrome however analysis showed no statistically significant difference between both groups.(16,19)

Ophthalmologic Pathology (3 Cohort study designs, N = 51, Very Low Certainty of Evidence)

Six (6) studies reported on eye problems as outcome. Incidence of retinopathy was investigated in 3 studies and pooled analysis showed retinopathy was significantly less common among screen detected individuals than the clinically detected group (OR 0.09, 95% CI 0.01 to 0.64, p=0.02). Other ophthalmologic complaints such as photophobia, nystagmus, subnormal ocular fundi and vision were investigated however analysis showed no statistically significant difference between the two groups.(14,19,21)

Hypoglycemia (2 Cohort study designs, N = 66, Very Low Certainty of Evidence)

Two (2) studies reported on the incidence of hypoglycemia and pooled analysis showed statistically fewer cases among the screen detected group compared to the unscreened group (OR 0.10, 95% CI 0.02 to 0.50, p=0.005).(18,21)

Neurologic Pathology

Two (2) studies investigated the incidence neurologic problems such as epilepsy, psychomotor development and neurologic symptoms however analysis showed no statistically significant difference between the two groups.(5)

Motor/Muscular Problem

There were 3 studies which investigated the incidence of myopathy, myoglobinuria and rhabdomyolysis. Only 1 study found a statistically significant difference as it reported fewer cases of myopathy among the screen detected group compared to the clinically detected group (40% vs 82.4%, p = 0.03).(5)

CPT1D, CPT2D, GA2

We found 1 retrospective cohort which studied the outcome of 1 screened and 1 clinically detected CPT1D patient. The screened patient exhibited normal development however the clinically detected patient suffered from developmental delay and other complications such as recurrent hepatic failure, nephromegaly, hemolytic anemia, and rhabdomyolysis.(15) 1 retrospective cohort reported 5 patients with CPT2D, 3 were identified by newborn screening. Mortality, cardiomyopathy and hepatopathy was seen only among the screened patients however both screened and clinically detected patients suffered from myopathic symptoms.(18) Lastly, a retrospective cohort reported on 7 GA2 patients, 2 of which were detected during newborn screening. Myopathy, metabolic acidosis and transaminitis was seen only among the clinically detected patients but both groups suffered from hypoglycemia and elevated creatine kinase.(22) Due to very low number of sample size, no definitive conclusions can be drawn regarding the benefit of early screening and subsequent intervention for CPT1D, CPT2D and GA2. The characteristics and outcomes of each study are found in Appendix B.

The overall certainty of evidence across all outcomes is very low. All studies included were of weak methodological quality due to their cohort design. Risk of bias was serious in all outcomes except for mortality due to performance bias. In all the studies, the outcome assessors knew whether the participants had been screened or clinically detected. Imprecision was also an issue due to uncertainty in the magnitude of effect and low number of events.

Summary of Outcomes of Screening versus No Screening for LCHADD / MTPD,

Outcomes	No. of Studies (no. of participants)	Interpretation	OR (95% CI)	Certainty of Evidence
LCHADD / MTPD				
Mortality	5 (107)	Not significant	0.36 [0.13, 1.02]	Very Low
Cardiomyopathy	4 (57)	Benefit	0.28 [0.09, 0.85]	Very Low
Hepatopathy	2 (19)	Benefit	0.06 [0.01, 0.60]	Very Low
Retinopathy	2 (30)	Benefit	0.09 [0.01, 0.64]	Very Low
Hypoglycemia	2 (39)	Benefit	0.10 [0.02, 0.51]	Very Low
CPT1D				
Developmental Delay	1 (2)	Not significant	0.11 [0.00, 10.27]	Very Low
CPT2D				
Mortality	1 (5)	Not significant	8.33 [0.22, 320.38]	Very Low
Cardiomyopathy	1 (5)	Not significant	3.00 [0.08, 115.34]	Very Low
Hepatopathy	1 (5)	Not significant	8.33 [0.22, 320.38]	Very Low
Hypotonia/ Myopathy	1 (5)	Not significant	0.12 [0.00, 4.61]	Very Low
GA2				
Myopathy	1 (7)	Not significant	0.07 [0.00, 2.33]	Very Low
Metabolic acidosis	1 (7)	Not significant	0.60 [0.02, 20.98]	Very Low
Hepatopathy	1 (7)	Not significant	0.28 [0.01, 8.76]	Very Low
Elevated Creatine Kinase	1 (7)	Not significant	1.50 [0.06, 40.63]	Very Low
Hypoglycemia	1 (7)	Not significant	11.00 [0.28, 433.80]	Very Low

CPT1/2D, GA2

4.8.3 Diagnostic Performance of Screening Tests

Test accuracy estimates of screening for LCHADD and MTPD with tandem mass spectrometry measurement of acylcarnitines in dried blood was investigated in 1 systematic review. This review included 10 studies which screened a total of 3,951,358 newborns. 23 cases of LCHADD/MTPD were identified. The only measure of test

accuracy that was consistently reported was positive predictive value (PPV). PPV ranged from 0% (0 true positives and 28 false positives from 276,565 babies screened) to 100% (13 true positives from 2,037,824 babies screened). It was not possible to calculate sensitivity, specificity, or negative predictive value as there was no systematic follow-up of babies who had screened negative.(4)

For the other diseases of interest, screening test performance for FAOD in NBS programs in Germany and US which screened a total of 1,064,336 newborns reported the following results. For CPT1D the false positive rate (FPR) is <0.001% with a positive predictive value (PPV) of 0%. For CPT2D the reported FPR ranges from <0.001% to 0.002% with a PPV of 14%. Lastly, for GA2 the reported FPR is 0.001% to 0.002% with a PPV of 33%.(2)

4.8.4 Cost Implication

Parameter	Cost
Screening intervention	
Expanded Newborn Screening/ ENBS Filter Card	Php 1,500
Confirmatory Test	
Urine organic acids	Php 10,400
Plasma Acylcarnitine	Php 14,800
DNA Analysis	£200 (14,042.81 Php)
Cost of Treatment	
L-Carnitine	Php 2,125/ bottle
LCFAOD Medical Food (Monogen, Enfaport)	Php 4,550/ can

An analysis of the economic impact of expanded newborn screening pilot which included LCHADD was done in England in 2013. Result of the study showed that for LCHADD the cost of long-term health and social care impact (no screening £636,641 vs screening £113,268), cost of managing the disease (no screening £81,900 vs screening £56,578) and life-time QALYs gained (no screening 17.80 vs screening 41.20) favors screening.(9) Key findings of this health economic evaluation demonstrated that screening for LCHADD is potentially cost saving and resulted in increased quality of life compared to no screening. However, LCHADD was not recommended to be included in the expansion of the NBS due to uncertainties in the definition, treatment and outcomes between LCHADD and the spectrum of conditions associated with MTPD which resulted in difficulty in parameterization of the economic model and was subject to uncertainty.(9)

A study was done in Canada to estimate the cost-effectiveness of expanding the Ontario newborn screening program to screen for each disease independently and for hypothetical bundles of up to 21 metabolic diseases. Results of the health economic evaluation showed that screening for the following fatty acid oxidation disorders yielded the following incremental cost effectiveness ratio (ICER) per life year gained. MCADD is the most cost-effective disease to screen alone has an ICER of only \$253,161 per LY (life years) gained while GA2 is the least cost-effective disease to screen for. The total

incremental cost of screening for just GA2 is \$2.5 million, but each infant who receives a diagnosis of GA2 only gains 0.033 LY. This results in an ICER of \$142.5 million per LY gained. The ICER for CPT1D, CPT2D, and LCHADD are \$1,192,814, \$1,243,125 and \$1,063,830 per LY gained respectively. By using a threshold of \$100,000 per LY gained, the authors find moderate evidence to support the adoption of screening for PKU and 14 additional diseases which include all FAODs with the exception of GA2.(10)

4.8.5 Equity, Acceptability, and Feasibility

The Philippines have been testing for these fatty acid oxidation disorders since the introduction of the expanded newborn screening in 2015. It is generally acceptable among families as they acknowledge the benefits of early screening. A qualitative study was done in Canada to investigate the experiences of caregivers of children with immunodeficiency (IMD), its impact on the child and family life, and interactions with the health care system. One of the reported challenges is encountering systems and health care providers unfamiliar with the child's disease. This is a particular concern when the child is received in an emergency room and the caregivers commonly expressed their distress and dissatisfaction with the inconsistency in care received from the emergency department during a crisis. Parents particularly worry about the capability of emergency professionals to manage the needs of patients with a rare (and unfamiliar to the provider) condition. Another concern of caregivers is the accessibility to a specialist metabolic clinic, with metabolic health care providers, dietitians and multiple healthcare providers from various disciplines involved in disease-specific care. Most participants were highly satisfied with the care that their child receive. They view the metabolic specialists as supportive and the dietitians as highly engaged. Aside from access to specialized care from IMD specific services, the accessibility to the special medical food or supplement is also an identified concern.(11)

A 3-month public consultation was hosted on the UK NSC website and direct email was sent to 16 key stakeholders. They received comments from 5 stakeholders namely the British Inherited Metabolic Disease Group, Genetic Alliance UK, Royal College of Midwives, Royal College of Paediatrics and Child Health Sheffield Children's Hospital. While only 1 stakeholder supported the review's conclusions, majority disagreed as their predominant opinion is that LCHADD should be added to the newborn bloodspot screening program. The common reasons cited are: (1) the low incidence of isolated LCHADD during the pilot study was not reflective of the clinical cases seen at metabolic centers that manage these conditions, (2) clinical and molecular heterogeneity should not be a bar to screen for the combined conditions (LCHADD and MTPD), particularly since infants with isolated LCHADD represent the majority of patients and the phenotype associated with it is considered to be well established, and (3) evidence suggesting benefit from early treatment should be a worthwhile reason to screen and (4) different standards should be used in the evaluation of the evidence for screening programs on rare conditions.(12)

4.8.6 Recommendations from Other Groups

UK National Screening Committee for fatty acid oxidation disease currently does not recommend newborn screening for LCHADD and MTPD. In this report the reviewers

examined 4 key questions relating to the effectiveness and appropriateness of newborn screening using tandem mass spectrometry for LCHADD and MTPD. These key questions examined the disease prevalence in the UK, genotypic – phenotypic association, test accuracy and outcomes of early intervention. Findings of the evidence review showed there were no studies from the UK on the number of people born with LCHADD/MTPD, different patient outcomes do not seem to be linked to their specific mutation, screening test indicates that false positives are common and there was lack of systematic follow up of babies who screened negative and lastly, there is uncertainty if early detection and intervention will result to better outcome compared to those symptomatically diagnosed. Given the result of the studies included within this review and the areas for further research, the UK NSC concluded that systematic population screening for LCHADD and MTPD cannot be recommended at this time.(13) We found no recommendations or CPGs for CPT1D, CPT2D and for GA2 however numerous countries have already included these diseases in their newborn screening program.(2,23-27) These are listed in Appendix F.

References

1. Newborn Screening Reference Center. Fact Sheets: Information for doctors about the disorders included in the Expanded Newborn Screening Panel. Oct 9, 2020.
2. Lindner M, Hoffmann GF, Matern D. Newborn screening for disorders of fatty-acid oxidation: experience and recommendations from an expert meeting. *J Inherit Metab Dis.* 2010 Oct;33(5):521-6. doi: 10.1007/s10545-010-9076-8.
3. Data from the Clinical Genetics and Research Unit – Institute of Human Genetics Census (June 2021)
4. Stinton C, Fraser H, Geppert J, Johnson R, Connock M, Johnson S, Clarke A, Taylor-Phillips S. Newborn Screening for Long-Chain 3-Hydroxyacyl-CoA Dehydrogenase and Mitochondrial Trifunctional Protein Deficiencies Using Acylcarnitines Measurement in Dried Blood Spots-A Systematic Review of Test Accuracy. *Front Pediatr.* 2021 Mar 19;9:606194. doi: 10.3389/fped.2021.606194.
5. Fraser H, Geppert J, Johnson R, Johnson S, Connock M, Clarke A, Taylor-Phillips S, Stinton C. Evaluation of earlier versus later dietary management in long-chain 3-hydroxyacyl-CoA dehydrogenase or mitochondrial trifunctional protein deficiency: a systematic review. *Orphanet J Rare Dis.* 2019 Nov 15;14(1):258. doi: 10.1186/s13023-019-1226-y.
6. Wilcken B. Fatty acid oxidation disorders: outcome and long-term prognosis. *J Inherit Metab Dis.* 2010 Oct;33(5):501-6. doi: 10.1007/s10545-009-9001-1
7. Pannier A. Long-Chain Fatty Acid Oxidation Disorders: Managed Care and Specialty Pharmacy Implications. *Am J Manag Care.* 2020;26:S139-S145.
8. Vockley J, Burton B, Berry GT, Longo N, Phillips J, Sanchez-Valle A, Tanpaiboon P, Grunewald S, Murphy E, Humphrey R, Mayhew J, Bowden A, Zhang L, Cataldo J, Marsden DL, Kakkis E. UX007 for the treatment of long chain-fatty acid oxidation disorders: Safety and efficacy in children and adults following 24weeks of treatment. *Mol Genet Metab.* 2017 Apr;120(4):370-377. doi: 10.1016/j.ymgme.2017.02.005.
9. Bessey A, Chilcott J, Pandor A, Paisley S. The Cost-Effectiveness of Expanding the UK Newborn Bloodspot Screening Programme to Include Five Additional Inborn Errors of Metabolism. *Int J Neonatal Screen.* 2020 Nov 20;6(4):93. doi: 10.3390/ijns6040093.
10. Cipriano LE, Rupar CA, Zaric GS. The cost-effectiveness of expanding newborn screening for up to 21 inherited metabolic disorders using tandem mass spectrometry: results from a decision-analytic model. *Value Health.* 2007 Mar-Apr;10(2):83-97. doi: 10.1111/j.1524-4733.2006.00156.x.
11. Siddiq S, Wilson BJ, Graham ID, Lamoureux M, Khangura SD, Tingley K, Tessier L, Chakraborty P, Coyle D, Dyack S, Gillis J, Greenberg C, Hayeems RZ, Jain-Ghai S, Kronick JB, Laberge AM, Little J, Mitchell JJ, Prasad C, Siriwardena K, Sparkes R, Speechley KN, Stockler S, Trakadis Y, Wafa S, Walia J, Wilson K, Yuskin N, Potter BK; Canadian Inherited Metabolic Diseases Research

- Network (CIMDRN). Experiences of caregivers of children with inherited metabolic diseases: a qualitative study. *Orphanet J Rare Dis.* 2016 Dec 7;11(1):168. doi: 10.1186/s13023-016-0548-2.
- 12. UK National Screening Committee. Newborn screening for mitochondrial trifunctional protein (MTP) disorders, including long-chain 3-hydroxyacyl-CoA dehydrogenase (LCHAD) deficiency. Consultation Responses. June 2019.
 - 13. Fraser H, Geppert J, Johnson R, Clarke A, Johnson S, Connock M, Taylor-Phillips S, and Stinton C. Screening for mitochondrial trifunctional protein disorders, including long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency. UK National Screening Committee. April 2019.
 - 14. Immonen T, Turanlahti M, Paganus A, Keskinen P, Tyni T, Lapatto R. Earlier diagnosis and strict diets improve the survival rate and clinical course of long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency. *Acta Paediatr.* 2016 May;105(5):549-54. doi: 10.1111/apa.13313.
 - 15. Kang E, Kim YM, Kang M, Heo SH, Kim GH, Choi IH, Choi JH, Yoo HW, Lee BH. Clinical and genetic characteristics of patients with fatty acid oxidation disorders identified by newborn screening. *BMC Pediatr.* 2018 Mar 8;18(1):103. doi: 10.1186/s12887-018-1069-z.
 - 16. Lund AM, Hougaard DM, Simonsen H, Andresen BS, Christensen M, Dunø M, Skogstrand K, Olsen RK, Jensen UG, Cohen A, Larsen N, Saugmann-Jensen P, Gregersen N, Brandt NJ, Christensen E, Skovby F, Nørgaard-Pedersen B. Biochemical screening of 504,049 newborns in Denmark, the Faroe Islands and Greenland--experience and development of a routine program for expanded newborn screening. *Mol Genet Metab.* 2012 Nov;107(3):281-93. doi: 10.1016/j.ymgme.2012.06.006.
 - 17. Sykut-Cegielska J, Gradowska W, Piekutowska-Abramczuk D, Andresen BS, Olsen RK, Ołtarzewski M, Pronicki M, Pajdowska M, Bogdańska A, Jabłońska E, Radomyska B, Kuśmierska K, Krajewska-Walasek M, Gregersen N, Pronicka E. Urgent metabolic service improves survival in long-chain 3-hydroxyacyl-CoA dehydrogenase (LCHAD) deficiency detected by symptomatic identification and pilot newborn screening. *J Inherit Metab Dis.* 2011 Feb;34(1):185-95. doi: 10.1007/s10545-010-9244-x
 - 18. Spiekerkoetter U, Lindner M, Santer R, Grotzke M, Baumgartner MR, Boehles H, Das A, Haase C, Hennermann JB, Karall D, de Klerk H, Knerr I, Koch HG, Plecko B, Röschinger W, Schwab KO, Scheible D, Wijburg FA, Zschocke J, Mayatepek E, Wendel U. Management and outcome in 75 individuals with long-chain fatty acid oxidation defects: results from a workshop. *J Inherit Metab Dis.* 2009 Aug;32(4):488-97. doi: 10.1007/s10545-009-1125-9.
 - 19. Karall D, Brunner-Krainz M, Kogelnig K, Konstantopoulou V, Maier EM, Möslinger D, Plecko B, Sperl W, Volkmar B, Scholl-Bürgi S. Clinical outcome, biochemical and therapeutic follow-up in 14 Austrian patients with Long-Chain 3-Hydroxy Acyl CoA Dehydrogenase Deficiency (LCHADD). *Orphanet J Rare Dis.* 2015 Feb 22;10:21. doi: 10.1186/s13023-015-0236-7. PMID: 25888220; PMCID: PMC4407779.
 - 20. De Biase I, Viau KS, Liu A, Yuzyuk T, Botto LD, Pasquali M, Longo N. Diagnosis, Treatment, and Clinical Outcome of Patients with Mitochondrial Trifunctional Protein/Long-Chain 3-Hydroxy Acyl-CoA Dehydrogenase Deficiency. *JIMD Rep.* 2017;31:63-71. doi: 10.1007/8904_2016_558.
 - 21. Fahnehjelm KT, Liu Y, Olsson D, Amréen U, Haglind CB, Holmström G, Halldin M, Andreasson S, Nordenström A. Most patients with long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency develop pathological or subnormal retinal function. *Acta Paediatr.* 2016 Dec;105(12):1451-1460. doi: 10.1111/apa.13536.
 - 22. Maguolo A, Rodella G, Dianin A, Nurri R, Monge I, Rigotti E, Cantalupo G, Salviati L, Tucci S, Pellegrini F, Molinaro G, Lupi F, Tonin P, Pasini A, Campostrini N, Ion Popa F, Teofoli F, Vincenzi M, Camilot M, Piacentini G, Bordugo A. 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 Aug 5;24:100632. doi: 10.1016/j.ymgmr.2020.100632.
 - 23. Loeber JG, Platis D, Zetterström RH, Almashanu S, Boemer F, Bonham JR, Borde P, Brincat I, Cheillan D, Dekkers E, Dimitrov D, Fingerhut R, Franzson L, Groselj U, Hougaard D, Knapkova M, Kocova M, Kotori V, Kozich V, Kremezna A, Kurkijärvi R, La Marca G, Mikelsaar R, Milenkovic T, Mitkin V, Moldovanu F, Ceglarek U, O'Grady L, Ołtarzewski M, Pettersen RD, Ramadza D, Salimbayeva D, Samardzic M, Shamsiddinova M, Songailienė J, Szatmari I, Tabatadze N, Tezel B, Toromanovic A, Tovmasyan I, Usurelu N, Vevere P, Vilarinho L, Vogazianos M, Yahyaoui R, Zeyda M, Schielen PCJ. Neonatal Screening in Europe Revisited: An ISNS Perspective on the

- Current State and Developments Since 2010. *Int J Neonatal Screen.* 2021 Mar 5;7(1):15. doi: 10.3390/ijns7010015. - EUROPE
- 24. Waisbren SE. Expanded newborn screening: information and resources for the family physician. *Am Fam Physician.* 2008 Apr 1;77(7):987-94. - USA
 - 25. Marsden D, Bedrosian CL, Vockley J. Impact of newborn screening on the reported incidence and clinical outcomes associated with medium- and long-chain fatty acid oxidation disorders. *Genet Med.* 2021 May;23(5):816-829. doi: 10.1038/s41436-020-01070-0. Epub 2021 Jan 25. – ASIA
 - 26. Wilson K, Kennedy SJ, Potter BK, Geraghty MT, and P Chakraborty. Developing a National Newborn Screening Strategy for Canada. *Health Law Review* 2010. 18(2): 31-39. - CANADA
 - 27. Krotoski D, Namaste S, Raouf RK, El Nekhely I, Hindi-Alexander M, Engelson G, Hanson JW, Howell RR; MENA NBS Steering Committee. Conference report: second conference of the Middle East and North Africa newborn screening initiative: Partnerships for sustainable newborn screening infrastructure and research opportunities. *Genet Med.* 2009 Sep;11(9):663-8. doi: 10.1097/GIM.0b013e3181ab2277. – MIDDLE EAST

4.9 Newborn Screening for Biotidinase Deficiency

RECOMMENDATIONS

Among asymptomatic, apparently healthy newborns, we recommend against the screening of biotidinase deficiency. (**Strong recommendation, very low certainty of evidence.**)

Considerations

The consensus panel considered the following when formulating this recommendation:

- The very low burden of disease among Filipinos coupled with the high cost of confirmatory testing of the disease and the very low certainty of evidence led the consensus panel to decide against the screening of biotidinase deficiency.

4.9.1 Burden of disease

Currently, there is no published data on the exact prevalence of biotinidase deficiency in the Philippines except for the prevalence (1 in 111,127) reported by Padilla in 2012 based on the Filipino newborn population registered in the California Newborn Screening Program.(1) According to the Metabolic Registry of the Institute of Human Genetics, there was only 1 case of biotinidase deficiency reported in the Philippines since screening was started in 2014 resulting to an estimated prevalence of 1 in 1,748,857, which is lower compared to the worldwide prevalence of 1 in 61,067 for the combined profound and partial biotinidase deficiency reported by Wolf in 1991 which also vary widely by country.(2,3)

Biotinidase deficiency is an autosomal recessive disorder which has an inability to recycle endogenous biotin. Clinical manifestations of biotinidase deficiency usually appear in untreated patients between ages one week and ten years (median age of 3 months) with profound biotinidase deficiency (less than 10% of normal enzyme activity) presenting as neurologic abnormalities (e.g. seizures, hypotonia, ataxia, visual problems, hearing impairment, developmental delay) and skin manifestations (e.g., alopecia, skin rash, candidiasis) while those with partial biotinidase deficiency (10% to 30% of normal enzyme activity) exhibits hypotonia, skin rash, and hair loss, particularly during times of stress or infection.(4-6)

Patients with biotinidase deficiency diagnosed prior to development of symptoms (e.g., diagnosed by newborn screening) and treated with biotin remain asymptomatic and are developmentally at par with age while those untreated may develop symptoms and sequelae which can be irreversible.(4-6) Once biotin treatment is started, skin and neurologic manifestations resolves immediately, however if visual impairment, hearing impairment and developmental delay already occurred, they can be irreversible despite biotin therapy.(4-6)

4.9.2 Benefits and Harms of Screening Tests

Mortality

There were no studies found on mortality.

Complications (1 Observational Study, Very Low Certainty of Evidence)

There were no systematic reviews, meta-analysis or randomized controlled trials found on complications or morbidity. There was only one study found that assessed complications of biotinidase deficiency among those screened and clinically detected. It was an observational cohort done in Europe that compared screened children to symptomatic children with respect to three (3) variables namely presence of residual impairment (e.g. hearing impairment, visual disorder, and developmental delay), social adaptation, and behavioral disorders. There were 37 children (24 males, 13 females) with profound biotinidase deficiency (PBD) enrolled in the study. The mean age at diagnosis in symptomatic children was 17.5 months (range 2–53 months, median 10 months) and the mean age at starting biotin supplementation was 23 months (range 3–60 months, median 15 months). The mean age at diagnosis in the children diagnosed by newborn screening, was 11 days (range 5– 42 days, median 8 days), and the mean age at start of biotin supplementation was 40 days (range 10–158 days, median 22 days). The follow-up time covered was from 5 months to 18 years (median 6 years 6 months). The study demonstrated better outcomes in screening detected patients compared with clinically detected patients. Only children without newborn screening show residual symptoms with a higher risk of profound visual ($p<0.001$) and hearing ($p=0.004$) impairment. Children without screening more frequently showed speech delay (after 17 months; $p=0.022$), delayed onset of walking (after 18 months; $p=0.002$) and delayed onset of sitting (after 9 months; not significant) than those who were screened (See Table 1). There were no significant intergroup differences established on the Child Behavior Checklist and Vineland Adaptive Behavior Scales.(4) See Appendix B for GRADE Evidence Profile.

Table 1. WEBER 2004 Study Outcome in Patients with Profound BD: relevance to NB

Parameter	With NBS (Asymptomatic)(n=25)	Without NBS (Clinically detected) (n=12)	P Value
Visual Impairment	0 (0%)	4* (33%)	$p<0.001$
Hearing impairment	0 (0%)	4* (33%)	$p=0.004$
Delayed Speech (1st 3 words after 17 mos)	4 (16%)	6 (50%)	$p=0.022$
Delayed Onset of Walking (after 18 mos)	0 (0%)	5 (42%)	$p=0.002$

Delayed Onset of Sitting (after 9 mos)	1 (4%)	3 (25%)	NS
--	--------	---------	----

* Two patients had both optic atrophy and hearing impairment

Adverse Events

There were no harms or adverse events reported related to newborn screening as well as biotin therapy.

4.9.3 Diagnostic Performance of Screening Tests

Five observational studies (N=3,686,573) with newborn screening as the screening test for detection of biotinidase deficiency showed a pooled sensitivity of 0.93 (95% CI 0.67-0.99) with substantial heterogeneity ($I^2=83.76\%$) and pooled specificity of 1.0 with high heterogeneity ($I^2=97.76\%$). (11-15) The sensitivity and specificity values were derived based on the assumption that there were no false negatives reported as declared by the authors. Due to the high heterogeneity and serious risk of bias, the pooled data was rated with a very low overall methodological quality of evidence. See Appendix C for GRADE Evidence Profile.

4.9.4 Cost Implication

Based on the cost utility analysis done by Caroll et al. in the US, newborn screening for biotinidase deficiency saves money and results in an increase in quality-adjusted life-years (QALY) arriving at the conclusion that screening improved outcomes and reduced costs relative to not screening. (8) Despite the evaluation of false positive results which showed an increased cost in ruling out disease combined with the low prevalence of the disease, screening remained to be cost-saving. (8) This was further supported by the cost-effectiveness analysis study done by Vallejo-Torre et al. in Spain which reported that newborn screening for biotinidase deficiency is likely to be cost-effective compared with clinical detection with a screening cost effective probability of 70% at standard threshold values. (9) Since Spain has a higher biotinidase deficiency prevalence compared to the worldwide prevalence, an analysis was made assuming the worldwide prevalence and the conclusion remained the same that screening was found to be more cost-effective than clinical detection. (9) Table 1 shows the result of cost-benefit analysis studies done by Caroll et al. and Vallejo-Torre et al. Table 2 shows the estimated cost of screening, confirmatory test and eventual management per patient in Philippine peso.

Table 1. Results of Cost-Benefit Analysis Studies

Study	Incremental Cost ¹ (a negative number means that screening saves money)	Effectiveness (average effectiveness in QALY)	Incremental Effectiveness ² (a positive number means there is a gain in QALY)	Incremental Cost Effectiveness Ratio ³ (the incremental cost per QALY gained)
Carrol & Downs (US)	(-) \$13	77.19008	0.0005	Dominates
Vallejo-Torres (Spain)	(+) \$1.24	22.573	0.00005	Dominates*

1 - Difference in average cost between testing and not testing

2 - Difference in effectiveness between testing and not testing

3 - "Dominates" means that screening or testing saves money and improves outcomes relative to not screening/testing

* - In the base case computation, the incremental cost per QALY gained was \$24, 677 but the disability cost was not included and they assumed that all patients who had partial BD did not develop symptoms; when disability cost was factored in and it was assumed that those diagnosed with partial BD may develop symptoms the results showed that Screening "Dominates"

Table 2. Estimated Cost of Screening, Confirmatory Test and Eventual Management

Parameter	Screening intervention	Confirmatory Test	Confirmatory Test
	Screening (ENBS)	(Urine Organic Acids and/or Plasma Acylcarnitine)	Gene Mutation Analysis
Unit cost of screening intervention	P1,750	P8,300 for Urine Organic Acid P12, 650 for Plasma Acylcarnitine	250 US Dollars; approximately P20,000-P25,000 inclusive of processing and shipping fee
Cost of Management	Biotin Supplementation: P13 per tablet (5000 mcg/tablet) at a dose of 1-2 tablets daily; P13-P26 x 365 = P4,745 - P9,490 in 1 year		

*Values taken from the Philhealth package and existing rate (as of August 2021) of the Institute of Human Genetics for biochemical and molecular genetic testing.

4.9.5 Equity, Acceptability, and Feasibility

In a study done in New England, patients diagnosed by newborn screening (homocystinuria galactosemia, maple syrup urine disease and biotinidase deficiency) put the families at risk for increased stress and parent-child dysfunction due to false positive screening results.(10) However, patients diagnosed by newborn screening showed less parental stress and less difficulty in meeting their children's need compared with clinically detected patients.(10) There were no specific studies for biotinidase deficiency found on

patient's values and preference, equity, acceptability and feasibility on the literature search done.

4.9.6 Recommendations from Other Groups

The current recommendation of the UK National Screening Committee (NSC) is unavailable since they are currently undergoing revision. The 2012 UK NSC does not recommend newborn screening for biotinidase deficiency.(16) According to the 2012 review committee, the evidence needed to support a screening program for biotinidase deficiency in the UK is not available as demonstrated by the following statements:

1. The condition is known to be rare, but there was a lack of UK prevalence data.
2. There was limited understanding of how the type of deficiency (partial or profound) or genotype is linked with symptomatic outcomes.
3. The screening test is relatively simple and involves newborn DBS sampling, however, the sensitivity and number of false negatives was uncertain as follow-up testing was not performed for screen negatives.
4. A highly effective and safe treatment is available and all children with profound deficiency are treated, however, there was uncertainty around the management of partial deficiency; most children appear to be treated but there was variability in the dose given.

There were no recommendations found from other groups during the time that the literature search was conducted. However, the following countries has included biotinidase deficiency in their newborn screening program: Austria, Denmark, Faroe Islands, Greenland, Norway, Quebec, Greece, Brazil, Minnesota, UAE, Canada, US (particularly Virginia and California), Netherlands, Hungary, 4 cities in central Anatolia (Sivas, Tokat, Malatya, Erzincan) and several countries in Europe (Austria, Belgium Flanders, Belgium Wallonia, Germany, Italy, Spain, Sweden, Switzerland). Majority of the countries who included biotinidase deficiency in their newborn screening program demonstrated that screening for profound biotinidase deficiency and early treatment with biotin result in excellent outcomes for older adolescents and adults with the disease. These outcome results indicate that newborn screening for biotinidase deficiency is one of the most successful newborn screening program and provide further support that newborn screening for biotinidase deficiency is an excellent example of preventing symptoms and ensuring successful outcomes with early and continual treatment.

References

1. Padilla, CD. 2012. Enhancing Case Detection of Selected Inherited Disorders through Expanded Newborn Screening in the Philippines.
2. Wolf, B. 1991. Worldwide survey of neonatal screening for biotinidase deficiency. *J. Inherit. Metab. Dis.* 1991; 14:923–927.
3. Metabolic Registry of Institute of Human Genetics-National Institute of Health
4. Weber P, Scholl S, ER Baumgartner. 2004. Outcome in patients with profound biotinidase deficiency: relevance of newborn screening. *Developmental Medicine & Child Neurology.* 2004; 46: 481–484.
5. Moslinger D, Stockler-Ipsiroglu S, Scheibenreiter S, Teifenthaler M, Muhl A, Seidl R, Strobl W, Plecko B, Suormala T, Baumgartner ER. 2001. Clinical and Neuropsychological Outcome of 33 Patients with Biotinidase Deficiency Ascertained by National Neonatal Screening and Family Studies in Austria. *Eur J Pediatr.* 2001; 160: 277-282.

6. Wolf, B. 2015. Why screen newborns for profound and partial biotinidase deficiency. *Molecular Genetics and Metabolism*. 2015; 114:382-387.
7. Burton BK, Roach ES, Wolf B, Weissbecker KA. 1987. Sudden death associated with biotinidase deficiency. *Pediatrics*. 1987; 79:482–483
8. Carroll A, Downs S. 2006. Comprehensive Cost-Utility Analysis of Newborn Screening Strategies. *PEDIATRICS*. 2006; 117(5).
9. Vallejo-Torres L, Castilla I, Couce ML, Pérez-Cerdá C, Martín-Hernández E, Pineda M, Campistol J, Arrospide A, Morris S, Serrano-Aguilar P. 2015. Cost-Effectiveness Analysis of a National Newborn Screening Program for Biotinidase Deficiency. *PEDIATRICS*. 2015; 136(2).
10. Waisbren SE, Read CY, Ampola M, Brewster TG, Demmer L, Greenstein R, Ingham CL, Korson M, Msall M, Puesch S, Seashore M, Shih VE, Levy HL. 2002. Newborn screening compared to clinical identification of biochemical genetic disorders. Short Report. *J. Inherit. Metab. Dis.* 2002; 25:599-600.
11. Lund AM, Hougaard DM, Simonsen H, Andresen BS, Christensen M, Dunø M, Skogstrand K, Olsen RKJ, Jensen UG, Cohen A, Larsen N, Saugmann-Jensen P, Gregersen N, Brandt NJ, Christensen E, Skovby F, Nørgaard-Pedersen B. 2012. Biochemical screening of 504,049 newborns in Denmark, the Faroe Islands and Greenland—experience and development of a routine program for expanded newborn screening. *Mol Genet Metab*. 2012; 107:281–293.
12. Kwon, C. and Farrel, P. 2000. The Magnitude and Challenge of False Positive Newborn Screening Test Results. *Arch PediatrAdolesc Med.* 2000; 154:714-718
13. Burlina AB, Sherwood WG, Marchioro MV, Dalla Bernardina B, Gaburro D. 1988. Neonatal screening for biotinidase deficiency in north eastern Italy. *Eur J Pediatr.* 1988; 147:317-318.
14. Wiltink RC, Kruijshaar ME, Minkelen RV, Onkenhout W, Verheijen FW, Kemper EA, Spronzen FJ, Ploeg AT, Niezen-Koning KE, Saris JJ, Williams M. 2016. Neonatal screening for profound biotinidase deficiency in the Netherlands: consequences and considerations. *European Journal of Human Genetics.* 2016; 24:1424–1429.
15. Heard G, Wolf B, Jefferson LG, Weissbecker KA, Nance WE, Secor McVoy JR, Napolitano A, Mitchell PL, Lambert FW, Linyear AS. 1986. Neonatal Screening for BD: Results of a 1 year pilot study. *J PEDIATR.* 1986; 108:40-46.
16. Bazian. 2017. Newborn screening for biotinidase deficiency External review against programme appraisal criteria for the UK National Screening Committee. UK NSC external review – Newborn screening for biotinidase deficiency. July 2017.

Included Studies (5 Diagnostic Accuracy, 1 Outcome)

1. Lund AM, Hougaard DM, Simonsen H, Andresen BS, Christensen M, Dunø M, Skogstrand K, Olsen RKJ, Jensen UG, Cohen A, Larsen N, Saugmann-Jensen P, Gregersen N, Brandt NJ, Christensen E, Skovby F, Nørgaard-Pedersen B. 2012. Biochemical screening of 504,049 newborns in Denmark, the Faroe Islands and Greenland—experience and development of a routine program for expanded newborn screening. *Mol Genet Metab.* 2012; 107:281–293.
2. Kwon, C. and Farrel, P. 2000. The Magnitude and Challenge of False Positive Newborn Screening Test Results. *Arch Pediatr Adolesc Med.* 2000; 154:714-718
3. Burlina AB, Sherwood WG, Marchioro MV, Dalla Bernardina B, Gaburro D. 1988. Neonatal screening for biotinidase deficiency in north eastern Italy. *Eur J Pediatr.* 1988; 147:317-318.
4. Wiltink RC, Kruijshaar ME, Minkelen RV, Onkenhout W, Verheijen FW, Kemper EA, Spronzen FJ, Ploeg AT, Niezen-Koning KE, Saris JJ, Williams M. 2016. Neonatal screening for profound biotinidase deficiency in the Netherlands: consequences and considerations. *European Journal of Human Genetics.* 2016; 24:1424–1429.
5. Heard G, Wolf B, Jefferson LG, Weissbecker KA, Nance WE, Secor McVoy JR, Napolitano A, Mitchell PL, Lambert FW, Linyear AS. 1986. Neonatal Screening for BD: Results of a 1 year pilot study. *J PEDIATR.* 1986; 108:40-46.
6. Weber P, Scholl S, ER Baumgartner. 2004. Outcome in patients with profound biotinidase deficiency: relevance of newborn screening. *Developmental Medicine & Child Neurology.* 2004; 46: 481–484.

4.10 Newborn Screening for Beta-Ketothiolase Deficiency

RECOMMENDATIONS

Among asymptomatic, apparently healthy newborns, we recommend against the screening of beta-ketothiolase deficiency. (*Strong recommendation, very low certainty of evidence*).

Considerations

The consensus panel considered the following when formulating this recommendation:

- The very low burden of disease among Filipinos coupled with the high cost of confirmatory testing of the disease led the consensus panel to decide against the screening of beta-ketothiolase deficiency.

4.10.1 Burden of disease

Beta-ketothiolase deficiency (BKTD) also known as mitochondrial acetoacetyl-CoA thiolase deficiency is an autosomal recessive disease of isoleucine and ketone body utilization.(1) Worldwide, a total of 250 confirmed cases of Beta-ketothiolase deficiency have been reported with varying ethnicity.(2) Reported prevalence varies widely between geographic regions from 1 per 313,000 newborns in Minnesota, USA to 1 per 1,419, 156 births in China.(3,4) In the Philippines, no case of BKTD has been reported since the introduction of expanded newborn screening in 2014.(5)

Patients with BKTD clinically manifests between 6 and 36 months of age with acute attacks of ketoacidosis, vomiting, dehydration, lethargy, and seizure commonly triggered by infection, stress and fasting. Newborn manifestation is said to be very rare. Diagnosis can be confirmed by enzyme assay or ACAT1 gene DNA genomic analysis. Management include sufficient oral or intravenous glucose supplementation and blood glucose monitoring for acute crisis and avoidance of fasting with mild fat and protein restriction during asymptomatic periods.(6) Patients with this deficiency generally have a favorable outcome however, early recognition and management may prevent the occurrence of recurrent, severe ketoacidotic crisis and subsequent irreversible neurological complications (ataxia, hypotonia, choreoathetosis, dystonia, developmental delay, coma) or death.(7)

4.10.2 Benefits and Harms of Screening Tests

There were no direct studies on benefit or harm of Beta-kethothiolase deficiency screening versus no screening among healthy and/or asymptomatic newborns in terms of mortality and other patient-related outcomes.

Only one study made by Lin et al in 2021 provided some evidence on the difference in outcomes between BKTD patients identified through newborn screening and those previously diagnosed or selectively screened. This study is a multicenter retrospective cohort which included data from 18 provinces or municipalities in China between January 2009 and May 2020. A total of 16,088,190 newborns were screened and 14 patients were

identified through NBS. In total, twenty-nine patients were genetically diagnosed with BKTD. Clinical outcomes of the cohort of 29 BKTD patients from this study showed that patients identified through newborn screening had fewer cases of mortality (0% vs 30.0%), episodes of metabolic decompensation (21.4% vs 73.3 %), and neurologic impairment (0% vs 46. 7%). This study concluded that patients who did not undergo NBS had a delayed diagnosis and poorer prognosis. However, it is important to note that the authors of the study reported that data for the 15 patients identified through selective screening were incomplete and thus definite conclusions or recommendations cannot be made.

Grunert *et al* reviewed the clinical course and outcome of all patients (244 identified) with BKTD published to date and showed that 8.9% of patients had died. Majority of the deaths occurred during the patient's first metabolic crisis when the diagnosis was still unknown. This study suggests that early detection through early screening could potentially avoid severe metabolic crises and death among patients with BKTD.(2)

The overall certainty of data for the benefit of BKTD screening on the outcomes presented is deemed very low as it is based only one retrospective study with high risk of bias and potential for publication bias. Due to lack of additional and higher quality studies, inconsistency and imprecision cannot be thoroughly evaluated. See Appendix B for GRADE Evidence Profile.

No randomized controlled trials, CPGs or systematic reviews were found comparing outcomes of patients with treatment versus no treatment among those with BKTD. Grunert *et al* in their report made a narrative summary on dietary treatment given for 163 out of 244 confirmed BKTD patients.(2) Eighty-six patients (86/163; 52.7%) were on either a low isoleucine or low-protein diet defined as protein intake of 0.5-2 g/kg/day. Forty-seven patients (47/163; 28.8%) followed a diet restricted in protein and fat. None of the patients received a special amino acid mixture. There is still no evidence to date substantiating the relationship between protein restriction for prevention of neurological manifestation or developmental impairment among BKTD patients.

4.10.3 Diagnostic Performance of Screening Tests

Based on two retrospective observational studies by Estrella *et al*. (2014) and Lin *et al*. (2021), the pooled sensitivity of expanded newborn screening (NBS) in diagnosing BKTD is 0.52 (95% CI 0.35-0.68). The above studies were appraised to have high risk of bias since confirmatory diagnostic test were done on screen-positive samples only with non-blinding of interpreters. Clinical follow-up or further tests on screen-negative cases were not explicitly reported. The true disease status of screen-negative cases is not known hence specificity of the test cannot be derived. Inconsistency and imprecision cannot be thoroughly evaluated since both studies also differ in confirmatory test done thus the overall certainty evidence was deemed very low. See Appendix C for GRADE Evidence Profile.

4.10.4 Cost Implication

There are no published or unpublished cost-effectiveness studies on newborn screening for Beta-Ketothiolase Deficiency. Diagnosis is made by conducting secondary metabolic analysis including plasma acylcarnitine profile (showing elevated Tiglylcarnitine (C5:1) and 2 methyl – 3 hydroxy butyrate Carnitine, C5OH) and urine organic acid analysis.(9) Confirmatory test for BKTD (Enzyme assay or ACAT1 gene mutation analysis), although not available locally, could be done through transport of samples to overseas laboratories. Table 1 shows the estimated screening and additional testing cost per patient based on existing Philhealth package and the Institute of Human Genetics rates for biochemical and molecular genetic testing.

Table 1. Estimated annual cost of screening for Beta-Ketothiolase deficiency

	Regular Rate (PhP)	PWD Rate (PhP)
Expanded NBS	1750	-
Urine Metabolic Screen	4500	3600
Comprehensive Urine Metabolic Profile	10400	8320
Organic Acid Analysis Screen	8300	6640
Plasma Acylcarnitine Quantitation	12,650	10,120
DNA Extraction	2500	-
DNA Analysis	12500 (~250 USD)	-

The cornerstone of management of Beta-ketothiolase deficiency is prevention of ketoacidosis which can be done by avoidance of fasting and glucose supplementation during periods of stress or infection. Majority of patients are advised mild protein and fat restriction for life. L-carnitine supplementation (50-200 mg/kg/day) can be given in select patients with proven low carnitine levels only.(9,10) The estimated cost of management is shown in the table below.

Table 2. Estimated Cost of Management	
Amino Acid Modified Infant Formula Isoleucine, Leucine & Valine Free	~PHP 2348/can (400 g)
L-Carnitine (Carnicor), 1g/10ml oral suspension, 10 bottles per box	PHP 875.16/box

4.10.5 Equity, Acceptability, and Feasibility

No qualitative studies were found on patient's values and preference, equity, acceptability and feasibility specifically for Beta-Ketothiolase deficient patients. According to a public engagement study done among Canadian Citizens, expanded newborn screening is generally deemed beneficial and accepted as part of routine screening. Harm brought about by screening such as anxiety, stigma or unwanted knowledge were generally thought to be minimal in comparison to the consequences of delayed diagnosis especially for the vulnerable pediatric patients.(11) Other studies that explored on parental views on newborn screening have shown that one of the major motivators for parental acceptance of expansion of newborn screening is the healthcare providers' ability to impart

information especially on the succeeding tests and management that the parent can facilitate to maximize benefit to the child's health.(12) Parents of children diagnosed with biochemical genetic disorders through newborn screening also expressed lower levels of stress as children diagnosed earlier had fewer functional and developmental problems compared to those diagnosed only upon onset of clinical symptoms.(13) In the Philippines, centralized, regional follow-up teams which include a physician, nurse, counselor and administrative staff have been established since the expansion of the newborn screening program last 2014 in order to ensure clinical follow-up and proper counseling of patients and their family.(14)

4.10.6 Recommendations from Other Groups

The U.S. Advisory Committee on Heritable Disorders in Newborns and Children and the U.S. Department of Health and Human Services secretary (HHS) established the Recommended Uniform Screening Panel (RUSP) as guidance for each of their state universal newborn screening (NBS) program. The RUSP currently includes Beta-ketothiolase Deficiency as one of the 34 core conditions to be included in the newborn screening test as an expert panel deemed that it met three qualifications: a) It can be identified at a time (24-48 hours after birth) at which it would not ordinarily be detected clinically, b) a test with appropriate sensitivity and specificity is available for it and c) benefits of early detection, timely intervention, and efficacious treatment of the condition have been demonstrated.(15) To date, all states in the U.S include BKTD in their NBS panel. Table 3 shows the other countries that include BKTD in their newborn screening program. Majority of the countries demonstrated that screening for Beta-ketothiolase deficiency is beneficial as most adverse complications were related to the first ketoacidotic episodes when the diagnosis had not yet been established hence early screening is recommended.(15-18)

Table 3. Countries that include Beta-Ketothiolase Deficiency in the NBS Program

Region	Countries
Asia Pacific	Australia, New Zealand, China, India (regional),, Japan, South Korea, Malaysia, Myanmar, New Zealand, Singapore, Sri Lanka, Taiwan, Thailand and Vietnam
Europe	Germany, France, Hungary, Iceland, Israel Italy, North Macedonia, Norway Poland, Slovakia, Slovenia, Spain (pilot), Sweden, Switzerland, Ukraine (pilot)
North America and Canada	Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, District of Columbia, Florida, Georgia, Guam, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, US Virgin Islands, Utah, Vermont, Virginia, Washington, West Virginia, Wisconsin, Wyoming

Latin America	Costa Rica, Chile, Uruguay, Colombia, Brazil (regional), Argentina (Regional)
Middle East and North Africa	Cyprus, Iran, Lebanon, United Arab Emirates, Qatar, Saudi Arabia, Tunisia

References

1. Nguyen KN, Abdelkreem E, Colombo R, Hasegawa Y, Can NTB, Bui TP, et al. Characterization and outcome of 41 patients with beta-ketothiolase deficiency: 10 years' experience of a medical center in northern Vietnam. *J Inherit Metab Dis.* 2017;40:395–401
2. Grünert SC, Sass JO. 2-methylacetoacetyl-coenzyme A thiolase (beta- ketothiolase) deficiency: one disease - two pathways. *Orphanet J Rare Dis.* 2020 Apr 28;15(1):106.
3. Fukao T, Scriver CR, Kondo N; t2 Collaborative Working Group. The clinical phenotype and outcome of mitochondrial acetoacetyl-CoA thiolase deficiency (beta-ketothiolase or T2 deficiency) in 26 enzymatically proved and mutation- defined patients. *Mol Genet Metab.* 2001 Feb;72(2):109-14.
4. Lin, Y., Yang, Z., Yang, C. et al. C4OH is a potential newborn screening marker—a multicenter retrospective study of patients with beta-ketothiolase deficiency in China. *Orphanet J Rare Dis* 2021; 16 (224) <https://doi.org/10.1186/s13023-021-01859-5>
5. Metabolic Registry Data
6. Fukao T, Sasai H, Aoyama Y, Otsuka H, Ago Y, Matsumoto H, Abdelkreem E. Recent advances in understanding beta-ketothiolase (mitochondrial acetoacetyl-CoA thiolase, T2) deficiency. *J Hum Genet.* 2019 Feb;64(2):99-111. doi: 10.1038/s10038-018-0524-x. Epub 2018 Nov 5. PMID: 30393371
7. Abdelkreem E, Akella RRD, Dave U, Sane S, Otsuka H, Sasai H, Aoyama Y, Nakama M, Ohnishi H, Mahmoud S, Abd El Aal M, Fukao T. Clinical and Mutational Characterizations of Ten Indian Patients with Beta-Ketothiolase Deficiency. *JIMD Rep.* 2017;35:59-65
8. Estrella J, Wilcken B, Carpenter K, Bhattacharya K, Tchan M, Wiley V. Expanded newborn screening in New South Wales: missed cases. *J Inherit Metab Dis.* 2014 Nov;37(6):881-7
9. Fukao, T. Beta-ketothiolase deficiency. *Orphanet Encyclopedia*, September 2001 <https://www.orpha.net/data/patho/GB/uk-T2.pdf>
10. Grünert SC, Sass JO. 2-methylacetoacetyl-coenzyme A thiolase (beta- ketothiolase) deficiency: one disease- two pathways. *Orphanet J Rare Dis.* 2020 Apr 28;15(1):106. doi: 10.1186/s13023-020-01357-0
11. Hayeems RZ, Miller FA, Bombard Y, Avard D, Carroll J, Wilson B, et al. Expectations and values about expanded newborn screening: A public engagement study. *Health Expectations.* 2013;18(3):419–29.
12. Nicholls SG, Southern KW. Parental decision-making and acceptance of newborn bloodspot screening: an exploratory study. 2013;8(11): e79441. doi:10.1371/journal.pone.0079441
13. Waisbren SE. Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA.* 2003;290(19):2564.
14. Padilla C. Expanding newborn screening and the initiation of regional follow-up in the Philippines. *Ann Transl Med* 2015;3(S2):AB014. doi: 10.3978/j.issn.2305-5839.2015.AB014
15. American College of Medical Genetics Newborn Screening Expert Group. Newborn screening: toward a uniform screening panel and system--executive summary. *Pediatrics.* 2006 May;117(5 Pt 2):S296-307. doi: 10.1542/peds.2005-2633I. PMID: 16735256
16. Skrinska V, Khneisser I, Schielen P, Loeber G. Introducing and expanding newborn screening in the MENA region. *Int J Neonatal Screen.* 2020;6(1):12.
17. Borrajo G. Newborn screening in Latin America : A brief overview of the state of the art. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics.* 2021;187(3):322-328.
18. Loeber J, Platis D, Zetterström R, Almashanu S, Boemer F, Bonham J et al. Neonatal Screening in Europe Revisited: An ISNS Perspective on the Current State and Developments Since 2010. *International Journal of Neonatal Screening.* 2021;7(1):15.

4.11 Newborn Screening for Holocarboxylase Synthetase Deficiency

RECOMMENDATIONS

Among asymptomatic, apparently healthy newborns, we recommend against the screening of holocarboxylase synthetase deficiency. (**Strong recommendation, very low certainty of evidence.**)

Considerations

The consensus panel considered the following when formulating this recommendation:

- The very low burden of disease among Filipinos coupled with the high cost of confirmatory testing of the disease led the consensus panel to decide against the screening of holocarboxylase synthetase deficiency.

4.11.1 Burden of disease

Holocarboxylase synthetase (HCLS) deficiency is an autosomal recessive disorder of biotin metabolism which leads to multiple carboxylase deficiency with manifestations which include but not limited to emesis, hypotonia, lethargy, seizures, metabolic ketolactic acidosis, hyperammonemia, developmental delay, skin rash and alopecia in newborns within hours to weeks after birth.(1) Being considered as one of the rarest inborn error of metabolism, exact prevalence worldwide is currently unknown but prevalence at birth is presently estimated to be less than 1/20,000.(2) According to the 2020 Philippine Newborn Screening Report, the prevalence of HCLS is 1/1,069,667 since inception of routine testing in the country.(3)

In the absence of early recognition and treatment, patients with HCLS deficiency often have a severe natural course resulting to long-term neurological sequelae and significant intellectual disability.(1) Early intervention with biotin and lifelong use satisfactorily improved the outcome of both symptomatic and asymptomatic neonates.(2)

4.11.2 Benefits and Harms of Screening Tests

Carrying out a comprehensive search of literature, there were no clinical practice guidelines, systematic reviews, meta-analysis or randomized clinical trials found to address the benefits and harm of screening for HCLS deficiency. However, several case reports and case series from Asian countries (e.g. China, Vietnam, Pakistan and Japan), Samoa, Ghana, Germany, Italy and the United States of America (USA) were documented involving 25 patients.(4-14) Only one case underwent newborn screening and the rest were screened using several metabolic panel after becoming symptomatic and unresponsive to clinical management. Several cases immediately presented with classic HCLS manifestations during the neonatal period but majority of them presented with a chronic course prior to diagnosis. Recognition of HCLS deficiency through enzyme determination and DNA sequence analysis led to prompt initiation of biotin supplementation with usual dose range of 5-20 mg/day which significantly improved the dermatologic lesions and neurologic complications in most cases. (Appendix)

4.11.3 Diagnostic Performance of Screening Tests

In an observational study by Lund et al., expanded newborn screening (NBS) showed an overall false positive rate of 0.038%, positive predictive value of 37% and specificity of 99.99%.⁽¹⁵⁾ In newborns diagnosed with holocarboxylase synthetase deficiency, three newborns had true positive results and no report of false negative reading hence calculated test sensitivity was 100%. However, no specific data was documented for false positive and true negative results for HCLS deficiency hence specificity cannot be derived.

The above study was appraised with unclear and high risk of bias in the patient selection and reference standard criteria exemplified by non-disclosure regarding avoidance of case control design and inappropriate exclusions and non-blinding of interpreters as they were already aware of the positive results prior to the conduct of the mutational analysis. With only one included study, inconsistency and imprecision cannot be thoroughly evaluated hence the overall certainty of evidence was thus deemed very low.

Once a newborn tests positive for HCLS deficiency, demonstration of deficient HCLS activity in leukocytes or fibroblast extracts and/or mutation analysis will be carried out as confirmatory diagnostics. As soon as the diagnosis is certain, patients are started on free biotin supplementation with the goal to improve clinical status of asymptomatic newborns and to prevent progression of symptoms in asymptomatic counterparts.⁽¹⁶⁾

4.11.4 Cost Implication

There is no available cost-effective analysis study of newborn screening focusing specifically on HCLS deficiency. However, in a study by Tiwana et.al., expanded NBS in Texas involving classical organic acid disorder in general was noted to be cost-effective with an incremental cost-effectiveness ratio for screened group at about \$1, 800 quality-adjusted life-years translating to improved patient outcomes by preventing avoidable morbidity and mortality and having better quality of life without incurring extremely high treatment costs.⁽¹⁷⁾ The following table summarizes the costing parameters based on the Philhealth package and existing rate (as of August 2021) of the Institute of Human Genetics for biochemical and molecular genetic testing.

	Regular Rate (PhP)	PWD Rate (PhP)
Expanded NBS	1, 500	-
Organic Acid Analysis Screen	8, 300	6, 640
Comprehensive Urine Metabolic Profile	10, 400	8, 320
Urine Metabolic Screen	4, 500	3, 600
DNA Extraction	2, 500	-
DNA Sequencing	12, 500 (~250 USD)	-

Likewise, the lifetime supplementation of biotin as primary treatment with a dose of 5-20 mg/day currently costs Php13 per 5000 mcg/tablet with an estimated annual cost of Php 4, 745 to 18, 980.

4.11.5 Equity, Acceptability, and Feasibility

In the 2003 study by Waisbren et al., expanded NBS showed better health outcomes for affected population with better overall parental stress index in the screened group compared to the clinically identified patients. However, one drawback documented from the above study was the effect of false-positive screening result which placed families at risk for increased stress and parent-child dysfunction.(18) Improved communication with parents regarding the need for repeat screening to mitigate the effects of false-positive results was supported by a study by Gurian et al. in 2006.(19) In a public engagement Canadian study, majority (77-88%) of subjects endorsed screening without explicit consent for treatable disorders with 62% of them supporting unpressured choice for screening for untreatable disorders as participants valued treatment-related benefits for the affected infant and informational benefits for the families.(20) Public perception toward expanded NBS was favorable in a 2015 survey conducted involving Canadian families where all respondents preferred to avoid false-positive results and overdiagnosis but were willing to accept these to achieve a moderate clinical benefit.(21) De Luca in her 2018 study documented that most participants were enthusiastic about expanded NBS but participants with more years of educations were cautious regarding the extensive costs of diagnosing and management of rare disorders.(22) There were no specific studies for HLCS deficiency and local studies found during the literature search.

4.11.6 Recommendations from Other Groups

As of press time, there are no available recommendations regarding screening of holocarboxylase synthetase deficiency in asymptomatic newborns. However, screening for holocarboxylase synthetase deficiency is presently included in the routine expanded newborn screening in several areas such as the European Union which include Netherlands, Norway and Denmark among others, Australia, USA and Asian nations which include China, Qatar, Korea, Japan and the Philippines.(23, 24)

References

1. Donati TR, Blackburn PR, Atwal PS. Holocarboxylase Synthetase Deficiency Pre and Post Newborn Screening. *Molecular Genetics and Metabolism Reports*. 2016;7:40–4. Available from: doi: 10.1016/j.ymgmr.2016.03.007.
2. Wolf B. Orphanet: Holocarboxylase Synthetase Deficiency [Internet]. Available from: <https://www.orpha.net> [Accessed 17th August 2021].
3. Newborn Screening Reference Center. NBS statistics [Internet]. Available from: <https://newbornscreening.ph> [Accessed 17th August 2021].
4. Xiong Z, Zhang G, Luo X, Zhang N, Zheng J. Case Report of Holocarboxylase Synthetase Deficiency (Late-Onset) in 2 Chinese patients. *Medicine (Baltimore)*. 2020;99(18):e19964. Available from: doi: 10.1097/MD.00000000000019964.
5. Zheng Z, Yuan G, Zheng M, Lin Y, Zheng F, Jiang M, et al. Clinical, Biochemical and Genetic Analysis of a Chinese Han Pedigree with Holocarboxylase Synthetase Deficiency: A Case Report. *BMC Med Genet*. 2020;21:155. Available from: doi: <https://doi.org/10.1186/s12881-020-01080-4>.
6. De Castro M, Zand DJ, Licherter-Konecki U, Kirmse B. Severe Neonatal Holocarboxylase Synthetase Deficiency in West African Siblings. *JIMD Rep*. 2015;20:1–4. Available from: doi:10.1007/8904_2014_367.
7. Hui J, Law E, Chung C, Fung S, Yuen P, Tang N. The First Reported HLCS Gene Mutation Causing Holocarboxylase Synthetase Deficiency in a Vietnamese Patient. *World J Pediatr*. 2012;8(3):278–80. Available from: doi: 10.1007/s12519-011-0301-9.
8. Wilson CJ, Myer M, Darlow BA, Stanley T, Thomson G, Baumgartner ER, et al. Severe Holocarboxylase Synthetase Deficiency with Incomplete Biotin Responsiveness Resulting in

- Antenatal Insult in Samoan Neonates. *J Pediatr.* 2005;147(1):115–8. Available from: doi: 10.1016/j.jpeds.2005.03.006.
- 9. Yokoi K, Ito T, Maeda Y, Nakajima Y, Kurono Y, Sugiyama N, et al. A Case of Holocarboxylase Synthetase Deficiency with Insufficient Response to Prenatal Biotin Therapy. *Brain Dev.* 2009;31(10):775–8. Available from: doi: 10.1016/j.braindev.2008.12.016.
 - 10. Van Hove JLK, Josefsberg S, Freehauf C, Thomas JA, Thuy LP, Barshop BA, et al. Management of a Patient with Holocarboxylase Synthetase Deficiency. *Mol Genet Metab.* 2008;95(4):201–5. Available from: doi:10.1016/j.ymgme.2008.09.006.
 - 11. Cadieux-Dion M, Gannon J, Newell B, Nopper AJ, Jenkins J, Heese B, et al. Delayed Diagnosis of Holocarboxylase Synthetase Deficiency in Three Patients with Prominent Skin Findings. *Pediatr Dermatol.* 2021;38(3):655–8. Available from: doi: 10.1111/pde.14586.
 - 12. Santer R, Muhle H, Suormala T, Baumgartner ER, Duran M, Yang X, et al. Partial Response to Biotin Therapy in a Patient with Holocarboxylase Synthetase Deficiency: Clinical, Biochemical, and Molecular Genetic Aspects. *Mol Genet Metab.* 2003;79(3):160–6. Available from: doi: 10.1016/s1096-7192(03)00091-x.
 - 13. Fatima S, Aamir M, Bibi A. Multiple Carboxylase Deficiency Organic Acidemia as a Cause of Infantile Seizures. *J Coll Physicians Surg Pak.* 2021;3(1):95–7. Available from: doi: 10.29271/jcpsp.2021.01.95.
 - 14. Morrone A, Malvagia S, Donati MA, Funghini S, Ciani F, Pela I, et al. Clinical Findings and Biochemical and Molecular Analysis of Four Patients with Holocarboxylase Synthetase Deficiency. *Am J Med Genet.* 2002;111(1):10–8. Available from: doi: 10.1002/ajmg.10532.
 - 15. Lund A, Hougaard D, Simonsen H, Andresen B, Christensen M, Dunø, M, Skogstrand K, et al. Biochemical Screening of 504,049 Newborns in Denmark, the Faroe Islands and Greenland — Experience and Development of a Routine program for Expanded Newborn Screening. *Mol Genet Metab.* 2012;107(3): 281–293. Available from: doi: <https://doi.org/10.1016/j.ymgme.2012.06.006>.
 - 16. National Center for Advancing Translational Sciences. Holocarboxylase Synthetase Deficiency [Internet]. Available from: <https://rarediseases.info.nih.gov/diseases/2721/holocarboxylase-synthetase-deficiency> [Accessed 17th August 2021].
 - 17. Tiwana SK, Rascati KL, Park H. Cost-Effectiveness of Expanded Newborn Screening in Texas. *Value in Health.* 2012;15(5):613–21. Available from: doi: <https://doi.org/10.1016/j.jval.2012.02.007>
 - 18. Waisbren SE. Effect of Expanded Newborn Screening for Biochemical Genetic Disorders on Child Outcomes and Parental Stress. *JAMA.* 2003;290(19):2564 -72. Available from: doi: 10.1001/jama.290.19.2564.
 - 19. Gurian EA. Expanded Newborn Screening for Biochemical Disorders: The Effect of a False-Positive Result. *Pediatrics.* 2006;117(6):1915–21. Available from: doi: 10.1542/peds.2005-2294.
 - 20. Hayeems RZ, Miller FA, Bombard Y, Avard D, Carroll J, Wilson B, et al. Expectations and Values About Expanded Newborn Screening: A Public Engagement Study. *Health Expect.* 2013;18(3):419–29. Available from: doi: 10.1111/hex.12047
 - 21. Miller FA, Hayeems RZ, Bombard Y, Cressman C, Barg CJ, Carroll JC, et al. Public Perceptions of the Benefits and Risks of Newborn Screening. *Pediatrics.* 2015;136(2):e413-23. Available from: doi: 10.1542/peds.2015-0518.
 - 22. DeLuca JM. Public Attitudes toward Expanded Newborn Screening. *J Pediatr Nurs.* 2018;38:e19-e23. Available from: doi: 10.1016/j.pedn.2017.10.002.
 - 23. Pourfarzam, M., & Zadhoush, F. Newborn Screening for inherited metabolic disorders; news and views. *Journal of Research in Medical Sciences : The Official Journal of Isfahan University of Medical Sciences.* 2013; 18(9): 801–808.
 - 24. Martínez-Morillo E, García B, Menéndez F. Challenges for Worldwide Harmonization of Newborn Screening Programs, *Clinical Chemistry.* 2016; 62 (5): 689–698. Available from <https://doi.org/10.1373/clinchem.2015.240903>

4.12 Newborn Screening for Isovaleric Acidemia

RECOMMENDATION

Among asymptomatic, apparently healthy newborns, we recommend against the screening of isoaleric acidemia. (**Strong recommendation, very low certainty of evidence.**)

Considerations

The consensus panel considered the following when formulating this recommendation:

- The very low burden of disease among Filipinos coupled with the high cost of confirmatory testing of the disease led the consensus panel to decide against the screening of isoaleric acidemia.

4.12.1 Burden of disease

Isovaleric acidemia (IVA) is caused by a defect in leucine catabolism due to deficiency of the mitochondrial isoaleryl CoA dehydrogenase (IVD), leading to the accumulation of derivatives of isoaleryl CoA, including isoalerylcarnitine (C5). Introduction of tandem mass spectrometry (MS/MS) for newborn screening (NBS) allowed the detection of elevated levels of C5 in dried blood spots. In urine, the elevation of IVG confirms the metabolic diagnosis of IVA.(1)

The phenotypic spectrum is broad. “Classic” IVA” present with early disease onset (EO, <29 days) among untreated newborns who may succumb to death due to neonatal hyperammonemic encephalopathy and severe metabolic acidosis. While in the late onset group (LO, ≥29 days), the first clinical symptoms occur only after the newborn period and is often followed by poor neurocognitive outcome due to delayed detection and intervention. The “mild” variant with an attenuated disease was identified only in the NBS era, and the most common cause is a missense variant of IVD gene (c.941C>T, p.Ala311Val).(2)

Early diagnosis through newborn screening (NBS) and treatment with a protein restricted diet and supplementation with L-carnitine and glycine are effective in promoting normal development in severely affected individuals.(1)

With the identification of individuals with mild IVA by NBS, estimated birth prevalences of IVA increased from 1 in 280 000 newborns to 1 in 90 000 to 100 000 newborns worldwide, suggesting the phenotype detected by NBS is different and may include individuals that would not have presented clinically.(2) In the Philippines, there are 3 confirmed cases of IVA from the 3,209,001 babies screened since 2014, giving a prevalence of 1.1 in 1,000,000.(3)

4.12.2 Benefits and Harms of Screening Tests

Benefits of Screening for IVA

Data on the benefits of screening versus no screening on mortality, presence of metabolic decompensations, and neurocognitive impairments were pooled from three longitudinal, observational studies.(2,4,5) In a systematic, cost-effectiveness analysis done in Thailand, outcomes, including QALYs, were compared between their “pre-expanded newborn screening program”- where only phenylketonuria is screened – versus the “expanded newborn screening program using MS/MS” where six prioritized diseases were screened, including IVA, followed by pre-symptomatic treatment.(6) In all studies, there were weak methodological study designs, with serious risk of bias across all outcomes due to absence of blinding, clinical heterogeneity (since outcome analysis included both mild and classic/severe phenotypes), use of non-contemporaneous controls and low event rates. Therefore, the overall certainty of evidence is rated as very low.

Mortality (2 Observational studies, N=197, screened=36, unscreened=161, Very Low Certainty of Evidence) (

Mortality in screened individuals tended to be lower (2.8%, 1/36) than in the pre-screening era (18.6%; 30/161), OR 0.12 (95% CI 0.16 to 0.95). 2,4)

Metabolic Decompensations (2 Observational studies, N=26, screened=13, unscreened=13, Very Low Certainty of Evidence)

Among screened IVA individuals 4/13 (30.8%) experienced at least one metabolic decompensation, while 61.5% of unscreened patients (8/13) had metabolic decompensation/s. The computed OR is 0.28, with 85% CI 0.055 to 1.41, crossing the line of no effect.(4,5) Therefore, NBS but does not reliably prevent the manifestation of total neonatal and recurrent metabolic decompensations, but the crises of the screened individuals were milder and had good clinical outcomes, as observed in the Polish cohort.(5)

Neurocognitive Impairment (2 Observational studies, N=209, screened=115, unscreened=94, Very Low Certainty of Evidence)

Screening showed benefit on long-term neurocognitive outcomes of individuals. Around 21/94 (22.3%) of the NBS cohort were assessed to have neurocognitive impairment, whereas 45/115 (39.1%) of unscreened individuals had some form of neurocognitive delay. (2,5)

Quality-Adjusted Life Year (1 Systematic review, N=23, Very Low Certainty of Evidence)

IVA patients noticeably benefit from early detection which extends their QALY by 3.69.(6)
See Appendix C for GRADE Evidence Profile.

Harm of Screening for IVA

There were no systematic studies on the harms directly related to screening of IVA in newborns.

It is known that false positive screening results may affect parental stress and the parent-child relationship.(7) Tangeraas, et.al., narrated that in screened positive newborns for any disorder, their families maybe left in a state of uncertainty during the confirmatory process.(8) Unfortunately in IVA, the positive predictive value of a raised C5 acylcarnitine may be quite low resulting to significant FP results. The FP cases may be due to pivaloylcarnitine, a compound isobaric to, and interfering with the measurement of C5. Pivaloylcarnitine can be present in blood due to pivalate-containing antibiotics and more recently, the inclusion of neopentanoate as an emollient in some moisturizing creams used as nipple balms by nursing mothers.(9)

Because of NBS, an additional biochemically mild and potentially asymptomatic form of IVA was discovered.(1) Long-term outcomes in these individuals is excellent, which is supported by the lack of metabolic decompensations, lower rates of hospitalizations, and normal psychomotor development and cognitive function, regardless of dietary management. The results of Mutze and colleagues therefore suggest that newborns with mild IVA identified by NBS are at high risk of overtreatment. Strict diet applied during this formative period of life might have a lifelong negative impact on eating behavior.(2)

4.12.3 Diagnostic Performance of Screening Tests

Diagnostic Performance of NBS in diagnosing IVA

Performance of screening for IVA with MSMS, with C5 as primary analyte, was conducted in Norway (2020) and Sweden (2020).(8,10) In the Swedish study, the confirmatory tests included biochemical and genetic tests; while a rapid DNA analysis using Sanger sequencing was performed as confirmatory with supplementary diagnostic biochemical test in the Norwegian study.

In both studies, a second-tier method of separation of C5 isomers (pivaloylcarnitine, isovalerylcarnitine and 2-methylbutyrylcarnitine) was employed using a simple separation step, LC-MSMS.(8,10,11) Pivaloylcarnitine is a pivmecillinam metabolite isobaric to isovalerylcarnitine (C5) and the cause of many false-positive results in first-tier screening for IVA. Betalactam antibiotics are esterified with pivalic acid for better enteral absorption. The high percentage of NBS false positive results for IVA in countries where pivalic acid containing antibiotics are being prescribed could be attributed to this.(11) More recently maternal use of pivalic ester pro-drugs or pivalic acid derivatives used as emollients in some nipple creams have also contributed to a significant number of FPs.(9)

Both studies screened a total of 1,461,703 newborns. Eight cases of IVA were detected and confirmed, giving a combined prevalence of 1 in 182,713. PPV ranged from 33.3% to 100%, with a mean of 66.7%. Remarkably, in the Swedish cohort, the second-tier test with LC-MSMS was implemented after just over 7000 babies had been screened. Before that, there was a high number of false positives (and no true positive cases) for IVA due

to maternal treatment with pivalinic acid containing antibiotics. After implementation of the second tier, there has only been one false positive recall, resulting in a PPV of 86%.(10)

During the study period, missed cases in any of the screened disorders were reported to the study group, but none were reported for IVA. Based on both studies, NBS has a 100% sensitivity, 99.99% specificity, and 100% NPV for IVA screening.(8,10)

Additionally, three observational studies from 3 Asian countries and 1 from Germany, which used only one tier screening step with MSMS, were reviewed.(12-15) One-tier screening test is the algorithm being followed in the Philippines. The sensitivity and specificity of one-tier testing was comparable from the two-tier testing.(12-15) However, data from these studies could not be pooled due to missing information and lack of systematic follow-up of screen-negative babies.

See Appendix C for GRADE Evidence Profile.

4.12.4 Cost Implication

There are no locally published economic evaluation studies that estimate the cost-effectiveness/ benefit/ utility of newborn screening for IVA.

The current estimated cost for the screening for IVA via MSMS using dried blood samples and the cost of its confirmatory test (urine organic acid analysis, plasma acylcarnitine and DNA Analysis) in the Philippines is shown in Table 1, as well as the estimated unit cost of therapeutic products for IVA.(16)

Table 1. Estimated cost of screening for and managing Isovaleric Acidemia

Parameter	Cost
Screening intervention	
Expanded Newborn Screening and MSMS	Php 1,500
Confirmatory Test	
Urine organic acids	Php 8,300
Plasma Acylcarnitine	Php 12,650
DNA Analysis via Next Generation Sequencing (NGS)	\$250* = Php 12,249 USD 1 = Php 49.98
Medical Food	
Amino acid based leucine-free powdered formula (IVA Anamix Early Years/ IVA Anamix Infant/ LMD) • For 0-1 year old	Php 7,200/can
Amino acid based leucine- free powdered formula (IVA Anamix Next /LMD) • For 1 year and above	Php 10,556/can
Medications	
L-Carnitine (Carnicor), 1g/10ml oral suspension,	Php 875.16/box

10 bottles per box	
Glycine 500mg per tablet, 100's	Php 840/bottle

An evaluation of the cost-effectiveness of expanding NBS in the UK, including IVA, concluded that screening dominated over not screening. It was estimated to increase both QALYs and reduce overall costs – including the cost of long-term health and social care impact (no screening £262,377 vs screening £48,313), cost of managing the disease (no screening £171,859 vs screening £69,704) and life-time QALYs gained (no screening 29.90 vs screening 39.29).(17) However, it stated that greatest decision uncertainty are those associated with IVA, due to the question regarding it's true clinical birth prevalence and the specificity of the test. There is the problem of over-detection estimated for IVA is due to the mild, asymptomatic phenotypes, not normally diagnosed in unscreened individuals. The NSC recommends that IVA screening should have a lower test threshold than that initially used in their pilot in order to reduce potential over-detection.(17)

According to modelled survival estimate for IVA (Appendix D), mortality associated with non-screened IVA occurs in the first two years of life, primarily in the neonatal period, and only among patients with the classic phenotype (i.e. those not carrying the low-risk c.932C>T mutation). The model assumes that mortality after the first seven days is avoided in the screen-detected group and that survival is unaffected in the low risk subgroup. Overall, the model estimated a survival of approximately 95% in the screen-detected cohort of IVA patients and 80% in the symptomatically- detected (non-screened) cohort.(17)

While a cost-analysis study done by Thiboonboon and colleagues based on the Thai population suggests that early medical treatment indeed improves the health outcome in IVA patients resulting in 3.69 QALY higher than those with late detection. However, their data showed an ICERs which are above the agreed economic threshold used in Thailand. Further, they demonstrated that the costs of giving lifetime care in IVA patients who were detected early (THB 3,728,014) are higher than those in clinically diagnosed patients (THB 3,409,629), where 1 I\$ = 17.79 THB in 2013. Reasons for this included lower incidence of these disorders which makes it less favorable and the accrued lifetime maintenance and preventive treatment these saved patients will require. They concluded that screening for additional IEMs in Thailand using MS/MS is economically unattractive. They recommend continuing their program of screening for PKU only, while prioritizing early treatment for some disorders, including IVA due to their favorable outcomes.(6)

4.12.5 Equity, Acceptability, and Feasibility

It is well-established that the early and asymptomatic detection of inherited metabolic disorders by NBS can offer very significant health benefits for the affected child and the family. However, there is also evidence that FP results reported by NBS programs can have a significant negative impact on families and in some cases these effects may endure for many years.(9) In the survey by Gurian, parents in the false-positive group attained higher total scores on the parenting stress index (PSI) than did parents in the normal-screened group. They also scored higher on the parent-child dysfunction subscale and the difficult child subscale. Additionally, while all mothers worry about their child's

health, mothers in the FP group reported more worry about their child's future (mean: 2.4; SD: 1.4), compared with mothers in the normal-screened group (mean: 1.9; SD: 1.0; P = 0.04). The anxiety also reflects as parents being more overprotective and more focused on physical symptoms of their child, that could be associated with an increase in the frequency of hospital attendances during the first 6 months of life among FP children (mean: 0.17 hospitalization vs 0.08 hospitalization; P = 0.09).(7)

Therefore, families would benefit if FP results can be avoided by improving the PPV of the screening methodology. This is especially relevant in IVA, wherein the contribution made by pivaloylcarnitine to raised C5 acylcarnitine results is well-established. This can be done through the introduction of second tier testing for C5 isomers following an initial elevated C5.(8,10,11) In addition, decreasing the time needed to settle the FP cases may diminish the burden of FP cases on both families and the health system. Lastly, in a fraction of FP cases, parents expressed high levels of understanding and confidence towards the screening system even when their child was exposed to unnecessary testing, provided that follow-up was scheduled without delay and access to the metabolic team was given while waiting.(8)

4.12.6 Recommendations from Other Groups

There was only one (1) CPG that could answer our research question. This is based on preliminary data from the European registry and network for Intoxication type Metabolic Diseases (E-IMD).(18) They recommended that NBS can help prevent serious outcomes of IVA patients, and that most of the IVA patients who have been diagnosed early, and fewer patients with a late diagnosis, develop normally. But these were based on low grades of recommendations (Table 2).

Table 2. Recommendations from European registry and network for Intoxication type Metabolic Diseases (E-IMD)(18)

Group	Recommendation	Grade of Recommendation
E-IMD (2014)	NBS, though patients can be either symptomatic or asymptomatic, can help prevent serious outcome of IVA patients.	D*
	Most of the IVA patients who have been diagnosed early (< 6 weeks of life) and fewer patients with a late diagnosis develop normally. If diagnosed early, IVA outcome is much better than if diagnosed late.	D*

* Evidence level 3 and 4; or Extrapolated evidence from studies rated as 2+

The first countries that introduced IVA to their NBS programs were Australia and Germany. Since then, it has been implemented in NBS programs in about 50 countries worldwide, but with varying coverage.¹

Table 3. Countries with published experience in NBS for IVA(1)

Region (# of countries)	Countries
Asia Pacific (11)	Australia, China, India, Japan, Malaysia, New Zealand, Philippines, Singapore, South Korea, Thailand, Taiwan
Europe (22)	Austria, Belgium, Czech Republic, Denmark, Estonia, Germany, Greece, Hungary, Iceland, Italy, Liechtenstein, Macedonia, Netherlands, Norway, Poland, Portugal, Russia, San Marino, Spain, Sweden, Switzerland, United Kingdom
North America (2)	United States, Canada
South America (9)	Argentina, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Mexico, Uruguay, Venezuela
Africa (1)	South Africa
Middle East (5)	Kuwait, Lebanon, Saudi Arabia, Qatar, United Arab Emirates

The apparent benefits demonstrated for patients pre-symptomatically diagnosed with IVA through NBS and treated early make it an ideal candidate for NBS programs. However, an additional “mild” IVA phenotype with only slight biochemical abnormalities and a potentially asymptomatic has been detected by NBS. At present, there is little information on the long-term outcomes of this patient cohort, and if they are really at risk for severe catabolic episodes.(1)

There is extreme heterogeneity in the treatment strategies used for mild IVA – with many individuals receiving intensive treatment at least in infancy, while others did not. This supports the belief that the overall excellent clinical outcome is not relevantly affected by metabolic control. Mutze and colleagues supports the assumption of “mild IVA” to be a benign variant, therefore may not require metabolic treatment and eventually may be excluded from NBS panels.(2) In relation to that, several authors recommend that longitudinal studies of IVA-screened individuals should be established to allow for a better understanding of the long-term outcome and clinical spectrum including definition of the “mild” phenotype and to provide the basis for management recommendations and counseling.(1,2,8) Results from additional studies may also lead to adjustment of NBS cut-off levels in order to not detect individuals with benign variants.(1)

Due to the known interference of the different C5 isomers, especially pivaloylcarnitine, in the screening results for IVA, improvement in the PPV could be achieved by implementing of second-tier analyses prior to referral.(7,9,10)

It is interesting to note that while many studies have shown that MS/MS is cost-effective in their specific country settings, the data comes mostly from high-income countries. Hence, this cannot be directly applied to other countries due to generalizability and transferability issues, and differences in multiple factors such as demography, epidemiology of disease, health infrastructure, clinical practice, and healthcare cost. The

study by Thiboonboon, et al. is the first study that comprehensively evaluated the cost-effectiveness of MS/MS in low- and middle-income countries (LMICs). While they acknowledge the benefits of early management for IVA provided by NBS, they concluded that implementing the MS/MS screening program in Thailand does not meet the criteria for cost-effectiveness based on their current local threshold recommendations.(6)

References

1. Schlune A, Riederer A, Mayatepek E, Ensenauer R. Aspects of Newborn Screening in Isovaleric Acidemia. International Journal of Neonatal Screening. 2018;4(1):7.
2. Mütze U, Henze L, Gleich F, Lindner M, Grünert S, Spiekerkoetter U et al. Newborn screening and disease variants predict neurological outcome in isoformic aciduria. Journal of Inherited Metabolic Disease. 2021;44(4):857-870.
3. Newborn Screening Reference Center 2020, accessed 31 July 2021, https://newbornscreening.ph/index.php?option=com_content&view=article&id=83:nbsstatistics&c_atid=32:statistics&Itemid=57
4. Couce M, Aldamiz-Echevarría L, Bueno M, Barros P, Belanger-Quintana A, Blasco J et al. Genotype and phenotype characterization in a Spanish cohort with isoformic acidemia. Journal of Human Genetics. 2016;62(3):355-360.
5. Szymbańska E, Jezela-Stanek A, Bogdańska A, Rokicki D, Ehmke vel Emczyńska-Selig E, Pajdowska M et al. Long Term Follow-Up of Polish Patients with Isovaleric Aciduria. Clinical and Molecular Delineation of Isovaleric Aciduria. Diagnostics. 2020;10(10):738.
6. Thiboonboon K, Leelahavarong P, Wattanasirichaigoon D, Vatanavicharn N, Wasant P, Shotelersuk V et al. An Economic Evaluation of Neonatal Screening for Inborn Errors of Metabolism Using Tandem Mass Spectrometry in Thailand. PLOS ONE. 2015;10(8):e0134782.
7. Gurian E. Expanded Newborn Screening for Biochemical Disorders: The Effect of a False-Positive Result. PEDIATRICS. 2006;117(6):1915-1921.
8. Tangeraas T, Sæves I, Klingenberg C, Jørgensen J, Kristensen E, Gunnarsdottir G et al. Performance of Expanded Newborn Screening in Norway Supported by Post-Analytical Bioinformatics Tools and Rapid Second-Tier DNA Analyses. International Journal of Neonatal Screening. 2020;6(3):51.
9. Bonham J, Carling R, Lindner M, Franzson L, Zetterstrom R, Boemer F et al. Raising Awareness of False Positive Newborn Screening Results Arising from Pivalate-Containing Creams and Antibiotics in Europe When Screening for Isovaleric Acidemia. International Journal of Neonatal Screening. 2018;4(1):8.
10. Sörensen L, von Döbeln U, Åhlman H, Ohlsson A, Engvall M, Naess K et al. Expanded Screening of One Million Swedish Babies with R4S and CLIR for Post-Analytical Evaluation of Data. International Journal of Neonatal Screening. 2020;6(2):42.
11. Janzen N, Steuerwald U, Sander S, Terhardt M, Peter M, Sander J. UPLC-MS/MS analysis of C5-acylcarnitines in dried blood spots. 2021.
12. Lin Y, Chen D, Peng W, Wang K, Lin W, Zhuang J et al. Newborn screening for isoformic acidemia in Quanzhou, China. Clinica Chimica Acta. 2020;509:25-29.
13. Yunus Z, Rahman S, Choy Y, Keng W, Ngu L. Pilot study of newborn screening of inborn error of metabolism using tandem mass spectrometry in Malaysia: outcome and challenges. Journal of Pediatric Endocrinology and Metabolism. 2016;29(9).
14. Lim J, Tan E, John C, Poh S, Yeo S, Ang J et al. Inborn Error of Metabolism (IEM) screening in Singapore by electrospray ionization-tandem mass spectrometry (ESI/MS/MS): An 8year journey from pilot to current program. 2021.
15. Ensenauer R, Fingerhut R, Maier E, Polanetz R, Olgemöller B, Rösninger W et al. Newborn Screening for Isovaleric Acidemia Using Tandem Mass Spectrometry: Data from 1.6 Million Newborns. Clinical Chemistry. 2011;57(4):623-626.
16. National Institutes of Health-Institute of Human Genetics. ENBS Medications Costing. 2021. Manila.

17. Bessey A, Chilcott J, Pandor A, Paisley S. The Cost-Effectiveness of Expanding the UK Newborn Bloodspot Screening Programme to Include Five Additional Inborn Errors of Metabolism. *Int J Neonatal Screen*. 2020 Nov 20;6(4):93. doi: 10.3390/ijns6040093.
18. European Registry and network for introduction type metabolic diseases (E-IMD) Isovaleric Acidemia: Quick Reference Guide 2014, accessed 22 Spetember 2021, http://www.e-imd.org/rc/e-imd/htm/Article/2014/e-imd-20140716-085102-695/src/htm_fullText/en/IVA%20guideline_Quick%20reference%20guide_Ensenauer_201408.pdf

4.13 Screening for Developmental Delay

RECOMMENDATIONS

1. Among asymptomatic, apparently healthy children born preterm*, we recommend screening for developmental delay at 3-5 months, 12 months, 24 months corrected age and at 36-48 months of age. (*Strong recommendation, low certainty of evidence*).
2. Among asymptomatic, apparently healthy children who have any of the following risk factors: maternal alcohol use during pregnancy, gestational diabetes, gestational hypertension or maternal obesity, we recommend screening for developmental delay at 9-, 18- and 24-30 months. (*Strong recommendation, low certainty of evidence*).
3. Among asymptomatic, apparently healthy children who were exposed to maternal cigarette smoking during pregnancy, there is insufficient evidence to recommend for the screening of developmental delay. (*Low certainty of evidence*).
4. Among asymptomatic, apparently healthy children whose mothers were anemic, there is insufficient evidence to recommend for the screening of developmental delay. (*Low certainty of evidence*).

*Preterm defined as less than 37 weeks age of gestation

Considerations

The consensus panel considered the following when formulating these recommendations:

- Overall, the panel discussed the significance of these risk factors. Based on the first PHEX, universal screening for asymptomatic, apparently healthy children was not recommended. Therefore, these high risk populations were identified as possible risk factors for developmental delay.
- The panel also discussed the need to include those with stormy neonatal courses in the next update.
- Prematurity
 - The consensus panel agreed that along with the high prevalence of premature births, the screening for developmental delay will lead to early intervention and therefore, improved outcome. These benefits outweigh the risks of screening.
 - The ages at which these children should be screened was an issue discussed by the panel. In the end, the frequency for screening was based on the 2017 NICE Guidelines on Developmental Follow-up of Children and Young People Born Preterm.
- Exposure to maternal alcohol use during pregnancy, gestational diabetes, gestational hypertension, maternal obesity:
 - The ages at which these children should be screened was discussed. The final recommendation is based on the AAP Guidelines.
- Exposure to maternal cigarette smoking during pregnancy:
 - Although the panel agrees that maternal cigarette smoking during pregnancy has been associated with poor outcomes for both mothers and

their babies, the studies revealed insignificant results for screening for developmental delay, in particular.

- Exposure to maternal anemia during pregnancy:
 - The panel decided that the lone study was not enough basis to make a recommendation.

4.13.1 Burden of disease

Developmental disabilities are a group of developmental lifelong conditions due to impairment in physical, learning, language or behavior areas. Children diagnosed with developmental disabilities typically require services to address behavioral and developmental challenges. Current reports on the prevalence and trends of developmental disabilities among children in the United States show an overall increase in the prevalence of any developmental disability from 2009 to 2017, with a current estimate of 17.8% among children 3 to 17 years old.(1) In the Philippines, there are 1.4 million persons with disability (PWD), comprising 1.57% of the 92.1 million total population.² An estimated 1,014,332 children below 5 years have some form of developmental disabilities.³ One of five PWDs are school-aged children and adolescents aged 5–19 years, of whom children aged 10–14 constitute the largest proportion across different age groups (7.2%).(2)

The Global Burden of Disease Study of 2016 estimated the years lived with disabilities (YLDs) among children with developmental disorders and showed an increase from 3.8 million in 1990 to 3.9 million in 2016. In the Philippines, a similar trend was seen from 53,972 (1990) to 63,694 (2016) with 550.6 YLDs per 100,000 as of 2016. The increasing burden reflects the absence of any systematic global initiative to curtail this burden among the increasing number of children who are surviving the first five years of life. It is also likely that improving neonatal rates, particularly for those born prematurely is contributing to this burden.(3)

Early identification of children at risk of having developmental disabilities facilitates parental counseling and early intervention services to prevent delays, stimulate emerging skills, and create a more supportive and protective environment. It also prevents over servicing children who are typically developing, and facilitates targeted interventions for those with a higher risk of developmental delay.(4)

4.13.2 Prognostic Outcomes

Prematurity

The 2017 National Institute for Health and Care Excellence (NICE) Guideline on Developmental follow-up of children and young people born preterm reviewed 19 studies that looked into the association between gestational age and various developmental disorders. Results of these studies could not be pooled due to differences in the inclusion/exclusion criteria for participants, ages of participants at the time of assessment, confounders adjusted for in multivariate analysis models, outcome definitions and measurement tools, and consistency of results.(5)

Cerebral palsy

Evidence from 4 studies (1 low quality evidence, 3 moderate quality evidence) showed an increase in the risk of cerebral palsy for preterm infants.(6-9)

Intellectual disability

Low to high quality evidence from 7 studies showed an increase in the risk of intellectual disability for preterm infants. Low quality of evidence from 1 study, moderate quality of evidence from 3 studies and high quality of evidence from 1 study showed a significantly increased risk of developmental delay in children born preterm as compared to term controls. Moderate quality evidence from 1 study showed no significant increased risk of developmental delay.(8,10-15)

Speech and/or language disorder

Low to moderate quality evidence from 3 studies showed mixed results. Two studies (1 low quality evidence and 1 moderate quality evidence) showed an increase in the risk of speech and/or language delay in children born preterm as compared to term controls. Low quality evidence from 1 study showed no association between prematurity and serious language impairment.(11,16-17)

Attention deficit hyperactivity disorder

Low to moderate quality evidence from 5 studies showed mixed results. 3 studies (2 low quality evidence and 1 moderate quality evidence) showed a significant increase in the risk of attention deficit hyperactivity disorder in children born preterm. Low quality evidence from 2 studies showed no association between prematurity and attention deficit hyperactivity disorder.(13,16,18-20)

Autism spectrum disorder

Evidence from 2 studies (1 low quality evidence, 1 high quality evidence) showed a significant increase in the risk of autism spectrum disorder for preterm children.(13,21)

Specific learning disability

Moderate quality evidence from 1 study showed an increase in the risk of learning disorders in reading and mathematics among children born before 26 weeks of gestation compared to full term controls assessed at 11 years.(22)

Composite outcome

High quality evidence from 1 study showed a significant increase in the risk of neurodevelopmental disorder (developmental delay, cerebral palsy, blindness, or deafness) comparing children born at 22–26 weeks of gestation as compared to those born 27–28 weeks of gestation assessed at 2–3 years corrected age.(23-24)

Maternal smoking

Low quality evidence from 2 studies showed conflicting results. One prospective cohort showed non-significant increased risk of gross motor delay (OR 1.4, 95% CI 0.10 to 20.9).(25) The same study showed non-significant slightly decreased risk of fine motor delay (OR 0.90, 95% CI 0.20 to 3.80), non-significant decreased risk of communication delay (OR 0.70, 95% CI 0.20 to 2.90), and that problem solving delays occurred equally in those children exposed and non-exposed to maternal smoking during pregnancy (OR 1.0, 95% CI 0.10 to 4.60). On the other hand, low quality evidence from the other prospective study showed non-significant increased risk of suspected developmental delay (PR 1.29, 95% CI 0.86 to 1.92).(26)

Maternal alcohol use

Low quality evidence from 1 study showed a significantly increased risk of intellectual disability (OR 1.81, 95% CI 1.53 to 2.14) with prenatal alcohol use.(27) However, low quality evidence from a case-control study showed significantly decreased risk of autism spectrum disorder (ASD) and other developmental disabilities (DD) in children whose mother drank alcohol during the third trimester (ASD: OR 0.40, 95% CI 0.30 to 0.70; DD: OR 0.70, 95% CI 0.5 to 0.9).(28)

Maternal anemia

Low quality evidence from a prospective study showed non-significant increased risk of suspected developmental delay (PR 1.48, 95% CI 0.95 to 2.29).(26)

Gestational diabetes

Low quality evidence from 2 prospective and 1 case-control study showed non-significant increased risk of ASD (HR 1.86, 95% CI 0.92 to 3.76; HR 1.10, 95% CI 0.77 to 1.56; and OR 1.52, 95% CI 0.82 to 2.83) in children whose mothers had gestational diabetes.(29-31) Low quality evidence from 1 prospective study showed non-significant increased risk of ADHD (HR 0.99, 95% CI 0.50 to 1.94).(29) Low quality evidence from 1 retrospective study showed significantly increased risk of delay in communication skills (OR 2.17, 95% CI 1.28 to 3.66). The same study showed non-significant increased risk of fine motor (OR 1.11, 95%CI 0.63 to 1.95), gross motor (OR 1.40, 95% CI 0.83 to 2.35), and problem solving skills delay (OR 1.47, 95% CI 0.85 to 2.53), but non-significant decreased risk in personal social skills delay (OR 0.94, 95% CI 0.4 to 1.85).(32) Low quality evidence from 1 case-control study showed significantly increased risk of developmental delay (OR 2.33, 95% CI 1.08 to 5.05).(31) Low quality evidence from 1 prospective cohort study showed non-significant increased risk of suspected developmental delay (HR 1.29, 95% CI 0.86 to 1.92).(26)

Maternal obesity

A systematic review of 41 studies (meta-analysis done on 32 of the 41 studies included 6 case-control and 26 cohort studies involving a total of 36 cohorts) showed that pre-pregnancy obesity significantly increased the risk of compromised neurodevelopmental outcomes (OR 1.51, 95% CI 1.35 to 1.69), which included attention deficit hyperactivity disorder (ADHD) (OR 1.62, 95% CI 1.23 to 2.14), ASD (OR 1.36, 95% CI 1.08 to 1.70) and developmental delay OR (1.58, 95% CI 1.39 to 1.79).(33) Low quality evidence from

1 retrospective cohort study showed significantly increased risk in fine motor skills delay (OR 1.66, 95% CI 1.08 to 2.57). This study also showed non-significant slight increase in risk of delay in communication skills (OR 1.12, 95% 0.63 to 2.0), gross motor skills (OR 1.13, 95% CI 0.69 to 1.86), and problem solving skills 1.01 (0.60-1.70). There is a non-significant protective effect of maternal obesity in risk of delay in personal-social skills (OR 0.85, 95% CI 0.54 to 1.35).(32) Moderate quality evidence from 1 retrospective cohort study showed significant slight increase in risk of delay in cognitive development (RR 1.04, 95% CI 1.02 to 1.07) and physical development (RR 1.04, 95% 1.01 to 1.08). Communication skill development was not found to be at risk (RR 1.0, 95% CI 0.99 to 1.02) while there was non-significant increase in risk for delay in social-emotional (RR 1.02, 95% CI 0.98 to 1.08) and adaptive (RR 1.02, 95% CI 0.97 to 1.07) development. Risk for global developmental delay was significantly increased (RR 1.05, 95% 1.01 to 1.08).(34)

Gestational hypertension

Low quality evidence from 1 retrospective study showed significant increase in risk in delay in gross motor skills (OR 2.33, 95% CI 1.17 to 4.26). There is a non-significant increase in risk in delay in communication skills (OR 1.81, 95% CI 0.83 to 3.93) and problem solving skills (OR 1.75, 95% CI 0.85 to 3.62). Risk for fine motor (OR 0.52, 95% CI 0.16 to 1.68) and personal social (OR 0.95, 95% CI 0.34 to 2.67) skill delay were found to be decreased although non-significant.(32) Low to moderate quality evidence from 2 prospective cohort studies showed significantly increased risk for ASD (HR 1.29, 95% CI 1.24 to 1.34; OR 1.71, 95% CI 1.30 to 2.25).(30,35) On the other hand, low quality evidence from 1 case-control study showed that increased risk for ASD was not significant (OR 2.84, 95% CI 0.94 to 8.56).(31) Moderate quality evidence from 1 prospective study showed significantly increased risk of ADHD (HR 1.24, 95% CI 1.20 to 1.28).(35) Low quality evidence from 1 case-control study showed non-significant increased risk for developmental delay (OR 3.58, 95% CI 0.93 to 13.78).(31) Low quality evidence from 1 prospective cohort also showed non-significant increased risk for suspected developmental delay (OR 1.28, 95% CI 0.85 to 1.93).(26)

4.13.3 Diagnostic Performance of Screening Tests

Diagnostic accuracy of developmental screening tests

The Canadian Task Force on Preventive Health Care (CTFPHC) considered indirect evidence on the accuracy of screening tools used to assess developmental delay in children aged one to four years with no known developmental concerns. (36) Four studies investigated the diagnostic accuracy of the Ages and Stages Questionnaire (ASQ) and the Parents' Evaluation of Developmental Status (PEDS) as developmental screening tests compared to various reference tests (Table 1).

Table 1. Index tests and reference tests used in studies

Study	Indexed Test	Reference Test
Gollenberg 2009	ASQ	BSID-II
Limbos 2011	ASQ and PEDS	BSID, WPPSI, VABS, PLS
Rydz 2006	ASQ	BDI
Steenis 2015	ASQ	BSID-III

*ASQ: Ages and Stages Questionnaire; BDI: Battelle Developmental Inventory; BSID: Bayley Scales of Infant and Toddler Development; PEDS: Parents' Evaluation of Developmental Status; PLS: Preschool Language Scale; VABS: Vineland Adaptive Behavior Scales; WPPSI: Wechsler Preschool and Primary Scale of Intelligence

Table 2 shows the diagnostic accuracy of the developmental screening tests. There was low quality of evidence due to biases across studies, indirect evidence, and imprecision. ASQ was found to have moderate sensitivity but low specificity across age groups. PEDS had low sensitivity or specificity depending on the age group when the cut-off of one atypical finding was used as a positive result. When a cut-off of two abnormal findings was used as a positive result, the specificity improved, but the sensitivity dropped to below acceptable levels.(36)

Table 2. Diagnostic accuracy of developmental screening tests

Screening Test	Study Design	# of Studies and Sample Size	Results	Certainty of Evidence (GRADE)
ASQ*	Cohort	4; 1,001	Sensitivity: 55.0% (95% CI: 47.1% to 66.7%) ^[1] Specificity: 86.0% (95% CI: 38.6% to 94.3%) Positive Likelihood Ratio 4.2 (95% CI: 1.1 to 8.2) ^[1] Negative Likelihood Ratio 0.61 (95% CI: 0.47 to 0.86)	LOW
PEDS	Cohort	1; 331	Sensitivity: 41.1% (95% CI: 24.7% to 59.3%) ^[1] Specificity: 89.3% (95% CI: 85.1% to 92.5%) Positive Likelihood Ratio 3.8 (95% CI: 95% CI: 2.3 to 6.4) ^[1] Negative Likelihood Ratio 0.66 (95% CI: 0.5 to 0.88)	VERY LOW

*Results reported for ASQ are medians

Diagnostic accuracy of Ages and Stages Questionnaire (ASQ) in correctly identifying intellectual disability in preterm children

The NICE Guideline on Developmental follow-up of children and young people born preterm reviewed 5 studies assessing the diagnostic value of the Ages and Stages Questionnaire (ASQ) compared to Wechsler Preschool and Primary Scale of Intelligence (WPPSI) or Bayley Scales of Infant and Toddler Development (BSID) in identifying intellectual disability among preterm children. Results could not be pooled due to differences in the cut-off points used for index tests and reference standards, gestational age at birth of participants, and ages of participants at the time of assessment.(5)

Low quality evidence from 1 study reviewing the diagnostic performance of ASQ on intellectual disability showed that a cut-off of ASQ 2 SD below the mean had moderate sensitivity, low specificity, and not useful positive or negative likelihood ratio among preterm children (GA 32-36 weeks), and moderate sensitivity and specificity, and moderately useful positive or negative likelihood ratio among extremely preterm children (GA <32 weeks) in diagnosing developmental delay (defined as 1 SD below the mean in Bayley-III) when assessed at 8, 18, or 30 months.(37)

Among preterm children (mean GA 25.4 weeks) assessed at 18-22 months corrected age, moderate quality evidence from 1 study reviewing the diagnostic performance of ASQ on intellectual disability showed that a cut-off of ASQ 1 SD below the mean gave high sensitivity, low specificity, and not useful positive or negative likelihood ratio compared to the reference standard of BSID-II 2 SD below the mean. A cut-off of ASQ 2 SD below the mean gave a low sensitivity and specificity, and not useful positive and negative likelihood ratio compared to the reference standard of BSID II 1 or 2 SD below the mean.(38)

Among preterm children (GA <31 weeks) assessed at 18 months corrected age, low quality evidence from 1 study reviewing the diagnostic performance of ASQ on intellectual disability showed that a cut-off of ASQ 2 SD below the mean had a low sensitivity, high specificity, moderately useful positive likelihood ratio and not useful negative likelihood ratio compared to the reference standard of Bayley MDI 1 SD below the mean.(39)

Among preterm children (GA 29-36 weeks) assessed at 12 months corrected age, moderate to low quality evidence from 1 study reviewing the diagnostic performance of ASQ on intellectual disability showed that a cut-off of ASQ 1 SD below the mean had low sensitivity and specificity, and not useful positive and negative likelihood ratio compared to the BSID MDI <85 as reference standard. A cut-off of ASQ 1.5 SD and 2 SD below the mean both had moderate specificity, low sensitivity, and not useful positive or negative likelihood ratios. When BSID-II PDI <85 was used as the reference standard, the cut-offs of ASQ 1 SD, 1.5 SD, and 2 SD below the mean all had high specificity and moderately useful positive likelihood ratios, but low specificity and not useful negative likelihood ratios.(40)

Among preterm children (GA 29-36 weeks) assessed at 24 months corrected age, moderate to low quality evidence from 1 study reviewing the diagnostic performance of

ASQ on intellectual disability showed that a cut-off of ASQ 1 SD below the mean had high sensitivity, low specificity, not useful positive likelihood ratio but moderately useful negative likelihood ratio compared to BSID-II MDI <85 as reference standard. A cut-off of ASQ 1.5 SD and 2 SD below the mean both had moderate sensitivity, moderate or close to moderate specificity, and not useful positive likelihood ratio. A cut-off of ASQ 1.5 SD below the mean had moderately useful negative likelihood ratio compared to BSID-II MDI <85. When BSID-II PDI <85 was used as the reference standard, the cut-offs of ASQ 1 SD and 1.5 SD below the mean had low sensitivity and specificity, and not useful positive and negative likelihood ratio. A cut-off of ASQ 2 SD below the mean had high specificity but low sensitivity, and not useful positive or negative likelihood ratio.(40)

Moderate quality evidence from 1 study showed that ASQ score <270 had moderate sensitivity and specificity, not useful positive likelihood ratio, and moderately useful negative likelihood ratio compared to an IQ score <70 on WPPSI-III among preterm children (GA ≤35 weeks) assessed at 5 years of age. An ASQ score <280 had moderate sensitivity, low specificity, and not useful positive and negative likelihood ratio compared to an IQ score <85 on WPPSI-III.(41)

4.13.4 Cost Implication

The estimated annual cost of developmental screening in the Philippines is PhP 450.308 per patient, as detailed in Table 3.

Table 3. Estimated annual cost of developmental screening

Procedure	Cost (PHP)
Developmental screening	240.75
Training in the use of the developmental screening tool	146.00
Developmental screening complete set (PEDS – Filipino)	63.558
TOTAL	450.308

Philhealth has a Z Benefit Package of Children with Developmental Disability, which covers for initial assessment and rehabilitation therapy (Table 4).(42)

Table 4. Estimated cost of assessment and treatment from Philhealth package

Description	Amount	
Assessment by a medical specialist to three allied health professional or rehabilitation therapist	P3,626 -P5,276	1 per cycle year for a maximum of 3 cycles
Rehabilitation therapy	P5,000 per tranche	Maximum of 9 tranches per year

In private institutions and therapy centers, estimated cost for assessment and treatment ranges from Php 3,450.00 to 16,071.85, as shown in Table 5.

Table 5. Estimated cost of assessment and treatment in private institutions

Procedure	Cost Per Session (PHP)
Developmental Assessment	1,500–8,571.85
Physical Therapy	600–2,500
Occupational Therapy	600–2,500
Speech Therapy	750–2,500
TOTAL	3,450–16,071.85

4.13.5 Equity, Acceptability, and Feasibility

Stressors in early childhood can disrupt neurologic, metabolic, and immunologic systems, leading to poor developmental outcomes. These stressors can result from vulnerabilities related to socio-economic disparities. Although health, social care, and education programs that serve young children and their families and communities offer opportunities to foster sensitive relationships and environments, these systems also fragment their efforts because of limitations that limit the age ranges they can serve and the types of services they can provide. Exposure to hardship and stressors such as poverty, lack of protection and stability in the home setting, and lack of access to quality early education can have a negative effect on the development of young children.(43)

Developmental delay screening tools are also not inclusive, primarily because of the common use of the English language. Improved translation and interpretation resources may decrease disparities but there are limited available Filipino translations of assessment and screening tools used by rehabilitation professions.(44)

Implementing a universal screening intervention program in the Philippines will need to have a thorough understanding of the rehabilitation process and the issues surrounding it. Universal screening for developmental delay will necessitate a primary care approach, primarily community-based. This entails training barangay health workers to become community-based rehabilitation workers who can screen community members, regardless of age, who may have developmental delays. This offers a more widespread and cost effective screening process.(45,46)

4.13.6 Recommendations from Other Groups

The Canadian Task Force on Preventive Health Care recommends against screening for developmental delay using standardized tools in children aged 1 to 4 years with no apparent signs of developmental delay and whose parents and clinicians have no concerns about development. It does not apply to children who present with signs suggestive of possible developmental delay, those whose parents express concern that could indicate developmental delay or those whose development is being closely monitored because of identified risk factors, such as premature birth or low birth weight.(36)

The US Preventive Services Task Force addressed screening for speech and language delays, and autism spectrum disorder, but it did not address general developmental screening. It found insufficient evidence to recommend screening for speech and language delay and disorders in children aged 5 years or younger whose parents or clinicians do not have specific concerns about their speech, language, hearing, or development. There was inadequate evidence on the accuracy of screening instruments for speech and language delay for use in primary care settings, and on the accuracy of surveillance (active monitoring) by primary care clinicians to identify children for further evaluation for speech and language delays and disorders. Risk factors that have been reported to be associated with speech and language delay and disorders include male sex, family history of speech and language impairment, low parental education level, and perinatal risk factors (e.g., prematurity, low birth weight, and birth difficulties). (47)

The American Academy of Pediatrics recommends that all children should receive periodic developmental screening using^[11] a standardized test. In the absence of established risk factors or parental or provider concerns, a general developmental screen continues to be recommended at the 9-, 18-, and 30-month visits. (48)

References

1. Zablotsky B, Black L, Maenner M, Schieve L, Danielson M, Bitsko R, et al. Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. *Pediatrics*. 2019;144(4).
2. National Statistics Office: 2010 Census of Population and Housing.
3. Olusanya B, Global Research on Developmental Disabilities Collaborators. Developmental disabilities among children younger than 5 years in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*. 2018;6:e1100-21.
4. Faruk T, King C, Muhit M, Islam M, Jahan I, Baset K et al. Screening tools for early identification of children with developmental delay in low- and middle-income countries: a systematic review. *BMJ Open*. 2020;10(11):e038182.
5. National Institute for Health and Care Excellence. Developmental follow-up of children and young people born preterm [Internet]. [London]: NICE; 2017 [updated 2017 Aug; cited 2021 Sep 2]. (Clinical guideline [NG72]). Available from: <https://www.nice.org.uk/guidance/ng72>
6. Odd DE, Lingam R, Emond A, Whitelaw A. Movement outcomes of infants born moderate and late preterm. *Acta Paediatr*. 2013;102(9):876-882.
7. Hirvonen M, Ojala R, Korhonen P, et al. Cerebral palsy among children born moderately and late preterm. *Pediatrics*. 2014;134(6):e1584-e1593.
8. Petri尼 JR, Dias T, McCormick MC, Massolo ML, Green NS, Escobar GJ. Increased risk of adverse neurological development for late preterm infants. *J Pediatr*. 2009;154(2):169-176.
9. Sukhov A, Wu Y, Xing G, Smith LH, Gilbert WM. Risk factors associated with cerebral palsy in preterm infants. *J Matern Fetal Neonatal Med*. 2012;25(1):53-57.
10. Woythaler MA, McCormick MC, Smith VC. Late preterm infants have worse 24-month neurodevelopmental outcomes than term infants. *Pediatrics*. 2011;127(3):e622-e629.
11. Serenius F, Källén K, Blennow M, et al. Neurodevelopmental outcome in extremely preterm infants at 2.5 years after active perinatal care in Sweden. *JAMA*. 2013;309(17):1810-1820.
12. Larroque B, Ancel PY, Marchand-Martin L, et al. Special care and school difficulties in 8-year-old very preterm children: the Epipage cohort study. *PLoS One*. 2011;6(7):e21361.
13. Singh GK, Kenney MK, Ghandour RM, Kogan MD, Lu MC. Mental Health Outcomes in US Children and Adolescents Born Prematurely or with Low Birthweight. *Depress Res Treat*. 2013;2013:570743.
14. Helderman JB, O'Shea TM, Kuban KC, et al. Antenatal antecedents of cognitive impairment at 24 months in extremely low gestational age newborns. *Pediatrics*. 2012;129(3):494-502.

15. Hillemeier MM, Morgan PL, Farkas G, Maczuga SA. Perinatal and socioeconomic risk factors for variable and persistent cognitive delay at 24 and 48 months of age in a national sample. *Matern Child Health J.* 2011;15(7):1001-1010.
16. Rabie NZ, Bird TM, Magann EF, Hall RW, McKelvey SS. ADHD and developmental speech/language disorders in late preterm, early term and term infants. *J Perinatol.* 2015;35(8):660-664.
17. Wolke D, Samara M, Bracewell M, Marlow N; EPICure Study Group. Specific language difficulties and school achievement in children born at 25 weeks of gestation or less. *J Pediatr.* 2008;152(2):256-262.
18. Burnett A, Davey CG, Wood SJ, et al. Extremely preterm birth and adolescent mental health in a geographical cohort born in the 1990s. *Psychol Med.* 2014;44(7):1533-1544.
19. Rogers CE, Lenze SN, Luby JL. Late preterm birth, maternal depression, and risk of preschool psychiatric disorders. *J Am Acad Child Adolesc Psychiatry.* 2013;52(3):309-318.
20. Johnson S, Hollis C, Kochhar P, Hennessy E, Wolke D, Marlow N. Psychiatric disorders in extremely preterm children: longitudinal finding at age 11 years in the EPICure study. *J Am Acad Child Adolesc Psychiatry.* 2010;49(5):453-63.e1.
21. Kuzniewicz MW, Wi S, Qian Y, Walsh EM, Armstrong MA, Croen LA. Prevalence and neonatal factors associated with autism spectrum disorders in preterm infants. *J Pediatr.* 2014;164(1):20-25.
22. Johnson S, Wolke D, Hennessy E, Marlow N. Educational outcomes in extremely preterm children: neuropsychological correlates and predictors of attainment. *Dev Neuropsychol.* 2011;36(1):74-95.
23. Kent AL, Wright IM, Abdel-Latif ME; New South Wales and Australian Capital Territory Neonatal Intensive Care Units Audit Group. Mortality and adverse neurologic outcomes are greater in preterm male infants. *Pediatrics.* 2012;129(1):124-131.
24. Toome L, Varendi H, Männamaa M, Vals MA, Tänavsuu T, Kolk A. Follow-up study of 2-year-olds born at very low gestational age in Estonia. *Acta Paediatr.* 2013;102(3):300-307.
25. Moore BF, Shapiro AL, Wilkening G, Magzamen S, Starling AP, Allshouse WB, Adgate JL, Dabelea D. Prenatal Exposure to Tobacco and Offspring Neurocognitive Development in the Healthy Start Study. *J Pediatr.* 2020 Mar;218:28-34.e2.
26. de Moura DR, Costa JC, Santos IS, et al. Risk factors for suspected developmental delay at age 2 years in a Brazilian birth cohort. *Paediatr Perinat Epidemiol.* 2010;24(3):211-221.
27. O'Leary C, Leonard H, Bourke J, D'Antoine H, Bartu A, Bower C. Intellectual disability: population-based estimates of the proportion attributable to maternal alcohol use disorder during pregnancy. *Dev Med Child Neurol.* 2013 Mar;55(3):271-7.
28. Singer AB, Aylsworth AS, Cordero C, Croen LA, DiGuiseppi C, Fallin MD, Herring AH, Hooper SR, Pretzel RE, Schieve LA, Windham GC, Daniels JL. Prenatal Alcohol Exposure in Relation to Autism Spectrum Disorder: Findings from the Study to Explore Early Development (SEED). *Paediatr Perinat Epidemiol.* 2017 Nov;31(6):573-582.
29. Li M, Fallin MD, Riley A, Landa R, Walker SO, Silverstein M, Caruso D, Pearson C, Kiang S, Dahm JL, Hong X, Wang G, Wang MC, Zuckerman B, Wang X. The Association of Maternal Obesity and Diabetes With Autism and Other Developmental Disabilities. *Pediatrics.* 2016 Feb;137(2):e20152206.
30. Cordero C, Windham GC, Schieve LA, Fallin MD, Croen LA, Siega-Riz AM, Engel SM, Herring AH, Stuebe AM, Vladutiu CJ, Daniels JL. Maternal diabetes and hypertensive disorders in association with autism spectrum disorder. *Autism Res.* 2019 Jun;12(6):967-975.
31. Krakowiak P, Walker CK, Bremer AA, Baker AS, Ozonoff S, Hansen RL, Hertz-Pannier I. Maternal metabolic conditions and risk for autism and other neurodevelopmental disorders. *Pediatrics.* 2012 May;129(5):e1121-8.
32. Girchenko P, Tuovinen S, Lahti-Pulkkinen M, Lahti J, Savolainen K, Heinonen K, Pyhälä R, Reynolds RM, Hämäläinen E, Villa PM, Kajantie E, Pesonen AK, Laivuori H, Räikkönen K. Maternal early pregnancy obesity and related pregnancy and pre-pregnancy disorders: associations with child developmental milestones in the prospective PREDO Study. *Int J Obes (Lond).* 2018 Jun;42(5):995-1007.
33. Sanchez CE, Barry C, Sabhlok A, Russell K, Majors A, Kollins SH, Fuemmeler BF. Maternal pre-pregnancy obesity and child neurodevelopmental outcomes: a meta-analysis. *Obes Rev.* 2018

- Apr;19(4):464-484. doi: 10.1111/obr.12643. Epub 2017 Nov 22. PMID: 29164765; PMCID: PMC6059608.
- 34. Duffany KO, McVeigh KH, Kershaw TS, Lipkind HS, Ickovics JR. Maternal Obesity: Risks for Developmental Delays in Early Childhood. *Matern Child Health J*. 2016 Feb;20(2):219-30. doi: 10.1007/s10995-015-1821-z. PMID: 26694046.
 - 35. Wang H, László KD, Gissler M, Li F, Zhang J, Yu Y, Li J. Maternal hypertensive disorders and neurodevelopmental disorders in offspring: a population-based cohort in two Nordic countries. *Eur J Epidemiol*. 2021 May;36(5):519-530.
 - 36. Screening and treatment for developmental delay in early childhood (ages 1-4): systematic review [Internet]. Canadian Task Force on Preventive Health Care. 2015.
 - 37. Schonhaut L, Armijo I, Schönstedt M, Alvarez J, Cordero M. Validity of the ages and stages questionnaires in term and preterm infants. *Pediatrics*. 2013;131(5):e1468-e1474.
 - 38. Woodward BJ, Papile LA, Lowe JR, et al. Use of the Ages and Stages Questionnaire and Bayley Scales of Infant Development-II in neurodevelopmental follow-up of extremely low birth weight infants. *J Perinatol*. 2011;31(10):641-646.
 - 39. Skellern CY, Rogers Y, O'Callaghan MJ. A parent-completed developmental questionnaire: follow up of ex-premature infants. *J Paediatr Child Health*. 2001;37(2):125-129.
 - 40. Simard MN, Luu TM, Gosselin J. Concurrent validity of ages and stages questionnaires in preterm infants. *Pediatrics*. 2012;130(1):e108-e114.
 - 41. Halbwachs M, Muller JB, Nguyen The Tich S, et al. Usefulness of parent-completed ASQ for neurodevelopmental screening of preterm children at five years of age. *PLoS One*. 2013;8(8):e71925.
 - 42. Philhealth Circular No. 2017-0029. [cited 9 September 2021]. Available from: https://www.philhealth.gov.ph/circulars/2017/TS_circ2017-0029.pdf
 - 43. Black M, Walker S, Fernald L, Andersen C, DiGirolamo A, Lu C. Early childhood development coming of age: science through the life course. *The Lancet*. 2017;389(10064):77-90.
 - 44. Rodrigues K, Hambidge S, Dickinson M, Richardson D, Davidson A. Developmental screening disparities for languages other than English and Spanish. *Academic Pediatrics*. 2016;16(7):653-9.
 - 45. Periquet A. Community-based rehabilitation in the Philippines. *International disability studies*. 1989;11(2):95-6.
 - 46. Duque R, Leon Cd. Issues and perspectives: Occupational Therapy in the Philippines: Where have we been and where are we headed. *Philippine Journal of Occupational Therapy*. 2006;2:57-67.
 - 47. Siu AL; US Preventive Services Task Force. Screening for Speech and Language Delay and Disorders in Children Aged 5 Years or Younger: US Preventive Services Task Force Recommendation Statement. *Pediatrics*. 2015;136(2):e474-e481.
 - 48. Lipkin PH, Macias MM; Council on Children with Disabilities, Section on Developmental and Behavioral Pediatrics. Promoting Optimal Development: Identifying Infants and Young Children With Developmental Disorders Through Developmental Surveillance and Screening. *Pediatrics*. 2020;145(1):e20193449.

4.14 Screening for Autism Spectrum Disorder

RECOMMENDATIONS

Among asymptomatic, apparently healthy children, we recommend for the screening of autism spectrum disorder between the ages of 18 to 24 months old using the M-CHAT R/F. (**Strong recommendation, moderate certainty of evidence.**)

Considerations

The consensus panel considered the following when formulating this recommendation:

- Although the M-CHAT R/F has not been validated in Filipino as of press time, the very high burden of disease along with the benefit of early intervention for those who screen positive for ASD led the panel to vote for this recommendation.
- The panel agreed that the M-CHAT R/F needs to be validated in Filipino and possibly other dialects especially since this is a parent-reported tool.
- The panel discussed the harm of over-labeling patients and emphasized the importance of proper training for the healthcare workers who will be administering, interpreting and explaining these results to the parents of the patients.

4.14.1 Burden of disease

The CDC estimates that 1 in 54 children is diagnosed with autism spectrum disorder (ASD). (1) There is a high variability of ASD prevalence between countries, with estimates ranging from 0.8/1000 in Bangladesh, to 93/1000 in Japan. However, there has been a marked increase in ASD prevalence worldwide over the last 10 years.(2) ASD is a neurodevelopmental disorder marked by impairments in social interaction and communication, and/or restricted, repetitive patterns of behavior.(3) While the mean age at diagnosis of ASD is 5 years old,(4) symptoms can be observed as early as infancy.(5-6) The prognosis for ASD is variable, ranging from normal intelligence and social communication skills, albeit with residual ASD manifestations, to difficulty with integration in society and work. Treatment for ASD is highly individualized and tailored to the strengths and weaknesses of the patient. In young children diagnosed early, early intensive behavioral and developmental interventions are most commonly recommended.(7)

4.14.2 Benefits and Harms of Screening Tests

There were no direct studies found on universal ASD screening versus no screening among apparently healthy children on adaptive functioning or impact on familial psychosocial dynamics. Instead, we looked into studies on the effectiveness of early intervention among children diagnosed with ASD.

Effect of early intervention on adaptive functioning

A Cochrane review published in 2018, composed of 1 RCT (9) and 4 CCTs,(10-13) evaluated the effects of early intensive behavioral intervention (EIBI) on children less than 6 years old. We attempted to update this review using the same search strategies,

however no new studies which fit the original inclusion criteria were found. See Appendix A for the search strategy and Appendix B, figure 1 for the PRISMA diagram.

All 5 studies included children less than 6 years old with a diagnosis of ASD. This Cochrane review found that treatment with EIBI resulted in a net benefit, with a mean adaptive behavior score which was 9.58 points higher (95% CI 5.57 to 13.60, P< 0.001) on the Vineland Adaptive Behaviors Scale (VAB-S), compared to children who received treatment as usual. Evidence was downgraded by two levels as 4 out of 5 included studies were controlled clinical trials. Publication bias could not be ascertained due to the low number of studies. No adverse events were identified among the studies. The Cochrane review did not investigate the impact of an ASD diagnosis on family psychosocial dynamics. No data was available for the subgroups identified in the guideline question. Table 1 shows the summary evidence regarding the benefits of EIBI on children less than 6 years old with a diagnosis of ASD. See Appendix D for the GRADE Evidence Profile and Appendix E, figure 1 for the forest plot of the pooled effect of EIBI on adaptive behavior.

Table 1. Effects of EIBI on children less than 6 years old with a diagnosis of ASD

Outcomes	No. of Studies (no. of participants)	Effect Estimate (95% CI)	Interpretation	Certainty of Evidence
Adaptive behavior (measured by composite score on VAB-S; higher score equated with better outcomes)	5 (202)	9.58 points higher than TAU group (5.57 – 13.6)	Benefit	⊕⊕○○ Low
Adverse events	No adverse events were reported in any study			
Impact of diagnosis on family psychosocial dynamics	Outcome was not investigated in any study			

4.14.3 Diagnostic Performance of Screening Tests

A systematic review by the United States Preventive Services Task Force (USPSTF) in 2016 identified 18 unique studies which assessed various ASD screening tools in asymptomatic children aged 12-36 months.(14) ASD screening tools included the CHAT, M-CHAT-F, M-CHAT R/F, CSBS-ITC, ITC, FYI, ESAT, SACS, and YACHT-18.(15-18,23-31) We updated this systematic review using the same search strategy, and identified 8

additional studies on ASD screening, comprising the PDQ and SCQ, M-CHAT-F, PDDST, and the M-CHAT R/F. See Appendix 1 for the search strategy, Appendix B, figure 2 for the PRISMA diagram, and Appendix 3 for the characteristics of included studies.

Subgroup data was available for 1 study on the M-CHAT-F and the study on PDQ. (32-33) No adverse effects were identified in any of the studies. No studies on the impact of diagnosis on family psychodynamics were identified.

The M-CHAT-F is a 23-item checklist accomplished by parents, based on observation of their child's current skills and behaviors. Based on 6 cross-sectional (cohort type accuracy) studies, which used the DSM-IV and various additional behavior tests (VABS, CARS, ADOS, MSEL) as the reference standard, the pooled sensitivity of the M-CHAT-F was 0.83 (95% CI: 0.55 to 0.95) and the pooled specificity was 0.95 (95% CI: 0.84 to 0.99). (19,21,23,34-35) All 6 studies involved apparently healthy children aged 4-41 months old. No adverse effects nor harms were reported during the screenings. No subgroups by age, sex, race, nor ethnicity were reported. However, the quality of evidence for both sensitivity and specificity was low due to common issues in risk of bias (additional behavioral tests to confirm the diagnosis of ASD differed between and within studies) and inconsistency across the studies. See Appendix D, Table 1 for the GRADE Evidence profile and Appendix E, figure 2 for the forest plots on sensitivity and specificity.

In 2014, the M-CHAT-F was revised into the M-CHAT R/F by dropping 3 items, reorganizing the remaining items, simplifying language, and adding examples. These changes were made in order to decrease false positives while maintaining a high sensitivity. Based on a 2014 study by Robin and others, the sensitivity of the M-CHAT R/F is estimated at 0.85 (95% CI 0.79 – 0.92), while the specificity is estimated at 0.99 (95% CI 0.99 – 0.99).(24) Quality of evidence is moderate due to risks of bias specifically in patient selection and patient flow and timing. See Appendix D, Table 2 for the GRADE Evidence profile.

The PDQ is a 10 item parent-report questionnaire on social referencing and communication in young children. Based on a 2015 study by Zahorodny and others, the PDQ has an estimated sensitivity and specificity of 84.62% and 99.84%, respectively.(32) Quality of evidence is low. There was very serious imprecision in the study, as it did not state 95% confidence intervals for neither sensitivity nor specificity. See Appendix D, Table 3 for the GRADE Evidence profile.

4.14.4 Cost Implication

A Canadian study done in 2018 evaluated the cost effectiveness of universal screening, high-risk screening, and surveillance monitoring for ASD.(40) They concluded that universal screening at 18 and 24 months was not a cost-effective approach and did not ensure that children with ASD would receive appropriate interventions earlier. With universal screening, every additional child diagnosed at 3 years old would cost an additional 757,116.9 CAD (Php 300,000,000). In contrast, high-risk screening would cost an additional 41,651 CAD (Php 1,700,000) for every child diagnosed at age 3. Moreover,

they suggested that shortening waiting times to official diagnosis would be more effective in ensuring earlier intervention.

One local evaluation of the economic impact of ASD diagnosis and treatment found that there was a significant financial burden on at least 55% of the families who participated in the study. With limited PhilHealth coverage for children with ASD, most families paid for ASD-related services out-of-pocket. Overall costs per family, including diagnosis, follow-up, and interventions in the first year following diagnosis ranged from free of charge to P506,278, depending on where the family sought consult. The authors observed that there was a declining follow-up rate for developmental monitoring, which they attributed partly to mounting costs from ASD follow-up and treatment.(41)

Table 3. Costs of screening, confirmatory tests, and treatment for ASD for Filipino families

Procedure	Cost (PHP)
ASD screening test (e.g. M-CHAT R/F administration)	0 - 1,235
ASD confirmatory tests (ADI-R, ADOS)	13,978.30
Professional fees (general pediatrician, developmental pediatrician)	0 - 10,000 (per consult)
Applied Behavior Analysis (e.g. EIBI)	600 - 150,000 (per year)

4.14.5 Equity, Acceptability, and Feasibility

A Philippine study done in 2015 explored parental perceptions of autism and their health-seeking behaviors.(42) Among parents, there was an apparent decreased stigma towards autism. They understood that both genes and environment played a role in the development of autism, and rejected myths which attributed autism to a curse or parental sin. Reasons for delayed consultation and diagnosis included lack of finances, and lack of knowledge on whom to first consult, and prolonged waiting period prior to the scheduled appointment. Importantly, the authors observed a mean age of initial symptom recognition at 24.42 months and shortened time to diagnosis (mean = 1.83 months) following initial consult and screening. However, they stressed that earlier diagnosis did not necessarily mean earlier interventions.

4.14.6 Recommendations from Other Groups

In 2001, 2007, and again in 2019, The American Academy of Pediatrics recommended universal screening for ASD in children at 18 months and 24 months, using screening tools such as the M-CHAT R/F, SCQ, or STAT.(43-45) This recommendation was based on the availability of numerous validated screening tools, various early intervention strategies, and the consideration that lack of parental concern did not necessarily mean their children were asymptomatic. The Society for Developmental and Behavioral Pediatrics endorses universal screening at 18 months and 24 months as well.(46)

However, in 2016, US Preventive Services Task Force found insufficient evidence on benefits and harms of universal ASD screening. Hence, they were unable to make a recommendation for or against this.(47)

References

1. Maenner MJ, Shaw KA, Baio J, EdS1, Washington A, Patrick M, et al. Prevalence of Autism spectrum disorder among children aged 8 years - Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2016. MMWR Surveill Summ. 2020;69(4):1–12.
2. Chiarotti F, Venerosi A. Epidemiology of autism spectrum disorders: A review of worldwide prevalence estimates since 2014. Brain Sci. 2020;10(5):274
3. Diagnostic and statistical manual of mental disorders (dsm-5(r)). 5th ed. American Psychiatric Publishing; 2013.
4. van 't Hof M, Tisseur C, van Berckelaer-Onnes I, van Nieuwenhuyzen A, Daniels AM, Deen M, et al. Age at autism spectrum disorder diagnosis: A systematic review and meta-analysis from 2012 to 2019. Autism. 2021;25(4):862–73.
5. Zwaigenbaum L, Bryson S, Rogers T, Roberts W, Brian J, Szatmari P. Behavioral manifestations of autism in the first year of life. Int J Dev Neurosci. 2005;23(2–3):143–52.
6. Werner E, Dawson G, Osterling J, Dinno N. Brief report: Recognition of autism spectrum disorder before one year of age: A retrospective study based on home videotapes. J Autism Dev Disord. 2000;30(2):157–62.
7. Coury D. Medical treatment of autism spectrum disorders. Curr Opin Neurol. 2010;23(2):131–6.
8. Reichow, B., Hume, K., Barton, E. E., & Boyd, B. A. (2018). Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). Cochrane Database of Systematic Reviews. doi:10.1002/14651858.CD009260.pub3 (<https://doi.org/10.1002/14651858.CD009260.pub3>)
9. Smith T, Groen AD, Wynn JW. Randomized trial of intensive early intervention for children with pervasive developmental disorder. American Journal of Mental Retardation 2000;105(4):269-85. [DOI: 10.1352/0895-8017(2000)105%3C0269:RTOIEI%3E2.0.CO;2; PUBMED: 10934569]
10. Cohen H, Amerine-Dickens M, Smith T. Early intensive behavioral treatment: replication of the UCLA model in a community setting. Journal of Developmental and Behavioral Pediatrics 2006;27(2 Suppl):S145-55.
11. Howard JS, Stanislaw H, Green G, Sparkman CR, Cohen HG. Comparison of behavior analytic and eclectic early interventions for young children with autism after three years. Research in Developmental Disabilities 2014;35(45):3326-44.
12. Magiati I, Charman T, Howlin P. A two-year prospective follow-up study of community-based early intensive behavioural intervention and specialist nursery provision for children with autism spectrum disorders. Journal of Child Psychology and Psychiatry 2007;48(8):803-12.
13. Remington B, Hastings RP, Kovshoff H, degli Epinosa F, Jahr E, Brown T, et al. Early intensive behavioral intervention: outcomes for children with autism and their parents after two years. American Journal of Mental Retardation 2007;112(6):418-38.
14. McPheeters ML, Weitlauf A, Vehorn A, et al. Screening for Autism Spectrum Disorder in Young Children: A Systematic Evidence Review for the U.S. Preventive Services Task Force [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2016 Feb. (Evidence Syntheses, No. 129.) References.
15. Baird G, Charman T, Baron-Cohen S, et al. A screening instrument for autism at 18 months of age: a 6-year followup study. J Am Acad Child Adolesc Psychiatry. 2000 Jun;39(6):694–702
16. VanDenHeuvel A, Fitzgerald M, Greiner B, et al. Screening for autistic spectrum disorder at the 18-month developmental assessment: a population-based study. Ir Med J. 2007 Sep;100(8):565–7.
17. Chlebowski C, Robins DL, Barton ML, et al. Large-scale use of the modified checklist for autism in low-risk toddlers. Pediatrics. 2013 Apr;131(4):e1121–7.
18. Miller JS, Gabrielsen T, Villalobos M, et al. The each child study: systematic screening for autism spectrum disorders in a pediatric setting. Pediatrics. 2011 May;127(5):866–71.
19. Kamio Y, Inada N, Koyama T. A nationwide survey on quality of life and associated factors of adults with high-functioning autism spectrum disorders. Autism. 2013;17(1):15–26.

20. Nygren G, Sandberg E, Gillstedt F, et al. A new screening programme for autism in a general population of Swedish toddlers. *Res Dev Disabil.* 2012 Jul-Aug;33(4):1200–10.
21. Canal-Bedia R, Garcia-Primo P, Martin-Cilleros MV, et al. Modified checklist for autism in toddlers: cross-cultural adaptation and validation in Spain. *J Autism Dev Disord.* 2011 Oct;41(10):1342–51
22. Inada N, Koyama T, Inokuchi E, et al. Reliability and validity of the Japanese version of the Modified Checklist for autism in toddlers (M-CHAT). *Research in Autism Spectrum Disorders.* 2011;5(1):330–6.
23. Robins DL, Fein D, Barton ML, et al. The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. *J Autism Dev Disord.* 2001 Apr;31(2):131–44
24. Robins DL, Casagrande K, Barton M, Chen C-MA, Dumont-Mathieu T, Fein D. Validation of the modified checklist for Autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics.* 2014;133(1):37–45.
25. Pierce K, Carter C, Weinfeld M, et al. Detecting, studying, and treating autism early: the one-year well-baby check-up approach. *J Pediatr.* 2011 Sep;159(3):458–65 e1-6.
26. Wetherby AM, Brosnan-Maddox S, Peace V, et al. Validation of the infant-toddler checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism.* 2008;12(5):487–511.
27. Ben-Sasson A, Habib S, Tirosh E. Feasibility and Validity of Early Screening for Identifying Infants With Poor Social-Communication Development in a Well-Baby Clinic System. *J Pediatr Nurs.* 2013 Nov 25
28. Turner-Brown LM, Baranek GT, Reznick JS, et al. The First Year Inventory: A longitudinal followup of 12-month-old to 3-year-old children. *Autism.* 2013;17(5):527–40.
29. Dietz C, Swinkels S, van Daalen E, et al. Screening for autistic spectrum disorder in children aged 14–15 months. II: population screening with the Early Screening of Autistic Traits Questionnaire (ESAT). Design and general findings. *J Autism Dev Disord.* 2006 Aug;36(6):713–22.
30. Barbaro J, Dissanayake C. Prospective identification of autism spectrum disorders in infancy and toddlerhood using developmental surveillance: the social attention and communication study. *J Dev Behav Pediatr.* 2010 Jun;31(5):376–85.
31. Honda H, Shimizu Y, Nitto Y, et al. Extraction and Refinement Strategy for detection of autism in 18-month-olds: a guarantee of higher sensitivity and specificity in the process of mass screening. *J Child Psychol Psychiatry.* 2009 Aug;50(8):972–81.
32. Zahorodny W, Shenouda J, Mehta U, Yee E, Garcia P, Rajan M, et al. Preliminary evaluation of a brief autism screener for young children. *J Dev Behav Pediatr.* 2018;39(3):183–91.
33. Pandey J, Verbalis A, Robins DL, Boorstein H, Klin AMI, Babitz T, et al. Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. *Autism.* 2008;12(5):513–35.
34. Sturner R, Howard B, Bergmann P, et al. Autism Screening With Online Decision Support by Primary Care Pediatricians Aided by M-CHAT/F. *Pediatrics.* 2016;138(3):e20153036
35. Baduel S, Guillou Q, Afzali MH, Foudon N, Kruck J, Rogé B. The French version of the modified-checklist for autism in toddlers (M-CHAT): A validation study on a French sample of 24 month-old children. *J Autism Dev Disord.* 2017;47(2):297–304.
36. Janvier YM, Harris JF, Coffield CN, Louis B, Xie M, Cidav Z, et al. Screening for autism spectrum disorder in underserved communities: Early childcare providers as reporters. *Autism.* 2016;20(3):364–73.
37. Samango-Sprouse CA, Stapleton EJ, Aliabadi F, Graw R, Vickers R, Haskell K, et al. Identification of infants at risk for autism spectrum disorder and developmental language delay prior to 12 months. *Autism.* 2015;19(3):327–37.
38. Khawaja MK, Hazzard AP, Robins DL. Sociodemographic barriers to early detection of autism: Screening and evaluation using the M-CHAT, M-CHAT-R, and Follow-Up. *J Autism Dev Disord.* 2015;45(6):1797–808.
39. Hardy S, Haisley L, Manning C, Fein D. Can screening with the Ages and Stages Questionnaire detect autism? *J Dev Behav Pediatr.* 2015;36(7):536–43.
40. Yuen T, Carter MT, Szatmari P, Ungar WJ. Cost-effectiveness of universal or high-risk screening compared to surveillance monitoring in autism spectrum disorder. *J Autism Dev Disord.* 2018;48(9):2968–79.

41. Quilendrino MIO, Castor MAR, Mendoza NRN, Vea JR, Castillo-Carandang NT. The Cost of Autism: An Analysis of The Economic Impact of Autism on the Filipino Family. Unpublished.
42. Quilendrino MIO, Castor MAR, Mendoza NRNP, Vea JR, Castillo-Carandang NT. Parents' perceptions of autism and their health-seeking behaviors. *Clin Epidemiol Glob Health*. 2015;3:S10–5.
43. Committee on Children With Disabilities. Technical report: the pediatrician's role in the diagnosis and management of autistic spectrum disorder in children. *Pediatrics*. 2001;107(5):E85.
44. Johnson CP, Myers SM, American Academy of Pediatrics Council on Children With Disabilities. Identification and evaluation of children with autism spectrum disorders. *Pediatrics*. 2007;120(5):1183–215.
45. Lipkin PH, Macias MM, COUNCIL ON CHILDREN WITH DISABILITIES, SECTION ON DEVELOPMENTAL AND BEHAVIORAL PEDIATRICS. Promoting optimal development: Identifying infants and young children with developmental disorders through developmental surveillance and screening. *Pediatrics*. 2020;145(1):e20193449.
46. Stancin T. SDBP response to USPSTF draft recommendation on screening for ASD in young children [Internet]. Available from: http://sdgp-video.s3.amazonaws.com/wp-content/uploads/2020/11/20130751/USPSTF_response.pdf
47. Siu AL, US Preventive Services Task Force (USPSTF), Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, et al. Screening for autism spectrum disorder in young children: US Preventive Services Task Force recommendation statement: US preventive services task force recommendation statement. *JAMA*. 2016;315(7):691–6.

4.15 Screening for Learning Disorders

Among asymptomatic, apparently healthy children, there is insufficient evidence to recommend for or against the screening of specific learning disorders (reading disability) in the primary health care setting. (*Very low certainty of evidence*).

Considerations

The consensus panel considered the following when formulating this recommendation:

- Based on the studies presented in this review, there is no unified standard tool identified to screen for learning disorders, specifically reading disabilities.
- The panel also considered whether it is more appropriate for the primary educators of these children rather than physicians to screen them for the reading disabilities.

4.15.1 Burden of disease

Specific learning disability (SLD) is a neurodevelopmental disorder that impedes the ability to learn or use specific academic skills (e.g., reading, writing and arithmetic), which is the foundation of other academic learning.(1) Among all learning disorders, the most common is reading disorder also named as dyslexia.(2) Epidemiological studies report prevalence rates of 4.9% to 7.5% for reading disability. (3-6)

In the Philippines, 6.4% of our 10-14 years old and 2.8% of 15-19 years old cannot read and write. The basic literacy rate, or the ability of a person to read and write with understanding, among our 10 years old and above population is relatively high at 96.5% and is higher among our 15-19 years old group with 98.3%.(7) This is also evident in a study done by UNICEF, where they found that a typical Filipino Grade 5 student is able to read a range of everyday texts fluently and begin to engage their meaning.(8) However, comparing it with other ASEAN countries, this reading literacy is 12 points lower than the average score of all participating countries and three bands lower than the highest proficiency band.(8)

In a larger study including 15-year-old students from 79 countries, Filipino students ranked very low and had an overall reading literacy of 340, which is 147 points lower than the average score of the cohort. Around 80% of Filipino students had proficiency level below level 2, meaning majority of our students cannot identify the main idea in moderate length texts and have difficulty relating the text to their outside knowledge.(9)

If a learning disability goes undetected, the child's poor scholastic performance will bring in adverse impact on health related quality of life by causing poor self-esteem, disturbed peer and family relationships, and unease social interactions.(10)

4.15.2 Benefits and Harms of Screening Tests

The main benefit of universal screening for dyslexia risk is that it could prevent the reading problems associated with early common, but often under-identified reading disability. Reading intervention in early elementary school clearly reduces the risk for a reading

problem in general.(11) The most important issue is the reliability and validity of decisions made by professionals and families based on the screening method. A possible harm of screening would be wasting of resources to provide intervention or additional assessment to students who did not need the extra support for false positive errors. However, the consequences of not detecting the problem is perhaps more serious which could mean lack of access to early reading intervention.(11)

Early diagnosis, indeed, is a main requisite when dealing with subjects suspected to suffer from reading disabilities and it has been stated that early diagnosis is essential to maximize the rehabilitative outcome in dyslexic children.(12-13) There is little debate as to whether the early identification of students is a useful mechanism by which students who are at-risk for reading problems, including dyslexia, can be routed to appropriate next steps such as intensive early interventions (in pre-school or kindergarten) or more in-depth diagnostic testing for diagnoses of reading disabilities. Screening and intervention can focus on early literacy skills as well as areas of self-regulation and executive control that can hinder the development of reading, writing, language processing, and comprehension.(11)

There were no studies identified on the direct effect of screening for learning disability. To evaluate the effectiveness of these screening tools indirectly, different studies on the effect of early intervention were identified and assessed.

In an evidence review done in 2015, among people with learning disability, different interventions were assessed for its effectiveness in terms of reducing the challenging behavior and improve adaptive functioning.(14) Challenging behavior as defined by the Royal College of Psychiatrists (2007), is behavior of such an intensity, frequency and duration as to threaten the quality of life and/or the physical safety of the individual or others and is likely to lead to responses that are restrictive, aversive or result in exclusion. These would include: aggressive behavior (such as verbal abuse, threats and physical violence), destructive behavior (such as breaking or destroying furniture and other objects and setting fires), disruptive behavior (such as repetitive screaming, smearing feces, setting off fire alarms when there is no fire, calling the emergency services when there is no emergency), self-injurious behavior (including self-biting, head banging), sexually harmful behavior (including sexual assaults, rape and stalking).(14)

In terms of using an educational intervention versus any control, based on very low-quality evidence from a single study ($n = 294$), it is found that these educational interventions are more effective than control in reducing the severity of behavior that challenges at end of treatment. However, the precision of this estimate was poor. In the same study, educational intervention is more effective than control in increasing both social and communicative adaptive functioning at end of treatment. (Table 1)

Table 1: Summary of findings table for educational intervention versus any control

Outcomes	Comparative Risks (CI 95%) Educational intervention versus any control	No of Participants	Quality of Evidence (GRADE)
Behavior that challenges(severity) post treatment Change score	Mean behavior in Intervention group was 0.19 SD higher (0.42 lower to 0.04 higher)	294 (1 study)	Very low
Adaptive Functioning - social (post treatment)	Mean adaptive functioning in the Intervention group was 0.76 SD higher (0.52 lower to 1 higher)	294 (1 study)	Very low
Adaptive Functioning - communication(post treatment)	Mean adaptive functioning in the Intervention group was 0.94 SD higher (0.7 lower to 1.19 higher)	294 (1 study)	Very low

Source: National Collaborating Centre for Mental Health (UK). (2015). Challenging behaviour and learning disabilities: prevention and interventions for people with learning disabilities whose behaviour challenges.. Table 59. Summary of findings table for educational intervention versus any control. Page 177-178

In comparing home based vs. center based early behavioral intervention, the review found that with very low-quality evidence (1 study, n=44), it is inconclusive as to the effectiveness of home-based when compared with centre-based early behavioral intervention in reducing the severity of behavior that challenges at the end of treatment. However, home-based early behavioral intervention was less effective than the centre-based early behavioral intervention in increasing the social and communicative adaptive functioning (very low quality of evidence with 1 study, n=56 and the precision of the estimate for communicative adaptive functioning was poor.) (Table 2).

Table 2: Summary of findings table for home-based versus centre-based early behavioral intervention

Outcomes	Comparative Risks (CI 95%) Behavioral intervention: Centre-based early - Home based early	No of Participants	Quality of Evidence (GRADE)
Behavior that challenges(severity) post treatment	Mean behavior in Intervention group was 0.11 SD lower (0.7 lower to 0.48 higher)	44 (1 study)	Very low ^{1,2}
Adaptive Functioning -social (post treatment)	Mean adaptive functioning in Intervention group was 0.63 SD higher (1.17 to 0.09 lower)	56 (1 study)	Very low ^{1,2}
Adaptive Functioning - communication(post treatment)	Mean adaptive functioning in Intervention group was 0.46 SD higher (1 lower to 0.07 higher)	55 (1 study)	Very low ^{1,2}

¹Crucial limitation of r1 criterion or some limitations for multiple criteria sufficient to lower one's confidence in the estimate of effect

² Optimal information size not met, Small single study

Source: National Collaborating Centre for Mental Health (UK. (2015). Challenging behaviour and learning disabilities: prevention and interventions for people with learning disabilities whose behaviour challenges. Table 60. Summary of findings table for home-based versus any center-based early behavioral intervention. Page 178

With very low quality of evidence and poor precision of estimate (2 studies, n =117) (Low evidence was further downgraded to Very low due to imprecision), interventions involving parent education, support and skills training, is more effective than control in reducing the severity of behavior that challenges at the end of treatment and up to 52-week follow-up. The intervention is inconclusive in improving the adaptive function at the end of the treatment (n=28) but is found to be more effective at 52 week follow up (2 studies, n =119). It is also with low quality evidence (1 study, n =68) that this intervention is more effective in improving communicative adaptive functioning at 26 week follow up. (Table 3).

Table 3: Summary of findings table for parent education, support and skills training versus any control

Outcomes	Comparative Risks (CI 95%) Control Parent Education, support and skills training	No of Participants	Quality of Evidence (GRADE)
Behavior that challenges(severity) post treatment	Mean behavior that challenges in the intervention group was 0.4 SD lower (0.93 lower to 0.12 higher)	57 (1 study)	Very Low ^{2,3,4}
Behavior that challenges(severity) follow-up (26-52 weeks)	Mean behavior that challenges in the intervention group was 0.37 SD lower (0.79 lower to 0.05 higher)	117 (2 studies)	Very Low ^{2,3,4}
Adaptive Functioning - global (post treatment)	Mean adaptive functioning in Intervention group was 0.25 SD higher (0.27 lower to 0.77 higher)	55 (1 study)	Very Low ^{1,4}
Adaptive Functioning - global follow-up (26-52 weeks)	Mean adaptive functioning in Intervention group was 0.52 SD higher (0.15 to 0.88 higher)	119 (2 studies)	Very Low ^{2,3,4}
Adaptive Functioning - communication follow-up (mean 26 weeks)	Mean adaptive functioning in Intervention group was 0.75SD higher (0.26 to 1.25 higher)	68 (1 study)	Very Low ^{1,4}

¹ Optimal information size not met; small single study

² Most information is from studies at moderate risk of bias

³Optimal information size not met

⁴ Imprecision

Adapted from: National Collaborating Centre for Mental Health (UK). (2015). Challenging behaviour and learning disabilities: prevention and interventions for people with learning disabilities whose behaviour challenges. Table 65. Summary of findings table for parent education, support and skills training versus any control. Page 181

In terms of functional academics, different interventions focusing on improving reading skills have been studied. Meta-analysis and systematic reviews on phonics training, repeated reading, spelling intervention, intensive early training and various interventions were appraised using AMSTAR¹⁵ with most of the studies fulfilling the criteria (APPENDIX D)

A meta-analysis done in 2018 reports that phonics training can improve the literacy in English-speaking poor readers and can improve certain literacy related skills in poor readers.(16) Phonics training probably improved irregular word reading accuracy, mixed/regular word reading fluency, and non-word reading fluency, may improve mixed/regular word reading accuracy, non-word reading accuracy, spelling, letter-sound knowledge, and phonological output and may slightly improve reading comprehension. The quality of evidence was low to moderate for the included studies since there were wide confidence intervals for the outcomes affecting the precision of the data.

In another review involving 25 studies, repeated reading as an intervention made a positive effect on reading fluency by increasing the correct words per minute post

intervention (1.41, 95% CI 0.99-1.41).(17) A larger effect is observed in elementary students compared to secondary students, however the study has large heterogeneity.

In a meta-analysis involving 34 studies, it was shown that learners with dyslexia or spelling deficits who take part in a spelling intervention show better reading and spelling performance compared with children who received regular school practice or no spelling instruction.(18) Another review with a different intervention reported intensive early reading interventions improved early literacy in alphabetic language in 25 studies among struggling readers in kindergarten through third grade comparing it with a “business as usual” learning control group.(19)

The latest meta-analysis on different dyslexia treatment published in 2021 was composed of 1862 subjects in 40 studies evaluating reading performances showed a modest effect (mean standard difference 0.38, 95% CI 0.31-0.46) compared to no intervention. However, this study comprised of low powered, small true effect studies had substantial heterogeneity.(20)

Table 4. Studies on Different Interventions for Reading Disabilities

Author, year	McArthur, 2018 ¹⁶	Lee, 2015 ¹⁷	Galuschka, 2020 ¹⁸	Wanzek, 2018 ¹⁹	Toffalini, 2021 ²⁰
Country	Canada/UK/USA/Aus	USA	Germany	USA	Italy
Number of patients, studies	14 studies	34 studies	34 studies SR 28 studies Meta	25 studies	n=1862 40 studies
Intervention	Phonics Training	Repeated Reading	Spelling Intervention	Intensive early reading	Various interventions
Control	No or alternative training		Present	Business as usual	Present
Outcomes	Literacy related skills Word/nonword reading accuracy, fluency, comprehension, spelling	Reading fluency	Reading and spelling performance	Early literacy in alphabetic language	reading performance
Effect	Summary of Findings (Appendix E)	1.41 [0.99-1.41] Large effect Larger effect in elementary students compared to secondary students	0.58 [0.4-0.76]	0.39 Publication bias 0.28	0.38 [0.31-0.46] Modest effect

4.15.3 Diagnostic Performance of Screening Tests

Identifying learning disability requires a multistep process and use of tools and resources that are translated and adapted to the local cultural context, and the participation of a multidisciplinary team of experts.(21) No studies are found as to specific screening tools for learning disability.

In a status update on dyslexia laws in the US, 18 states have implemented universal screening for dyslexia or are in the process of completing pilot programs to formalize universal screening procedures.(22) However, there is no standardized screening protocol being used or endorsed specifically for reading. No single test or tool can measure all reading skills. The National Center on Intensive Intervention does independent, standardized reviews of screening measures to help institutions select an appropriate screening tool.(23) (APPENDIX F)

4.15.4 Cost Implication

Since learning disability is under the umbrella of developmental disability, it can be covered by Philhealth with Z Benefits for Children with Developmental Disability (Philhealth Circular No. 2017-0029). Under the package, initial assessment can be as low as P3,626 to as high as P15,828. Once diagnosed, the package will also cover as much as nine tranches of therapy amounting to P5,000 each tranche.(24)

Table 5: Estimated Cost of Assessment and Treatment from Philhealth Package

Description	Amount	Description
Initial assessment by a medical specialist to three allied health professional or rehabilitation therapist	P3,626 -P5,276	1 per cycle year for a maximum of 3 cycles
Rehabilitation therapy	P5,000 per tranche	Maximum of 9 tranches per year

Source: Philhealth Circular No. 2017-0029 (24)

In communication with a local reading intervention specialists (Council of Reading Intervention Specialists), the rate for literacy assessment ranges from P2,500 (informal) to P10,000 (formal/standardized) while the literacy intervention would cost around P990 to P1,100 per hour.(25)

4.15.5 Equity, Acceptability, and Feasibility

The Department of Education through the Bureau of Learning Delivery-Teaching and Learning Division administers the Revised Philippine Informal Reading Inventory (Phil-IRI), a classroom-based assessment tool aiming to measure and describe the learners' reading performance in both English and Filipino languages in oral reading, silent reading and listening comprehension.(26) These three types of assessment aim to determine the learner's independent, instructional and frustration levels. This however is administered by teachers to grades 3-6 students and not part of the usual well child consult of a physician.

The Philippines does not have a comprehensive program to properly identify and assess children at-risk of learning difficulty/disability. Since there is no standardized screening tool available, foreign made tests are being used, which poses a question on the validity and applicability of these tests. Since there is lack of laws, policies and programs, and experts who can administer screening tools, patients are mostly referred to private specialists or institutions which are usually not affordable for poor families.(27) Socioeconomic status affects access to both assessment and treatment.

This pandemic also greatly affected screening and treatment for children with disability. In a study by UNICEF, among children with disabilities, as much 51.8 % are unable to access to education/learning service, 50.7% are unable to access child development centers and 49.5% are unable to access habilitation and rehabilitation services.(28)

4.15.6 Recommendations from Other Groups

The National Joint Committee on Learning Disabilities strongly supports comprehensive assessment and evaluation of students with learning disabilities by a multidisciplinary team for the identification and diagnosis of students with learning disabilities.(29)

American Academy of Pediatrics states that based on some research evidence, as well as consensus, all children should have annual developmental surveillance as well as

formal developmental screening at ages 30 and 48 months for early identification of risk factors for learning disability.(30)

American Academy of Family Physician recommends: (31)

- Efforts should be made to identify children with reading disabilities and to implement interventions at an early age (Evidence rating C)
- Children with problems in school performance should be evaluated for reading difficulties (Evidence rating C)
- Children with reading difficulties should receive individualized instruction emphasizing phonologic awareness (Evidence rating B)

*Evidence rating A = consistent, good-quality patient-oriented evidence; B = inconsistent or limited-quality patient-oriented evidence; C = consensus, disease-oriented evidence, usual practice, expert opinion, or case series

Children with speech and language disorders may be at increased risk of learning and literacy disabilities including difficulty in reading. The United States Preventive Services Task Force did a review on speech and language delay in general and they conclude that the current evidence is insufficient to assess the balance of benefits and harms of screening for speech and language delay and disorders in children aged 5 years or younger (I statement).(32) An update literature surveillance report was done in April 2019, however, there are still no new evidences to support an updated systematic review.(33)

References

1. American Psychiatric Association A. Diagnostic and statistical manual of mental disorders. Washington, DC: American Psychiatric Association; 1980 Jan 1.
2. Sharma P, Mahapatra A. Specific Learning Disorder and Dyslexia: Introduction and Need for Awareness in Nepal. Post-Graduate Medical Journal of NAMS. 2020 Jan 27;19(02).
3. Cai L, Chen Y, Hu X, Guo Y, Zhao X, Sun T, Wu Y, Li X. An Epidemiological Study on the Children with Chinese Developmental Dyslexia. Journal of Developmental & Behavioral Pediatrics. 2020 Mar 11.
4. Di Folco C, Guez A, Peyre H, Ramus F. Epidemiology of developmental dyslexia: A comparison of DSM-5 and ICD-11 criteria. medRxiv. 2020 Jan 1.
5. Kausar N, Farhat N, Maqsood F, Qurban H. Specific learning disorder among primary school children of Sarai Alamgir. JPMA. 2021;71(1193).
6. Lin Y, Zhang X, Huang Q, Lv L, Huang A, Li A, Wu K, Huang Y. The Prevalence of Dyslexia in Primary School Children and Their Chinese Literacy Assessment in Shantou, China. International Journal of Environmental Research and Public Health. 2020 Jan;17(19):7140.
7. Statistics AP. Functional Literacy, Education and Mass Media Survey. Philippine Statistics Authority; 2021.
8. UNICEF. Southeast Asia Primary Learning Metrics: National Report of the Philippines. 12. [Internet]. Unicef.org. 2021 [cited 9 September 2021]. Available from: <https://www.unicef.org/philippines/media/2556/file/Southeast%20Asia%20Primary%20Learning%20Metrics%202019%20National%20Report%20of%20the%20Philippines.pdf>
9. Department of Education. PISA 2018: National Report of the Philippines. <https://www.deped.gov.ph/wp-content/uploads/2019/12/PISA-2018-Philippine-National-Report.pdf>
10. Karande S, Bhosrekar K, Kulkarni M, Thakker A. Health-related quality of life of children with newly diagnosed specific learning disability. Journal of tropical pediatrics. 2009 Jun 1;55(3):160-9.
11. Petscher, Y., Fien, H., Stanley, C., Gearin, B., Gaab, N., Fletcher, J.M., & Johnson, E. (2019). Screening for Dyslexia. Washington, DC: U.S. Department of Education, Office of Elementary

- and Secondary Education, Office of Special Education Programs, National Center on Improving Literacy. Retrieved from improvingliteracy.org.
12. Aleci C. Can Self-reported Ocular Motor and Perceptive Alterations Predict a Reading Disability? A Pilot Study on the Analytic Anamnestic Protocol. *Cureus*. 2019 Sep;11(9).
 13. Peterson RL, Pennington BF. Developmental dyslexia. *Lancet*. 2012;379:1997–2007.
 14. National Institute for Health and Care Excellence. Challenging behaviour and learning disabilities: prevention and interventions for people with learning disabilities whose behaviour challenges [Internet]. Nice.org.uk. 2021 [cited 9 September 2021]. Available from: <https://www.nice.org.uk/guidance/ng11/resources/challenging-behaviour-and-learning-disabilities-prevention-and-interventions-for-people-with-learning-disabilities-whose-behaviour-challenges-pdf-1837266392005>
 15. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *bmj*. 2017 Sep 21;358.
 16. McArthur G, Sheehan Y, Badcock NA, Francis DA, Wang HC, Kohnen S, Banales E, Anandakumar T, Marinus E, Castles A. Phonics training for English-speaking poor readers. *Cochrane Database of Systematic Reviews*. 2018(11).
 17. Lee J, Yoon SY. The effects of repeated reading on reading fluency for students with reading disabilities: A meta-analysis. *Journal of learning disabilities*. 2017 Mar;50(2):213-24.
 18. Galuschka K, Görgen R, Kalmar J, Haberstroh S, Schmalz X, Schulte-Körne G. Effectiveness of spelling interventions for learners with dyslexia: A meta-analysis and systematic review. *Educational Psychologist*. 2020;55(1):1-20.
 19. Wanzek J, Stevens E, Williams K, Scammacca N, Vaughn S, Sargent K. Current Evidence on the Effects of Intensive Early Reading Interventions. *Journal of Learning Disabilities*. 2018;51(6):612-624.
 20. Toffalini E, Giofrè D, Pastore M, Garretti B, Fraccadori F, Szűcs D. Dyslexia treatment studies: A systematic review and suggestions on testing treatment efficacy with small effects and small samples. *Behavior Research Methods*. 2021 Mar 10:1-9.
 21. Hayes AM, Dombrowski E, Shefcyk AH, Bulat J. Learning Disabilities Screening and Evaluation Guide for Low-and Middle-Income Countries. Occasional Paper. RTI Press Publication OP-0052-1804. RTI International. 2018 Apr.
 22. Academic Screening Tools Chart [Internet]. Charts.intensiveintervention.org. 2021 [cited 6 September 2021]. Available from: <https://charts.intensiveintervention.org/ascreening>
 23. Dyslexia Laws in the USA: A 2018 Update - International Dyslexia Association [Internet]. International Dyslexia Association. 2021 [cited 9 September 2021]. Available from: <https://dyslexiaida.org/dyslexia-laws-in-the-usa-an-update/>
 24. Philhealth Circular No. 2017-0029. [cited 9 September 2021]. Available from: https://www.philhealth.gov.ph/circulars/2017/TS_circ2017-0029.pdf
 25. Jocson, C. Screening for reading disability. [Internet]. E-mail from Cecilia Q. Jocson (ciljocs@yahoo.com) to Fides Roxanne Castor (fides_castor@yahoo.com) 2021 September 26 [cited 2021 September 30]
 26. Revised Philippine Informal Reading Inventory (Phil-IRI) - TeacherPH [Internet]. TeacherPH. 2021 [cited 24 September 2021]. Available from: <https://www.teacherph.com/revised-phil-iri/>
 27. Valladolid VC. Towards Developing a Reading Disability Identification Model for Philippine Public Schools. *Educational Measurement and Evaluation Review*. 2014;5(1):1-.
 28. UNICEF in partnership with Council for the Welfare of Children-Subcommittee on Children with Disabilities. Situation of children with disabilities in the context of COVID 19 2020. Available from: <https://www.unicef.org/philippines/media/2476/file/Situation%20of%20Children%20with%20Disabilities%20in%20the%20Context%20of%20COVID-19.pdf>
 29. Comprehensive Assessment and Evaluation of Students with Learning Disabilities: A Paper Prepared by the National Joint Committee on Learning Disabilities. *Learning Disability Quarterly*. 2011;34(1):3
 30. Rimrodt SL, Lipkin PH. Learning disabilities and school failure. *Pediatrics in Review-Elk Grove*. 2011 Aug 1;32(8):315.

31. Hamilton S, Glascoe FP. Evaluation of children with reading difficulties. *American Family Physician*. 2006 Dec 15;74(12):2079-84.
32. Siu AL. Screening for speech and language delay and disorders in children aged 5 years or younger: US Preventive Services Task Force recommendation statement. *Pediatrics*. 2015 Aug 1;136(2):e474-81.
33. Literature surveillance report: Speech and language delay and disorders in children age 5 and Younger: Screening: United States Preventive services taskforce [Internet]. Literature Surveillance Report: Speech and Language Delay and Disorders in Children Age 5 and Younger: Screening | United States Preventive Services Taskforce. 2019 [cited 2021Sep6]. Available from: <https://www.uspreventiveservicestaskforce.org/uspstf/document/literature-surveillance-report/speech-and-language-delay-and-disorders-in-children-age-5-and-younger-screening>

5. RESEARCH IMPLICATIONS

Many research questions from the identified clinical questions in this CPG were unanswered in terms of benefits and harms of screening, equity, applicability, and feasibility. Direct evidence is lacking to provide a definitive list of conditions to be recommended for inclusion in the newborn screening due to the rarity of the disorders and lack of comparative studies among those screened versus unscreened newborns.

For the congenital metabolic disorders, establishing direct evidence through clinical trials may be problematic. Because of this challenge, establishing the diagnostic performance of tests as indirect evidence can be adequate. However, the burden of these diseases are low worldwide and conducting studies on these cases may not always be available.

On the other hand, for the developmental disorders, there is still no specific screening tool identified for the specific diseases such as learning disorders and developmental delay. Psychometric properties of some standardized tools used to detect developmental delays among apparently healthy children have not been established. These tests have also not been validated in the vernacular as of press time.

Few cost-effectiveness studies are available for screening the diseases included in this CPG, but most of them are conducted in Western countries. In fact for the congenital disorders, there are no cost-effective studies mentioned. This is possibly due to the very small number of participants in the country who would qualify for these studies. However for the developmental disorders specifically for reading disabilities and Autism Spectrum Disorder, the burden of disease has been increasing and cost-effectiveness research is still not adequately investigated.

Social science research also plays a vital role in examining the impact of the diseases. Although there are some local studies already examining the motivations among the parents in entering their children in screening programs for these developmental disorders, these need to be further updated as these studies can become determinants for government to push for or against the screening of these disorders. Qualitative studies may also provide more information on the probable harm of mislabeling.

Many research questions emerged from collating the evidence for this CPG and can be explored further. In particular, there were fourteen (14) metabolic disorders in the Expanded Newborn Screening that were not included in the evidence reviews. These disorders are: (1) Galactosemia, (2) Maple Syrup Urine Disease, (3) Phenylketonuria, (4) Glutaric aciduria Type 1, (5) Propionic aciduria, (6) Methylmalonic aciduria, (7) 3-methylcrotonyl carboxylase deficiency, (8) Carnitine uptake deficiency, (9) Very long chain acyl-CoA Dehydrogenase (VLCAD) Deficiency, (10) Medium chain acyl-CoA Dehydrogenase (MCAD) Deficiency, (11) Citrullinemia Type 1, (12) Argininosuccinic Aciduria, (13) Congenital Hypothyroidism and (14) Congenital Adrenal Hyperplasia. Although these conditions are already present in the ENBS package, they were not prioritized because they have a higher incidence compared to the sixteen (16) disorders tackled in this PHEX. Filling in these gaps can provide a clearer picture of the impact of

screening programs using previously mentioned tests and may influence the recommendations for updating this guideline.

6. DISSEMINATION AND IMPLEMENTATION

A full copy of this document will be sent to the Department of Health for transmittal and publication. The Disease Prevention and Control Bureau will transmit copies of this CPG to the Philippine Health Insurance Corporation (PHIC) and health maintenance organizations (HMOs) and NGOs involved in a periodic health examination. The recommendations and the evidence summaries will be posted in the PHEX web based application.

All strong recommendations in this guideline can be used for monitoring and auditing practices in institutions. This can be converted to key performance indicators and it can also be used in creating clinical pathways.

The DOH planned to develop a simplified version of this CPG and made it available in the format that will be ready for reproduction and dissemination to the patients in different health care settings. It will also be available for interested parties who might visit the DOH website.

7. APPLICABILITY ISSUES

The PHEX Task Force accentuates some caveats of this CPG using equity and applicability lenses. Comprehensive history taking, physical examination, and monitoring are essential parts of evaluating risk factors and the probability of developing diseases. This CPG does not necessarily supersede the consumers' (i.e., health professionals, hospital administrators, employers, payors, patients) values, settings, and circumstances.

Although this CPG intends to influence the direction of health policies for the general population, it should not be the sole basis for recreating or abolishing practices that aim to improve the health conditions of many Filipinos, particularly those part of the workforce.

8. UPDATING OF THE GUIDELINES

The recommendations herein shall hold until such time that new evidence on screening, diagnosing or managing various risk factors and diseases emerges and contingencies dictate updating this Philippine Guidelines on Periodic Health Examination. The CPGs will be updated every 3-5 years or earlier if new significant evidence becomes available.

9. APPENDICES

Search Strategy, Characteristics of Included Studies, Forest Plots, GRADE Evidence, and Cost-effectiveness Studies for the Research Questions

1. Pulse Oximetry Screening for Critical Congenital Heart Disease

APPENDIX A. Search Yield

Summary of Search: (done 1st to 2nd week of July, 2021, updated September 16, 2021)

1. Search from CPG database

a. NICE (National Institute for Health and Care Excellence)

#	Query	RESULTS	Comments
1	“Neonate”	3	Manual search - none
2	“Newborn”	3	Manual search - none
3	“Pulse oximetry”	1	Manual search - none

b. NICE Evidence Search

1	“pulse oximetry neonates”	86	Manual search (after 2015) - 4
---	---------------------------	----	--------------------------------

	Title	URL	Source	Date Published
1	Recommendations for the adult cardiac sonographer performing echocardiography to screen for critical congenital heart disease in the newborn (Wasserman et al)	https://www.asecho.org/wp-content/uploads/2021/03/PIIS0894731720307793.pdf	American Society of Echocardiography	2021
2	What is the accuracy of pulse oximetry for detection of critical congenital heart defects (CHDs) in asymptomatic newborns?	https://www.cochranelibrary.com/cca/doi/10.1002/cc.a.2759/full	Cochrane Clinical Answers	
	Pulse oximetry screening for critical congenital heart defects (Plana et al)	https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011912.pub2/full	Cochrane Library	2018
3	Pulse oximetry screening in newborns to enhance detection of critical congenital heart disease (Narvey, Wong & Fournier) --- Practice Point	https://www.cps.ca/en/documents/position/pulse-oximetry-screening	Canadian Paediatric Society	2017
4	Role of pulse oximetry in examining newborns for congenital heart disease: a scientific statement from the American Heart Association and American Academy of Pediatrics. (Mahle et al)	https://pubmed.ncbi.nlm.nih.gov/19581492/	American Heart Association /AAP	2009

c. US Preventive Services Task Force (USPSTF)

#	QUERY	RESULTS
1	Newborn screening	129
	Filter: Published	102
	Filter: Pediatric	27
	Yield for:	
	• Pulse Oximetry Screening	0
	• Critical Congenital Heart Disease	0

d. Canadian Task Force for Preventive Health Care (CTFPHC)

#	QUERY	RESULTS
1	Critical Congenital Heart Disease	0
2	Newborn screening	6
	Yield for:	
	• Critical congenital heart disease	0
		0

	• Pulse oximetry	
--	------------------	--

2. PubMed Search

#	Query	RESULTS	Comments
1	guideline [Publication Type] OR practice guideline [Publication Type] or recommendation*[Title] OR standard* [Title] OR standard* [Title] OR guideline* [Title]) AND pulse oximetry	94	
2	Filter: Age birth-1 month, 2015-2021	18	Manual search - 1 guideline (but in Spanish)
3	guideline [Publication Type] OR practice guideline [Publication Type] or recommendation*[Title] OR standard* [Title] OR standard* [Title] OR guideline* [Title]) AND critical congenital heart disease Plus filter: Age birth-1 month, 2015-2021	21	Manual search - 1 guideline (but in Spanish)
4	guideline [Publication Type] OR practice guideline [Publication Type] or recommendation*[Title] OR standard* [Title] OR standard* [Title] OR guideline* [Title]) AND cchd screen Plus filter: Age birth-1 month, 2015-2021	2	Manual search - 1 guideline (but in Spanish)
5	asymptomatic newborns AND pulse oximetry	72	
6	Filter: 2015-2021	32	Manual search – 1 cochrane review (Plana et al, 2018), (2 cross sectional studies, 7 prospective cohorts, 2 retrospective cohort, 2 descriptive)
7	asymptomatic newborns AND congenital heart disease Filter: 2015-2021	155	Manual search – 1 cochrane review (Plana et al, 2018), 5 prospective cohort, 3 descriptive, 2 cross sectional, 3 retrospective cohort (1 in Chinese)*
	“asymptomatic newborns” AND “congenital heart disease” Filter: 2015-2021	13	Manual search – 0 guidelines, 1 cochrane review (Plana et al, 2018), 5 prospective cohort, 1 descriptive, 2 cross sectional*
8	((Pulse oximeter) OR (CCHD screen)) AND ((neonate) OR (Full term) OR (preterm)) AND (prevention) Filter: 2015-2021	28	Manual search – 0 guidelines, 1 Prospective
9	((((systematic review[Publication Type]) OR (guideline[Publication Type])) OR (practice guideline[Publication Type])) OR (evidence review[Publication Type])) AND (pulse oximetry[Title]) Filter: 2015-2021	13	Manual search – 1 guideline , 1 cochrane review (Plana et al, 2018, Spanish guideline), 4 articles
10	((((systematic review[Publication Type]) OR (guideline[Publication Type])) OR (practice guideline[Publication Type])) OR (evidence review[Publication Type]))) AND (pulse oximetry) Filter: 2015-2021	187	
11	Filter: Age Group newborn to 1 month, Clinical trial, meta analysis, RCT, Review, Systematic Review	39	Manual search 1 cochrane review (Plana et al, 2018) 6 studies and articles (with 2 other systematic reviews)**

12	(asymptomatic newborns) AND (pulse oximetry) Filter: 2015-2021	32	Manual search – 1 cochrane review (Plana et al, 2018), 17 articles* **
13	Asymptomatic neonates AND pulse oximetry Filter: 2015-2021, newborns-1 month	27	Manual search – 1 cochrane review (Plana et al, 2018), 15 studies and articles * ***
14	"pulse oximetry screening" Filter: 2015-2021, Child birth – 18 years, Newborn, Infant	91	
15	("pulse oximetry" OR "pulse oximeter") AND ("CCHD" OR "critical congenital heart disease" OR "CHD" OR "congenital heart disease") Filter: 2015-2021, Child birth – 18 years, Newborn, Infant	126	
16	(Infant, Newborn [MeSH] OR neonate* OR infant* OR newborn)* AND (Heart Defects, Congenital [MeSH] OR Heart Valve Diseases [MeSH] OR tetralogy fallot* OR cyanotic heart OR congenital heart OR congenital cardiac OR aortic coarctation OR valve diseases OR hypoplastic syndrome OR pulmonary atresia OR interruption of the aortic arch OR valve stenosis OR pulmonary atresia) AND (oximetry [MeSH] OR oximetry OR pulse oximeter*OR oxygen saturation OR O2 saturation) Filter: March 2017-2021	369	(search method of Cochrane review - Plana et al)

*With some overlap to query #6

** With some overlap to query #9

*** With some overlap to query #14

HERDIN Search done July 2021

Search term: “Pulse Oximetry” --- 0 completed studies

1 ongoing study: The Philippine Multicenter Pulse Oximetry Screening for Critical Congenital Heart Disease: A Pilot Study

APPENDIX B. Summary of Included Studies: Beyond Cochrane Review (April 2017 – September 2021)

Almawazini, 2017			
Methodological Quality	prospective cross-sectional		
Item	Judgement	Risk of Bias	Applicability Concerns
Patient Characteristics and Setting			
	Country: Saudi Arabia Setting: Pediatric and Neonatology Department, King Fahad Hospital Study Period: February 2016 to February 2017 Inclusion Criteria: All live birth newborn infants delivered and admitted to nursery unit for observation without any associated medical problem Exclusion Criteria: Preterm <34 weeks, diagnosed with CHD by fetal echocardiography, syndromic newborns, newborns with signs of sepsis, or who were admitted to NICU after delivery because of other medical problems, patients referred from other hospitals N screened: 2961 out of 3300 live birth newborn infants Gestational Age: more than or equal to 34 weeks		
Index Tests			
	Pulse oximetry screening was performed by using a pulse oximeter with an adhesive sensor placed on the baby's skin (Masimo Corporation) Screening Protocol: Site of Testing: Right hand (pre ductal), then on either foot (post ductal) Test timing: >24 hours HOL Oxygen Saturation: functional Threshold: <95% and difference in values >3% <i>"The screening was performed first in the right hand, then in either foot by using one pulse oximeter. The baby passed the screening if the oxygen saturation was ≥95% in the right hand and either foot and the difference in values was ≤3%. A failed screening result was defined as an oxygen saturation of <90% in the right hand and either foot in the initial or repeated screens. If the oxygen saturation was >90% and <95% in the right hand and either foot or there was a >3% difference between the right hand and either foot, then the screening was repeated after one hour, and the same process as described above was followed."</i>		
Target Condition and reference standard(s)			
	Target Condition: Critical Congenital Heart Disease (Critical TOF, interrupted aortic arch, HLHS, TGA, PA and TAPVR needing urgent intervention) Reference Standard: Positive screen: 2D Echo using Philips IE33 Ultrasound Negative Screen: no mention		
Flow and Timing			
	Duration of follow up: not stated Loss to follow up: none (n = 2961 analyzed)		
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	Unclear	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
<i>Domain 2: Index Test Pulse Oximetry</i>			

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	Yes	Unclear	Low
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Unclear	Unclear	
Did all patients receive a reference standard?	unclear		
Were all patients included in the analysis?	Yes		

Cloete, 2019 (prospective)	
Methodological Quality	Prospective
Patient Characteristics and Setting	Country: New Zealand Setting: 2 of New Zealand's 20 districts, three hospitals four primary maternity units District A: Metropolitan Hospital and its neighboring Primary Maternity Unit, District B: Region Level 1 and Level 2 Hospital, District C: Three primary maternal Units Study Period: May 2016 to April 2018 Inclusion Criteria: well newborn infants > or equal to 35 weeks Exclusion Criteria: infants with prenatal diagnosis of a congenital anomaly and other infants admitted to the NICU within 2 hours from birth N screened: 16 644 Gestational Age: more than or equal to 35 weeks
Index Tests	Pulse oximetry was performed by midwives or nurses with identical handheld oximeters (Masimo; Radical SET, version 5 with reusable sensors) Screening protocol: Site of testing: post-ductal (foot)Test timing: between 2 and 24 hours Oxygen saturation: functional Threshold: <95% <i>"Infants achieving an oxygen saturation of 95% or greater passed the test and required no further evaluation provided that they remained clinically well. Results below 90% warranted a referral to the nearest paediatric service for telephonic advice and/or clinical assessment. Saturation between 90% and 94% were regarded as an inconclusive result, and therefore, repeat testing had to be performed 1-2 hours later. Three consecutive results in the inconclusive range also warranted a paediatric referral."</i>

Target Condition and reference standard(s)	Target condition: critical congenital heart disease defined as cardiac related death or cardiac interventions in the first 28 days after birth Reference standard: Positive Screen: referral to the nearest pediatric service for telephonic advice and assessment Negative screen: Review of cardiac center databases to identify all infants, not identified on antenatal ultrasound screening that underwent a cardiac catheter and/or cardiac intervention, mortality data from New Zealand Ministry of Health's Mortality Collection and the Perinatal and Maternal Mortality Review Committee				
Flow and Timing	Duration of follow up: not stated Loss to follow up: none				
Item	Judgement	Risk of Bias	Applicability Concerns		
<i>Domain 1: Patient Selection</i>					
Was a consecutive or random sample of patients enrolled?	Unclear				
Was a case-control design avoided?	Yes	High	Low		
Did the study avoid inappropriate exclusions?	No				
<i>Domain 2: Index Test Pulse Oximetry</i>					
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low		
If a threshold was used, was it pre-specified?	Yes				
<i>Domain 3: Reference Standard</i>					
Is the reference standards likely to correctly classify the target condition?	Yes	Unclear	Low		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear				
<i>Domain 4: Flow and Timing</i>					
Were there at least 28 days of appropriate follow up between index test and reference standard?	Unclear	High			
Did all patients receive a reference standard?	No				
Were all patients included in the analysis?	No				
Diller, 2018					
Methodological Quality	Retrospective observational				

Patient Characteristics and Setting	Country: USA Setting: Large tertiary birth hospital with 2 delivery campuses in metropolitan Atlanta, Georgia Study Period: Jan 1 2013 to Dec 31, 2016 Inclusion Criteria: Term newborns born within study period Exclusion Criteria: Infants with a prenatal diagnosis of CCHD and infants transferred to NICU before screening N Screened: 77, 148 Gestational Age: Term newborns (age in weeks not explicitly stated)		
Index Tests	Type of pulse oximeter used not stated. Screening protocol: Site of Testing: Right hand (pre ductal) or either foot (post ductal) Timing: more than or equal to 24 hours Oxygen saturation: functional Threshold: <95%		
Target Condition and reference standard(s)	Target Condition: Critical Congenital Heart Disease (HLHS, Pulmonary Atresia, TOF, TAPVR, TGA, truncus arteriosus, tricuspid atresia, coarctation of the aorta, DORV, Ebstein Anomaly, interrupted arch, and any other single ventricle physiology), considered to be significant because of need for intervention or cardio ff up to prevent significant morbidity or mortality Reference Standard: Negative Screen: Surgical records reviewed at Childrens Healthcare of Atlanta (the sole pediatric cardiac surgical center in the area) Positive screen: Birth records screening data regarding any 2D echo results (as well as other significant non cardiac diagnoses)		
Flow and Timing	Records identified infants who underwent surgical intervention for CCHD within the first 6 months of life. Lost to Follow up: 44 (n analyzed = 77148 included in analysis of 77184 collected pulse oximetry results)		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	No (retrospective review of records)	High	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear	Unclear	Low
If a threshold was used, was it pre-specified?	Yes		

Domain 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes (confirmation of CCHD based on Echo, surgical records)	Unclear	Low
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Domain 4: Flow and Timing			
Were there at least 28 days of appropriate follow up between index test and reference standard?	No	Unclear	
Did all patients receive a reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Gopalakrishnan, 2021	
Methodological Quality	Prospective
Patient Characteristics and Setting	Country: India
	Setting: Post-natal ward of a tertiary care multispecialty referral hospital
	Study Period: Sept 2016 to March 2019
	Inclusion Criteria: All stable and asymptomatic intramural neonates delivered via normal or caesarian delivery
	Exclusion Criteria: Infants with antenatal ultrasound/echo diagnosis of CHD, requiring NICU admission, with major congenital malformations
	N screened: 1855
	Gestational Age: all applicable gestational ages
Index Tests	Pulse oximetry was performed by trained nursing staff with a Masimo Radical-7 handheld oximeter with reusable probes
	Screening protocol: AAP/AHA Protocol
	Site of testing: right hand and either foot in a non specified order
	Test timing: >24 hours
	Oxygen saturation: functional
	Threshold: <95%
	<i>"A positive pulse oximetry screen was defined as an SpO2 <90% in either the right hand or foot or an SpO2 between 90% and 94% in either site or a >3% difference between the two sites (repeated twice at 1-h intervals). A negative pulse oximetry screen was defined as an SpO2 ≥ 95% in the right hand or foot and ≤ 3% difference between the two sites."</i>

Target Condition and reference standard(s)	Target Condition: Critical Congenital Heart Disease, defined as duct dependent CHD that are life-threatening without treatment in the neonatal period or infancy such as pulmonary atresia with intact ventricular septum, pulmonary stenosis, tetralogy of fallot, total anomalous pulmonary venous return, transposition of the great arteries (TGA), tricuspid atresia, truncus arteriosus or duct-dependent systemic defects such as coarctation of aorta, interrupted aortic arch, hypoplastic left heart syndrome and aortic stenosis		
	Reference Standard:		
	Positive screen: thorough clinical examination and confirmatory 2D Echocardiography by cardiologist		
	Negative screen: clinical follow up at 6, 10 and 14 weeks and analyzed for symptoms, as well as analysis of hospital admission and discharge register		
Flow and Timing	Duration of follow up: 6, 10 and 14 weeks		
	Lost to Follow up: (examined 1855/2751 neonates delivered) excluded 896 infants due to antenatal diagnosis (n = 8), NICU admission (n=478), Major congenital malformations (n=4) , missed screening (n=380), refused consent (n = 6)		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	High	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	Unclear	Low
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Yes	High	
Did all patients receive a reference standard?	No		

Were all patients included in the analysis?	Yes		
---	-----	--	--

Hamilcikan, 2017			
Methodological Quality	Prospective		
Patient Characteristics and Setting	<p>Country: Turkey Setting: Third level neonatal center with 24/7 echocardiography service with a pediatric cardiologist on call Study Period: Oct 1, 2015 to October 31, 2016 Inclusion Criteria: All term and late preterm infants Exclusion Criteria: Infants admitted to the neonatal ICU and were not monitored by a pulse oximeter N screened: 4335 Gestational Age: >34 weeks</p>		
Index Tests	<p>Pulse oximetry was performed by the same nurse for each group with a Masimo Radical-7 handheld oximeter with reusable sensors and disposable adhesive straps</p> <p>Screening protocol:</p> <p>Site of testing: right hand/wrist and either foot in a non specified order</p> <p>Test timing: <24 hours (early discharge group) and >24 hours</p> <p>Oxygen saturation: functional</p> <p>Threshold: <95%</p> <p><i>"The PO screening protocol used followed the AAP (2012): positive after one quality signal in a pre- or postductal SpO₂ reading of <90%. The screening was also considered positive after two repeated measurements, with a 1-h interval between them, of either <95% for both limbs or with an absolute difference of >3% between the pre- and postductal readings. Measurements were not performed when the infant was crying or moving..."</i></p>		
Target Condition and reference standard(s)	<p>Target Condition: Critical Congenital Heart Disease, defined as HLHS, PA with intact ventricular septum, bicuspid atresia, as well as infants who are dying or requiring medical intervention within the first 28 days of life as a result of coarctation of the aorta, AV stenosis, PS, TOF, DORV, Epstein Anomaly, or PA with VSD</p> <p>Reference Standard:</p> <p>Positive screen: High quality echocardiography with interpretation by a clinician with expertise in the diagnosis of CHD</p> <p>Negative screen: No additional evaluation if clinically well and who show no signs of possible CHD</p>		
Flow and Timing	<p>Duration of follow up: no mention</p> <p>Lost to Follow up: 282 (analyzed n = 4236/4518) Reasons: 226 not approached for consent, 56 no consent</p>		
Item	Judgement	Risk of Bias	Applicability Concerns

Domain 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes	Low	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Domain 2: Index Test Pulse Oximetry			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes	Low	Low
Domain 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No	Unclear	Low
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Domain 4: Flow and Timing			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Unclear	Unclear	
Did all patients receive a reference standard?	No		
Were all patients included in the analysis?	Yes		

Hu, 2017 (prospective, PO ± auscultation)	
Methodological Quality	Multicenter prospective observational study
Patient Characteristics and Setting	Country: China Setting: 15 hospitals - 8 urban and 7 suburban hospitals, 5 tertiary and 10 secondary hospitals. Ten hospitals had echocardiography available on site Study Period: 13 hospitals (between July 1, 2012 and Dec 31, 2014), 2 hospitals between February 13, 2013 to December 31, 2014 Inclusion Criteria: consecutive asymptomatic newborn infants irrespective of gestational age Exclusion Criteria: newborn infants with prenatally diagnosed CHD N screened: 167,190 Gestational Age: all gestational ages

Index Tests	<p>*Clinical auscultation, pulse oximetry performed by same clinician (retrieved from Zhao, 2014)</p> <p>Pulse oximetry was performed with a new generation RAD-5v pulse oximeter (Masimo, Irvine, CA, USA) with a multisite reusable sensor (LNOP YI, Masimo)</p> <p>Screening protocol: Site of testing: pre ductal (right hand) and post ductal (either foot) Test timing: longer than 24 hours Oxygen saturation: functional Threshold: <95%</p> <p><i>"We used measurement criteria proposed by the US Secretary of Health and Human Services to implement screening. The clinician repeated pulse oximetry testing 4 h later if the first pulse oximeter oxygen saturation (SpO_2) measurement was between 90% and 95%. Screening was deemed positive if an SpO_2 of less than 95% was obtained both on the right hand and on either foot on two measures, separated by 4 h; a difference between the two extremities was more than 3% on two measures, separated by 4 h; or any measure was less than 90%."</i></p>		
Target Condition and reference standard(s)	<p>Main Target Condition: Critical Congenital Heart Disease (defects causing death or needing intervention before 28 days of age)</p> <p>Secondary: serious congenital heart disease (defects needing intervention before 6 months), significant (defects lasting longer than 6 months of age but not critical or serious) non significant (defects not physically appreciable and not persistent after 6 months of age)</p> <p>Reference standard:</p> <p>Positive screen: Echocardiography within 24 hours</p> <p>Negative screen: clinical follow up at 6 weeks of age, in combination with feedback from parents about cardiac symptoms like cyanosis, tachypnea and feeding difficulty, follow up at one year (for some patients)</p>		
Flow and Timing	<p>Duration of follow up: clinical follow up at 6 weeks of age</p> <p>Lost to follow up: none (n= 167, 190 screened)</p>		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	Unclear	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
<i>Domain 2: Index Test Pulse Oximetry</i>			

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	No	High	Low
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Yes	Unclear	
Did all patients receive a reference standard?	No		
Were all patients included in the analysis?	Yes		

Narayen, 2018 (prospective)	
Methodological Quality	Prospective
Patient Characteristics and Setting	<p>Country: Netherlands Setting: Multiregional, 75 regional community midwifery practices, 11 regional hospitals, 3 academic hospitals Study Period: July 2015 to December 2016</p> <p>Inclusion Criteria: Infants with a gestational age more than or equal to 35 weeks who were not admitted to the pediatric department, and with no clinical indication for pulse oximetry monitoring</p> <p>Exclusion Criteria: With prenatal diagnosis of CCHD and/or with symptoms directly after birth</p> <p>N screened: 23,959</p> <p>Gestational Age: more than or equal to 35 weeks</p> <p>(Infants antenatally diagnosed with CCHD were assessed in a secondary analysis.)</p>

Index Tests	<p>Pulse oximetry was performed by a nurse or midwife (who underwent 1 day training session) using a handheld oximeter with a reusable sensor (Nellcor PM10N) and disposable adhesive sensor wraps (Medtronic, Dublin Ireland)</p> <p>Screening protocol:</p> <p>Site of testing: pre-ductal (right hand) and post-ductal (foot) in no particular order</p> <p>Test timing: < 24 hours (1st HOL), > 24 hours (Day 2 or 3 of life). Timing was adapted to coincide with the regular home visits of community midwives after birth</p> <p>Oxygen saturation: functional</p> <p>Threshold: <95%</p> <p><i>"The first pulse oximetry screening after birth was considered positive if (1) the pre- or postductal SpO₂ reading was <90%; and (2) 2 independent measurements, with at least a 1-hour interval, revealed a SpO₂ <95% for both limbs or an absolute difference of >3% between the pre- and postductal readings. When the first SpO₂ screening was normal (SpO₂ ≥95% in either limb and <3% difference between both limbs), the pre- and postductal SpO₂ measurements were repeated on day 2 or 3 of life, either in the maternity ward or at home during the follow-up visit of the community midwife. This second SpO₂ screening was considered positive if SpO₂ <95% in both limbs or if a >3% difference between limbs was present"</i></p> <p><i>"Screening was considered positive after 2 instead of 3 abnormal readings"</i></p>		
Target Condition and reference standard(s)	<p>Target condition: all congenital heart defects that lead to death or require surgical or catheter intervention within the first 28 days of life, including hypoplastic left heart syndrome, pulmonary atresia with intact ventricular septum ,simple transposition of the great arteries, interrupted aortic arch, critical aortic or pulmonary valve stenosis, critial tetralogy of Fallot, or total anomalous venous return</p> <p>Reference standard:</p> <p>Positive screen without a non cardiac explanation: echocardiography and follow up</p> <p>Negative screen: follow up, records review of the Center for Congenital Heart Disease Amsterdam-Leiden, Mortality registries (including all surgical and catheter interventions in newborns and infants with congenital heart disease), examined until after 3 months of last included study participant</p>		
Flow and Timing	<p>Duration of follow up: not stated explicitly</p> <p>Lost to follow up: 20769 screened in first hour of life out of 23959 included infants, 14381 screened at day 2 or 3</p>		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	Unclear	Low
Was a case-control design avoided?	Yes		

Did the study avoid inappropriate exclusions?	No		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	Unclear	Low
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Unclear		
Did all patients receive a reference standard?	Yes	Unclear	
Were all patients included in the analysis?	Yes		

Nuntnarumit, 2017	
Methodological Quality	Prospective cohort
Patient Characteristics and Setting	<p>Country: Thailand</p> <p>Setting: across 2 hospitals, a university hospital and a regional hospital (with ratio of healthy newborns three to four times higher than the university hospital)</p> <p>Study Period: February 2010 to January 2011</p> <p>Inclusion Criteria: Healthy newborn infants with gestational age of more than or equal to 35 weeks</p> <p>Exclusion Criteria: sick newborns who required observation and monitoring, or those with congenital anomalies requiring initial echocardiography evaluation</p> <p>N screened: 10, 603</p> <p>Gestational Age: more than or equal to 35 weeks</p>

Index Tests	Pulse oximetry was performed using Radical Masimo (Masimo Radical 7) in infants in quiet alert or in sleeping state Screening protocol: Site of testing: right wrist (pre ductal) and either foot (post ductal) Test timing: > 24 hours until before discharge Oxygen saturation: functional Threshold: <95% "Definition of passed (negative) pulse oximetry screening result was preductal SpO2 more than or equal to 95% and difference between pre- and postductal SpO2 of less than or equal to 3%. Pulse oximetry screening would be considered failed (positive) POS if pre- or postductal SpO2 was less than 90% or if pre- and postductal SpO2 was 90 to 95% or the difference between pre- and postductal was more than 3%, after three times of repeated measurements with 1 h apart."		
Target Condition and reference standard(s)	Target Condition: Critical Congenital Heart Disease defined (CRIT.CHD) as CHDs requiring surgery or catheterization within 1 year of age Reference standard: Positive screen: echocardiography (also performed if CHD is suspected from PE) Negative screen: questionnaires approximately 1 month after discharge regarding general health or any problem related to CRIT.CHD		
Flow and Timing	Duration of follow up: 1 month Lost to follow up: 10603 infants met inclusion criteria out of 11407 live births, 804 newborns excluded (not meeting inclusion criteria) (78.3% at 1 month)		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	Low	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	No	High	Low
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		

Domain 4: Flow and Timing			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Unclear		
Did all patients receive a reference standard?	No	Unclear	
Were all patients included in the analysis?	Yes		

Paranka, 2018	
Methodological Quality	Prospective Cohort
Patient Characteristics and Setting	Country: Colorado, USA Setting: 34 sites in the neonatal ICU, 24 (70%) were located at or below 2000 feet, 5 (15%) were located between 4700 and 6000 feet and 5 (15%) were located above 6000 ft in altitude Study Period: November 2012 to February 2016 Inclusion Criteria: infants between 35 to 44 weeks Exclusion Criteria: infants on supplemental oxygen N screened: 6144 Gestational Age: 35 to 44 weeks
Index Tests	Pulse oximetry was performed with a handheld oximeter with a reusable probe Screening protocol: Site of testing: pre-ductal (right hand) and post-ductal (foot) Test timing: after 24 h and within 96 hours of birth Oxygen saturation: functional Threshold: <95% <i>"As per AAP Screening Algorithm, a screen result was considered positive if any oxygen saturation measure was <90%, oxygen saturation was <95% in both extremities on 3 measures, each separated by 1 h, or there was a >3% absolute difference in oxygen saturation between the right hand and foot on 3 measures, each separated by 1 h. Any screening that was ≥95% in either extremity with ≤3% absolute difference in oxygen saturation between the upper and lower extremity was considered a negative test."</i> <i>"Infants born above 4700 feet was screened using the AAP algorithm and then using an algorithm adjusted for barometric pressure if the neonate did not pass the screen...The calculated FiO2 (fraction of inspired oxygen) was administered to the neonate using an oxygen hood and the screen was repeated one time."</i> <i>"One high altitude site used an algorithm with a lower threshold saturation of ≥93% with a ≤3% difference as opposed to the AAP saturation of ≥95% with the ≤3% difference to determine which infants got an echocardiogram."</i>
Target Condition and reference standard(s)	Target Condition: Critical Congenital Heart Disease (not defined) Reference Standard: Positive screen: echocardiogram Negative screen: Follow up 1 and 2 months after birth (some patients got echocardiogram with significant physical exam)

Flow and Timing	Duration of follow up: 1 to 2 months after birth Lost to follow up: 6144 assigned a study code, 35 infants not screened		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	Low	Unclear
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	No	High	Unclear
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	No	High	
Did all patients receive a reference standard?	No		
Were all patients included in the analysis?	Yes		

Schwartz, 2021	
Methodological Quality	Retrospective Cohort
Patient Characteristics and Setting	Country: Maryland, USA
	Setting: Holy Cross Hospital, a large community hospital with nearly 10,000 deliveries a year
	Study Period: Sept 1, 2012 to September 1, 2020
	Inclusion Criteria: newborns admitted to the well infant nursery who received routine POS
	Exclusion Criteria: Newborns who did not receive POS due to parent refusal, missed screen, died before screen, transferred to another hospital before 24 hours, with prior antenatal diagnosis of CCHD from 2D Echo, with prior cardiac workup before POS due to development of symptoms
	N screened: 65403
	Gestational Age: all applicable gestational ages
Index Tests	Screening protocol: AAP/AHA Protocol
	Site of testing: right hand and either foot in a non specified order

	Test timing: >24 hours		
	Oxygen saturation: functional		
	Threshold: <95%		
Target Condition and reference standard(s)	Target Condition: Critical Congenital Heart Disease, defined as any of the POS core conditions, which includes hypoplastic left heart syndrome, pulmonary atresia, tetralogy of Fallot, total anomalous pulmonary venous connection, d-transposition of the great arteries, tricuspid atresia, truncus arteriosus, aortic coarctation, Ebstein anomaly, double outlet right ventricle, interrupted aortic arch, single ventricle (not otherwise specified), and other critical cyanotic lesions not otherwise specified		
	Reference Standard:		
	Positive screen: medical chart and records review of those who failed POS, those who needed a second rescreen to pass POS and those who had failed POS who required transfer to a bigger hospital (Children's National Hospital). All POS data are encoded in the OZ database.		
	Negative screen: review of database OZ charts of patients with abnormal echocardiography findings but passed POS, review of surveillance of 3 major cardiac referral centers in the area which includes mortalities from undetected CCHD, as well as review CCHD identified on birth and death certificates by the Birth Defects Reporting System		
Flow and Timing	Duration of follow up: not stated		
	Lost to Follow up: (examined 64780/65414 neonates delivered) excluded infants who had refused parental consent (n = 5), mortality <24 HOL (n=6), Infants were also not screened due to POS performed < 24HOL (physician override) (n=12), as well as missed screens (n = 611)		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	No	High	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear	Unclear	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	Yes	Unclear	Low

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Unclear		
Did all patients receive a reference standard?	Yes	Unclear	
Were all patients included in the analysis?	Yes		

APPENDIX C. Summary of Included Studies: Cochrane Review (Jan 2015 to March 2017)
[Retrieved from Appendices] ²⁰

Gomez-Rodriguez, 2015			
Methodological Quality	Cross sectional prospective study		
Patient Characteristics and Setting	<p>Country: Mexico Setting: Department of Neonatology, UMAE 48-Instituto Mexicano del Seguro Social (IMSS), León, Gto Study period: July 2010 to April 2011 Inclusion criteria: newborns > 6 hours of age in whom no CHD was suspected; only tested consecutive newborns who were available during the working hours of investigators Exclusion criteria: newborns with lung disease no informed consent N screened: 1037 Gestational age: mean (SD): 38.9 (1.1) weeks Prevalence of CCHD: 1.9 per 1000 live births</p>		
Index Tests	<p>Pulse oximetry was performed with a Rad-5 handheld pulse oximeter with multisite sensor Screening protocol: Site of testing: left lower extremity (post-ductal) Test timing: within 24 hours (mean age at pulse oximetry screening 12 hours - range 6 to 48 hours) Oxygen saturation: functional Threshold: < 95% Measurement was taking during 2minutes until the reading remained the same in 2 determinations</p>		
Target Condition and reference standard(s)	<p>Target condition: critical congenital heart disease (no definition included) Reference standard(s): Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: clinical records of all follow-up at 6 months</p>		
Flow and Timing	Duration of follow-up: 6 months Loss to follow-up: none (n analyzed: 1037)		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	No	Unclear	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	No	Unclear	Low

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Yes	Low	
Did all patients receive a reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Jones, 2015			
Methodological Quality	Retrospective observational study		
Patient Characteristics and Setting	Country: United Kingdom Setting: neonatal intensive care unit, Northwick Park Hospital, Harrow, Middlesex (level-2 neonatal unit without on-site access to pediatric echocardiography) Study period: September 1, 2011, to August 31, 2013 Inclusion criteria: all newborns admitted to the neonatal unit during the study period Exclusion criteria: Antenatal diagnosis of CCHD Admitted to neonatal intensive care unit after birth Live birth cohort, n = 11,233 (973 neonatal unit admissions) N screened: 10,260 Gestational age: not stated Prevalence of CCHD: 0.2 per 1000 live births		
Index Tests	Type of pulse oximeter not stated Screening protocol: Site of testing: both pre-ductal and post-ductal Test timing: within 24 hours Oxygen saturation: not stated Threshold: ≤ 95% (or pre-ductal and post-ductal difference > 3%)		
Target Condition and reference standard(s)	Target condition: critical congenital heart disease defined as CHD resulting in death or requiring surgical intervention or therapeutic catheterization within the first 28 days of life Reference standard(s): Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: National Congenital Heart Disease Audit		
Flow and Timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 10,260)		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	Low	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<i>Domain 2: Index Test Pulse Oximetry</i>			

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	No	Unclear	Low
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Yes	Unclear	
Did all patients receive a reference standard?	Unclear		
Were all patients included in the analysis?	Yes		

Klausner, 2017	
Methodological Quality	Retrospective observational study
Patient Characteristics and Setting	Country: USA Setting: 4 Yale-New Haven Health System hospitals in Connecticut Study period: January 1 and December 31, 2014 Inclusion criteria: all newborns delivered during the study period Exclusion criteria: Live-born infants who died before CCHD screening Antenatal screening Live birth cohort, n = 10,589 (171 [1.6%] underwent an echocardiogram before screening, and 98 [0.9%] were not screened; 96 were missed in error and parents refused in 2 instances) N screened: 10,320 Gestational age: 9584 (90.5%) were term (> 37 weeks) Prevalence of CCHD: 0 per 1000 live births
Index Tests	Type of pulse oximeter not stated Screening protocol: Site of testing: both pre-ductal and post-ductal Test timing: longer than 24 hours Oxygen saturation: not stated Threshold: < 95%
Target Condition and reference standard(s)	Target condition: critical congenital heart disease defined as structural defect associated with hypoxemia in the newborn period that requires surgical intervention before 1 year and, without intervention, can lead to significant morbidity and mortality Reference standard(s): Reference standard used for positive pulse oximetry results: echocardiogram Reference standard used for negative pulse oximetry results: follow-up
Flow and Timing	Of 10,316 infants with negative pulse oximetry at the time of birth, possible to review post discharge records of only 52.1% (n = 5367)

Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	Low	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	Unclear	Low
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Yes	Unclear	
Did all patients receive a reference standard?	Unclear		
Were all patients included in the analysis?	Yes		

Ozalkaya, 2016	
Methodological Quality	Retrospective Observational Study
Patient Characteristics and Setting	Country: Turkey Setting: Bursa Sevket Yilmaz Training and Research Hospital Study period: between January 2014 and December 2014 Inclusion criteria: asymptomatic newborns Exclusion criteria: Referred within first 24 hours of life or admitted to neonatal intensive care unit Perinatal CCHD Live birth cohort, n = 10,200 (excluded: hospitalized = 1100, referred = 890, perinatal CCHD = 2) N screened: 8208 Gestational age: not stated Prevalence of CCHD: 1 per 1000 live births

Index Tests	<p>Pulse oximetry was performed with a Nellcor pulse oximeter.</p> <p>Screening protocol: Site of testing: both pre-ductal and post-ductal Test timing: longer than 24 hours Oxygen saturation: functional Threshold: Screening test was considered positive in newborns whose saturation with pulse oximetry was less than or equal to 95% and/or who had a difference < 3% between right lower and right extremities</p>		
Target Condition and reference standard(s)	<p>Target condition: CCHD defined as congenital heart disease requiring catheter-based or surgical intervention within the first month of life, or causing high mortality and morbidity in the first weeks of life</p> <p>Reference standard(s): Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: echocardiography</p>		
Flow and Timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 8208)		
Item	Judgement	Risk of Bias	Applicability Concerns
<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	Low	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	Yes	Low	Low
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Yes	Low	
Did all patients receive a reference standard?	Unclear		
Were all patients included in the analysis?	Yes		

Methodological Quality				Prospective Observational Study			
Patient Characteristics and Setting		Country: South Africa Setting: Mowbray Maternity Hospital (MMH), a busy level-2 maternity hospital in the Western Cape Province, SA Study period: May 19 to September 19, 2014 Inclusion criteria: All neonates > 6 hours old with no clinical signs of cardiovascular disease were eligible Exclusion criteria: “unwell” infants, those < 6 hours old, those born to mothers < 14 years of age or unable to give informed verbal consent (owing to illness, illiteracy, or language barriers); all infants with a prenatal diagnosis of CHD or any signs of CHD, including a heart murmur ($\geq 3/6$) or significant dysmorphic features Livebirth cohort, n = 2256 (1220 mothers not approached) N screened: 1001 (44%) Prevalence of CCHD: 1 per 1000 live births					
Index Tests		Pulse oximetry was performed with Nellcor pulse oximeters. Screening protocol: Site of testing: right hand and any foot Test timing: longer than 24 hours Oxygen saturation: functional Threshold: < 95%					
Target Condition and reference standard(s)		Target condition: CCHD, which leads to death or needs surgical intervention before 28 days Reference standard(s): Reference standard used for positive pulse oximetry results: echocardiography Reference standard used for negative pulse oximetry results: not stated					
Flow and Timing		Duration of follow-up: no physical follow-up Loss to follow-up: none (n analyzed: 1001)					
Item		Judgement		Risk of Bias		Applicability Concerns	
<i>Domain 1: Patient Selection</i>							
Was a consecutive or random sample of patients enrolled?	No	High	Low	High	Low	High	Low
Was a case-control design avoided?	Yes						
Did the study avoid inappropriate exclusions?	Unclear						
<i>Domain 2: Index Test Pulse Oximetry</i>							
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low	Low	Low	Low	Low
If a threshold was used, was it pre-specified?	Yes						
<i>Domain 3: Reference Standard</i>							
Is the reference standards likely to correctly classify the target condition?	No	High	High	High	High	High	High

Were the reference standard results interpreted without knowledge of the results of the index tests?	unclear		
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Yes	Unclear	
Did all patients receive a reference standard?	No		
Were all patients included in the analysis?	No		

Zuppa, 2015			
Methodological Quality	Prospective		
Patient Characteristics and Setting	Country: Italy Setting: Agostino Gemelli General Hospital Study period: 2 years (from 2009 to 2010) Inclusion criteria: all newborns admitted to the nursery. These newborns by definition were considered healthy or were under observation for maternal disease, mild prematurity, or low birth weight Exclusion criteria: Newborns with syndrome Total number of newborn infants included: N screened: 5750 Prevalence of CCHD: 0.2 per 1000 births		
Index Tests	Pulse oximetry was performed with an Ohmeda 3900 pulse oximeter Screening protocol: Site of testing: post-ductal (foot) Test timing: longer than 24 hours Oxygen saturation: functional Threshold: < 95% <i>"The measurement was performed by a professional nurse in all newborns admitted to the nursery, between the 48th and 72nd hours of life, before discharge. The probe detector was placed on one of the two legs, making sure that the newborn was quiet and with warm ends. The measurement was performed in presence of stable, continuous and free of artefacts pulse wave, for at least 3 minutes. In case of positive screening, a second check was carried out by medical staff after 15 to 30 min."</i>		
Target Condition and reference standard(s)	Target condition: CCHD defined as severe cardiac alterations that require cardiac surgery during the first year of life Reference standard(s): Reference standard used for positive pulse oximetry results: electrocardiographic and echocardiography (echocardiograph "HP Sonos 4500, Agilent Technologies" [Andover,MA], a multifrequency probe [5 to 12 MHz], suitable for study of the neonatal heart). Evaluation was performed by 2- dimensional analysis [2-D], analysis of M-mode, and Doppler ultrasound) Reference standard used for negative pulse oximetry results: not stated		
Flow and Timing	Duration of follow-up: not stated Loss to follow-up: none (n analyzed: 5751)		
Item	Judgement	Risk of Bias	Applicability Concerns

<i>Domain 1: Patient Selection</i>			
Was a consecutive or random sample of patients enrolled?	Yes	Low	Low
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
<i>Domain 2: Index Test Pulse Oximetry</i>			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Low	Low
If a threshold was used, was it pre-specified?	Yes		
<i>Domain 3: Reference Standard</i>			
Is the reference standards likely to correctly classify the target condition?	Unclear	High	Low
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
<i>Domain 4: Flow and Timing</i>			
Were there at least 28 days of appropriate follow up between index test and reference standard?	Yes	Unclear	
Did all patients receive a reference standard?	Unclear		
Were all patients included in the analysis?	Unclear		

APPENDIX D. Included Studies Bias Assessment Table

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Almawazini 2017	?	?	+	?	+	+	+
Cloete 2019	-	?	?	-	+	+	+
Diller 2018	-	?	?	?	+	+	+
Gomez-Rodriguez 2015	?	+	?	+	+	+	+
Gopalakrishnan, 2021	-	+	?	-	+	+	+
Hamilcikan 2017	+	+	?	?	+	+	+
Hu 2017	?	+	-	?	+	+	+
Jones 2016	+	+	?	?	+	+	+
Klausner 2017	+	+	?	?	+	+	+
Narayen 2018	?	+	?	?	+	+	+
Nuntnarumit 2017	+	+	-	?	+	+	+
Ozalkaya 2015	+	+	+	+	+	+	+
Paranka 2018	+	+	-	-	?	+	?
Schwartz, 2021	-	?	?	?	+	+	+
Slitine 2020	-	+	?	-	+	+	+
Van Niekerk 2016	-	+	-	?	+	+	+
Zuppa 2015	+	+	-	?	+	+	+

- High	? Unclear	+ Low
---	---	--

APPENDIX E. GRADE Evidence Profile: Pulse Oximetry Screening on Mortality

Pulse Oximetry Screening compared to No Screening for CCHD in Asymptomatic Newborns

Bibliography: Abouk 2017, Banait 2018

Certainty assessment							Summary of findings				
Participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With No Screening	With Pulse Oximetry Screening		Risk with No Screenin g	Risk difference with Pulse Oximetry Screenin g
Mortality (follow-up: 12 months)											
18435321 (2 observational studies)	serious ^a	not serious	not serious	not serious	none	⊕⊕⊕○ Moderate	1911/15531561 (0.0%)	281/2903760 (0.0%)	not estimable	1 per 10,000	

CI: confidence interval

Explanations

a. Limited information on the baseline characteristics of the populations studied. With risk for selection bias in allocation between screened and unscreened cohorts. Abouk: only age was provided; CCHD prevalence may vary between states

APPENDIX F. GRADE Evidence Profile: Pulse Oximetry Screening for Critical Congenital Heart Disease in Asymptomatic Newborns

Question: Should Pulse Oximetry Screening be used to screen for Critical Congenital Heart Disease in Asymptomatic Newborns?

Sensitivity	0.71 (95% CI: 0.53 to 0.85)	Prevalences	0.044% ⁽⁷⁾	
Specificity	1.00 (95% CI: 1.00 to 1.00)			

Outcome	No of studies (No of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 10,000 patients tested	Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias		
True positives (patients with Critical Congenital Heart Disease)	17 studies 418219 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious ^b	serious ^c	none ^d	3 (2 to 4)	⊕⊕○○ Low
								1 (0 to 2)	
False negatives (patients incorrectly classified as not having Critical Congenital Heart Disease)	17 studies 418219 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious ^b	not serious	none ^d	9996 (9996 to 9996)	⊕⊕⊕○ Moderate
								0 (0 to 0)	

Explanations

a. With unclear issues on conduct of reference standard (issues on partial verification and differential verification bias) and flow, timing.

- b. high heterogeneity between all studies but noted improved sensitivity with subgroup analyses
- c. there was a high confidence interval noted across studies
- d. Publication bias cannot be excluded but deemed to not be sufficient to downgrade evidence quality

APPENDIX G. Characteristics of Included Studies

Study	Population	Antenatal Screening	Index Test			Reference Standard	
			Limb	Test timing	Altitude	Positive	Negative
Almawazini	preterm	excluded	pre and post	> 24h	no	2D Echo	no mention
Cloete et al	preterm	excluded	post	<24h	no	pediatric referral	review of cardiac center databases
Diller et al	term	excluded	pre and post	>24h	no	birth records regarding 2D echo results	records review
Gomez Rodriguez et al	term	excluded	post	<24h	no	2D Echo	clinical follow up
Gopalakrishnan et al	preterm	excluded	pre and post	>24h	no	2D Echo	clinical follow up, record review
Hamilcikan et al	preterm	included	pre and post	<24h	no	2D Echo	none
Hu et al	preterm	excluded	pre and post	>24h	no	2D Echo	clinical follow up
Jones	preterm	excluded	pre and post	<24h	no	2D Echo	National Congenital Heart Disease Audit
Klausner et al	preterm	excluded	pre and post	>24h	no	2D Echo	clinical follow up
Narayen et al	preterm	excluded	pre and post	>24h	no	2D Echo and follow up	clinical follow up, records review, mortality registry
Nuntnarumit et al	preterm	included	pre and post	>24h	no	2D Echo	follow up using questionnaires
Ozalkaya et al	term	excluded	pre and post	>24h	no	2D Echo	2D Echo
Paranka et al	preterm	included	pre and post	>24h	yes	2D Echo	clinical follow up
Schwartz et al	preterm	excluded	pre and post	>24h	no	2D Echo	clinical follow up, records review, birth and death reports
Slitine et al	preterm	excluded	pre and post	>24h	no	2D Echo	no mention
Van Niekerk et al	term	excluded	pre and post	>24h	no	2D Echo	no mention
Zuppa et al	preterm	excluded	post	>24h	no	2D Echo	no mention

APPENDIX H. Subgroup Analysis

	Total no. of Studies	Sensitivity (95% CI)	I²	Specificity (95% CI)	I²
<i>Population</i>					
Included Preterm Newborns	13	77% (64-86)	58.51	100%	99.5
Excluded Preterm Newborns (Term Only)	4	58% (17-90)	72.34	100%	98.17
<i>Antenatal Diagnosis</i>					
Included	3	84% (73-94) ^a	0	100%	0
Excluded	14	72% (49-87)	57.52	100%	99.5
<i>Test Timing</i>					
Less than 24H (<24H)	4				
Longer than 24H (>24H)	13	68% (43-72)	70	100%	99.7
<i>Index Test Limb</i>					
Foot Only (Post Ductal)	3	cannot be computed	cannot be computed	100%	0
Foot and Right Hand (Pre and Post Ductal)	14	72% (53-85)	64.37	100%	99.7
<i>Altitude</i>					
Screening Included High Altitudes	1				
Screening Excluded High Altitudes	16	75% (55-88)	56.89	100%	99.6

2. Newborn Screening for Cystic Fibrosis

APPENDIX A. Search Yield

1. Pre-Appraised Clinical Practice Guideline Search

Organization	Guidelines	Year
National Institute for Health and Care Excellence (NICE)	Cystic Fibrosis Diagnosis and Management NICE Guideline NG78	2017
UK National Screening Committee (UK NSC)	UK NSC Screening Recommendation	2017
US Preventive Services Task Force (USPSTF)	Newborn Screening for Cystic Fibrosis	2003
Canadian Task Force for Preventive Health Care (CTFPHC)	Screening for Cystic Fibrosis	1991

2. Journal Search (Cochrane, PubMed, MEDLINE, Google Scholar, JSTOR, HERDIN)

#	Query	Results
1	"Cystic Fibrosis" [MeSH]	152,792
2	"Newborn Screening" OR "NBS"	39,331
3	"Immunoreactive Trypsinogen" OR "IRT"	7,589
4	"Sweat Test"	9,176
5	"Genetic Testing" OR "DNA Analysis"	128,084
6	"Asymptomatic" AND ("Newborn" OR "Neonate")	28,341
7	"Mortality" OR "Death"	3,637,322
8	"Morbidity" OR "Complications"	5,170,959
9	#1 AND #2	2,281
10	#1 AND #3	1,168
11	#1 AND #4	10,810
12	#1 AND #5	12,603
13	#9 OR #10 OR #11 OR #12	19,513
14	#13 AND #6	1,158
15	#14 AND #7	3,729
16	#14 AND #8	2,251
17	#15 OR #16	74

3. UK NSC Strategy Search

a. Cochrane

#	Query	Results
1	MeSH descriptor: [Cystic Fibrosis] this term only	6,252
2	(cystic next fibrosis):ti,ab,kw	5,747
3	(cf or mucoviscidosis):ti,ab	5,231
4	((fibrocystic or fibrosis) near/3 pancreas):ti,ab	12
5	(cystic and pancre*):ti,ab	544
6	#1 or #2 or #3 or #4 or #5	8,504
7	MeSH descriptor: [Neonatal Screening] this term only	1,611
8	((neonat* or newborn*) near/5 screen*):ti,ab	517
9	MeSH descriptor: [Mass Screening] this term only	9,266
10	MeSH descriptor: [Infant, Newborn] explode all trees	21,122
11	#9 and #10	211
12	#7 or #8 or #11	1,908
13	#6 and #12 Publication Year from 2017 to 2021	2

b. Journal Databases (PubMed, MEDLINE, Google Scholar, JSTOR, HERDIN)

#	Query	Results
1	'newborn screening'/de	1,210,193
2	((neonat* OR newborn*) NEAR/2 screen*):ab,ti	478,571
3	'mass screening'/de	362,858
4	'newborn'/de	3,184,434
5	#3 AND #4	27,519
6	#1 OR #2 OR #5	445,922
7	ceas*:ab,ti OR cessation:ab,ti OR stop:ab,ti OR stopped:ab,ti OR continu*:ab,ti OR discontinu*:ab,ti	9,638,055
8	appropriate*:ab,ti OR inappropriate*:ab,ti OR unnecessary:ab,ti OR question*:ab,ti	11,017,747
9	harmful:ab,ti OR harm*:ab,ti OR adverse:ab,ti	5,933,551
10	benefit*:ab,ti AND (risk*:ab,ti OR harm*:ab,ti)	5,023,766
11	'side effect'/exp	2,734,033
12	(side NEAR/1 effect*):ab,ti	8,574,361
13	overdiagnosis:ab,ti OR 'over diagnosis':ab,ti	1,462,098
14	'patient safety'/exp	4,377,516
15	'risk assessment'/de	3,297,739
16	'risk benefit analysis'/exp	322,026
17	#7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16	11,095,385
18	#6 AND #17	1,774,517
19	'cystic fibrosis'/de	1,388,449
20	'cystic fibrosis':ab,ti	1,331,946
21	cf:ab,ti OR mucoviscidosis:ab,ti 5	1,370,137
22	((fibrocystic OR fibrosis) NEAR/3 pancreas):ab,ti	330,309
23	cystic:ab,ti AND pancre*:ab,ti	294,140
24	#19 OR #20 OR #21 OR #22 OR #23	1,787,209
25	#18 AND #24	78,745
26	#18 AND #24 AND [2017-2021]/py	759

APPENDIX B. Characteristics of Included Studies

Study ID	Patients (n)	Interventions	Outcomes	Method
Tridello 2017	Patients with CF in Verona, Italy (n = 485)	Cystic Fibrosis Newborn Screening vs. Clinical Diagnosis	30-year Survival, Overall Survival, Cumulative Incidence of CF-Related Death	Multi-Center Retrospective Cohort Study
Schlüter 2020	Patients with CF in the UK (n = 2216)	Cystic Fibrosis Newborn Screening vs. Clinical Diagnosis	Forced Vital Capacity	Single-Center Retrospective Cohort Study
van der Ploeg 2015	Asymptomatic Newborns in the Netherlands (n = 145,499)	IRT, PAP	Sensitivity, Specificity	Multi-Center Cohort Study
Palomaki 2003	Non-Hispanic Caucasians in the USA (n = 2,198)	DNA Testing	Sensitivity, Specificity	Multi-Center Cohort Study
Beauchamp 2019	At-Risk Couples in the USA (n = 115,571)	CFTR Gene Sequencing	Sensitivity, Specificity	Single-Center Cohort Study
Vernooij-van Langen 2012	Asymptomatic Newborns in the Netherlands (n = 372,713)	IRT-PAP, IRT-DNA-seq, IRT-PAP-DNA-seq	Sensitivity, Specificity	Multi-Center Cohort Study
Sadik 2019	Asymptomatic Newborns in Eastern Andalusia, Spain (n = 68,502)	IRT-IRT, IRT-PAP-IRT	Sensitivity, Specificity	Multi-Center Cohort Study
Massie 2012	Asymptomatic Newborns in Victoria, Australia (n = 139,695)	IRT-DNA	Sensitivity, Specificity	Single-Center Cohort Study

APPENDIX C. Summary of Findings Table of Cystic Fibrosis Newborn Screening

Question: For asymptomatic newborns, should cystic fibrosis screening be done?

Patient or Population: Asymptomatic Newborns

Setting: Outpatient Setting

Intervention: Cystic Fibrosis Newborn Screening

Comparison: Clinical Diagnosis

Benefits and Harms

№ of Studies	Study Design	Risk of Bias	Certainty Assessment			Other Considerations	№ of Patients		Effect		Certainty
			Inconsistency	Indirectness	Imprecision		NBS	Clinical Diagnosis	Relative (95% CI)	Absolute (95% CI)	
30-year Survival											
1	observational study	serious ^a	serious ^b	not serious	serious ^{b,c}	none	342 participants	143 participants	Rate Ratio 1.38 (1.10 to 1.74)	--	⊕○○○ VERY LOW
Overall Survival (Severe Symptoms)											
1	observational study	serious ^a	serious ^b	not serious	serious ^{b,c}	none	95 participants	34 participants	HR 2.23 (1.23 to 4.03)	276 more per 1,000 (from 65 fewer to 474 more)	⊕○○○ VERY LOW
Overall Survival (Moderate Symptoms)											
1	observational study	serious ^a	serious ^b	not serious	serious ^{b,c}	none	80 participants	40 participants	HR 3.86 (1.21 to 12.38)	290 more per 1,000 (from 67 fewer to 300 more)	⊕○○○ VERY LOW
Overall Survival (Weak Symptoms)											
1	observational study	serious ^a	serious ^b	not serious	serious ^{b,c}	none	149 participants	58 participants	HR 0.99 (0.44 to 4.15)	88 more per 1,000 (from 293 fewer to 205 more)	⊕○○○ VERY LOW

Nº of Studies	Study Design	Certainty Assessment					Other Considerations	NBS	Clinical Diagnosis	Relative (95% CI)	Effect Absolute (95% CI)	Certainty
		Risk of Bias	Inconsistency	Indirectness	Imprecision							
Cumulative Incidence of CF-Related Death (Severe Symptoms)												
1	observational study	serious ^a	serious ^b	not serious	serious ^{b,c}	none	95 participants	34 participants	Rate Ratio 0.46 (0.20 to 1.061)	--	⊕○○○	VERY LOW
Cumulative Incidence of CF-Related Death (Moderate Symptoms)												
1	observational study	serious ^a	serious ^b	not serious	serious ^{b,c,d}	none	80 participants	40 participants	Rate Ratio 0.15 (0.04 to 0.55)	--	⊕○○○	VERY LOW
Cumulative Incidence of CF-Related Death (Weak Symptoms)												
1	observational study	serious ^a	serious ^b	not serious	serious ^{b,c,d}	none	149 participants	58 participants	Rate Ratio 0.17 (0.05 to 0.56)	--	⊕○○○	VERY LOW
FVC (5-6 years old)												
1	observational study	not serious	serious ^b	not serious	serious ^{b,d}	none	1444 participants	772 participants	MD 1.29 higher (-0.12 to 2.70)	(0.16 lower to 2.74 higher)	⊕○○○	VERY LOW

CI: Confidence interval; HR: Hazard Ratio; MD: Mean Difference

Explanations

- a. Some subjects were excluded from the final analysis.
- b. There were no other studies for comparison.
- c. There were less than 30 events recorded per outcome.
- d. Confidence intervals crossed the threshold.

Diagnostic Performance

Outcome	No of Studies (No of Patients)	Study Design	Certainty Assessment					Publication bias	Effect per 10,000 Patients Tested	Test Accuracy CoE
			Risk of Bias	Indirectness	Inconsistency	Imprecision				
Immunoreactive Trypsinogen Testing										
True Positives (patients with CF)									2	
False Negatives (patients incorrectly classified as not having CF)	1 (145,499)	cross-sectional (cohort type accuracy study)	serious ^a	not serious	serious ^b	serious ^b	none		(2 to 2) 0 (0 to 0)	⊕○○○ VERY LOW
True Negatives (patients without CF)									9897 (9892 to 9902)	
False Positives (patients incorrectly classified as having CF)	1 (145,499)	cross-sectional (cohort type accuracy study)	not serious	not serious	serious ^b	serious ^b	none		101 (96 to 106)	⊕⊕○○ LOW
Pancreatitis-Associated Protein Testing										
True Positives (patients with CF)									2	
False Negatives (patients incorrectly classified as not having CF)	1 (145,499)	cross-sectional (cohort type accuracy study)	serious ^a	not serious	serious ^b	serious ^b	none		(2 to 2) 0 (0 to 0)	⊕○○○ VERY LOW
True Negatives (patients without CF)									8997 (8830 to 9144)	
False Positives (patients incorrectly classified as having CF)	1 (145,499)	cross-sectional (cohort type accuracy study)	not serious	not serious	serious ^b	serious ^b	none		1001 (854 to 1168)	⊕⊕○○ LOW
DNA Testing										
True Positives (patients with CF)									2	
False Negatives (patients incorrectly classified as not having CF)	1 (2,198)	cross-sectional (cohort type accuracy study)	not serious	serious ^c	serious ^b	serious ^b	none		(2 to 2) 1 (1 to 1)	⊕○○○ VERY LOW
True Negatives (patients without CF)									9938 (9868 to 9988)	
False Positives (patients incorrectly classified as having CF)	1 (2,198)	cross-sectional (cohort type accuracy study)	not serious	serious ^c	serious ^b	serious ^b	none		60 (10 to 130)	⊕○○○ VERY LOW

Outcome	No of Studies (No of Patients)	Study Design	Risk of Bias	Indirectness	Inconsistency	Imprecision	Publication bias	Effect per 10,000 Patients Tested	Test Accuracy CoE
CFTR Gene Sequencing									
True Positives (patients with CF)									
False Negatives (patients incorrectly classified as not having CF)	1 (115,571)	cross-sectional (cohort type accuracy study)	not serious	serious ^c	serious ^b	serious ^b	none	3 (2 to 3) (-1 to 1)	⊕○○○ VERY LOW
True Negatives (patients without CF)	1 (115,571)	cross-sectional (cohort type accuracy study)	not serious	serious ^c	serious ^b	serious ^b	none	9998 (9718 to 9998) (-1 to 280)	⊕○○○ VERY LOW
IRT-IRT									
True Positives (patients with CF)									
False Negatives (patients incorrectly classified as not having CF)	1 (68,502)	cross-sectional (cohort type accuracy study)	not serious	not serious	serious ^b	serious ^b	none	2 (2 to 2) 0 (0 to 0)	⊕⊕○○ LOW
True Negatives (patients without CF)	1 (68,502)	cross-sectional (cohort type accuracy study)	not serious	not serious	serious ^b	serious ^b	none	9988 (9978 to 9988) 10 (10 to 20)	⊕⊕○○ LOW
IRT-PAP									
True Positives (patients with CF)									
False Negatives (patients incorrectly classified as not having CF)	1 (372,713)	cross-sectional (cohort type accuracy study)	serious ^d	not serious	serious ^b	serious ^b	none	2 (1 to 2) 0 (0 to 1)	⊕○○○ VERY LOW
True Negatives (patients without CF)	1 (372,713)	cross-sectional (cohort type accuracy study)	serious ^d	not serious	serious ^b	serious ^b	none	9988 (9986 to 9989) 10 (9 to 12)	⊕○○○ VERY LOW

Outcome	No of Studies (No of Patients)	Study Design	Certainty Assessment					Effect per 10,000 Patients Tested	Test Accuracy CoE
			Risk of Bias	Indirectness	Inconsistency	Imprecision	Publication bias		
IRT-PAP-IRT									
True Positives (patients with CF)	1 (68,502)	cross-sectional (cohort type accuracy study)	not serious	not serious	serious ^b	serious ^b	none	2 (2 to 2)	⊕⊕○○ LOW
False Negatives (patients incorrectly classified as not having CF)								0 (0 to 0)	
True Negatives (patients without CF)	1 (68,502)	cross-sectional (cohort type accuracy study)	not serious	not serious	serious ^b	serious ^b	none	9998 (9988 to 9998)	⊕⊕○○ LOW
False Positives (patients incorrectly classified as having CF)								0 (0 to 10)	
IRT-DNA									
True Positives (patients with CF)	1 (139,695)	cross-sectional (cohort type accuracy study)	not serious	not serious	serious ^b	serious ^b	none	2 (2 to 2)	⊕⊕○○ LOW
False Negatives (patients incorrectly classified as not having CF)								0 (0 to 0)	
True Negatives (patients without CF)	1 (139,695)	cross-sectional (cohort type accuracy study)	not serious	not serious	serious ^b	serious ^b	none	9988 (9938 to 9998)	⊕⊕○○ LOW
False Positives (patients incorrectly classified as having CF)								10 (0 to 60)	
IRT-DNA-seq									
True Positives (patients with CF)	1 (372,713)	cross-sectional (cohort type accuracy study)	serious ^d	not serious	serious ^b	serious ^b	none	2 (1 to 2)	⊕○○○ VERY LOW
False Negatives (patients incorrectly classified as not having CF)								0 (0 to 1)	
True Negatives (patients without CF)	1 (372,713)	cross-sectional (cohort type accuracy study)	serious ^d	not serious	serious ^b	serious ^b	none	9997 (9996 to 9997)	⊕○○○ VERY LOW
False Positives (patients incorrectly classified as having CF)								1 (1 to 2)	

Outcome	No of Studies (No of Patients)	Study Design	Certainty Assessment					Effect per 10,000 Patients Tested	Test Accuracy CoE
			Risk of Bias	Indirectness	Inconsistency	Imprecision	Publication bias		
IRT-PAP-DNA-seq			Sensitivity 95%, 95% CI (73.10–99.70%); Specificity 99.99%, 95% CI 99.99–100%; Pre-Test Probability of 0.016%						
True Positives (patients with CF)								2 (1 to 2)	
False Negatives (patients incorrectly classified as not having CF)	1 (372,713)	cross-sectional (cohort type accuracy study)	serious ^d	not serious	serious ^b	serious ^b	none	0 (0 to 1)	⊕○○○ VERY LOW
True Negatives (patients without CF)								9997 (9997 to 9998)	
False Positives (patients incorrectly classified as having CF)	1 (372,713)	cross-sectional (cohort type accuracy study)	serious ^d	not serious	serious ^b	serious ^b	none	1 (0 to 1)	⊕○○○ VERY LOW

CI: Confidence interval

Explanations

- a. Data from other studies were included in the analysis for sensitivity.
- b. There were no other studies for comparison.
- c. The study screened for carriers instead of CF(+) patients.
- d. Some subjects were excluded from the final analysis.

APPENDIX D. Cost-Effectiveness Study

Author	Year	Country	Population	Intervention	Control	Cost-Effective? (Y/N)
YES						
van der Ploeg	2015	Netherlands	Asymptomatic Newborns	<ul style="list-style-type: none"> • IRT-PAP • IRT-DNA • IRT-DNA-seq • IRT-PAP-DNA-seq 	No Screening	The different CF NBS strategies have cost-effectiveness ratios varying from €23,600 to €29,200 (₽1,391,875 to ₽1,722,151) per life-year gained making them economically justifiable in a country where €40,000 (₽2,359,502) per quality-adjusted life-year is acceptable.

APPENDIX E. Recommendations from Other Groups

Guideline	Population	Recommendations
UK National Screening Committee (UK NSC, 2017)	Asymptomatic Newborns	<ul style="list-style-type: none"> • It is recommended that CF NBS be done.
European Cystic Fibrosis Society (ECFS, 2018)	Asymptomatic Newborns	<ul style="list-style-type: none"> • There is clear evidence to support NBS for CF. • If the incidence of CF is <1/7000 births, careful evaluation is required as to whether NBS is valid. • Other factors in making the decision on whether to implement screening should include available healthcare resources and the ability to provide a clear pathway to treatment. • Infants identified with CF through a NBS program should have prompt access to specialist CF care that achieves ECFS standards.

3. Newborn Screening for Sickle Cell Disease

APPENDIX A. Search Yield

DATABASE	SEARCH STRATEGY / SEARCH TERMS	DATE AND TIME OF SEARCH	RESULTS	
			Yield	Eligible Articles
NICE (National Institute for Health and Care Excellence)	Sickle cell disease screening in newborn OR neonate	July 8, 2021	228	3 UK NSC Recommendation (1) HTA/Evidence Assessment (1) SR (1)
U.S. Preventive Services Task Force (USPSTF)	Sickle cell disease, hemoglobinopathies, screening	July 8, 2021	33	0
Canadian Task Force for Preventive Health Care (CTFPHC)	Sickle cell disease, hemoglobinopathies, screening	July 8, 2021	4	0
Guideline International Network	Sickle cell disease	July 8, 2021	4	0
Cochrane Library	sickle cell disease in Title Abstract Keyword OR hemoglobinopathies in Title Abstract Keyword AND newborn screening OR neonatal screening in Title Abstract Keyword - with Cochrane Library publication date Between Jan 2015 and Jul 2021, in Cochrane Reviews, Cochrane Protocols, Trials (Word variations have been searched)	July 8, 2021	48	0
	Adopted from UK NSC #1 MeSH descriptor: [Anemia, Sickle Cell] explode all trees 797 #2 MeSH descriptor: [Hemoglobin, Sickle] this term only 24 #3 ("sickle cell" or scd or "sc disease*"):ti,ab,kw 2687 #4 ('sickle anaemia' or 'sickle anemia'):ti,ab,kw 1565 #5 ("hbs disease" or "hb ss disease*"):ab,ti 7 #6 ("hemoglobin s" or "haemoglobin s*"):ab,ti 54 #7 #1 or #2 or #3 or #4 or #5 or #6 2699 #8 MeSH descriptor: [Neonatal Screening] this term only 134 #9 ((neonat* or newborn*) near/5 screen*):ti,ab 520	August 17, 2021	14	1 (duplicate of SR from NICE)

	#10 MeSH descriptor: [Mass Screening] this term only 3240 #11 MeSH descriptor: [Infant, Newborn] explode all trees 16573 #12 #10 and #11 65 #13 #8 or #9 or #12 635 #14 #7 and #13 14			
PubMed	("guideline"[Publication Type] OR "practice guideline"[Publication Type] OR "recommendation*"[Title] OR "standard*"[Title] OR "standard*"[Title] OR "guideline*"[Title]) AND (((("sickle"[All Fields] OR "sickled"[All Fields] OR "sickles"[All Fields] OR "sickling"[All Fields]) AND ("cells"[MeSH Terms] OR "cells"[All Fields] OR "cell"[All Fields])) OR "anemia, sickle cell"[MeSH Terms]) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields] OR ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields])) ((sickle cell) OR (sickle cell disease[MeSH Terms])) AND (newborn screening OR neonatal screening) Filters: Meta-Analysis, Systematic Review, from 2015/1/1 - 2021/7/8	July 8, 2021	3 1	1 1 (duplicate of SR in NICE)
Total			321	4
Disease Burden/Epidemiology				
PubMed	((sickle cell disease OR sickle cell) AND (epidemiology in asia OR burden of disease in asia)) AND (((((meta-analysis[ti] OR systematic review[ti] OR systematic literature review[ti] OR systematic scoping review[ti] OR systematic narrative review[ti] OR systematic qualitative review[ti] OR systematic evidence review[ti] OR systematic quantitative review[ti] OR systematic meta-review[ti] OR systematic critical review[ti] OR systematic mixed studies review[ti] OR systematic mapping review[ti] OR systematic cochrane review[ti] OR systematic search and review[ti] OR		3	2

	systematic integrative review[ti]) NOT comment[pt] NOT (protocol[ti] OR protocols[ti])) NOT MEDLINE[subset]) OR (Cochrane Database Syst Rev[ta] AND review[pt]) OR systematic review[pt])) AND ((meta-analysis[Filter] OR systematicreview[Filter])) AND (2015/1/1:2021/7/8[pdat])))			
Cost-effectiveness				
PubMed	((sickle cell disease OR sickle cell) AND (Newborn screening OR neonatal screening)) AND (cost effectiveness OR cost effective) Filters: from 2015/1/1 - 2021/7/15		14	4

APPENDIX B. Characteristics of Included Studies

Table B.1. Characteristics of Included Guidelines (1)

Study Author/Year /Area	Search Strategy	Study Selection/ PICO	Quality Assessment/ Data Extraction/ Data Analysis and Synthesis	Included Studies	Recommendation
UK NSC 2015	Page 6	P – newborns I – neonatal screening C – O – harms and benefits from screening, harms from screening or screening programme cessation	<u>Quality assessment</u> <u>Tools:</u> Not mentioned <u>Assessment by 2 reviewers:</u> first pass sift by information specialist and second pass sift by health research analyst <u>Data extraction</u> <u>Used of standardized form:</u> Not mentioned <u>Data extraction by 2 reviewers:</u> Not mentioned <u>Data analysis and synthesis:</u> Not mentioned	Balance of Harms and Benefits from Screening: No studies identified Screening program cessation: No studies identified Harms of screening (p. 5) Brosco 2006 – historical overview of universal NBS in US, review of published literature on whether universal screening has led to substantial morbidity and mortality from misguided medical treatment of false positives; comprehensive search was reported but other methods not reported	<p><u>Is there evidence to alter the current UKNSC recommendation to offer a national screening programme for sickle cell disease to newborn babies?</u></p> <p>The single relevant study identified suggested that the SCD screening programme in the US has not led to widespread harm. It does not provide sufficient evidence to suggest that the evidence supporting the national SCD screening programme needs to be reviewed in more depth or stopped.</p>

Table B.2. Characteristics of Included HTA/Systematic Review (1 HTA, 1 SR)

Study Author/Year/ Area	Search Strategy	Study Selection/ PICO	Quality Assessment/ Data Extraction/ Data Analysis and Synthesis	Included Studies	AMSTAR 2
Institute of Health Economics STE Report 2016 Alberta Canada	Appendix S.A Page 52-55 Appendix S.A.2 Page 55-57 Section 3.1 *Methodology Appendix T.A Page 118	Inclusion Criteria Page 138 Stage 1 Systematic reviews and HTAs Stage 2 Primary studies Exclusion Criteria Studies were excluded if they met any of the following criteria: <ul style="list-style-type: none"> •reviews that do not meet the criteria for systematic reviews; •conference abstracts, letters, news, and editorial comments, case reports; •tests that are not performed on the dried blood spots; •target conditions other than the seven core conditions defined for this project; •studies that covered a broad range of conditions but outcomes 	<u>Quality Assessment</u> Page 139-140 SR/HTA – AMSTAR Primary Studies for screening accuracy – QUADAS2 Treatment effectiveness studies – EPHPP <u>Data Extraction</u> One researcher extracted data according to data extraction forms <u>Data Analysis and Synthesis</u> Narrative and tabular summary Limitations in applicability and relevance of study results to Alberta context were identified and discussed <u>External Review</u> Draft report was reviewed	<u>See Appendix T.D</u> Page 169-178 <u>SR/HTAs (4 of which 2 are Cochrane reviews)</u> Study 51 – medium quality Study 48, 49, 50 – low quality Study 51 (2002/2012) Cochrane Penicillin prophylaxis for children with SCD Study 49 Cochrane No study on NBS for SCD Study 48 (2000) HTA Study 50 (2010) HTA – Quebec <u>Primary Studies (8)</u> 7 studies – high risk of bias 1 study – unclear rating	Moderate quality No explicit mention of review methods being established prior to conduct of review and if with significant deviations in protocol No pooling/meta-analysis done

		were not reported separately for the primary target (e.g., studies focused on primary immunodeficiency, but outcomes were not presented separately for SCID); or animal studies	by members of provincial EAG	<u>Excluded studies and reasons – Appendix TB</u> <u>Page 142</u>	
Runkel et al. 2020 Germany	Screening for sickle cell disease in newborns: a systematic review	P – Newborns I – SCD screening + treatment C – No SCD screening OR SCD screening with no treatment O – Mortality, Morbidity (pain, hospitalizations, infections), adverse events, health-related QOL	<u>Quality assessment Tools:</u> IQWiG method, Cochrane risk of bias tool <u>Assessment by 2 reviewers:</u> conducted by one person and checked by another <u>Data extraction</u> <u>Use of standardized form:</u> used standardized tables developed and used by IQWiG Data extraction by 2 reviewers: <u>Data analysis and synthesis:</u> treatment effects as odds ratios for binary outcomes, 95% CI and p values (Wald test), sensitivity	1 (Retrospective controlled cohort) Risk of bias: High - Treatment groups not studies in parallel - Unclear comparability of groups or adequate control for confounding factors - Non-blinding of patient/investigator	Moderate Quality No explicit mention of review methods being established prior to conduct of review and if with significant deviations in protocol No explanation for selection of studies No pooling/meta-analysis done

			analysis and subgroup analysis to examine impact of variables were planned		
--	--	--	--	--	--

Table B.3. Quality assessment results for the included HTA/Systematic Review for SCD using the AMSTAR 2 Tool

AMSTAR Questions	QA Rating	
	Runkel <i>et al.</i> 2020	Institute of Health Economics 2016
1. Did research question and inclusion criteria for review include PICO components	Yes	Yes
2. Did review contain explicit statement that review methods were established prior to conduct of the review and justify any significant deviations from the protocol?	No	No
3. Did review authors explain their selection of the study designs for inclusion in the review?	No	No
4. Did review authors use a comprehensive literature search strategy?	Partial Yes	Partial Yes
5. Did review authors perform study selection in duplicate?	Yes	No
6. Did review authors perform data extraction in duplicate?	Yes	Yes
7. Did the review authors provide a list of excluded studies and justify the exclusions?	Yes	Yes
8. Did the review authors describe the included studies in adequate detail?	Yes	Yes
9. Did the review authors use a satisfactory technique for assessing the risk of bias in individual studies that were included in the review	Yes (For NSRI only, no RCTs were found)	Yes
10. Did the review authors report in the sources of funding for the studies included in the review?	Yes	Yes
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	N/A (No meta-analysis conducted)	N/A
12. If meta-analysis was performed, did review authors assess impact of risk of bias in individual studies on the results of meta-analysis or evidence synthesis?	N/A (No meta-analysis conducted)	N/A
13. Did the review authors account for risk of bias in the individual studies when interpreting the results of the review?	No	No
14. Did review authors provide satisfactory explanation for any heterogeneity observed in the results of the review?	No	No
15. If with quantitative synthesis, did authors carry out adequate investigation of publication	N/A (No meta-analysis conducted)	N/A

bias and discussed impact on the results of the review?		
16. Did the review authors report any potential sources of conflict of interest or funding received?	Yes	Yes
Total score (out of 16)	9	8
Rating (high/moderate/low/critically low)	Moderate	Moderate

APPENDIX C: GRADE Evidence Profile of Included Studies on Benefits/Harm of SCD Screening

Table C.1. GRADE evidence profile describing outcomes of newborn screening versus no screening (if applicable) for Sickle Cell Disease

Certainty assessment							No of patients		Effect		Certainty	Importance	
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Screening for SCD	No screening	Relative (95% CI)	Absolute (95% CI)			
Mortality													
1	observational studies	very serious ^a	not serious	not serious ^b	not serious	none	395/500 (79.0%)	105/500 (21.0%)	OR 0.09 (0.04 to 0.22)	187 fewer per 1,000 (from 199 fewer to 155 fewer)		CRITICAL	
Combined Mortality													
1 ^c	observational studies	serious ^c	not serious	not serious	not serious	none	0/0	0/0	HR 2.40 (1.11 to 5.19)	2 fewer per 1,000 (from 5 fewer to 1 fewer)		CRITICAL	
Mortality (follow up: range 7.2 years to 9.4 years)													
1 ^d	observational studies	very serious ^d	not serious ^d	not serious	serious ^d	none	0/81 (0.0%) ^d	5/64 (7.8%) ^d	Overall mortality rate 1.8 (0.0 to 0.0)	- per 1,000 (from -- to --)		CRITICAL	
Mortality													
1 ^e	observational studies ^e	serious ^e	not serious	serious ^e	not serious	all plausible residual confounding would reduce the demonstrated effect	For all black children age 0 to 14 years, there was a 38% reduction in SCD death rates in 3 states (NYC, Texas, Georgia) between 1983-1986 and 1987-1990. This did not differ significantly from the 29% decline in SCD-related mortality in 6 other states with large black populations that began SCD screening during 1987-1990.						CRITICAL
Mortality													
1	observational studies	very serious ^f	not serious	not serious	serious ^d	none	In the 18 years before any State NBS (1970-1988) there were 13 deaths attributed to sickle cell diseases. The limited State NBS program conducted between 1988 and July 2000 missed testing five affected children who subsequently died. These results document a marked reduction in mortality since the introduction of NBS for hemoglobinopathies and suggest that the Connecticut NBS program, coupled with comprehensive follow-up care, greatly reduced mortality.						CRITICAL

Certainty assessment							Nº of patients		Effect		Certainty	Importance			
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Screening for SCD	No screening	Relative (95% CI)	Absolute (95% CI)					
1	observational studies	serious ^g	not serious	not serious ^b	not serious	none	7 deaths (1.8%) occurred during study period vs 17.6% in the same age from Jamaican Sickle Cell Cohort Study							 VERY LOW	CRITICAL

Mortality

1	observational studies	very serious ^h	not serious	not serious ⁱ	not serious	none	Among 244 patients with SCD followed-up after neonatal screening, there was 3.6% mortality rate and a 6.8% first-year mortality compared to 9.5% national infant mortality in children with SCD.				 VERY LOW	CRITICAL
---	-----------------------	---------------------------	-------------	--------------------------	-------------	------	--	--	--	--	--	----------

Hospitalizations (follow up: range 7.2 years to 9.4 years)

1	observational studies	very serious ^d	not serious	not serious	serious ^d	none	Patients experienced 513 hospitalizations, including 13 episodes of sepsis with or without meningitis and ten acute sequestration crises.				 VERY LOW	IMPORTANT
---	-----------------------	---------------------------	-------------	-------------	----------------------	------	---	--	--	--	--	-----------

Time from birth to first hospital admission

1	observational studies	serious ^g	not serious	not serious ^b	not serious	none	Study group has earlier time to first admission (1.68 years versus 1.72 and 2.55 years in earlier controls in Lee study), likely reflecting a lower threshold for admitting sick patients, earlier intervention and better outcomes. Overall proportion of patients admitted in the study group (72%) was lower than the number admitted in control group (85%).				 VERY LOW	IMPORTANT
---	-----------------------	----------------------	-------------	--------------------------	-------------	------	--	--	--	--	--	-----------

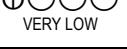
Acute splenic sequestration

1	observational studies	serious ^g	not serious	not serious ^b	not serious	none	Acute splenic sequestration was responsible for a significant number of clinical events - 188/584 (32%) , a reflection of parental education program, but is no longer a significant cause of mortality. Only 1/188 death occurred in this study representing 0.53% of total episodes demonstrating benefits of early detection and interventions.				 VERY LOW	IMPORTANT
---	-----------------------	----------------------	-------------	--------------------------	-------------	------	--	--	--	--	--	-----------

Acute chest syndrome

1	observational studies	serious ^g	not serious	not serious ^b	not serious	none	Acute chest syndrome (onset of acute respiratory signs/ symptoms with radiological evidence of a new pulmonary infiltrate) was the most common clinical (non-death) event. This accounted for almost 50% of all clinical events (289/584). It occurred in 142 patients. 43 patients had one episode, 42 patients had two episodes and 57 patients three or more episodes.				 VERY LOW	IMPORTANT
---	-----------------------	----------------------	-------------	--------------------------	-------------	------	---	--	--	--	--	-----------

Morbidity (at age 1.5 to 2.5 years)

1	observational studies	very serious ^h	not serious	not serious ⁱ	not serious	none	Out of 5 patients who underwent initial screening and on follow-up, 2 had anemia, splenomegaly between ages 1.5 to 2.5 years				 VERY LOW	IMPORTANT
---	-----------------------	---------------------------	-------------	--------------------------	-------------	------	--	--	--	--	--	-----------

Survival

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Screening for SCD	No screening	Relative (95% CI)	Absolute (95% CI)		
1	observational studies	very serious ^k	not serious	not serious	not serious	none	0	0	-	94 % higher (0.914 higher to 0.918 higher)	 VERY LOW	IMPORTANT

Survival at age 18

1	observational studies	very serious ^l	not serious	not serious ^m	not serious	none	0	0	-	93.9 % higher (0.903 higher to 0.962 higher)	 VERY LOW	IMPORTANT
---	-----------------------	---------------------------	-------------	--------------------------	-------------	------	---	---	---	--	--	-----------

CI: Confidence interval; OR: Odds ratio; HR: Hazard Ratio

Explanations

- a. High risk of bias due to comparison of non-concurrent groups in different decades, non-blinding of patients and investigators, and unclear control for confounding factors; Although review authors mentioned comparability of groups in terms of baseline characteristics seems likely, as newborns of the same region were included in both groups
- b. Note that newborn screening method used cord blood as sample instead of heel-prick
- c. Combined mortality (from acute splenic sequestration and pneumococcal septicemia/meningitis); Sources of bias: Not randomized trial, unclear how information from lost to follow-up patients was included or analyzed, unclear if outcomes assessors were aware of intervention status
- d. Bias cannot be assessed properly since only the abstract can be accessed. No confidence intervals reported.
- e. Sources of bias based on ROBINS-I: 1) Classification of intervention - non-differential misclassification of intervention status; 2) Post-intervention – bias due to missing data: exclusion of individuals with missing information about intervention status or other variables such as confounders (retrospective review of death certificates); Directness - number of children who underwent newborn screening was not mentioned
- f. Bias cannot be assessed properly since only the abstract can be accessed. No confidence intervals reported.
- g. Sources of bias: Not randomized trial, unclear how information from lost to follow-up patients was included or analyzed, unclear if outcomes assessors were aware of intervention status
- h. Sources of bias: moderate risk of selection bias and blinding, non-randomized, confounding
- i. Intervention: Penicillin prophylaxis, parent education, pneumococcal immunization and insecticide treated bed nets (malaria prophylaxis)
- j. Intervention: Penicillin prophylaxis and folic acid
- k. Sources of bias: selection bias, confounding, no blinding
- l. Sources of bias: Moderate selection bias, non-blinding, confounding, more children with mild phenotype were lost to follow-up compared to Hb SS and Hb SB⁰
- m. Screening/confirmatory test not reported; Intervention: penicillin prophylaxis until age 5 years, pneumococcal vaccination at age 2 and 5 years

APPENDIX D: Summary of Diagnostic Accuracy Studies

Table D.1. Summary of diagnostic accuracy findings from different methods of screening for SCD (as adopted from STE done in Alberta Canada)

Study	Newborns screened	TP	FP	TN	FN	Sensitivity	Specificity	PPV	NPV
High Performance Liquid Chromatography (HPLC)									
<i>Campbell 1999</i> <i>Index: HPLC</i> <i>Ref: IEF</i>	25,750	29	NR	NR	<i>0 SCD (5 non SCD samples were detected by IEF and not by HPLC)</i>	<i>99.45</i> Not seen in actual study May not be for SCD only <i>But based on computation, Sn = 100</i>	<i>99.99</i> Not seen in actual study. May not be for SCD only	NR	NR
<i>Eastman 1996</i> <i>Index: HPLC performance</i> <i>No Ref</i>	2,500,000	NR	2	NR	2 (BT)	99.81	100	99.81	100
<i>Lobitz 2014</i> <i>Index: HPLC</i> <i>Ref: CE</i>	34,084	14	0	NR	NR	NR	NR	NR	NR
<i>Panigraphi 2012</i> <i>Index: HPLC</i> <i>Ref HPLC B thalassemia short program</i>	1158	NR	0	NR	3	NR	NR	NR	NR
Isoelectric Focusing (IEF)									
<i>Campbell 1999</i> <i>Index: HPLC</i> <i>Ref: IEF</i>	25,750	NR	NR	NR	<i>0 SCD (2 non SCD samples detected by HPLC and not by IEF)</i>	99.78	100	100	99.99
<i>McGann 2013</i> <i>Index: IEF</i> <i>Ref: CE</i>	36, 453	NR	NR	NR	NR	*Concordance in 9 samples Discordance in 9 samples			

Tandem Mass Spectrometry (MS/MS)									
<i>Boemer 2011</i> Index: MS/MS Ref: Sequencing of <i>B globin gene</i>	43,736	NR	0	NR	NR	NR	NR	NR	NR
<i>Moat 2014</i> Index: MS/MS Ref: HPLC	13,249	7	8 (6 Hb SC, 2 Hb SS)	NR	0	100% (computed based on given data from study)	99.94% (computed based on given data from study)	NR	NR

*NR – Not reported

APPENDIX E: GRADE Evidence Profiles of Diagnostic Accuracy Studies

Campbell 1999 – HPLC vs IEF

Sensitivity	1.00 (95% CI: -- to --)				Prevalence	0.002%	1.5%	0.015%			
Outcome	No of studies (No of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 100,000 patients tested			Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 0.002%	pre-test probability of 1.5%	pre-test probability of 0.015%	
True positives (patients with SCD)	1 studies 25570 patients	cross-sectional (cohort type accuracy study)	not serious	not serious	not serious	serious ^a	none	2 (0 to 0)	1500 (0 to 0)	15 (0 to 0)	 MODERATE
False negatives (patients incorrectly classified as not having SCD)								0 (2 to 2)	0 (1500 to 1500)	0 (15 to 15)	
True negatives (patients without SCD)	1 studies 25570 patients	cross-sectional (cohort type accuracy study)	not serious	not serious	not serious	serious ^a	none	0 (0 to 0)	0 (0 to 0)	0 (0 to 0)	 MODERATE
False positives (patients incorrectly classified as having SCD)								99998 (99998 to 99998)	98500 (98500 to 98500)	99985 (99985 to 99985)	

Explanations

a. No confidence interval. Based on computation, detection of all 29 SCDs by both IEF and HPLC yields Sn = 100, Sp cannot be computed as there is no data for TN and FP in article.

Lobitz 2014 - HPLC vs CE

Sensitivity	-- (95% CI: -- to --)				Prevalence	0.002%			
Specificity	-- (95% CI: -- to --)								
Outcome	# of studies (No of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 100,000 patients tested	
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 0.002%	
True positives (patients with SCD)	1 studies 34084 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	0 (0 to 0)	
False negatives (patients incorrectly classified as not having SCD)								2 (2 to 2)	
True negatives (patients without SCD)	1 studies 34084 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	0 (0 to 0)	
False positives (patients incorrectly classified as having SCD)								99998 (99998 to 99998)	

Explanations

a. High risk of bias in terms of patient selection, unclear if threshold selection for index test was specified, reference standard no likely to classify target condition, and not interpreted without knowledge of index test, patients did not all receive reference standard and not all were included in analysis

Eastman 1996 – HPLC performance alone (no reference standard)

Sensitivity	1.00 (95% CI: -- to --)				Prevalence	0%			
Specificity	1.00 (95% CI: -- to --)								
Outcome	# of studies (No of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 100,000 patients tested	
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 0%	
True positives (patients with SCD)	1 studies 2500000 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	0 (0 to 0)	
False negatives (patients incorrectly								0 (0 to 0)	

Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 100,000 patients tested	Test accuracy CoE	
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias			
classified as not having SCD)										
True negatives (patients without SCD)	1 studies 2500000 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	100000 (0 to 0)		
False positives (patients incorrectly classified as having SCD)								0 (100000 to 100000)		

Explanations

a. Sources of bias - Non random sample of patients, no reference standard performed compared to HPLC

Panigraphi 2012 – HPLC versus HPLC beta thalassemia short program

Sensitivity	-- (95% CI: -- to --)				Prevalence	0.002%				
Specificity	-- (95% CI: -- to --)									
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested	Test accuracy CoE	
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias			
True positives (patients with SCD)	1 studies 1158 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	0 (0 to 0)		
False negatives (patients incorrectly classified as not having SCD)								0 (0 to 0)		
True negatives (patients without SCD)	1 studies 1158 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	0 (0 to 0)		
False positives (patients incorrectly classified as having SCD)								1000 (1000 to 1000)		

Explanations

a. Sources of bias - threshold for index test not specified, reference standard not interpreted without knowledge of index test, patients did not all receive reference standard and not all patients were included in analysis

McGann 2013 – IEF vs CE

Sensitivity	-- (95% CI: -- to --)				Prevalence	0.002%			
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested	Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 0.002%	
True positives (patients with SCD)	1 studies 36453 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	0 (0 to 0)	 LOW
False negatives (patients incorrectly classified as not having SCD)								0 (0 to 0)	
True negatives (patients without SCD)	1 studies 36453 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	0 (0 to 0)	 LOW
False positives (patients incorrectly classified as having SCD)								1000 (1000 to 1000)	

Explanations

a. Sources of bias – non-consecutive or random sample of patients enrolled, threshold for index test not specified, reference standard not interpreted without knowledge of index test, patients did not all receive reference standard and not all patients were included in analysis

Boemer 2011 – MS/MS vs Sequencing of whole B globin gene

Sensitivity	-- (95% CI: -- to --)				Prevalence	0.002%			
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested	Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 0.002%	
True positives (patients with SCD)	1 studies 43736 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	0 (0 to 0)	 LOW
False negatives (patients incorrectly								0 (0 to 0)	

Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested	Test accuracy CoE	
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias			
classified as not having SCD)										
True negatives (patients without SCD)	1 studies 43736 patients	cross-sectional (cohort type accuracy study)	very serious ^a	not serious	not serious	not serious	none	0 (0 to 0)		
False positives (patients incorrectly classified as having SCD)								1000 (1000 to 1000)		

Explanations

a. Sources of bias - unclear if consecutive or random samples and if inappropriate exclusions were avoided, threshold for index test not specified, unclear if reference standard able to correctly classify target condition and if reference standard not interpreted without knowledge of index test, patients did not all receive reference standard and not all patients were included in analysis

Moat 2014 – MS/MS vs HPLC

Sensitivity	-- (95% CI: -- to --)					Prevalence	0.002%		
Specificity	-- (95% CI: -- to --)								
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested	Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 0.002%	
True positives (patients with SCD)	1 studies 13.249 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	not serious	none	0 (0 to 0)	
False negatives (patients incorrectly classified as not having SCD)								0 (0 to 0)	
True negatives (patients without SCD)	1 studies 13.249 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	not serious	none	0 (0 to 0)	
False positives (patients incorrectly classified as having SCD)								1000 (1000 to 1000)	

Explanations

a. Sources of bias - unclear if index test was interpreted without knowledge of reference, and if reference standard were interpreted without knowledge of index test

APPENDIX F: Cost-effectiveness of SCD screening

Table F.1. Estimated cost inputs (CAN\$) for confirmatory tests, treatment, consultation and follow-up of patients screened for SCD in Alberta Canada.

Cost description	Cost Value (CAN\$)	Location	High liAssumption	Source
Confirmatory testing				
Follow-up testing (HPLC + CBC) and genetic confirmation	1,330.00	Laboratory Services (AHS)	Assume 1% of screen-positive patients receive test	
Treatment for sequelae				
Hospitalization, newborn sepsis, yearly	10,663.30	IHDA	Physician cost is assumed to be same as that of follow-up physician consultation	
Oral amoxicillin, yearly	337.63	Alberta drug benefit	Amoxicillin, DIN/PIN 02036347, 125 mg, unit cost \$0.4625, 125 mg twice daily	
Condition management				
Initial physician consultation, per visit	346.00	Schedule of Medical Benefits	1.5 hours per visit	
Follow-up physician consultation, per visit	123.60	Schedule of Medical Benefits	Every 3 months less than 1 year of age and every 6 months afterwards	
Genetic counselling, one time	163.92	ALIS, Expert Advisory Group	4 hours, wage \$40.98 per hour	
Follow-up evaluation				
Transcranial ultrasound, per test	325.18	Schedule of Medical Benefits	Once per year	
CBC, per test	17.94	Schedule of Medical Benefits	Every 3 months less than 1 year of age and every 6 months afterwards	

Estimated costs in Philippine peso may be based on exchange rates in the same period of publication of the Canada STE (March 2016): 1 CAN\$ = Php 35.70.

4. Newborn Screening for Thalassemia

APPENDIX A. Search Yield

DATABASE	SEARCH STRATEGY / SEARCH TERMS	DATE AND TIME OF SEARCH	RESULTS	
			Yield	Eligible
NICE (National Institute for Health and Care Excellence)	Thalassemia screening in newborn OR neonate	August 8, 2021	241	2 Evidence Assessment - 1 Guideline - 1
U.S. Preventive Services Task Force (USPSTF)	Thalassemia, hemoglobinopathies, screening	August 8, 2021	33	0
Canadian Task Force for Preventive Health Care (CTFPHC)	Thalassemia, screening	August 8, 2021	0	0
Guideline International Network	Thalassemia	August 8, 2021	1	0
Cochrane Library	(thalassemia in Title Abstract Keyword OR hemoglobinopathies in Title Abstract Keyword) AND (newborn screening OR neonatal screening in Title Abstract Keyword) - with Cochrane Library publication date Between Jan 2015 and Aug 2021, in Cochrane Reviews, Cochrane Protocols, Trials (Word variations have been searched) MeSH descriptor: [Thalassemia] explode all trees AND MeSH descriptor: [Neonatal Screening] explode all trees - with Cochrane Library publication date Between Jan 2015 and Aug 2021, in Cochrane Reviews, Cochrane Protocols, Trials (Word variations have been searched)	August 8, 2021	16 0	0
PubMed	("Thalassemia"[MeSH Terms] AND "Neonatal Screening"[MeSH Terms] AND ("guideline"[Publication Type] OR "practice guideline"[Publication Type] OR "recommendation*[Title] OR "standard*[Title] OR "standard*[Title] OR "guideline*[Title] AND 2015/01/01:2021/08/08[Date - Publication])) AND (2015/1/1:2021/8/8[pdat]) ((thalassemia[MeSH Terms]) OR (thalassemia)) AND (newborn screening OR neonatal screening) Filters: Meta-Analysis, Systematic Review, from 2015/1/1 - 2021/7/8	August 8, 2021	1 1 162	0 0 0

	<p>((("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields]) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields] OR ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields])))) AND (2015/1/1:2021/8/8[pdat])</p> <p>("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields]) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields] OR ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields])) AND 2015/01/01:2021/08/08[Date - Publication] AND ("outcome assessment, health care"[MeSH Terms] OR "outcome and process assessment, health care"[MeSH Terms] OR "Patient Reported Outcome Measures"[MeSH Terms] OR "Treatment Outcome"[MeSH Terms] OR "Numbers Needed To Treat"[MeSH Terms] OR "Early Diagnosis"[MeSH Terms] OR "Quality-Adjusted Life Years"[MeSH Terms] OR "Disease Management"[MeSH Terms] OR "Death"[MeSH Terms] OR "Clinical Decision Rules"[MeSH Terms] OR "Long Term Adverse Effects"[MeSH Terms])</p> <p>("Thalassemia"[MeSH Terms] AND "Neonatal Screening"[MeSH Terms]) AND (2010/1/1:2021/8/8[pdat])</p>		4	0
Disease Burden/Epidemiology				
Pubmed	((("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields]) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]))) AND (2010/1/1:2021/8/8[pdat]))	August 8, 2021	66	4 Goonasekera 2018, Wen 2019, Zhong 2020, Padilla 2021

	<p>Fields] AND "screening"[All Fields]) OR ("neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields] OR ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields])) AND ("epidemiologies"[All Fields] OR "epidemiology"[MeSH Subheading] OR "epidemiology"[All Fields] OR "epidemiology"[MeSH Terms] OR "epidemiology s"[All Fields] OR ("cost of illness"[MeSH Terms] OR ("cost"[All Fields] AND "illness"[All Fields]) OR "cost of illness"[All Fields] OR ("disease"[All Fields] AND "burden"[All Fields]) OR "disease burden"[All Fields]))) AND (2010/1/1:2021/8/8[pdat])</p> <p>("Thalassemia"[MeSH Terms] AND "Neonatal Screening"[MeSH Terms] AND ("Epidemiology"[MeSH Terms] OR ("Cost of Illness"[MeSH Terms] OR "Global Burden of Disease"[MeSH Terms]))) AND (2010/1/1:2021/8/8[pdat])</p>		0	
Cost-Analysis				
Pubmed	<p>((("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaealias"[All Fields] OR "thalassealias"[All Fields]) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields] OR ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields])) AND ("cost benefit analysis"[MeSH Terms] OR ("cost benefit"[All Fields] AND "analysis"[All Fields]) OR "cost benefit analysis"[All Fields] OR ("cost"[All Fields] AND "effectiveness"[All Fields]) OR "cost effectiveness"[All Fields] OR (("economics"[MeSH Subheading] OR "economics"[All Fields] OR "cost"[All Fields] OR "costs and cost analysis"[MeSH Terms] OR ("costs"[All Fields] AND "cost"[All Fields] AND "analysis"[All Fields]) OR "costs and cost analysis"[All Fields]) AND ("effect"[All Fields] OR "effecting"[All Fields] OR "effective"[All Fields] OR "effectively"[All Fields] OR "effectiveness"[All Fields] OR "effectivenesses"[All Fields] OR "effectives"[All Fields] OR "effectivities"[All Fields] OR</p>	5	1 Esmaeilzadeh 2016	

August 8,
2021

0

	<p>"effectivity"[All Fields] OR "effects"[All Fields])))) AND (2015/1/1:2021/8/8[pdat])</p> <p>("Thalassemia"[MeSH Terms] AND "Neonatal Screening"[MeSH Terms] AND ("Costs and Cost Analysis"[MeSH Terms] OR "Cost of Illness"[MeSH Terms] OR ("Cost Sharing"[MeSH Terms] OR "Cost Savings"[MeSH Terms]) OR ("Cost Control"[MeSH Terms] OR "Cost-Benefit Analysis"[MeSH Terms] OR "Health Care Costs"[MeSH Terms] OR "Health Expenditures"[MeSH Terms] OR "Cost Allocation"[MeSH Terms] OR "Direct Service Costs"[MeSH Terms] OR "Hospital Costs"[MeSH Terms]))) AND (2015/1/1:2021/8/8[pdat])</p>			
Diagnostic accuracy				
Pubmed	<p>((("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields]) AND ((("diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnostic"[All Fields] OR "diagnostical"[All Fields] OR "diagnostically"[All Fields] OR "diagnostics"[All Fields]) AND ("accuracies"[All Fields] OR "accuracy"[All Fields]))) AND (2010/1/1:2018/8/13[pdat]))</p> <p>((((thalassemia) OR (thalassaemia[MeSH Terms])) AND (2015/1/1:2021/8/16[pdat]))) AND (sensitivity OR specificity AND (2015/1/1:2021/8/16[pdat]))) AND (health screening AND (2010/1/1:2021/8/16[pdat])))</p>	August 13, 2021	7 17	2 1
Benefit or Harm				
Pubmed	<p>((("Pharmacoepidemiology"[MeSH Terms] OR "Beneficence"[MeSH Terms] OR "Risk Assessment"[MeSH Terms]) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR</p>	August 13, 2021	6 7	0 0

"thalassaemias"[All Fields] OR "thalassemias"[All Fields])) AND (2015:3000/12/12[pdat])		
(("benefit"[All Fields] OR "benefited"[All Fields] OR "benefiting"[All Fields] OR "benefits"[All Fields] OR "benefitted"[All Fields] OR "benefitting"[All Fields]) AND 2015/01/01:2021/12/31[Date - Publication] AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields])) AND (2015:2021[pdat]))	2	0
(("harm"[All Fields] AND 2015/01/01:2021/12/31[Date - Publication] AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields])) AND (2015:2021[pdat]))	2	0
(("complicances"[All Fields] OR "complicate"[All Fields] OR "complicated"[All Fields] OR "complicates"[All Fields] OR "complicating"[All Fields] OR "complication"[All Fields] OR "complications"[All Fields] OR "complications"[MeSH Subheading] OR "complications"[All Fields]) AND 2015/01/01:2021/12/31[Date - Publication] AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields])) AND (2015:2021[pdat]))	7	0
(("complicances"[All Fields] OR "complicate"[All Fields] OR "complicated"[All Fields] OR "complicates"[All Fields] OR "complicating"[All Fields] OR "complication"[All Fields] OR "complications"[All Fields] OR "complications"[MeSH Subheading] OR "complications"[All Fields]) AND 2015/01/01:2021/12/31[Date - Publication] AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields])) AND (2015:2021[pdat]))	28	0

	<p>(("feasibilities"[All Fields] OR "feasibility"[All Fields] OR "feasible"[All Fields] OR "feasiblty"[All Fields]) AND 2015/01/01:2021/12/31[Date - Publication] AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields])) AND (2015:2021[pdat])</p> <p>(("outcome"[All Fields] OR "outcomes"[All Fields]) AND 2015/01/01:2021/12/31[Date - Publication] AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields])) AND (2015:2021[pdat])</p> <p>((("early diagnosis"[MeSH Terms] OR ("early"[All Fields] AND "diagnosis"[All Fields]) OR "early diagnosis"[All Fields] OR ("early"[All Fields] AND "detection"[All Fields]) OR "early detection"[All Fields] AND ("thalassaemia"[All Fields] OR "thalassemia"[MeSH Terms] OR "thalassemia"[All Fields] OR "thalassaemias"[All Fields] OR "thalassemias"[All Fields]) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields])) AND (2015/1/1:2021[pdat]))</p>		
Google Scholar	<p>All words: thalassemia, outcome Exact phrase: newborn screening OR neonatal screening where my words occur: anywhere in the article Return articles dated between 2015-2021 Sort by date, everything</p> <p>All words: thalassemia, complication Exact phrase: newborn screening OR neonatal screening</p>	August 20, 2021	440 0 292 0

	where my words occur: anywhere in the article Return articles dated between 2015-2021 Sort by date, everything	August 22, 2021			
	All words: thalassemia Exact phrase: newborn screening OR neonatal screening Any of the words: benefit, harm where my words occur: anywhere in the article Return articles dated between 2015-2021 Sort by relevance, everything	August 23, 2021	1430	0	
Total					2,840

APPENDIX B. Grade of Diagnostic Studies

Study	Design	Participants	Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	Test accuracy
Uaprasert 2014	Observational case-control study	241	serious	not serious	not serious	not serious	publication bias strongly suspected	⊕⊕○○ Low
Wu 2015	Observational case-control study	1169	serious	not serious	not serious	not serious	publication bias strongly suspected	⊕⊕○○ Low
Yu 2017	Observational case-control study	1081	serious	not serious	not serious	not serious	publication bias strongly suspected	⊕⊕○○ Low
Yang 2018	Observational case-control study	1193	serious	not serious	not serious	not serious	publication bias strongly suspected	⊕⊕○○ Low
Allaf 2020	Observational case-control study	343,036	not serious	not serious	not serious	not serious	publication bias strongly suspected	⊕⊕⊕○ Moderate

APPENDIX C. QUADAS Tool of Diagnostic Studies

		Uaprasert 2014	Wu 2015	Yu 2016	Yang 2018	Allaf 2020
Patient selection	Was a consecutive or random sample of patients enrolled?	–	–	–	–	+
	Was a case-control design avoided?	–	–	–	–	–
	Did the study avoid inappropriate exclusions?	–	–	–	+	+
	Risk of bias	Low	Low	Low	Low	Low
Index test	Were the index test results interpreted without knowledge of the results of the reference standard?	–	–	–	–	–
	If a threshold was used, was it pre-specified?	–	+	–	+	–
	Risk of bias	Unclear	Unclear	High	Unclear	Unclear
Reference standard	Is the reference standard likely to correctly classify the target condition?	+	+	+	+	+
	Were the reference standard results interpreted without knowledge of the results of the index test?	–	–	–	–	–
	Risk of bias	Unclear	Unclear	Unclear	Unclear	Unclear
Flow and timing	Was there an appropriate interval between index test(s) and reference standard?	+	+	+	+	+

	Did all patients receive a reference standard?	+	+	+	+	+
	Did patients receive the same reference standard?	+	+	+	+	+
	Were all patients included in the analysis?	+	+	+	+	+
	Risk of bias	Low	Low	Low	Low	Low
	Total score	6	6	5	7	7

Legend:

⊕ Fulfilled

⊖ Not Fulfilled

⊘ Unclear

5. Newborn Screening for G6PD Deficiency

APPENDIX A. Search Yield

Search Strategy and Databases Used

Last Search Done: September 2, 2021

A. Clinical Practice Guidelines Databases

1. NICE Guidelines

QUERY	RESULTS	YIELD
G6PD Deficiency	1	0
Newborn Screening	28	0

- Only mentioned in guidelines for “Jaundice in Newborn Babies under 28 days”

2. US Preventive Services Task Force (USPTF)

QUERY	RESULTS	YIELD
G6PD Deficiency Filter: Pediatric	33	0
Newborn Screening Filter: Pediatric	33	0

3. UK National Screening Committee

QUERY	RESULTS	YIELD
G6PD Deficiency	0	0

4. Canadian Task Force for Preventive Health Care

QUERY	RESULTS	YIELD
G6PD Deficiency	0	0
Newborn Screening	6	0

5. Guidelines International Network

QUERY	RESULTS	YIELD
G6PD Deficiency	0	0
Newborn Screening	2	0
Newborn Jaundice	1	1

B. PubMed

Search #	Query	Search Details	Results
10	#6 AND #9	("glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR ("g6pd"[All Fields] AND "deficiency"[All Fields]) OR "g6pd deficiency"[All Fields] OR ("glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR "glucose 6 phosphate dehydrogenase deficiency"[All Fields]	274

		Fields])) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("haemolytic anaemia"[All Fields] OR "anemia, hemolytic"[MeSH Terms] OR ("anemia"[All Fields] AND "hemolytic"[All Fields]) OR "hemolytic anemia"[All Fields] OR ("hemolytic"[All Fields] AND "anemia"[All Fields]) OR (("haemolytic"[All Fields] OR "hemolytic"[All Fields] OR "hemolytics"[All Fields]) AND ("crisis"[Journal] OR "crisis"[All Fields])) OR ("kernicterus"[MeSH Terms] OR "kernicterus"[All Fields]))	
9	((Hemolytic Anemia) OR (Hemolytic Crisis)) OR (Kernicterus)	"haemolytic anaemia"[All Fields] OR "anemia, hemolytic"[MeSH Terms] OR ("anemia"[All Fields] AND "hemolytic"[All Fields]) OR "hemolytic anemia"[All Fields] OR ("hemolytic"[All Fields] AND "anemia"[All Fields]) OR (("haemolytic"[All Fields] OR "hemolytic"[All Fields] OR "hemolytics"[All Fields]) AND ("crisis"[Journal] OR "crisis"[All Fields])) OR ("kernicterus"[MeSH Terms] OR "kernicterus"[All Fields])	85,050
8	#6 and #7	("glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR ("g6pd"[All Fields] AND "deficiency"[All Fields]) OR "g6pd deficiency"[All Fields] OR ("glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR "glucose 6 phosphate dehydrogenase deficiency"[All Fields])) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]) AND ("fluoroimmunoassay"[MeSH Terms] OR "fluoroimmunoassay"[All Fields] OR "fluoroimmunoassays"[All Fields])	1
7	Fluoroimmunoassay	"fluoroimmunoassay"[MeSH Terms] OR "fluoroimmunoassay"[All Fields] OR "fluoroimmunoassays"[All Fields]	3,615
6	#1 AND #5	("glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR ("g6pd"[All Fields] AND "deficiency"[All Fields]) OR "g6pd deficiency"[All Fields] OR ("glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR "glucose 6 phosphate dehydrogenase deficiency"[All Fields])) AND ("neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields]	328

		AND "screening"[All Fields] OR "newborn screening"[All Fields])	
5	Newborn Screening	"neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]	39,741
4	#1 AND #3	"neonatal screening"[MeSH Terms] OR ("neonatal"[All Fields] AND "screening"[All Fields]) OR "neonatal screening"[All Fields] OR ("newborn"[All Fields] AND "screening"[All Fields]) OR "newborn screening"[All Fields]	21
3		(guideline [Publication Type] OR practice guideline [Publication Type] or recommendation*[Title] OR standard* [Title] OR standard* [Title] OR guideline* [Title])	240,120
2	(G6PD Deficiency) OR (Glucose 6 Phosphate Dehydrogenase Deficiency) [MeSH Terms]	"glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR ("g6pd"[All Fields] AND "deficiency"[All Fields]) OR "g6pd deficiency"[All Fields] OR ("glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR "glucose 6 phosphate dehydrogenase deficiency"[All Fields])	6,775
1	(G6PD Deficiency) OR (Glucose 6 Phosphate Dehydrogenase Deficiency)	"glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR ("g6pd"[All Fields] AND "deficiency"[All Fields]) OR "g6pd deficiency"[All Fields] OR ("glucosephosphate dehydrogenase deficiency"[MeSH Terms] OR ("glucosephosphate"[All Fields] AND "dehydrogenase"[All Fields] AND "deficiency"[All Fields]) OR "glucosephosphate dehydrogenase deficiency"[All Fields] OR "glucose 6 phosphate dehydrogenase deficiency"[All Fields])	6,775

Most of the articles identified were studies on incidence rates, screening for adults or older children in malaria endemic areas or screening for jaundiced neonates. Observational studies used in evidence summary were identified from citations from some of the review articles yielded in this search.

APPENDIX B. GRADE Table

Author(s): Melissa Dator, Kathryn Braganza, Mary Ann Abacan, Ian Cabaluna

Question: G6PD Screening compared to No Screening for Asymptomatic Newborns

Setting:

Bibliography: Mallouh A, Imseeh G, Abu-Osba Y, Hamdan J. Screening for glucose-6-phosphate dehydrogenase deficiency can prevent severe neonatal jaundice. 1992; 12:391-395; Cohan N, Karimi M et al. The efficacy of a neonatal screening programme in decreasing the hospitalization rate of patients with G6PD deficiency in Southern Iran. Journal of Medical Screening. 2010; 17 (2) 66-67; Meloni T, Forteleoni G. Marked Decline of Favism after Neonatal Glucose -6-Phosphate Dehydrogenase Screening and Health Education: The Northern Sardinian Experience. Acta Haematologica, 1992; 87(1-2): 29-31

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	G6PD Screening	No Screening	Relative (95% CI)	Absolute (95% CI)		
Hospitalization due to Hemolytic Anemia (follow-up: range 3 years to 10 years)												
2	observational studies	serious ^a	not serious	serious ^a	serious ^b	none	441/1502 (29.4%)	1061/1502 (70.6%)	not estimable		 Very low	
Hospitalization due to Hemolytic Anemia (follow-up: range 5 years to 5.2 years)												
1	observational studies	serious ^c	not serious	not serious ^a	serious ^b	none	42/33943 (0.1%)	7/10005 (0.1%)	OR 1.79 (0.80 to 3.99)	1 more per 1,000 (from 0 fewer to 2 more)	 Very low	
Kernicterus (follow-up: range 5 years to 5.2 years)												
1	observational studies	serious ^c	not serious	serious ^a	serious ^b	very strong association	0/33943 (0.0%)	7/33943 (0.0%)	OR 0.0197 (0.0010 to 0.3400)	0 fewer per 1,000 (from 0 fewer to 0 fewer)	 Very low	

CI: confidence interval; OR: odds ratio

Explanations

a. Groups taken from 2 different time points or areas

b. Wide confidence intervals

c. Flawed measurement between the exposure and control group. Control group did not undergo confirmatory testing for G6PD deficiency.

Question: Should G6PD Screening be used to screen for G6PD Deficiency in Asymptomatic Newborns?

Sensitivity	1.00 (95% CI: -- to --)			Prevalence	15%				
Specificity	0.99 (95% CI: -- to --)								
Outcome	№ of studies (№ of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested	Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 15%	
True positives (patients with G6PD Deficiency)	1 study 13376 patients	cross-sectional (cohort type accuracy study)	serious	not serious	serious	serious	none	150 (0 to 0)	⊕○○○ VERY LOW
False negatives (patients incorrectly classified as not having G6PD Deficiency)								0 (150 to 150)	
True negatives (patients without G6PD Deficiency)	1 study 13376 patients	cross-sectional (cohort type accuracy study)	serious	not serious	serious	serious	none	842 (0 to 0)	⊕○○○ VERY LOW
False positives (patients incorrectly classified as having G6PD Deficiency)								8 (850 to 850)	

6. Newborn Screening for Homocystinuria and Methionine Adenosyltransferase Deficiency

APPENDIX A. Search Yield

*as of August 31, 2021

Search databases (all crosschecked): PubMed, google scholar, Cochrane, EMBASE
References of identified Systematic Reviews scanned

HOMOCYSTINURIA/CBS Deficiency

homocystinuria: "homocystinuria"[MeSH Terms] OR "homocystinuria"[All Fields] OR "homocystinurias"[All Fields]	2,346
homocystinuria: "homocystinuria"[MeSH Terms] OR "homocystinuria"[All Fields] OR "homocystinurias"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields] diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading]	1,230
homocystinuria: "homocystinuria"[MeSH Terms] OR "homocystinuria"[All Fields] OR "homocystinurias"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields] diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading] newborns: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields]	339
homocystinuria: "homocystinuria"[MeSH Terms] OR "homocystinuria"[All Fields] OR "homocystinurias"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields] diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading] asymptomatic: "asymptomatic"[All Fields] OR "asymptomatically"[All Fields] OR "asymptomatics"[All Fields] newborns: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields]	13
screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields] diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading]	9
homocystinuria: "homocystinuria"[MeSH Terms] OR "homocystinuria"[All Fields] OR "homocystinurias"[All Fields] Cystathione beta-synthase: "cystathione beta-synthase"[MeSH Terms] OR ("cystathione"[All Fields] AND "beta-synthase"[All Fields]) OR "cystathione beta-synthase"[All Fields] OR ("cystathione"[All Fields] AND "beta"[All Fields] AND "synthase"[All Fields]) OR "cystathione beta synthase"[All Fields] CBS deficiency: "homocystinuria"[MeSH Terms] OR "homocystinuria"[All Fields] OR ("cbs"[All Fields] AND "deficiency"[All Fields]) OR "cbs deficiency"[All Fields] diagnostic: "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnostic"[All Fields] OR "diagnostical"[All Fields] OR "diagnostically"[All Fields] OR "diagnostics"[All Fields] accuracy: "accuracies"[All Fields] OR "accuracy"[All Fields]	
homocystinuria: "homocystinuria"[MeSH Terms] OR "homocystinuria"[All Fields] OR "homocystinurias"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR	3

"early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields]
diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading]
newborns: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields]
practice guidelines: "practice guideline"[Publication Type] .or. "practice guidelines as topic"[MeSH Terms] .or. "practice guidelines"[All Fields]

MAT

methionine adenosyltransferase deficiency: "Hypermethioninemia"[Supplementary Concept] OR "Hypermethioninemia"[All Fields] OR "methionine adenosyltransferase deficiency"[All Fields]	197
methionine adenosyltransferase deficiency: "Hypermethioninemia"[Supplementary Concept] OR "Hypermethioninemia"[All Fields] OR "methionine adenosyltransferase deficiency"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields]	93
diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading]	
newborns: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields] methionine adenosyltransferase deficiency: "Hypermethioninemia"[Supplementary Concept] OR "Hypermethioninemia"[All Fields] OR "methionine adenosyltransferase deficiency"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields]	59
diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading]	
newborns: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields] methionine adenosyltransferase deficiency: "Hypermethioninemia"[Supplementary Concept] OR "Hypermethioninemia"[All Fields] OR "methionine adenosyltransferase deficiency"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields]	0
diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading]	
practice guidelines: "practice guideline"[Publication Type] .or. "practice guidelines as topic"[MeSH Terms] .or. "practice guidelines"[All Fields]	

APPENDIX B. PRISMA: Homocystinuria

PRISMA - Homocystinuria

*De novo done after evaluation of available CPGs with AGREE II

*Systematic review available did not contain any included studies

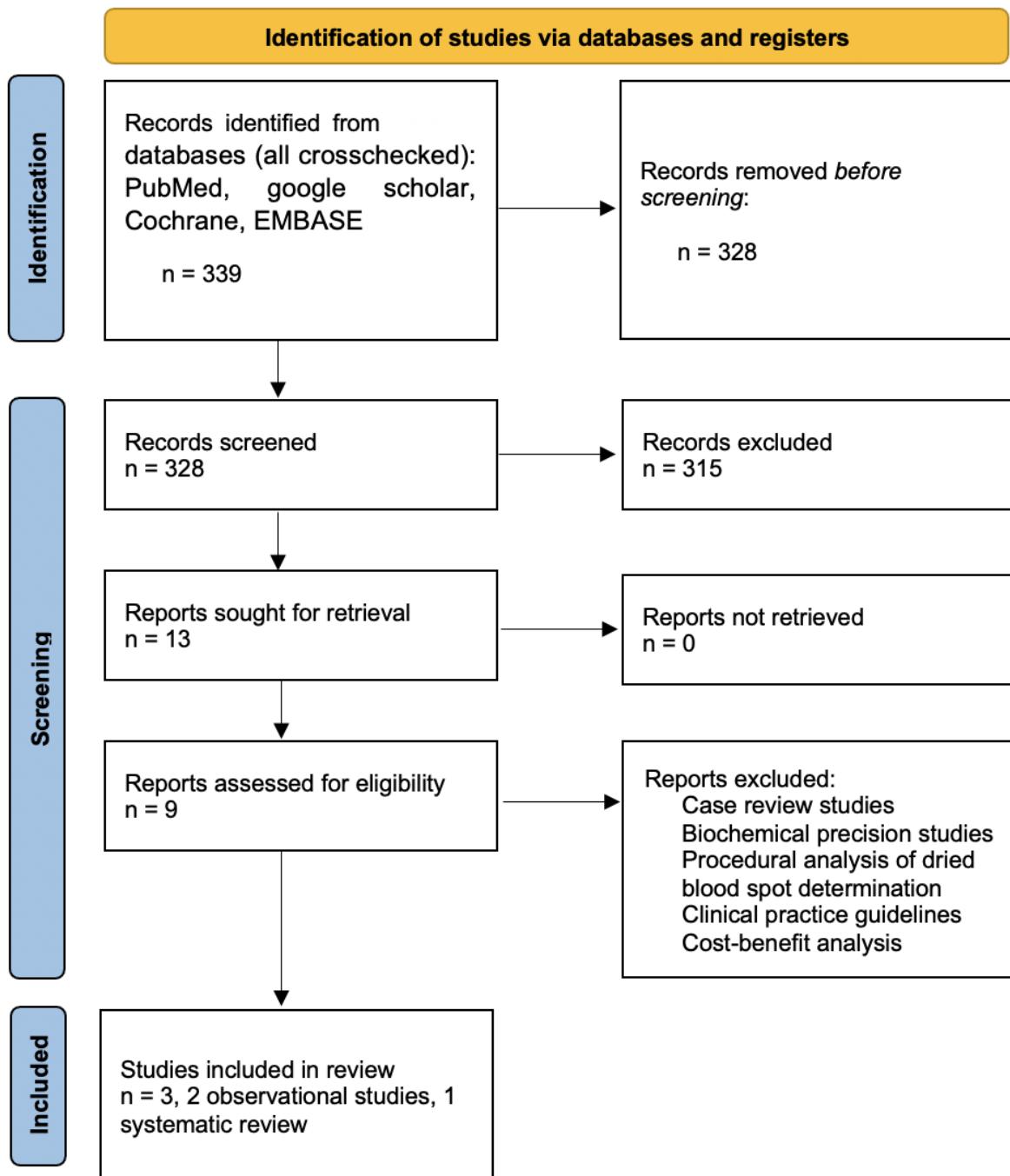


Figure 1. PRISMA for research question on screening for Homocystinuria among asymptomatic newborns

APPENDIX C. PRISMA: MAT I/III Deficiency

PRISMA – MAT I/III Deficiency

*De novo done after evaluation of available CPGs with AGREE II

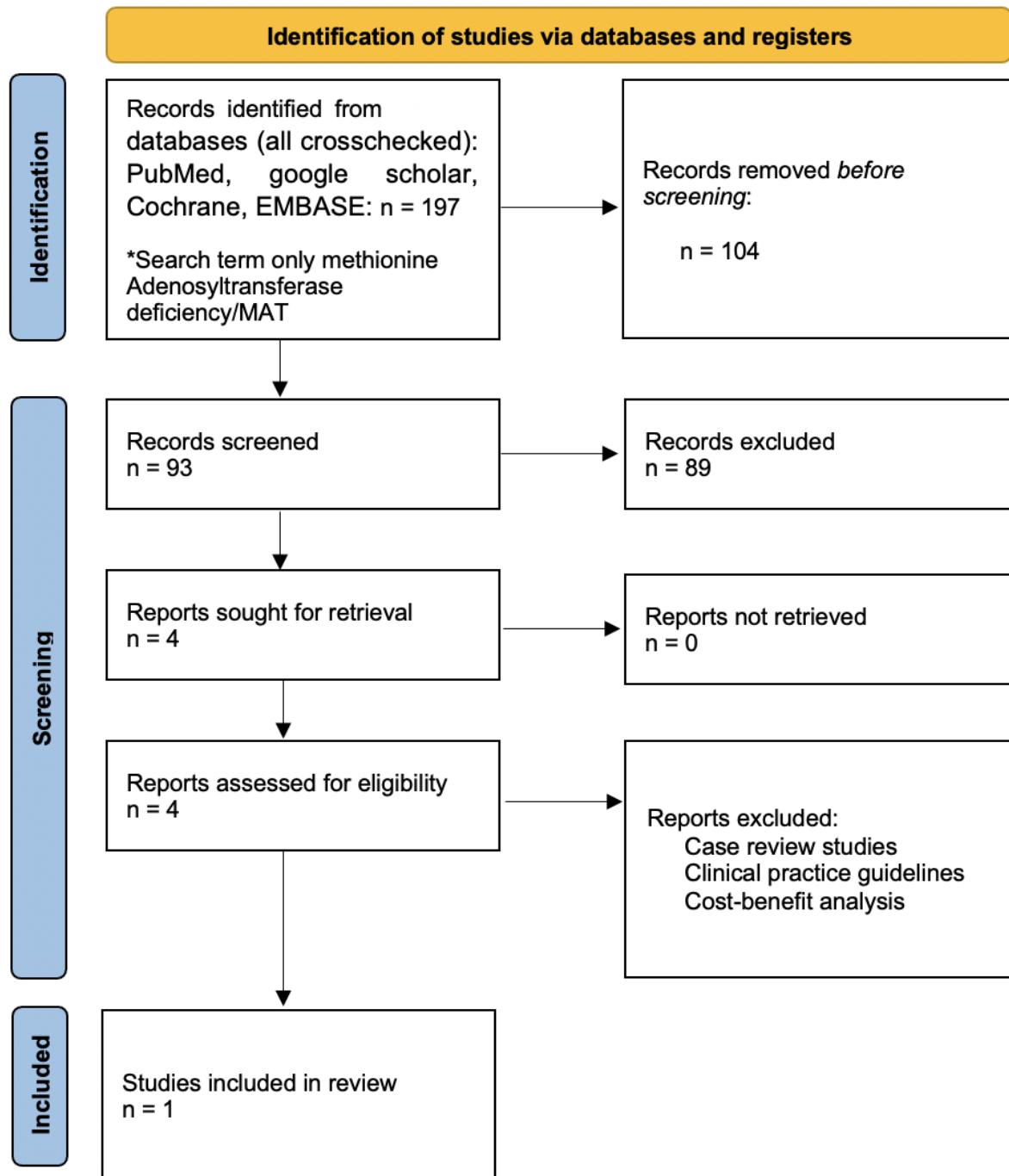


Figure 2. PRISMA for research question on screening for MAT deficiency among asymptomatic newborns

APPENDIX D. Characteristics of Included Studies

Table 6. Summary of Study Characteristics of Included Studies (HCU and MAT)

Study ID	Study Design	Setting	Index Test	Population	Reference Standard
1998 Yap	Retrospective	Ireland	Met on DBS	1.58M asymptomatic newborns	Plasma methionine and free homocysteine
2010 Gan-Schreier	Prospective	Qatar	tHcy on DBS Met on DBS	46,406 asymptomatic newborns	Molecular testing, mutational analysis
2013 Couce	Retrospective	Spain	Met on DBS	68,624 asymptomatic newborns	Molecular testing

APPENDIX E. QUADAS-2

Table 7. QUADAS-2 for Yap et al (1998)

Study ID Yap et al., 1998 Methodological quality			
Item	Authors' judgment	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	consecutive	Low, interpretation blinded to result of gold standard	
Was a case-control design avoided?	yes	low	
Did the study avoid inappropriate exclusions?	yes	low	
Could the selection of patients have introduced bias?	no	low	
Are there concerns that the included patients and setting do not match the review question?	no	low	
DOMAIN 2: Index Test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	low	
If a threshold was used, was it pre-specified?	Yes	low	
Could the conduct or interpretation of the index test have introduced bias?	no	low	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	no	low	
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes	low	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Not mentioned	unknown	
Could the reference standard, its conduct, or its interpretation have introduced bias?	Not mentioned	unknown	
Are there concerns that the target condition as defined by the reference standard does not match the question?	No	low	
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes	low	
Did all patients receive the same reference standard?	Yes	Low	
Were all patients included in the analysis?	Yes	low	
Could the patient flow have introduced bias?	No	Low	

Table 8. QUADAS-2 for Gan-Schreier et al (2010)

Study ID Gan-Schreier et al, 2010 Methodological quality			
Item	Authors' judgment	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	consecutive	Low, interpretation blinded to result of gold standard	
Was a case-control design avoided?	yes	low	
Did the study avoid inappropriate exclusions?	yes	low	
Could the selection of patients have introduced bias?	no	low	
Are there concerns that the included patients and setting do not match the review question?	no	low	
DOMAIN 2: Index Test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	low	
If a threshold was used, was it pre-specified?	Yes	low	
Could the conduct or interpretation of the index test have introduced bias?	no	low	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	no	low	
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes	low	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Not mentioned	unknown	
Could the reference standard, its conduct, or its interpretation have introduced bias?	Not mentioned	unknown	
Are there concerns that the target condition as defined by the reference standard does not match the question?	No	low	
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes	No	
Did all patients receive the same reference standard?	Yes	Low	
Were all patients included in the analysis?	Yes	low	
Could the patient flow have introduced bias?	No	No	

Table 9. QUADAS-2 for Couce et al (2013)

Study ID Couce et al, 2013 Methodological quality			
Item	Authors' judgment	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	consecutive	Low	
Was a case-control design avoided?	yes	low	
Did the study avoid inappropriate exclusions?	yes	low	
Could the selection of patients have introduced bias?	no	low	
Are there concerns that the included patients and setting do not match the review question?	no	low	
DOMAIN 2: Index Test			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	low	
If a threshold was used, was it pre-specified?	Yes	low	
Could the conduct or interpretation of the index test have introduced bias?	no	low	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?	no	low	
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes	low	
Were the reference standard results interpreted without knowledge of the results of the index tests?	Not mentioned	unknown	
Could the reference standard, its conduct, or its interpretation have introduced bias?	Not mentioned	unknown	
Are there concerns that the target condition as defined by the reference standard does not match the question?	No	low	
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes	Low	
Did all patients receive the same reference standard?	Yes	Low	
Were all patients included in the analysis?	Yes	low	
Could the patient flow have introduced bias?	No	No	

APPENDIX F. Summary of GRADE Quality of Evidence: Screening for Homocystinuria using Hcy on DBS

Table 10. Summary of GRADE Quality of Evidence: Screening for Homocystinuria using Hcy on DBS (Gan-Schreier)

Sensitivity	1.00 (95% CI: -- to --)				Prevalences	0.000005%	0.000003%	
Outcome	No of studies (No of patients)	Study design	Factors that may decrease certainty of evidence					Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	
True positives (patients with Homocystinuria)	1 study 14 patients (TP+FN)	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW
False negatives (patients incorrectly classified as not having Homocystinuria)	Total = 46,406							
True negatives (patients without Homocystinuria)	1 study 139 patients*	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW
False positives (patients incorrectly classified as having Homocystinuria)	Total = 46,406							

Explanations

*no true negatives reported

a. Downgraded for incomplete follow-up

b. Downgraded due to very low event rate, with a very large population size; Confidence intervals were not included as well in the study results.

APPENDIX G. Summary of GRADE Quality of Evidence: Screening for Homocystinuria using Met on DBS (Gan-Schreier)

Table 11. Summary of GRADE Quality of Evidence: Screening for Homocystinuria using Met on DBS (Gan-Schreier)

Sensitivity	0.5 (95% CI: -- to --)				Prevalences		0.000005%	0.000003%	
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Test accuracy CoE	
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias		
True positives (patients with Homocystinuria)	1 study 14 patients (TP+FN) Total = 46,406	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW	
False negatives (patients incorrectly classified as not having Homocystinuria)									
True negatives (patients without Homocystinuria)	1 study 0 patient* Total = 46,406	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW	
False positives (patients incorrectly classified as having Homocystinuria)									

Explanations

*no true negatives reported

a. Downgraded for incomplete follow-up

b. Downgraded due to very low event rate, with a very large population size; Confidence intervals were not included as well in the study results.

APPENDIX H. Summary of GRADE Quality of Evidence: Screening for Homocystinuria using Met on DBS (Yap)

Table 12. Summary of GRADE Quality of Evidence: Screening for Homocystinuria deficiency using Met on DBS (Yap)

Sensitivity	0.84 (95% CI: -- to --)				Prevalences	0.000005%	0.000003%	
Outcome	No of studies (No of patients)	Study design	Factors that may decrease certainty of evidence					Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	
True positives (patients with Homocystinuria)	1 study 25 patients (TP+FN)	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW
False negatives (patients incorrectly classified as not having Homocystinuria)	Total = 1,500,000							
True negatives (patients without Homocystinuria)	1 study 0 patient*	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW
False positives (patients incorrectly classified as having Homocystinuria)	Total = 1,500,000							

Explanations

*no true negatives reported

a. Downgraded for incomplete follow-up

b. Downgraded due to very low event rate, with a very large population size; Confidence intervals were not included as well in the study results.

APPENDIX I. Summary of GRADE Quality of Evidence: Screening for MAT Deficiency using Met on DBS

Table 13. Summary of GRADE Quality of Evidence: Screening for MAT deficiency using Met on DBS

Sensitivity	1.00 (95% CI: -- to --)	Prevalences	0.000036%	0.000009%	
Specificity	-- (95% CI: -- to --)				

Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	
True positives (patients with MAT deficiency)	1 study 18 patients (TP+FN)	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW
	Total = 68,624							
False negatives (patients incorrectly classified as not having MAT deficiency)	1 study 0 patient*	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW
True negatives (patients without MAT deficiency)	1 study 0 patient*	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW
False positives (patients incorrectly classified as having MAT deficiency)								

Explanations

*no true negatives reported

a. Downgraded for incomplete follow-up

b. Downgraded due to very low event rate, with a very large population size; Confidence intervals were not included as well in the study results.

APPENDIX J. AGREE II: Cystathionine beta-synthase deficiency

Table 14. AGREE II: Guidelines for the diagnosis and management of cystathionine beta-synthase deficiency

2016 Morris			
Guidelines for the diagnosis and management of cystathionine beta-synthase deficiency (HCY)			
Domain	Score (1-7)	Comments	
Domain 1. Scope and Purpose			
1. The overall objective/s of the guidelines is/are specifically described.	7	7	p. 51
2. The health question/s covered by the guideline is/are specifically described.	2	7	Vaguely presented The target population, screening, treatment and outcomes were described in the guideline.
3. The population to whom the guideline is meant to apply is specifically described	6	7	p. 50-51; the guideline aims to consider all patients with CBS deficiency, though some sections will only apply to subgroups.
Maximum possible score: 42, minimum possible score: 6, (36)	83%		
Domain 2. Stakeholder Involvement			
4. The guideline development group includes individuals from all relevant professional groups	5	7	p. 51; a Guideline Development Group (GDG) was convened, including pediatricians, adult physicians, dieticians, biochemists, a clinical geneticist and a statistician. *no specifics on the actual professionals involved
5. The views and preferences of the target population have been sought.	7	7	p. 51
6. The target users of the guideline are clearly defined.	7	7	p.51
Maximum possible score: 42, minimum possible score: 6, (36)	94.4%		
Domain 3. Rigour of Development			
7. Systematic methods were used to search for evidence.	7	5	p. 51-52 Electronic database and the time period identified however the full search strategy and terms used were not explicitly stated.
8. The criteria for selecting the evidence are clearly described.	2	3	Though they stated the methodology used for collecting the evidence base for this guideline, inclusion and exclusion criteria and characteristics of included studies were not explicitly stated.
9. The strength and limitations of the body of evidence are clearly described.	4	3	p. 51 Evidence levels were classified in accordance with the SIGN methodology (Table 1) and recommendations given in the guideline were graded depending on their level of evidence. However, methodological quality of the studies was not described.
10. The methods for formulating the recommendations are clearly described.	2	3	p. 51-52; mentioned as a general statement Quote: "Statements and supporting evidence were prepared and discussed with the full group at two further meetings". It is unclear whether decisions were reached through voting system, informal consensus, or formal consensus techniques (e.g., Delphi, Glaser techniques)
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	7	7	Benefits of early detection/screening and intervention were stated as well as side effect of treatment.
12. There is explicit link between the recommendations and the supporting evidence.	6	7	Each recommendation is linked to their supporting evidences in the paragraph below.
13. The guideline has been externally reviewed by experts prior to its publication.	7	7	Quote: "Further revision was made following comments from a patient support group representative (Tara Morrison, Director and Chair of HCU Network Australia) and two highly

			renowned external reviewers (James V. Leonard and Bridget Wilcken page 51
14. A procedure for updating the guideline is provided.	2	7	p. 69; envisioned revising the guidelines in 5 years to incorporate data from studies and other advances
Maximum possible score: 112, minimum possible score: 16, (96)	65%		
Domain 4. Clarity of Presentation			
15. The recommendations are specific and unambiguous.	7	7	Recommendations are specific and unambiguous
16. The different options for management of the condition or health issue are clearly presented.	7	7	Management options discussed in page 58-63
17. Key recommendations are easily identifiable.	7	7	Key recommendations easily recognizable in the article (<i>italicized</i>)
Maximum possible score: 42, minimum possible score: 6, (36)	100%		
Domain 5. Applicability			
18. The guideline describes facilitators and barriers to its application.	1	4	Not discussed thoroughly; Although the consensus was able to get feedback from key stakeholders, the barriers, guideline utilization, and quality indicators were not identified
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	1	1	Not discussed thoroughly; Dissemination/implementation of the guideline not stated
20. The potential resource implications of applying the recommendations have been considered.	1	2	Not discussed thoroughly; Though they were not able to find studies on QoL, no health economics evaluation or cost effectiveness analysis was done.
21. The guideline presents monitoring and/or auditing criteria.	1	3	Not discussed thoroughly; Quote: "The guidelines should be used to identify the areas in which data are lacking and to facilitate the design of international collaborative studies, using the same diagnostic measures and therapies." Page 69
Maximum possible score: 56, minimum possible score: 8, (48)	12.5%		
Domain 6. Editorial Independence			
22. The views of the funding body have not influenced the content of the guideline.	7	7	Quote: "The views of the funding bodies have not influenced the content of the guideline" page 69
23. Competing interests of guideline development group members have been recorded and addressed.	7	7	Guideline development group's declaration of competing interests explicitly stated Page 69
Maximum possible score: 28, minimum possible score: 4, (24)	100%		
OVERALL GUIDELINE ASSESSMENT Max – 322, Min – 46, (276)	239/322 (74%) 69.9%		
Rate the overall quality of this guideline.	4	6	
I would recommend this guideline for use	YES YES, with modifications NO		
Notes:	Rigour of Dev't was only at 65%, and Applicability was at 12.5%. Overall Guideline assessment was below 80%		

APPENDIX K. AGREE II: Cystathione Beta-synthase Deficiency

Table 15. AGREE II: Guidelines for the diagnosis and management of cystathionine beta-synthase deficiency

2015 Huemer			
Newborn screening for homocystinurias and methylation disorders: systematic review and proposed guidelines			
Domain	Score (1-7)	Comments	
Domain 1. Scope and Purpose			
1. The overall objective/s of the guidelines is/are specifically described.	7	7	p. 2 Quote: The aims above are pursued in this report with respect to guidelines for NBS for the homocystinurias and methylation disorders developed on the basis of a systematic review of the literature, current NBS practice and expert discussions."
2. The health question/s covered by the guideline is/are specifically described.	6	5	p.2; Expert consensus was targeted for each disease on the following key questions: a) is the natural course of the disease severe, b) is treatment generally beneficial, c) is early intervention more effective, d) are robust, valid and reliable methods, screening approaches and strategies available The target population, screening and outcome, were described in the guideline but information on outcomes is limited
3. The population to whom the guideline is meant to apply is specifically described	1	7	Not explicitly mentioned or defined, only The population to whom the guideline is meant to apply is specifically described.
Maximum possible score: 42, minimum possible score: 6, (36)	75%		
Domain 2. Stakeholder Involvement			
4. The guideline development group includes individuals from all relevant professional groups	2	1	p. 2; Implied, but not specified
5. The views and preferences of the target population have been sought.	1	1	Not stated in the consensus
6. The target users of the guideline are clearly defined.	1	4	The target users of the guideline are not explicitly stated but can be inferred
Maximum possible score: 42, minimum possible score: 6, (36)	11.1%		
Domain 3. Rigour of Development			
7. Systematic methods were used to search for evidence.	7	7	p. 2
8. The criteria for selecting the evidence are clearly described.	5	3	p. 2 The electronic database and key terms were stated however the time period and full search strategy was not reported
9. The strength and limitations of the body of evidence are clearly described.	1	2	Though they stated the methodology used for collecting the evidence base for this guideline, inclusion and exclusion criteria and characteristics of included studies were not explicitly stated.
10. The methods for formulating the recommendations are clearly described.	7	6	p. 2; the guidelines were established based on the literature evaluation and expert discussions following the GRADE approach. Expert consensus was targeted for each disease for each disease on key questions. Quote: "study design and level of evidence was evaluated for 103 publications according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group"

11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	5	4	Benefits of early screening and intervention were stated however side effects and risks not discussed.		
12. There is explicit link between the recommendations and the supporting evidence.	6	7	Each recommendation is linked to their supporting evidences in the paragraph above.		
13. The guideline has been externally reviewed by experts prior to its publication.	2	1	Implied but not explicitly stated; No mention of external reviewers of the guidelines prior to the publication		
14. A procedure for updating the guideline is provided.	1	1	Not stated		
Maximum possible score: 112, minimum possible score: 16, (96)	51.0%				
Domain 4. Clarity of Presentation					
15. The recommendations are specific and unambiguous.	5	7	Can be improved by specifying intended population; Recommendations are specific.		
16. The different options for management of the condition or health issue are clearly presented.	5	7	Different management options are discussed per condition		
17. Key recommendations are easily identifiable.	6	7	Specific key recommendations are grouped together per disease entity.		
Maximum possible score: 42, minimum possible score: 6, (36)	86.1%				
Domain 5. Applicability					
18. The guideline describes facilitators and barriers to its application.	1	1	Feedback from key stakeholders, the barriers, guideline utilization, and quality indicators were not identified		
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	1	1	Dissemination/implementation of the guideline not stated		
20. The potential resource implications of applying the recommendations have been considered.	1	1	No health economics evaluation or cost effectiveness analysis was done.		
21. The guideline presents monitoring and/or auditing criteria.	1	1	No statement to assess guideline implementation or criteria for assessing impact of guideline implementation		
Maximum possible score: 56, minimum possible score: 8, (48)	0				
Domain 6. Editorial Independence					
22. The views of the funding body have not influenced the content of the guideline.	1	1	Not indicated in the consensus		
23. Competing interests of guideline development group members have been recorded and addressed.	7	7	No conflict of interest as stated		
Maximum possible score: 28, minimum possible score: 4, (24)	50%				
OVERALL GUIDELINE ASSESSMENT					
Max – 322, Min – 46, (276)					
Rate the overall quality of this guideline.	4	4			
I would recommend this guideline for use	YES YES, with modifications NO				
Notes:	Only Domain 4 (Clarity of Presentation) passed with a grade of 86.1%. Over all Guideline Assessment was only at 48.8%.				

APPENDIX L. AGREE II: Inherited Methylation Disorders

Table 16. AGREE II: Consensus recommendations for the diagnosis, treatment and follow-up of inherited methylation disorders

2016 Baric			
Consensus recommendations for the diagnosis, treatment and follow-up of inherited methylation disorders (MAT)			
Domain	Score (1-7)	Comments	
Domain 1. Scope and Purpose			
1. The overall objective/s of the guidelines is/are specifically described.	6	7	p. 2 Quote: "A main aim of the European network and registry for homocystinurias and methylation defects (E-HOD) project is to provide recommendations for diagnosis and treatment for this rare group of disorders"
2. The health question/s covered by the guideline is/are specifically described.	1	7	Target population was implied but not explicitly defined. The target population, screening, treatment and outcomes were described in the guideline.
3. The population to whom the guideline is meant to apply is specifically described	3	7	Not explicitly mentioned, only described as a generally population of persons with elevated methionine The population to whom the guideline is meant to apply is specifically described
Maximum possible score: 42, minimum possible score: 6, (36)	69%		
Domain 2. Stakeholder Involvement			
4. The guideline development group includes individuals from all relevant professional groups	4	7	p. 2; general statement - To assure optimal expertise, experts within the E-HOD network and authors of publications related to patients with inherited methylation disorders have been invited to participate in this work. Quote: "To assure optimal expertise, experts within the E-HOD network and authors of publications related to patients with inherited methylation disorders have been invited to participate in this work."
5. The views and preferences of the target population have been sought.	1	1	Not mentioned/stated
6. The target users of the guideline are clearly defined.	1	4	Not clearly defined The target users of the guideline are not explicitly stated but can be inferred
Maximum possible score: 42, minimum possible score: 6, (36)	33.3%		
Domain 3. Rigour of Development			
7. Systematic methods were used to search for evidence.	6	2	p. 3-4 SIGN methodology was used for literature search however the electronic database, time period and full search strategy was not stated
8. The criteria for selecting the evidence are clearly described.	1	2	Not described thoroughly Though they stated the methodology used for collecting the evidence base for this guideline, inclusion and exclusion criteria and characteristics of included studies were not explicitly stated.
9. The strength and limitations of the body of evidence are clearly described.	6	1	p. 3 Methodological quality of the studies was not described
10. The methods for formulating the recommendations are clearly described.	1	1	It is unclear whether decisions were reached through voting system, informal consensus, or formal consensus techniques (e.g., Delphi, Glaser techniques)

11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	5	3	Benefits of treatment were stated however side effects and risks not discussed.
12. There is explicit link between the recommendations and the supporting evidence.	7	7	Each recommendation is linked to their supporting evidences in the paragraph above.
13. The guideline has been externally reviewed by experts prior to its publication.	2	1	Implied, but not described thoroughly; No mention of external reviewers of the guidelines prior to the publication
14. A procedure for updating the guideline is provided.	1	1	Not stated
Maximum possible score: 112, minimum possible score: 16, (96)	32.3%		
Domain 4. Clarity of Presentation			
15. The recommendations are specific and unambiguous.	6	7	Recommendations are specific and unambiguous
16. The different options for management of the condition or health issue are clearly presented.	7	7	p. 5-6 Different management options discussed per condition
17. Key recommendations are easily identifiable.	7	7	Specific key recommendations are grouped together per disease entity.
Maximum possible score: 42, minimum possible score: 6, (36)	97.2%		
Domain 5. Applicability			
18. The guideline describes facilitators and barriers to its application.	1	1	Feedback from key stakeholders, the barriers, guideline utilization, and quality indicators were not identified
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	1	1	dissemination/implementation of the guideline not stated
20. The potential resource implications of applying the recommendations have been considered.	1	1	No health economics evaluation or cost effectiveness analysis was done.
21. The guideline presents monitoring and/or auditing criteria.	1	1	No statement to assess guideline implementation or criteria for assessing impact of guideline implementation
Maximum possible score: 56, minimum possible score: 8, (48)	0%		
Domain 6. Editorial Independence			
22. The views of the funding body have not influenced the content of the guideline.	1	1	Not mentioned/indicated
23. Competing interests of guideline development group members have been recorded and addressed.	7	7	p. 14; no conflicts of interest
Maximum possible score: 28, minimum possible score: 4, (24)	50%		
OVERALL GUIDELINE ASSESSMENT			
Max – 322, Min – 46, (276)	161/322 (50%) 41.7%		
Rate the overall quality of this guideline.	3	4	
I would recommend this guideline for use	YES YES, with modifications NO		
Notes:	Only Domain 4 (Clarity of Presentation) passed with a grade of 97.2%. Over all Guideline Assessment was only at 46.2%.		

7. Newborn Screening for Tyrosinemia

APPENDIX A. Search Yield

*as of August 31, 2021

Search databases (all crosschecked): PubMed, google scholar, Cochrane, EMBASE
References of identified Systematic Reviews scanned

Tyrosinemia: "tyrosinaemia"[All Fields] OR "tyrosinemas"[MeSH Terms] OR "tyrosinemas"[All Fields] OR "tyrosinemia"[All Fields]	1,331
Tyrosinemia: "tyrosinaemia"[All Fields] OR "tyrosinemas"[MeSH Terms] OR "tyrosinemas"[All Fields] OR "tyrosinemia"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields] diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading]	702
newborn: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields] Tyrosinemia: "tyrosinaemia"[All Fields] OR "tyrosinemas"[MeSH Terms] OR "tyrosinemas"[All Fields] OR "tyrosinemia"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields] diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading]	265
Tyrosinemia: "tyrosinaemia"[All Fields] OR "tyrosinemas"[MeSH Terms] OR "tyrosinemas"[All Fields] OR "tyrosinemia"[All Fields] screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields] diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading] asymptomatic: "asymptomatic"[All Fields] OR "asymptomatically"[All Fields] OR "asymptomatics"[All Fields] newborns: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields]	11

<p>practice guidelines: "practice guideline"[Publication Type] .or. "practice guidelines as topic"[MeSH Terms] .or. "practice guidelines"[All Fields]</p> <p>newborn: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields]</p> <p>Tyrosinemia: "tyrosinaemia"[All Fields] OR "tyrosinemias"[MeSH Terms] OR "tyrosinemias"[All Fields] OR "tyrosinemia"[All Fields]</p> <p>screening: "diagnosis"[Subheading] OR "diagnosis"[All Fields] OR "screening"[All Fields] OR "mass screening"[MeSH Terms] OR ("mass"[All Fields] AND "screening"[All Fields]) OR "mass screening"[All Fields] OR "early detection of cancer"[MeSH Terms] OR ("early"[All Fields] AND "detection"[All Fields] AND "cancer"[All Fields]) OR "early detection of cancer"[All Fields] OR "screen"[All Fields] OR "screenings"[All Fields] OR "screened"[All Fields] OR "screens"[All Fields]</p> <p>diagnosis: "diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[Subheading]</p>	3
--	---

<p>tyrosinemia: "tyrosinaemia"[All Fields] OR "tyrosinemias"[MeSH Terms] OR "tyrosinemias"[All Fields] OR "tyrosinemia"[All Fields]</p> <p>asymptomatic: "asymptomatic"[All Fields] OR "asymptomatically"[All Fields] OR "asymptomatics"[All Fields]</p> <p>newborns: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields]</p> <p>tyrosine: "tyrosin"[All Fields] OR "tyrosinate"[All Fields] OR "tyrosinates"[All Fields] OR "tyrosine"[MeSH Terms] OR "tyrosine"[All Fields] OR "tyrosine's"[All Fields] OR "tyrosines"[All Fields]</p> <p>levels: "level"[All Fields] OR "levels"[All Fields]</p>	5
--	---

succinylacetone: "succinylacetone"[Supplementary Concept] OR "succinylacetone"[All Fields] accuracy: "accuracies"[All Fields] OR "accuracy"[All Fields]	6
tyrosinemia: "tyrosinaemia"[All Fields] OR "tyrosinemas"[MeSH Terms] OR "tyrosinemias"[All Fields] OR "tyrosinemia"[All Fields] asymptomatic: "asymptomatic"[All Fields] OR "asymptomatically"[All Fields] OR "asymptomatics"[All Fields] newborns: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields] succinylacetone: "succinylacetone"[Supplementary Concept] OR "succinylacetone"[All Fields] accuracy: "accuracies"[All Fields] OR "accuracy"[All Fields]	2

<p>succinylacetone: "succinylacetone"[Supplementary Concept] OR "succinylacetone"[All Fields] diagnostic: "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnostic"[All Fields] OR "diagnostical"[All Fields] OR "diagnostically"[All Fields] OR "diagnostics"[All Fields] accuracy: "accuracies"[All Fields] OR "accuracy"[All Fields] sensitivity: "hypersensitivity"[MeSH Terms] OR "hypersensitivity"[All Fields] OR "sensitive"[All Fields] OR "sensitively"[All Fields] OR "sensitives"[All Fields] OR "sensitivities"[All Fields] OR "sensitivity and specificity"[MeSH Terms] OR ("sensitivity"[All Fields] AND "specificity"[All Fields]) OR "sensitivity and specificity"[All Fields] OR "sensitivity"[All Fields]</p>	49
---	----

<p>specificity: "sensitivity and specificity"[MeSH Terms] OR ("sensitivity"[All Fields] AND "specificity"[All Fields]) OR "sensitivity and specificity"[All Fields] OR "specificity"[All Fields] OR "specific"[All Fields] OR "specifically"[All Fields] OR "specification"[All Fields] OR "specifications"[All Fields] OR "specificities"[All Fields] OR "specifics"[All Fields] OR "specificities"[All Fields] OR "specify"[All Fields]</p> <p>tyrosinemia: "tyrosinaemia"[All Fields] OR "tyrosinemas"[MeSH Terms] OR "tyrosinemias"[All Fields] OR "tyrosinemia"[All Fields]</p>	
<p>succinylacetone: "succinylacetone"[Supplementary Concept] OR "succinylacetone"[All Fields]</p> <p>diagnostic: "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnostic"[All Fields] OR "diagnostical"[All Fields] OR "diagnostically"[All Fields] OR "diagnostics"[All Fields]</p> <p>accuracy: "accuracies"[All Fields] OR "accuracy"[All Fields]</p> <p>sensitivity: "hypersensitivity"[MeSH Terms] OR "hypersensitivity"[All Fields] OR "sensitive"[All Fields] OR "sensitively"[All Fields] OR "sensitives"[All Fields] OR "sensitivities"[All Fields] OR "sensitivity and specificity"[MeSH Terms] OR ("sensitivity"[All Fields] AND "specificity"[All Fields]) OR "sensitivity and specificity"[All Fields] OR "sensitivity"[All Fields]</p> <p>specificity: "sensitivity and specificity"[MeSH Terms] OR ("sensitivity"[All Fields] AND "specificity"[All Fields]) OR "sensitivity and specificity"[All Fields] OR "specificity"[All Fields] OR "specific"[All Fields] OR "specifically"[All Fields] OR "specification"[All Fields] OR "specifications"[All Fields] OR "specificities"[All Fields] OR "specifics"[All Fields] OR "specificities"[All Fields] OR "specify"[All Fields]</p> <p>tyrosinemia: "tyrosinaemia"[All Fields] OR "tyrosinemas"[MeSH Terms] OR "tyrosinemias"[All Fields] OR "tyrosinemia"[All Fields]</p> <p>newborn: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "newborn"[All Fields] OR "newborns"[All Fields] OR "newborn's"[All Fields]</p> <p>neonate: "infant, newborn"[MeSH Terms] OR ("infant"[All Fields] AND "newborn"[All Fields]) OR "newborn infant"[All Fields] OR "neonatal"[All Fields] OR "neonate"[All Fields] OR "neonates"[All Fields] OR "neonatality"[All Fields] OR "neonatals"[All Fields] OR "neonate's"[All Fields]</p>	30

APPENDIX B. Characteristics of the Included Studies

Table 5. Summary of Study Characteristics of Included Studies in 2017 Geppert Systematic Review

SYSTEMATIC REVIEW on treatment				
Study ID	Study Design	Intervention	Outcomes	Population/Studies Included
2017 Geppert	Evaluation of pre-symptomatic nitisinone treatment on long-term outcomes in Tyrosinemia I patients	Nitisinone	Death Need for liver transplant Neurological crisis (post hoc analysis)	6 cohort studies, 1 cross-sectional study
Included Studies in Geppert 2017				
2014 Bartlett	Prospective cohort	Nitisinone	Need for liver transplant	N = 38
2015 McKiernan	Retrospective cohort	Nitisinone	Death Need for liver transplant	N = 17
2008 Santra	Retrospective cohort	Nitisinone	Acute/chronic liver failure/disease Renal disease	N = 21
2012 Larochelle	Cohort	Nitisinone	Death Need for liver transplant Neurological symptoms	N = 78
2015 Simoncelli	Retrospective cohort	Nitisinone	Need for liver transplant Neurological symptoms	N = 95
2014 Mayorandan	Retrospective cohort	Nitisinone	Hepatorenal outcomes	N = 168
2016 Van Ginkel	Cross-sectional	Nitisinone	Neurocognitive outcomes	N = 38

APPENDIX C. Characteristics of the Included Studies

Table 6. Summary of Study Characteristics of Included Studies in 2017 Stinton Systematic Review

SYSTEMATIC REVIEW on screening					
Study ID	Study Design	Setting	Index Test	Population/Studies Included	Reference Standard
2017 Stinton	Systematic Review of Test of Accuracy	England	Succinylacetone	5 observational studies on screening experience 5 case-control studies	variable
Included Studies in Stinton 2017					
2011 La Marca	Observational (retrospective)	Italy	Succinylacetone	136,075 newborns	DNA analysis
2012 Lund	Retrospective descriptive	Denmark	Succinylacetone	504,049 newborns	Urine organic acid analysis
2011 Morrissey	Retrospective descriptive	New York	Succinylacetone	500,000 newborns	Plasma amino acid analysis
2006 Sander	Retrospective descriptive	Germany	Succinylacetone	61,344 newborns	Plasma amino acid analysis
2013 Zytkovicz	Retrospective descriptive	England	Succinylacetone	518,687 newborns	DNA analysis
2004 Allard	Case-control	Quebec	Succinylacetone	4,002: 3 cases, 3199 controls	-
2011 Dhillon	Case-control	USA	Succinylacetone	~1,026: 6 cases, ~1,000 controls	-
2008 La Marca	Case-control	Italy	Succinylacetone	13,006: 6 cases, 13,000 controls	-
2012 Metz	Case-control	Europe	Succinylacetone	4,686: 3 cases, 4683 controls	-
2008 Turgeon	Case-control	North America	Succinylacetone	13,532: 11 cases, 13,521 controls	-

APPENDIX D. AMSTAR 2

2017 Geppert	Evaluation of pre-symptomatic nitisinone treatment on long-term outcomes in Tyrosinemia type 1 patients: a systematic review	<p>Notes:</p> <p>AMSTAR 2: MODERATE</p> <p>The applicability of findings to the screening context or clinical practice is limited as not all early-treated patients were identified by screening and late-treated groups included patients born prior to the availability of nitisinone.</p> <p>Was only able to find observational studies With overlapping population</p> <p>*No new studies after 2016</p>
--------------	--	--

Table 7. AMSTAR 2 Checklist: Evaluation of pre-symptomatic nitisinone treatment on long-term outcomes in Tyrosinemia type 1 patients: a systematic review

1. Did the research questions and inclusion criteria for the review include the components of PICO?			
<p>For Yes:</p> <p style="margin-left: 20px;">✓ Population</p> <p style="margin-left: 20px;">✓ Intervention</p> <p style="margin-left: 20px;">✓ Comparator group</p> <p style="margin-left: 20px;">✓ Outcome</p>			
<p>Optional (recommended)</p> <p style="margin-left: 40px;"><input type="checkbox"/> Timeframe for follow-up</p> <p style="margin-left: 40px;"><input type="checkbox"/> Yes</p> <p style="margin-left: 40px;"><input type="checkbox"/> No</p>			
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?			
<p>For Partial Yes:</p> <p>The authors state that they had a written protocol or guide that included ALL the following:</p> <ul style="list-style-type: none"> <input type="checkbox"/> review question(s) <input type="checkbox"/> a search strategy <input type="checkbox"/> inclusion/exclusion criteria <input type="checkbox"/> a risk of bias assessment 		<p>For Yes:</p> <p>As for partial yes, plus the protocol should be registered and should also have specified:</p> <ul style="list-style-type: none"> <input type="checkbox"/> a meta-analysis/synthesis plan, if appropriate, <i>and</i> <input type="checkbox"/> a plan for investigating causes of heterogeneity <input type="checkbox"/> justification for any deviations from the protocol <p style="text-align: right;"><input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No</p>	
3. Did the review authors explain their selection of the study designs for inclusion in the review?			
<p>For Yes, the review should satisfy ONE of the following:</p> <ul style="list-style-type: none"> <input type="checkbox"/> <i>Explanation for including only RCTs</i> ✓ OR Explanation for including only NRSI <input type="checkbox"/> <i>OR Explanation for including both RCTs and NRSI</i> <p style="text-align: right;"><input type="checkbox"/> Yes <input type="checkbox"/> No</p>			
4. Did the review authors use a comprehensive literature search strategy?			
<p>For Partial Yes (all the following):</p> <ul style="list-style-type: none"> <input type="checkbox"/> searched at least 2 databases (relevant to research question) <input type="checkbox"/> provided key word and/or search strategy <input type="checkbox"/> justified publication restrictions 		<p>For Yes, should also have (all the following):</p> <ul style="list-style-type: none"> ✓ searched the reference lists / bibliographies of included studies ✓ searched trial/study registries ✓ included/consulted content experts in the field <p style="text-align: right;"><input type="checkbox"/> Yes</p> <p style="text-align: right;"><input type="checkbox"/> Partial Yes <input type="checkbox"/> No</p>	

- (e.g. language)
- ✓ where relevant, searched for grey literature
 - ✓ conducted search within 24 months of completion of the review

5. Did the review authors perform study selection in duplicate?

For Yes, either ONE of the following:

- at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include
- OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.

Yes
 No

6. Did the review authors perform data extraction in duplicate?

For Yes, either ONE of the following:

- ✓ at least two reviewers achieved consensus on which data to extract from included studies
- OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.

Yes
 No

7. Did the review authors provide a list of excluded studies and justify the exclusions?

For Partial Yes:

- provided a list of all potentially relevant studies that were read in full-text form but excluded from the review

For Yes, must also have:

- ✓ Justified the exclusion from the review of each potentially relevant study

Yes
 Partial Yes
 No

8. Did the review authors describe the included studies in adequate detail?

For Partial Yes (ALL the following):

- described populations
- described interventions
- described comparators
- described outcomes
- described research designs

For Yes, should also have ALL the following:

- ✓ described population in detail
- ✓ described intervention in detail (including doses where relevant)
- ✓ described comparator in detail (including doses where relevant)
- ✓ described study's setting
- ✓ timeframe for follow-up

Yes
 Partial Yes
 No

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

RCTs

For Partial Yes, must have assessed RoB from

- unconcealed allocation, and
- lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)

For Yes, must also have assessed RoB from:

- allocation sequence that was not truly random, and
- selection of the reported result from among multiple measurements or analyses of a specified outcome

Yes
 Partial Yes
 No
 Includes only NRSI

NRSI							
For Partial Yes, must have assessed RoB:	For Yes, must also have assessed RoB:						
<input type="checkbox"/> from confounding, and <input type="checkbox"/> from selection bias	<input checked="" type="checkbox"/> methods used to ascertain exposures and outcomes, and selection of the reported result from among multiple measurements or analyses of a specified outcome						
10. Did the review authors report on the sources of funding for the studies included in the review?		<input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No <input type="checkbox"/> Includes only RCTs					
<p>For Yes</p> <table> <tr> <td><input type="checkbox"/> Must have reported on the sources of funding for individual studies included</td> <td><input type="checkbox"/> Yes in the review. Note:</td> </tr> <tr> <td>Reporting that the reviewers looked for this information</td> <td><input type="checkbox"/> No but it was not reported by study authors also qualifies</td> </tr> </table>				<input type="checkbox"/> Must have reported on the sources of funding for individual studies included	<input type="checkbox"/> Yes in the review. Note:	Reporting that the reviewers looked for this information	<input type="checkbox"/> No but it was not reported by study authors also qualifies
<input type="checkbox"/> Must have reported on the sources of funding for individual studies included	<input type="checkbox"/> Yes in the review. Note:						
Reporting that the reviewers looked for this information	<input type="checkbox"/> No but it was not reported by study authors also qualifies						
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?							
RCTs							
For Yes:	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted						
<input type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present. <input type="checkbox"/> AND investigated the causes of any heterogeneity							
For NRSI							
For Yes:	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted						
<input type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present <input type="checkbox"/> AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available <input type="checkbox"/> AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review							
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?							
For Yes:	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted						
<input type="checkbox"/> included only low risk of bias RCTs <input type="checkbox"/> OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.							
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?							
For Yes:	<input type="checkbox"/> Yes <input type="checkbox"/> No						
<input type="checkbox"/> included only low risk of bias RCTs <input checked="" type="checkbox"/> OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results							
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?							

For Yes:

- There was no significant heterogeneity in the results
 OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review

Yes
 No

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

For Yes:

- performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias

Yes
 No
 No meta-analysis conducted

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

For Yes:

- The authors reported no competing interests OR
 The authors described their funding sources and how they managed potential conflicts of interest

Yes
 No

APPENDIX E. AMSTAR 2

2017 Stinton	Newborn Screening for Tyrosinemia Type 1 using Succinylacetone – A systematic review of test accuracy	Notes: AMSTAR 2: Moderate No meta-analysis/pooling of studies done because of incomplete data, no follow-up of patients who screened negative *no new studies after 2016
--------------	---	--

Table 8. AMSTAR Checklist: Newborn Screening for Tyrosinemia Type 1 using Succinylacetone – A systematic review of test accuracy

<p>1. Did the research questions and inclusion criteria for the review include the components of PICO?</p> <p>For Yes:</p> <p>✓ Population Optional (recommended) <input type="checkbox"/> Timeframe for follow-up <input type="checkbox"/> Yes No</p> <p>✓ Intervention</p> <p>✓ Comparator group</p> <p>✓ Outcome</p>			
<p>2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?</p> <p>For Partial Yes:</p> <p>The authors state that they had a written protocol or guide that included ALL the following:</p> <ul style="list-style-type: none"> <input type="checkbox"/> review question(s) <input type="checkbox"/> a search strategy <input type="checkbox"/> inclusion/exclusion criteria <input type="checkbox"/> a risk of bias assessment <p>For Yes:</p> <p>As for partial yes, plus the protocol should be registered and should also have specified:</p> <ul style="list-style-type: none"> <input type="checkbox"/> a meta-analysis/synthesis plan, if appropriate, and <input type="checkbox"/> a plan for investigating causes of heterogeneity <input type="checkbox"/> justification for any deviations from the protocol <p><input type="checkbox"/> Yes Partial Yes No</p>			
<p>3. Did the review authors explain their selection of the study designs for inclusion in the review?</p> <p>For Yes, the review should satisfy ONE of the following:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Explanation for including only RCTs ✓ OR Explanation for including only NRSI <input type="checkbox"/> OR Explanation for including both RCTs and NRSI <p><input type="checkbox"/> Yes No</p>			
<p>4. Did the review authors use a comprehensive literature search strategy?</p> <p>For Partial Yes (all the following):</p> <ul style="list-style-type: none"> <input type="checkbox"/> searched at least 2 databases (relevant to research question) <input type="checkbox"/> provided key word and/or search strategy <input type="checkbox"/> justified publication restrictions (e.g. language) <p>For Yes, should also have (all the following):</p> <ul style="list-style-type: none"> ✓ searched the reference lists / bibliographies of included studies ✓ searched trial/study registries ✓ included/consulted content experts in the field ✓ where relevant, searched for grey literature ✓ conducted search within 24 months of completion of the review <p><input type="checkbox"/> Yes Partial Yes No</p>			
<p>5. Did the review authors perform study selection in duplicate?</p>			

For Yes, either ONE of the following:

- at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include
- OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.

Yes

No

6. Did the review authors perform data extraction in duplicate?

For Yes, either ONE of the following:

- at least two reviewers achieved consensus on which data to extract from included studies
- OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.

Yes

No

7. Did the review authors provide a list of excluded studies and justify the exclusions?

For Partial Yes:

- provided a list of all potentially relevant studies that were read in full-text form but excluded from the review

For Yes, must also have:

- Justified the exclusion from the review of each potentially relevant study

Yes

Partial Yes

No

8. Did the review authors describe the included studies in adequate detail?

For Partial Yes (ALL the following):

- described populations
 described interventions
 described comparators
 described outcomes
 described research designs

For Yes, should also have ALL the following:

- described population in detail
 described intervention in detail (including doses where relevant)
 described comparator in detail (including doses where relevant)
 described study's setting
 timeframe for follow-up

Yes

Partial Yes

No

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

RCTs

For Partial Yes, must have assessed RoB from

- unconcealed allocation, and
 lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)

For Yes, must also have assessed RoB from:

- allocation sequence that was not truly random, and
 selection of the reported result from among multiple measurements or analyses of a specified outcome

Yes

Partial Yes

No

Includes only NRSI

NRSI			
For Partial Yes, must have assessed RoB:	For Yes, must also have assessed RoB:		
<input checked="" type="checkbox"/> from confounding, and <input checked="" type="checkbox"/> from selection bias	<input type="checkbox"/> methods used to ascertain exposures and outcomes, <i>and</i> <input type="checkbox"/> selection of the reported result from among multiple measurements or analyses of a specified outcome		
10. Did the review authors report o		n the sources of funding for the studies included in the review?	
For Yes <input type="checkbox"/> Must have reported on the sources of funding for individual studies included Reporting that the reviewers looked for this information study authors also qualifies			
<input type="checkbox"/> Yes in the review. Note: <input type="checkbox"/> No but it was not reported by			
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?			
RCTs			
For Yes:	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted		
<input type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present. <input type="checkbox"/> AND investigated the causes of any heterogeneity			
For NRSI			
For Yes:	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted		
<input type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present <input type="checkbox"/> AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available <input type="checkbox"/> AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review			
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?			
For Yes:	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted		
<input type="checkbox"/> included only low risk of bias RCTs <input type="checkbox"/> OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.			
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?			
For Yes:	<input type="checkbox"/> Yes <input type="checkbox"/> No		
<input type="checkbox"/> included only low risk of bias RCTs <input type="checkbox"/> OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results			

14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

For Yes:

- There was no significant heterogeneity in the results
- OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review

Yes

No

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

For Yes:

- performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias

Yes

No

No meta-analysis conducted

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

For Yes:

- The authors reported no competing interests OR
- The authors described their funding sources and how they managed potential conflicts of interest

Yes

No

APPENDIX F. Summary of GRADE Quality of Evidence: Evaluation of pre-symptomatic treatment on long-term outcomes in Tyrosinemia I patients

Table 9. Summary of GRADE Quality of Evidence: Evaluation of pre-symptomatic nitisinone treatment on long-term outcomes in Tyrosinemia I patients

Certainty assessment							Impact	Certainty
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		
2	observational studies	serious ^a	not serious	very serious ^b	very serious ^c	none	No benefit of early nitisinone treatment on mortality was observed in this post hoc analysis (no deaths occurred in either the 24 screen-detected or five symptomatically presenting cases).	 VERY LOW
2	observational studies	serious ^a	not serious	very serious ^b	very serious ^c	none	Analysis indicated no significant difference in death rates between the early and late nitisinone-treated groups ($p = 0.49$ using Fisher exact test). In the Birmingham cohort, no death (0/12) occurred in the pre-symptomatically treated group after a follow-up time of 3-12.5 years, while two of five clinic-ally presenting infants died.	 VERY LOW
3	observational studies	serious ^a	not serious	very serious ^b	very serious ^c	none	Rates of liver transplantation were significantly lower amongst early and late nitisinone-treated patients compared to those who were never treated with nitisinone, and lower in those receiving early nitisinone treatment than late nitisinone treatment.	 VERY LOW
3	observational studies	serious ^a	not serious	very serious ^c	very serious ^c	none	After the exclusion of 26 TYR1 patients born prior to the nitisinone era, benefits were observed of early nitisinone-treatment on the need for liver transplantation (0/24 screened vs 3/5 symptomatically presenting patients needed liver transplantation)	 VERY LOW
2	observational studies	serious ^a	not serious	very serious ^{b,c}	very serious ^c	none	In the Birmingham cohort, some degree of learning difficulty was observed in 4/9 pre-symptomatically treated patients of school age. Authors of the multicentre, cross-sectional study did not identify any significant differences in reaction time or percentage of errors in any of the neuropsychological tasks The international cohort [18] observed impaired psychomotor development in 30/148 (20%), hyperactivity/attention deficit syndrome or behavioural disorders in 12/148 (8%) and learning/language difficulties or dyslexia in 2/148 (1.4%) of all nitisinone-treated TYR1 patients. None of which were associated with time at which nitisinone was initiated.	 VERY LOW
1	observational studies	serious ^a	not serious	very serious ^b	very serious ^c	none	While the proportion of patients with neurological crises differed significantly between the three groups in the initial analysis ($p < 0.001$, Fisher exact test), the post hoc analysis did not demonstrate a difference in the number of patients with neurological crises between screen-detected and symptomatically presenting patients with immediate nitisinone treatment after diagnosis (no neurological crisis occurred in either groups).	 VERY LOW

Explanations

a. Downgraded by 1 due to inability to control confounding factors

b. Downgraded by 2 due to indirectness in population, intervention, and differences in outcomes

c. Downgraded due to low event rates, and inability to generate estimates

APPENDIX G. Summary of GRADE Quality of Evidence: Screening of Tyrosinemia I using Succinylacetone

Table 10. Summary of GRADE Quality of Evidence: Systematic Review on Screening of Tyrosinemia I using Succinylacetone (Lund, Morrissey, Sander – confirmatory of SUAC in urine, plasma)

Sensitivity	1.00 (95% CI: -- to --)				Prevalences	0.000036%	0.000009%	
Outcome	No of studies (No of patients)	Study design	Factors that may decrease certainty of evidence					Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	
True positives (patients with Tyrosinemia)	3 studies 5 patients (TP+FN)	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW
False negatives (patients incorrectly classified as not having Tyrosinemia)	Total = 701,909							
True negatives (patients without Tyrosinemia)	3 studies 1 patient*	cross-sectional (cohort type accuracy study)	serious ^a	not serious	not serious	very serious ^b	none	⊕○○○ VERY LOW
False positives (patients incorrectly classified as having Tyrosinemia)	Total = 701,909							

Explanations

*No true negatives reported, 1 false positive reported across 5 studies

a. Downgraded for incomplete follow-up

b. Downgraded due to very low event rate, with a very large population size; Confidence intervals were not included as well in the study results.

APPENDIX H. Summary of GRADE Quality of Evidence: Screening of Tyrosinemia I using Succinylacetone

Table 11. Summary of GRADE Quality of Evidence: Systematic Review on Screening of Tyrosinemia I using Succinylacetone (La Marca, Zytkovicz – DNA analysis as confirmatory)

Sensitivity	1.00 (95% CI: -- to --)			Prevalences 0.000036% 0.000009%				
Outcome	No of studies (No of patients)	Study design	Factors that may decrease certainty of evidence					Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	
True positives (patients with Tyrosinemia)	2 studies 5 patients (TP+FN)	cross-sectional (cohort type accuracy study)	serious ^a	serious ^b	not serious	very serious ^c	none	⊕○○○ VERY LOW
False negatives (patients incorrectly classified as not having Tyrosinemia)	Total = 654,762							
True negatives (patients without Tyrosinemia)	2 studies 1 patient*	cross-sectional (cohort type accuracy study)	serious ^a	serious ^b	not serious	very serious ^c	none	⊕○○○ VERY LOW
False positives (patients incorrectly classified as having Tyrosinemia)	Total = 654,762							

Explanations

*No true negatives reported, 1 false positive reported across 5 studies

a. Downgraded for incomplete follow-up

b. Downgraded due to inclusion of newborns aged 1 month or more and symptomatic

c. Downgraded due to very low event rate, with a very large population size; Confidence intervals were not included as well in the study results.

APPENDIX I. Summary of GRADE Quality of Evidence: Screening of Tyrosinemia I using Succinylacetone

Table 12. Summary of GRADE Quality of Evidence: Systematic Review on Screening of Tyrosinemia I using Succinylacetone (case-control studies)

Sensitivity	1.00 (95% CI: 0.31 to 1.00)			Prevalences 0.000036% 0.000009%				
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	
True positives (patients with Tyrosinemia)	5 studies 36252 patients	case-control type accuracy study	serious ^a	very serious ^b	not serious	very serious ^c	none	⊕○○○ VERY LOW
False negatives (patients incorrectly classified as not having Tyrosinemia)								
True negatives (patients without Tyrosinemia)	5 studies 36252 patients	case-control type accuracy study	serious ^a	very serious ^b	not serious	very serious ^c	none	⊕○○○ VERY LOW
False positives (patients incorrectly classified as having Tyrosinemia)								

Explanations

*range – one study could not provide value for specificity due to incomplete data on controls

- a. Downgraded due to one study with incomplete report on controls
- b. Downgraded due to inclusion of symptomatic newborns
- c. Downgraded due to very low event rate, with a very large population size; Wide confidence intervals reported for sensitivity.

APPENDIX J. AGREE II: Newborn Blood Spot Screening for Tyrosinemia Type I

Table 13. AGREE II: Newborn Blood Spot Screening for Tyrosinemia Type I in the UK

2017 UK NSC			
Newborn Blood Spot Screening for Tyrosinemia Type I in the UK			
External review against programme appraisal criteria for the UK National Screening Committee (UK NSC)			
Domain	Score (1-7)	Comments	
Domain 1. Scope and Purpose			
1. The overall objective/s of the guidelines is/are specifically described.	7	7	p. 8, 16 : "This review will assess screening for tyrosinaemia type I, a disorder of amino acid metabolism." Page 4
2. The health question/s covered by the guideline is/are specifically described.	7	7	p. 8 Page 5 - 45
3. The population to whom the guideline is meant to apply is specifically described	2	6	Not explicitly defined The guideline to whom the population is meant to apply is not explicitly stated but can be inferred
Maximum possible score: 42, minimum possible score: 6, (36)	83.3%		
Domain 2. Stakeholder Involvement			
4. The guideline development group includes individuals from all relevant professional groups	2	7	Includes list of the members of the review group but no credentials mentioned Stakeholders listed in page 44
5. The views and preferences of the target population have been sought.	1	3	Not mentioned Although stakeholders were listed, it is not explicitly stated in the document if views and preferences of the target population have been sought
6. The target users of the guideline are clearly defined.	1	1	Not defined Target users for this guideline is not clearly defined
Maximum possible score: 42, minimum possible score: 6, (36)	25%		
Domain 3. Rigour of Development			
7. Systematic methods were used to search for evidence.	7	7	p. 8-9, 17-28 Methodology discussed in page 5 and presented in page 48-51
8. The criteria for selecting the evidence are clearly described.	7	7	p. 5, 17-28
9. The strength and limitations of the body of evidence are clearly described.	7	1	Described and summarized after every key question. p. 51-52 Strength and limitations not discussed
10. The methods for formulating the recommendations are clearly described.	1	1	Not clearly described The guideline did not describe the recommendation development process or its outcome (steps used in modified Delphi technique, voting procedures that were considered)
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	7	7	Benefits of early detection through screening and intervention described page 28-35, harm of screening and treatment discussed in page 37-39
12. There is explicit link between the recommendations and the supporting evidence.	7	7	Each criterion is linked to the corresponding evidence
13. The guideline has been externally	4	7	Mentioned in the methodology

reviewed by experts prior to its publication.			Page 48 Quote: "The draft update report was prepared by Bazian Ltd., and then adapted in line with comments from the National Screening Committee"
14. A procedure for updating the guideline is provided.	6	7	Quote "The UK NSC advises Ministers and the NHS in all four UK countries about all aspects of screening policy. Its policies are reviewed on a 3 yearly cycle. Review process is described in detail at http://www.screening.nhs.uk/policyreview page 1
Maximum possible score: 112, minimum possible score: 16, (96)	77%		
Domain 4. Clarity of Presentation			
15. The recommendations are specific and unambiguous.	1	6	Quote: "Newborn screening for tyrosinaemia type I could be considered."
16. The different options for management of the condition or health issue are clearly presented.	6	7	p. 44 Treatment options for tyrosinaemia type I are discussed in page 30-35
17. Key recommendations are easily identifiable.	2	7	Recommendations are not easily identified after summary of evidence Stated in key points page 2
Maximum possible score: 42, minimum possible score: 6, (36)	63.9%		
Domain 5. Applicability			
18. The guideline describes facilitators and barriers to its application.	4	1	p. 7, and inferred after every key question No description of barriers, guideline utilization, and quality indicators found in the article
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	1	5	The guideline described areas which could provide useful avenues for further research and enumerated which studies are needed to determine the feasibility of screening for tyrosinaemia type I in the UK page 48
20. The potential resource implications of applying the recommendations have been considered.	7	7	Cost-effectiveness analyses of screening was discussed page 40-42
21. The guideline presents monitoring and/or auditing criteria.	7	7	p. 43
Maximum possible score: 56, minimum possible score: 8, (48)	64.6%		
Domain 6. Editorial Independence			
22. The views of the funding body have not influenced the content of the guideline.	1	1	No explicit statement that the views or interests of the funding body have not influenced the final recommendations.
23. Competing interests of guideline development group members have been recorded and addressed.	1	1	Not mentioned
Maximum possible score score: 28, minimum 4, (24)	0%		
OVERALL GUIDELINE ASSESSMENT Max – 322, Min – 46, (276)	Total – 213/322 (66%) 60.5%		
Rate the overall quality of this guideline.	4	5	
I would recommend this guideline for use	YES YES, with modifications NO		
Notes:	Only Domain 1 passed (>80%), Rigour of Dev't only at 77%		

APPENDIX K. AGREE II: Diagnosis and treatment of Tyrosinemia I: A US and Canadian consensus group review and recommendations

Diagnosis and treatment of tyrosinemia type I: a US and Canadian consensus group review and recommendations			
Chinsky et al. (2017)			
Domain	Score (1-7)	Comments	
Domain 1. Scope and Purpose			
1. The overall objective/s of the guidelines is/are specifically described.	7	7	p. 2; to establish uniform recommendations for identification and treatment of affected children
2. The health question/s covered by the guideline is/are specifically described.	1	7	Key questions not specified Inferred in the introduction, page 1
3. The population to whom the guideline is meant to apply is specifically described	2	7	Population intended not explicitly identified, but inferred Inferred in the introduction, page 1
Maximum possible score: 42, minimum possible score: 6, (36)	69.4%		
Domain 2. Stakeholder Involvement			
4. The guideline development group includes individuals from all relevant professional groups	5	5	p. 2; General statement: the review and set of recommendations reflect the consensus opinions of a group of US and Canadian professionals comprising metabolic physicians, a psychologists, and a dietitian nutritionist, whose collective experience either directly or indirectly involves over 100 patients with HT-1. Though this consensus included US and Canadian professionals comprising metabolic physicians, a psychologist, and a dietitian nutritionist, other stakeholders should also have been consulted like pediatric subspecialty societies, family physicians, geneticists, patient support/advocacy groups etc.
5. The views and preferences of the target population have been sought.	1	1	No explicit statement in the consensus that the views and preferences of the target population have been sought
6. The target users of the guideline are clearly defined.	1	5	Not defined Not explicitly stated but can be inferred in the article
Maximum possible score: 42, minimum possible score: 6, (36)	33.3%		
Domain 3. Rigour of Development			
7. Systematic methods were used to search for evidence.	7	5	p. 2; methodology Electronic database and the time period identified however the full search strategy and terms used were not explicitly stated
8. The criteria for selecting the evidence are clearly described.	5	3	p. 2; methodology Though they stated their literature search included review of several published consensus approaches as well as recent methods used for similarly rare inborn errors of metabolism, inclusion and exclusion criteria and characteristics of included studies were not stated.
9. The strength and limitations of the body of evidence are clearly described.	2	6	Implied but not clearly described Recommendations were reviewed with AGREE II criteria as a guide. They used a hybrid approach blending evidence- and consensus-based processes for recommendations regarding inborn errors of metabolism.
10. The methods for formulating the recommendations are clearly described.	7	7	p. 2; methodology Quote: "Subsequently, recommendations were voted on and scored, with only recommendations reaching

			a level of agreement of 70% (strongly agree or agree) included in the present article"
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	7	7	Benefits and side effects of NTBC drug therapy and nutritional therapy stated page 6 - 9
12. There is explicit link between the recommendations and the supporting evidence.	7	7	Each recommendation is linked to their supporting evidences
13. The guideline has been externally reviewed by experts prior to its publication.	7	1	The review and set of recommendations reflect the consensus opinions of a group of US and Canadian professionals comprising metabolic physicians, a psychologists, and a dietitian nutritionist, whose collective experience either directly or indirectly involves over 100 patients with HT-1.
14. A procedure for updating the guideline is provided.	1	1	Not mentioned
Maximum possible score: 112, minimum possible score: 16, (96)	66.7%		
Domain 4. Clarity of Presentation			
15. The recommendations are specific and unambiguous.	7	7	Recommendations are specific and unambiguous
16. The different options for management of the condition or health issue are clearly presented.	7	7	Management options for the different clinical presentation of HT-1 discussed in page 6-13
17. Key recommendations are easily identifiable.	7	7	Key recommendations easily recognizable in the article
Maximum possible score: 42, minimum possible score: 6, (36)	100%		
Domain 5. Applicability			
18. The guideline describes facilitators and barriers to its application.	1	1	No statement regarding facilitators and barriers to implementing recommendations were sought (e.g., feedback from key stakeholders, pilot testing of guidelines before widespread implementation
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	1	7	Not outlined Quote: "This can only be accomplished by a robust NBS program, best achieved by screening neonatal blood for SA."
20. The potential resource implications of applying the recommendations have been considered.	1	1	Health economic evaluation, cost effectiveness analysis not done
21. The guideline presents monitoring and/or auditing criteria.	1	1	none
Maximum possible score: 56, minimum possible score: 8, (48)	33.3%		
Domain 6. Editorial Independence			
22. The views of the funding body have not influenced the content of the guideline.	7	7	Quote: "All discussions of both the content and writing of the manuscript were performed without the presence or input of any representative of Sobi, Inc." page 14
23. Competing interests of guideline development group members have been recorded and addressed.	7	7	Guideline development group's declaration of competing interests explicitly stated under the disclosure section Page 14
Maximum possible score: 28, minimum possible score: 4, (24)	100%		
OVERALL GUIDELINE ASSESSMENT			
Max – 322, Min – 46, (276)	213/322 (66%) 60.5%		
Rate the overall quality of this guideline.	5	5	
I would recommend this guideline for use	YES YES, with modifications NO		
Notes:	Although Domains 4 and 6 scored 100%, Rigour of Dev't was only at 66.7%. Overall Guide assessment was only at 63.4%		

APPENDIX L. Risk of Bias Summary

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Allard 2004	-	-	?	-	-	+	?
Dhillon 2011	-	-	?	-	?	+	?
Ia Marca 2008	-	-	?	-	-	+	?
Ia Marca 2011	?	?	?	-	-	+	-
Lund 2012	+	?	?	-	-	+	-
Metz 2012	-	-	-	-	?	+	-
Morrissey 2011	?	+	?	-	?	+	-
Sander 2006	+	+	?	-	-	+	-
Turgeon 2008	-	-	-	-	-	+	-
Zytkowicz 2013	-	?	?	-	-	+	-

- High
 ? Unclear
 + Low

Figure 1. Summary of risk of bias and applicability concerns: Newborn Screening for Tyrosinemia Type I using Succinylacetone – A systematic review of test accuracy

8. Newborn Screening for Fatty Acid Oxidation Disorders

APPENDIX A. Search Yield

Search strategies for MEDLINE, Cochrane Library and Google scholar was conducted in July 3, 2021 using the following keywords and MeSH.

LCHADD and MTP Deficiency

#	Keywords/MeSH	Yield
1	"Long chain 3-hydroxy acyl CoA dehydrogenase deficiency" OR "Trifunctional protein deficiency" OR "Trifunctional Protein Deficiency With Myopathy And Neuropathy" [Supplementary Concept]	122
2	"Mortality"[Mesh] OR "Infant Mortality"[Mesh] OR "Morbidity"[Mesh] OR "death" OR "complications" OR "sequela"	4,432,390
3	#1 AND #2	49

CPT1 and CPT2 Deficiency

#	Keywords/MeSH	Yield
1	"Carnitine palmitoyl transferase 1A deficiency" [Supplementary Concept] OR "Carnitine palmitoyl transferase 2 deficiency" [Supplementary Concept] OR "Carnitine palmitoyl transferase"	984
2	"Mortality"[Mesh] OR "Infant Mortality"[Mesh] OR "Morbidity"[Mesh] OR "death" OR "complications" OR "sequela"	4,432,390
3	#1 AND #2	136

MADD / GA2

#	Keywords/MeSH	Yield
1	"Multiple Acyl-CoA Dehydrogenase Deficiency" OR "Glutaric Aciduria" OR "Glutaric aciduria 2" [Supplementary Concept]	839
2	"Mortality"[Mesh] OR "Infant Mortality"[Mesh] OR "Morbidity"[Mesh] OR "death" OR "complications" OR "sequela"	4,432,390
3	#1 AND #2	208
4	#1 AND #2 Filters: Child: birth-18 years	147

COCHRANE LIBRARY SEARCH YIELD

#	Keywords/MeSH	Yield
1	Fatty Acid Oxidation	8
2	Carnitine palmitoyl transferase	17

3	Long chain 3-hydroxy acyl CoA dehydrogenase deficiency	3
4	Trifunctional protein deficiency	10
5	Multiple Acyl-CoA Dehydrogenase Deficiency OR Glutaric Aciduria type 2	1

APPENDIX B. Characteristics of Included Studies

Table 1. Summary of study characteristics for Test Accuracy

Study ID	Setting	Index Test	Index Test Specimen	Population	Sample Size	Reference standard	Reference Standard Specimen	Outcomes
Stinton 2021	Systematic Review (10 studies)	tandem mass spectrometry (TMS) LCHAD, TFP (n=23)	Dried blood spot	newborn	3,951,358	urine organic acids, blood acylcarnitine profiles, enzyme analysis in cultured fibroblasts or lymphocytes, mutation analysis, or at least 10-year follow-up	urine, blood, tissue	Results: PPV ranged from 0% (0 TP and 28 FP) from 276,565 babies screened to 100% (13 TP and 0 FP) from 2,037,824 babies screened SN, SP, and NPV could not be calculated as there was no systematic follow-up of babies who screened negative

Table 2. Summary of study characteristics for Benefit of Screening for LCHADD/ MTPD

Title/Author	Study design	Country	Number of patients	Intervention Group(s)	Control	Outcomes (Screened vs No Screening)
Fahnehjelm 2016	Cohort Study. Retrospective and prospective data collection Follow up: median 15 years (range 3 – 26 years)	Sweden	N=12 12/12 LCHADD	Screened (n=3)	No Screening (n=9)	Hypoglycemia (1/3 vs 7/9) Subnormal ocular fundii (2/3 vs 9/9)
Immonen 2016	Prospective Cohort (followed prospectively using data from retrospectively collected medical records) Follow up: range 1–11 years	Finland	16/16 LCHADD	Screened (n=1)	No Screening (n=15)	Mortality (0/1 vs 6/15) Cardiomyopathy (0/1 vs 4/10) Retinopathy (0/1 vs 9/10)

Karall 2015	Retrospective cohort Follow up: median 7.8 years (range 0.9 – 15.4 years)	Germany	N=14 14/14 LCHADD	Screened (n=9)	No Screening (n=5)	Cardiomyopathy (2/9 vs 5/5) Hepatopathy (1/9 vs 4/5) Retinopathy (3/9 vs 5/5)
Spiekerkotter 2009	Retrospective cohort	Metabolic Centers: Germany/ Switzerland/ Austria/ Netherlands	N=27 20/27 LCHADD 7/27 MTPD	Screened (n=10)	No Screening (n=17)	Mortality (2/10 vs 6/17) Cardiomyopathy (4/10 vs 8/17) Reye Syndrome (3/10 vs 6/17)* Myopathy (4/10 vs 14/17)* Hypoglycemia (4/10 vs 15/17)
De Biase 2017	Retrospective cohort Follow up: 10 years	USA	N=5 4/5 LCHADD 1/5 MTPD	Screened (n=3)	No Screening (n=2)	Arrhythmias (0/3 vs 1/2)* Retinopathy (2/3 vs 2/2) Neurologic symptoms (0/3 vs 1/2)* Myoglobinuria (0/3 vs 1/2)*
Kang 2018	Retrospective cohort Follow up: 10 years	South Korea	N=7 7/7 undifferentiated LCHADD/MTPD	Screened (n=1)	No Screening (n=6)	Mortality (1/1 vs 2/6)
Lund 2012	Case-control study Follow up: range 2 - 109 months	Denmark, Faroe Island, Greenland	N=5 5/5 undifferentiated LCHADD/MTPD	Screened (n=3)	No Screening (n=2)	Mortality (0/3 vs 1/2) Cardiomyopathy (0/3 vs 2/2) Hepatopathy (0/3 vs 1/2)
Sykut-Cegielska 2001	Retrospective cohort Follow up: 17 years	Poland	N=52 52/52 LCHADD	Screened (n=15)	No Screening (n=37)	Mortality (1/15 vs 13/37)

Table 3. Summary of study characteristics for Benefit of Screening for CPT1D and CPT2D

Title/Author	Study design	Country	Number of patients	Intervention Group(s)	Control	Outcomes (Screened vs No Screening)
Kang 2018	Retrospective cohort Follow up: 10 years	South Korea	N=2 CPT1D	Screened (n=0)	No Screening (n=1)	Developmental delay (0/1 vs 1/1*) The symptomatically screened patient suffered also from recurrent hepatic

						failure, nephromegaly, hemolytic anemia, and rhabdomyolysis
Spiekerkoetter 2009	Retrospective cohort	Metabolic Centers: Germany/ Switzerland/ Austria/ Netherlands	N=5 CPT2D	Screened (n=3)	No Screening (n=2)	Mortality (2/3 vs 0/2) Cardiomyopathy (1/3 vs 0/2) Hepatopathy (2/3 vs 0/2) Hypotonia/ Myopathy (1/3 vs 2/2)

Table 4. Summary of study characteristics for Benefit of Screening for GA2

Title/Author	Study design	Country	Number of patients	Intervention Group(s)	Control	Outcomes (Screened vs No Screening)
Maguolo 2020	Retrospective cohort Follow up: 5 years	Italy	N=7 GA2	Screened (n=2)	No Screening (n=5)	Myopathy/Rhabdomyolysis (0/2 vs 4/5) Metabolic Acidosis (0/2 vs 1/5) High AST/ALT (0/2) vs 2/5) High CK (1/2 vs 2/5) Hypoglycemia (1/2 vs 2/5)

APPENDIX C. GRADE Evidence Profiles

Grade Evidence Profile (LCHADD/MTPD)

Author(s): Reginald B. Balmeo, MD, Ian Theodore G. Cabaluna, MD, Kathryn Baltazar-Braganza, MD, Vaneza Leah Espino

Question: Newborn Screening compared to No screening for Asymptomatic newborn

Setting: Outpatient

Bibliography: Fraser H, Geppert J, Johnson R, Johnson S, Connock M, Clarke A, Taylor-Phillips S, Stinton C. Evaluation of earlier versus later dietary management in long-chain 3-hydroxyacyl-CoA dehydrogenase or mitochondrial trifunctional protein deficiency: a systematic review. *Orphanet J Rare Dis.* 2019 Nov 15;14(1):258. doi:10.1186/s13023-019-1226-y.

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Newborn Screening	No screening	Relative (95% CI)	Absolute (95% CI)		
Mortality												
5	observational studies	not serious	not serious	not serious	serious ^a	none	4/30 (13.3%)	28/77 (36.4%)	OR 0.36 (0.13 to 1.02)	193 fewer per 1,000 (from 294 fewer to 5 more)	 VERY LOW	
Cardiomyopathy												
4	observational studies	serious ^b	not serious	not serious	not serious	none	6/23 (26.1%)	19/34 (55.9%)	OR 0.28 (0.09 to 0.85)	297 fewer per 1,000 (from 456 fewer to 40 fewer)	 VERY LOW	
Hepatopathy												
2	observational studies	serious ^b	not serious	not serious	serious ^c	none	1/12 (8.3%)	5/7 (71.4%)	OR 0.06 (0.01 to 0.60)	584 fewer per 1,000 (from 690 fewer to 114 fewer)	 VERY LOW	
Retinopathy												
3	observational studies	serious ^b	not serious	not serious	not serious	none	5/13 (38.5%)	16/17 (94.1%)	OR 0.09 (0.01 to 0.64)	351 fewer per 1,000 (from 803 fewer to 30 fewer)	 VERY LOW	
Hypoglycemia												
2	observational studies	serious ^b	not serious	not serious	not serious	none	5/13 (38.5%)	22/26 (84.6%)	OR 0.10 (0.02 to 0.51)	491 fewer per 1,000 (from 747 fewer to 109 fewer)	 VERY LOW	

CI: Confidence interval; OR: Odds ratio

Explanations

a. Uncertainty in the magnitude of effect

b. Outcome assessors knew whether the participants had been screened or clinically detected

c. event rate very low

GRADE Evidence Profile (CPT1D)

Author(s): Reginald B. Balmeo, MD, Ian Theodore G. Cabaluna, MD, Kathryn Baltazar-Braganza, MD, Vaneza Leah Espino

Question: Newborn Screening for CPT1D compared to No screening for Asymptomatic newborn

Setting: outpatient

Bibliography: Kang E, Kim YM, Kang M, Heo SH, Kim GH, Choi IH, Choi JH, Yoo HW, Lee BH. Clinical and genetic characteristics of patients with fatty acid oxidation disorders identified by newborn screening. BMC Pediatr. 2018 Mar 8;18(1):103. doi: 10.1186/s12887-018-1069-z.

Certainty assessment							No of patients		Effect		Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Newborn Screening for CPT1D	No screening	Relative (95% CI)	Absolute (95% CI)		
Developmental Delay												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	0/1 (0.0%)	1/1 (100.0%)	OR 0.11 (0.00 to 10.27)	0 fewer per 1,000 (from 0 fewer to -)		Very low

CI: confidence interval; OR: odds ratio

Explanations

a. Outcome assessors knew whether the participants had been screened or clinically detected

b. Very wide CI with uncertainty in magnitude of effect

GRADE Evidence Profile (CPT2D)

Author(s): Reginald B. Balmeo, MD, Ian Theodore G. Cabaluna, MD, Kathryn Baltazar-Braganza, MD, Vaneza Leah Espino

Question: Newborn Screening for CPT2D compared to No screening for Asymptomatic newborn

Setting: outpatient

Bibliography: Spiekerkoetter U, Lindner M, Santer R, Grotzke M, Baumgartner MR, Boehles H, Das A, Haase C, Hennermann JB, Karall D, de Klerk H, Knerr I, Koch HG, Plecko B, Rösingher W, Schwab KO, Scheible D, Wijburg FA, Zschocke J, Mayatepek E, Wendel U. Management and outcome in 75 individuals with long-chain fatty acid oxidation defects: results from a workshop. J Inherit Metab Dis. 2009 Aug;32(4):488-97. doi: 10.1007/s10545-009-1125-9.

Certainty assessment							No of patients		Effect		Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Newborn Screening for CPT2D	No screening	Relative (95% CI)	Absolute (95% CI)		
Mortality												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	2/3 (66.7%)	0/2 (0.0%)	OR 8.33 (0.22 to 320.38)	0 fewer per 1,000 (from 0 fewer to 0 fewer)		Very low
Cardiomyopathy												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	1/3 (33.3%)	0/2 (0.0%)	OR 3.00 (0.08 to 115.34)	0 fewer per 1,000 (from 0 fewer to 0 fewer)		Very low
Hepatopathy												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	2/3 (66.7%)	0/2 (0.0%)	OR 8.33 (0.22 to 320.38)	0 fewer per 1,000 (from 0 fewer to 0 fewer)		Very low
Hypotonia/Myopathy												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	1/3 (33.3%)	2/2 (100.0%)	OR 0.12 (0.00 to 4.61)	0 fewer per 1,000 (from 0 fewer to -)		Very low

CI: confidence interval; OR: odds ratio

Explanations

- a. Outcome assessors knew whether the participants had been screened or clinically detected
- b. Very wide CI with uncertainty in magnitude of effect

GRADE Evidence Profile (GA2)

Author(s): Reginald B. Balmeo, MD, Ian Theodore G. Cabaluna, MD, Kathryn Baltazar-Braganza, MD, Vaneza Leah Espino

Question: Newborn Screening for GA2 compared to No screening for Asymptomatic newborn

Setting: outpatient

Bibliography: Maguolo A, Rodella G, Dianin A, Nurri R, Monge I, Rigotti E, Cantalupo G, Salvietti L, Tucci S, Pellegrini F, Molinaro G, Lupi F, Tonin P, Pasini A, Campostrini N, Ion Popa F, Teofoli F, Vincenzi M, Camilot M, Piacentini G, Bordugo A.

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 Aug 5:24:100632. doi: 10.1016/j.ymgmr.2020.100632.

Certainty assessment							No of patients		Effect		Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Newborn Screening for GA2	No screening	Relative (95% CI)	Absolute (95% CI)		
Myopathy/Rhabdomyolysis												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	0/2 (0.0%)	4/5 (80.0%)	OR 0.07 (0.00 to 2.33)	581 fewer per 1,000 (from -- to 103 more)		Very low
Metabolic Acidosis												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	0/2 (0.0%)	1/5 (20.0%)	OR 0.60 (0.02 to 20.98)	70 fewer per 1,000 (from 195 fewer to 640 more)		Very low
Hepatopathy												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	0/2 (0.0%)	2/5 (40.0%)	OR 0.28 (0.01 to 8.76)	243 fewer per 1,000 (from 393 fewer to 454 more)		Very low
Elevated Creatine Kinase												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	1/2 (50.0%)	2/5 (40.0%)	OR 1.50 (0.06 to 40.63)	100 more per 1,000 (from 362 fewer to 564 more)		Very low
Hypoglycemia												
1	observational studies	serious ^a	not serious	not serious	very serious ^b	none	1/2 (50.0%)	2/5 (40.0%)	OR 11.00 (0.28 to 433.80)	480 more per 1,000 (from 243 fewer to 597 more)		Very low

CI: confidence interval; OR: odds ratio

Explanations

- a. Outcome assessors knew whether the participants had been screened or clinically detected
- b. Very wide CI with uncertainty in magnitude of effect

APPENDIX D. Forest Plots

LCHADD/MTPD

Figure 1. Forest plot of comparison: Screening vs No screening, outcome: 1.1 Mortality

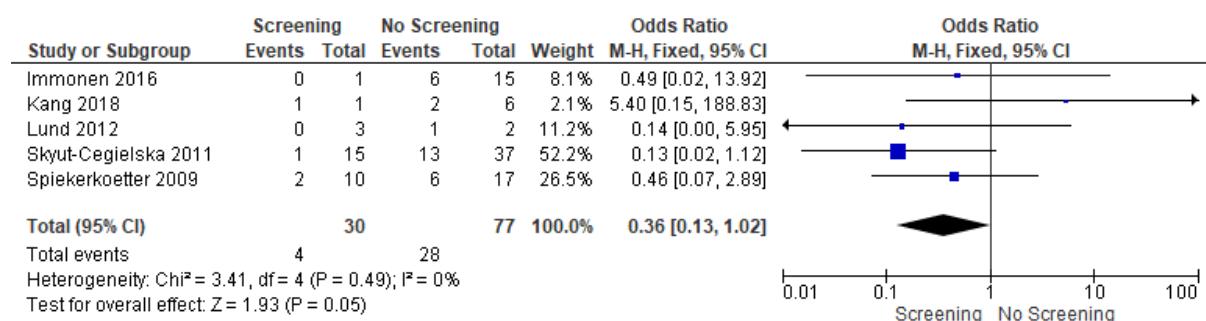


Figure 2. Forest plot of comparison: Screening vs No screening, outcome: 1.2 Cardiomyopathy

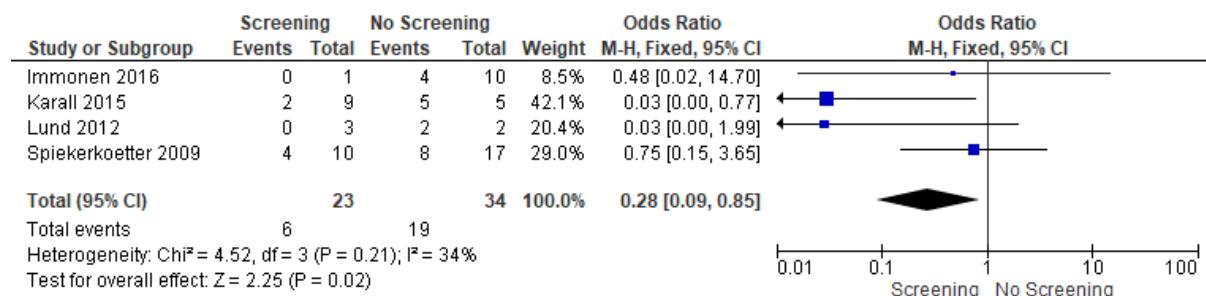


Figure 3. Forest plot of comparison: Screening vs No screening, outcome: 1.3 Hepatopathy

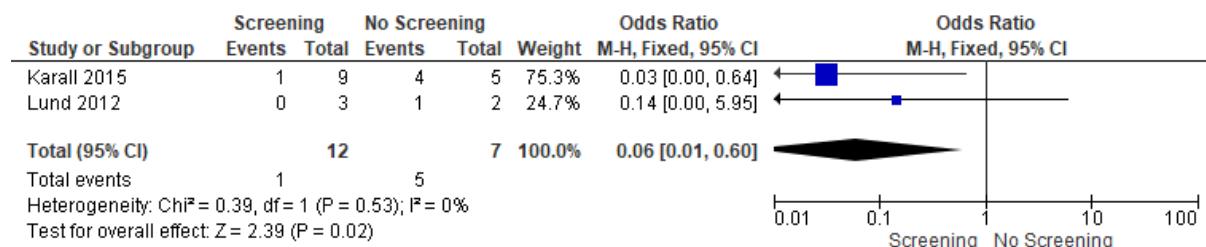


Figure 4. Forest plot of comparison: Screening vs No screening, outcome: 1.4 Retinopathy

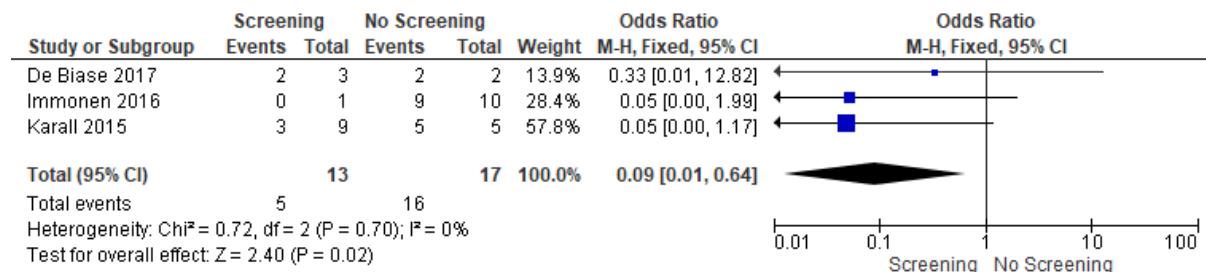
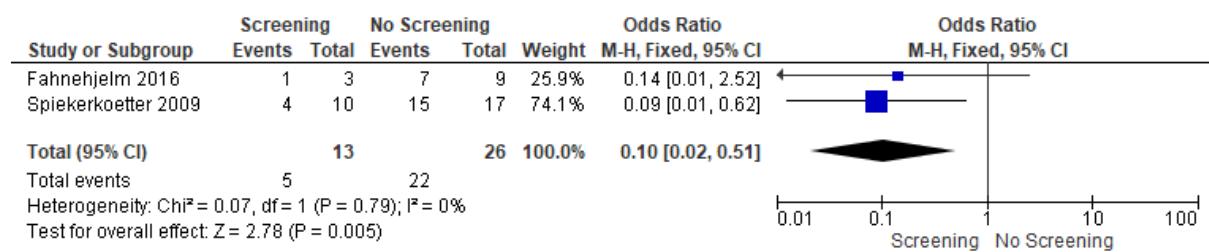
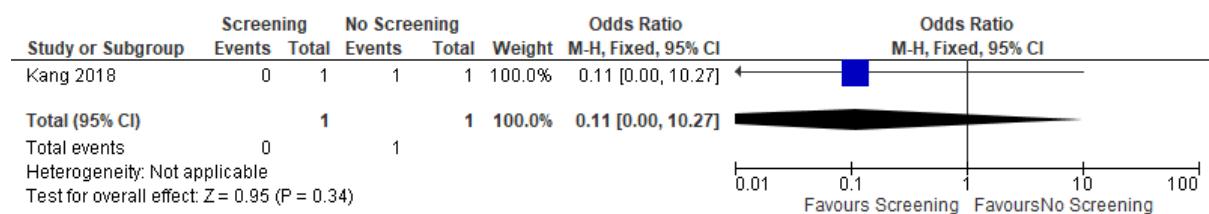


Figure 5. Forest plot of comparison: Screening vs No screening, outcome: 1.5 Hypoglycemia



CPT1D

Figure 6. Forest plot of comparison: Screening vs No screening, outcome: Developmental Delay



CPT2D

Figure 7. Forest plot of comparison: Screening vs No screening, outcome: Mortality

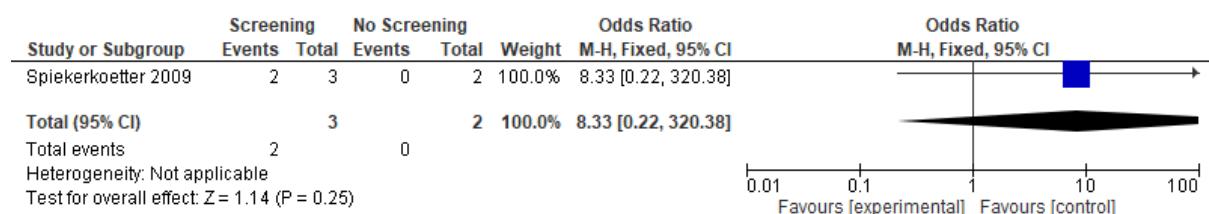


Figure 8. Forest plot of comparison: Screening vs No screening, outcome: Cardiomyopathy

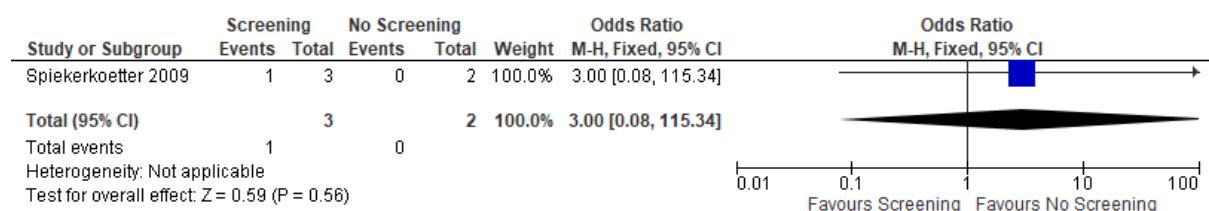


Figure 9. Forest plot of comparison: Screening vs No screening, outcome: Hepatopathy

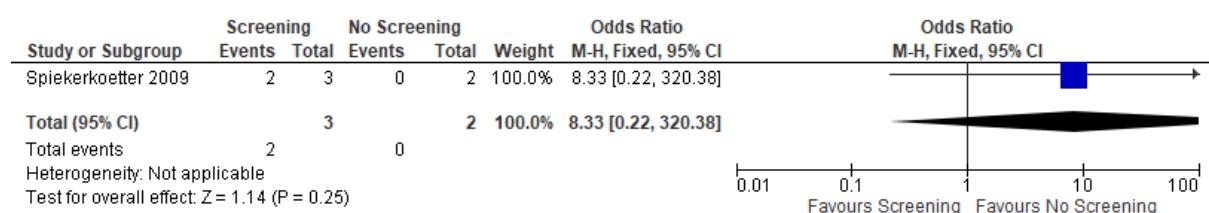
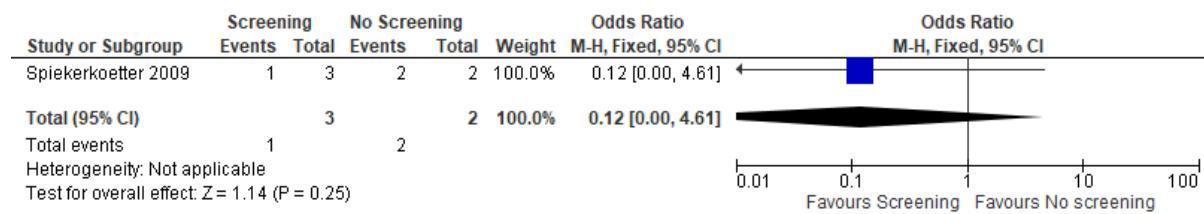


Figure 10. Forest plot of comparison: Screening vs No screening, outcome: **Hypotonia/Myopathy**



GA2

Figure 11. Forest plot of comparison: Screening vs No screening, outcome: **Myopathy**



Figure 12. Forest plot of comparison: Screening vs No screening, outcome: **Metabolic Acidosis**

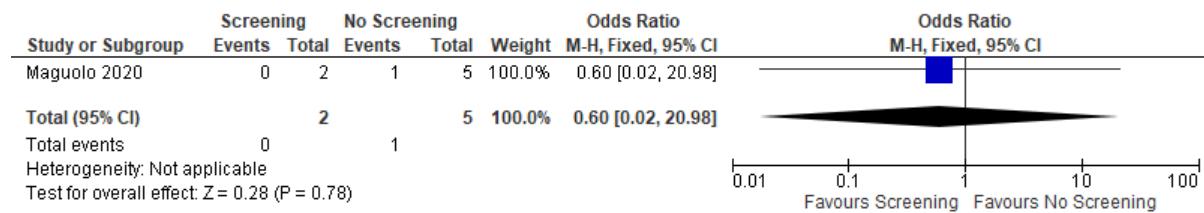


Figure 13. Forest plot of comparison: Screening vs No screening, outcome: **Hepatopathy**

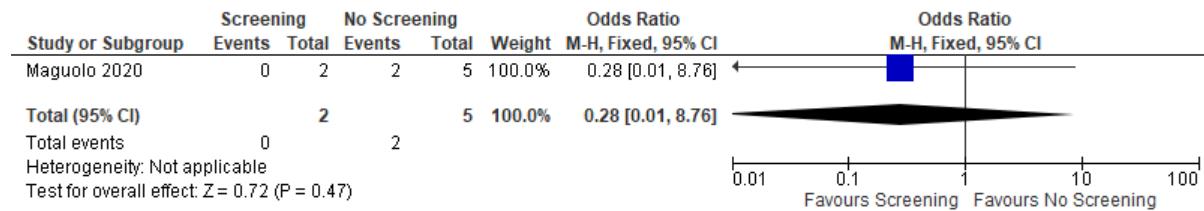


Figure 14. Forest plot of comparison: Screening vs No screening, outcome: **Elevated Creatine Kinase**

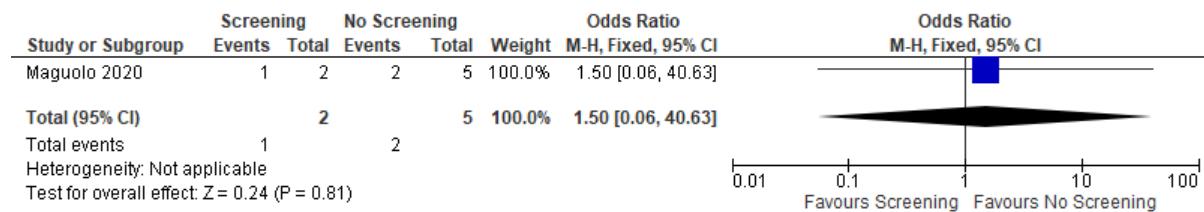
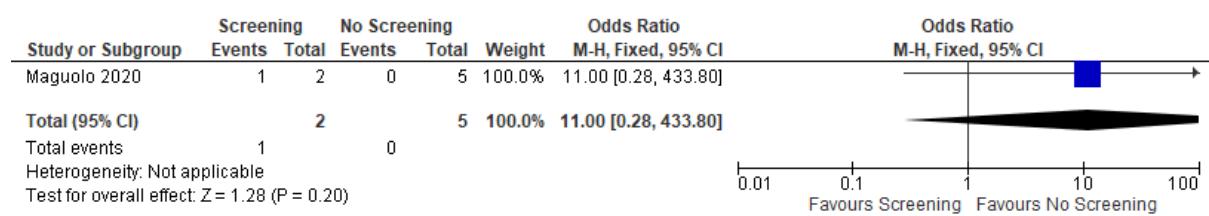


Figure 15. Forest plot of comparison: Screening vs No screening, outcome: **Hypoglycemia**



APPENDIX E. Cost-Effectiveness Analysis

Table 1: Incremental cost-effectiveness of each FAODs evaluated independently and a breakdown of incremental costs, and life-years gained per patient screened.¹²

<i>Fatty acid β-oxidation disorders</i>	<i>Incremental cost (\$*)</i>	<i>Incremental Life Years gained (\$)</i>	<i>Incremental Cost Effectiveness Ratio (ICER)</i>
<i>CPT1D</i>	19.92	1.67	1,192,814
<i>CPT2D</i>	19.89	1.60	1,243,125
<i>LCHADD</i>	20.00	1.88	1,063,830
<i>GA2/ MADD</i>	19.09	0.0134	142,462,687
<i>MCADD</i>	24.43	9.65	253,161

*1 Canadian dollar = 40.56 Php

APPENDIX F. Recommendations from Other Groups

Table 1. Recommendation from the UK National Screening Committee on **LCHADD and MTPD**

Group	Recommendation	Basis for recommendation
UK National Screening Committee (April 2019)	Screening babies for LCHAD and MTP deficiency is NOT RECOMMENDED	UK NSC does not recommend screening for LCHADD or MTPD until further studies provide definitive evidence on: 1. Prevalence of babies born with LCHAD/MTP deficiency in the UK 2. Association between genetic mutations and symptoms 3. Systematic follow up of babies who screened negative 4. Outcome of early detection and intervention

Table 2. Countries which have already included **LCHADD, MTPD, CPT1D, CPT2D and GA2** in their newborn screening program

FAOD	COUNTRIES
LCHADD	EUROPE (Austria, Croatia, Czech Republic, Denmark, Estonia, Finland, Germany, Hungary, Iceland, Israel, Italy, Netherlands, North Macedonia, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden), Australia, USA, Canada (Alberta, British Columbia, Ontario, Saskatchewan, New Brunswick, Nova Scotia, Prince Edward Island), Asia (China, Hong Kong, India, Japan, Lebanon, Philippines, Qatar, Singapore, South Korea, Taiwan)
MTPD	EUROPE (Austria, Croatia, Czech Republic, Denmark, Estonia, Finland, Germany, Hungary, Iceland, Israel, Italy, Netherlands, North Macedonia, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden), Australia, USA, Canada (Alberta, Ontario), Asia (China, Hong Kong, India, Japan, Philippines, Singapore, South Korea, Taiwan)
CPT1D, CPT2D	EUROPE (Austria, Czech Republic, Estonia, Finland, Germany, Hungary, Iceland, Netherlands (CPT1D only), North Macedonia, Norway, Poland, Portugal, Slovakia, Slovenia, Spain (pilot), Sweden), Australia, USA, Canada (New Brunswick, Nova Scotia, Prince Edward Island), Asia (China, Hong Kong, India, Japan, Lebanon, Philippines, Qatar, Singapore, South Korea, Taiwan)

**GA2/
MADD**

EUROPE: Austria, Belgium, Finland, Hungary, Spain, Iceland, Israel, Italy, North Macedonia, Poland, Portugal, Spain (pilot), Sweden, Ukraine (pilot), **Canada** (New Brunswick, Nova Scotia, Prince Edward Island), **Asia** (China, Hongkong, Japan, Lebanon, Philippines, Qatar, Singapore, Taiwan, Iran)

9. Newborn Screening for Biotidinase Deficiency

APPENDIX A. Search Yield

Summary of Search:

1. Search from CPG Database

a) NICE (National Institute for Health and Care Excellence)

#	QUERY	RESULTS
1	Organic acid disorders	453
	Yield for: ▪ Biotinidase deficiency	1 0 0 0
2	Newborn screening	2043
	Filter: evidence summary	538
	Yield for: ▪ Biotinidase deficiency	1 0 0 0

UK National Screening Committee. 2018. Newborn Screening for Biotinidase Deficiency. Retrieved from <https://view-health-screeningrecommendations.service.gov.uk/biotinidase-deficiency/> (The article was not available during the time the search was conducted since they have an ongoing review or revision of their recommendation)

b) U.S. Preventive Services Task Force (USPSTF)

#	QUERY	RESULTS
1	Organic acid disorders	0
2	Newborn screening	129
	Filter: Pediatric	8
	Yield	0

c) Canadian Task Force for Preventive Health Care (CTFPHC)

#	QUERY	RESULTS
1	Organic acid disorders	0
2	Newborn screening	6
	Yield	0

2. Search from Pubmed

#	QUERY	RESULTS
1	Search: guideline [Publication Type] OR practice guideline [Publication Type] or recommendation*[Title] OR standard* [Title] OR standard* [Title] OR guideline* [Title])	238,338
2	Search: biotinidase deficiency	633
3	Search: (#1) AND (#2)	3
4	Search: organic acid disorders	6652

5	Search: inborn error of metabolism	162,717
6	Search: (#4) OR (#5)	168,104
7	Search: newborn screening	39,375
8	Search: (#6) AND (#7)	5,206
9	Search: (#2) AND (#8)	127
	Yield	29

3. Search from Guidelines International Network

#	QUERY	RESULTS
1	Organic acid disorders	0
2	Biotinidase deficiency	0
3	Isovaleric academia	0
4	Holocarboxylase deficiency	0
5	Beta-ketothialase deficiency	0
6	Newborn screening	2
7	Inborn Error of Metabolism	1

4. Search from Google Scholar

#	QUERY	RESULTS
1	Biotinidase deficiency screening newborns clinical practice guidelines recommendations evidence	21700
2	Biotinidase deficiency screening newborns clinical practice guidelines recommendations evidence; filter: 2015-2021	11300
3	"Biotinidase deficiency" screening newborns clinical practice guidelines recommendations evidence; filter: 2015-2021	591
	Yield	9

Among the yield, the following articles were retrieved for further review:

Articles for Diagnostic Accuracy

1. Moslinger, D. et. al. 2001. Clinical and Neuropsychological Outcome of 33 Patients with BD Ascertained by National Neonatal Screening and Family Studies in Austria. Eur J Pediatr (2001) 160: 277-282.
2. Gannavarapu, S. 2015. Biotinidase deficiency: Spectrum of molecular, enzymatic and clinical information from newborn screening Ontario, Canada (2007–2014). YMGME-05944; No. of pages: 6; 4C.
3. Lund et al. 2012. Biochemical screening of 504,049 newborns in Denmark, the Faroe Islands and Greenland—experience and development of a routine program for expanded newborn screening. Mol Genet Metab 107:281–293)
4. Cowan TM, Kazerouni NN, Dharajiya N, et al. 2012. Increased incidence of profound biotinidase deficiency among Hispanic newborns in California. Molecular Genetics and Metabolism. 2012;106(4):485-7.
5. Kwon, C. and Farrel, P. 2000. The Magnitude and Challenge of False Positive Newborn Screening Test Results. Arch Pediatr Adolesc Med. 2000; 154:714-718.

6. Ohlsson, A., et. al. 2010. Profound biotinidase deficiency: a rare disease among native Swedes. *J Inherit Metab Dis* (2010) 33 (Suppl 3):S175–S180.
7. Carvalho, N. et. al. 2019. Novel mutations causing biotinidase deficiency in individual identified by the newborn screening program in Minas Gerais,Brazil. *Am J Med Genet.* 2019;1–5.
8. Neto, E. C., et. al. 2004. Newborn screening for biotinidasedeficiency in Brazil: biochemical and molecular characterizations. *Braz J Med Biol Res* 37(3) 200437: 295-299.
9. Loeber, J. G. 2007. Neonatal screening in Europe; the situation in 2004. *J Inherit Metab Dis* (2007) 30:430–438.
10. Burlina, A. B., et. al. 1988. Neonatal screening for biotinidase deficiency in north eastern Italy. *Eur J Pediatr* (1988) 147 : 317-318.
11. Wiltink, R., et. al. 2016. Neonatal screening for profound biotinidase deficiency in the Netherlands: consequences and considerations. *European Journal of Human Genetics* (2016) 24, 1424–1429.
12. Heard, G., et. el. 1986. Neonatal Screening for BD: Results of a 1 year pilot study. (*J PEDIATR* 1986;108:40-46).
13. Thodi, G., et. al. 2011. Mutational analysis for biotinidase deficiency of a Greek patients' cohort ascertained through expanded newborn screening.
14. Loukas. 2010.Expanded newborn screening in Greece: 30 months

of experience.

15. Thodi, G. et al. 2013. High incidence of partial biotinidase deficiency cases in newborns of Greek origin.
16. Sarafoglou, K. et al. 2009. High incidence of profound biotinidase deficiency detected in newborn screening blood spots in the Somalian population in Minnesota.
17. Al-Jasmi, F., et. al. 2015. Inborn Errors of Metabolism in the United Arab Emirates: Disorders Detected by Newborn Screening (2011–2014).
18. Carvaljo, N. et. al. 2019. Frequency of biotinidase gene variants and incidence of biotinidase deficiency in the NBS Program in Minas Gerais, Brazil.*J Med Screen* 0(0) 1–6.
19. Scriver, C. 2006. Community Genetics and Dignity in Diversity in the Quebec Network of Genetic Medicine.
20. Tangeraas, T., et. al. 2020. Performance of Expanded Newborn Screening in Norway Supported by Post-Analytical Bioinformatics Tools and Rapid Second-Tier DNA Analyses.*Int. J. Neonatal Screen.* 2020, 6, 51.

Articles for Outcome

1. Landau, Y., et. al. 2016. Long-term outcome of expanded newborn screening at Boston children's hospital: benefits and challenges in defining true disease.

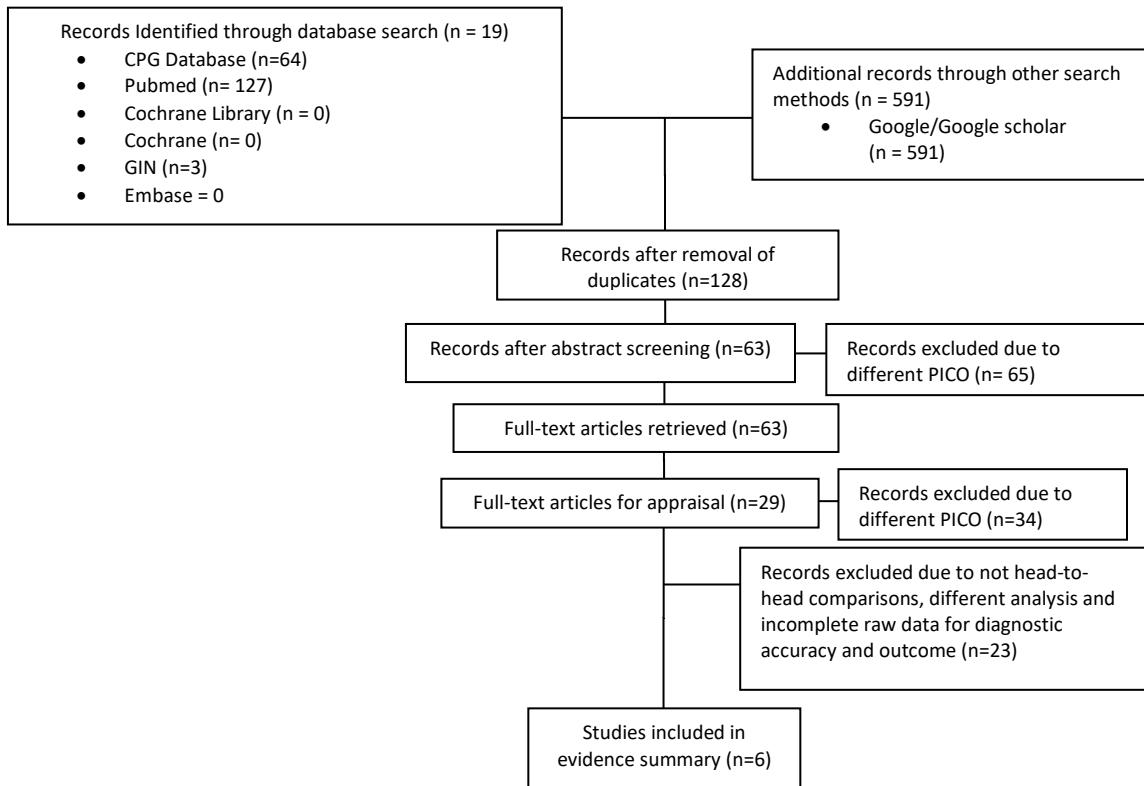
2. Moslinger, D., et.al., 2003. Molecular characterisation and neuropsychological outcome of 21 patients with profound biotinidase deficiency detected by newborn screening and family studies.
3. Laszlo, A., et. al. 2003. Neonatal screening for biotinidase deficiency in Hungary: Clinical, biochemical and molecular studies. *J. Inherit. Metab. Dis.* 26 (2003) 693-698.
4. Wiltink, R., et. al. 2016. Neonatal screening for profound biotinidase deficiency in the Netherlands: consequences and considerations. *European Journal of Human Genetics* (2016) 24, 1424–1429.
5. Waisbren, S. E. et. al. 2002. Newborn screening compared to clinical identification of biochemicalgenetic disorders. Short Report. *J. Inherit. Metab. Dis.* 25 (2002) 599-600.
6. Weber, P., et. al. 2004. Outcome in patients with profound biotinidase deficiency: relevance of newborn screening. *Developmental Medicine & Child Neurology* 2004, 46: 481–484.
7. Jay, A., et. al. 2015. Outcomes of individuals with profound and partial biotinidase deficiency ascertained by newborn screening in Michigan over 25 years. *Genetics in medicine | Volume 17 | Number 3 | March 2015.* Page 205-209.
8. Secor, J., et. al. 1990. Partial biotinidase deficiency: Clinical and biochemical features.(J PEDIATR 1990;116:78-83).
9. Wolf, B. 2017. Successful outcomes of older adolescents and adults with profound biotinidase deficiency identified by newborn screening. *Volume 19 | Number 4 | April 2017 | Genetics in medicine.* Page 396-402.
10. Wolf, B. 2015. Why screen newborns for profound and partial biotinidase deficiency. *Molecular Genetics and Metabolism* 114 (2015) 382-387.
11. Gannavarapu, S. 2015. Biotinidase deficiency: Spectrum of molecular, enzymatic and clinical information from newborn screening Ontario, Canada (2007–2014). YMGME-05944; No. of pages: 6; 4C.
12. Linder, et al. 2011. Efficacy and outcome of expanded newborn screening for metabolic diseases—report of 10 years from South-West Germany. *Orphanet J Rare Dis* 6:44)
13. Tanzer, F., et. al. 2009. Neonatal Screening for Biotidinidase Deficiency: Results of a 1-year Pilot Study in Four Cities in Central Anatolia. *Journal of Pediatric Endocrinology & Metabolism*, 22, 1113-1116 (2009).

Included Studies (5 Diagnostic Accuracy, 1 Outcome)

1. Lund AM, Hougaard DM, Simonsen H, Andresen BS, Christensen M, Dunø M, Skogstrand K, Olsen RKJ, Jensen UG, Cohen A, Larsen N, Saugmann-Jensen P, Gregersen N, Brandt NJ, Christensen E, Skovby F, Nørgaard-Pedersen B. 2012. Biochemical screening of 504,049 newborns in Denmark, the Faroe Islands and Greenland—experience and development of a routine program for expanded newborn screening. *Mol Genet Metab.* 2012; 107:281–293.

2. Kwon, C. and Farrel, P. 2000. The Magnitude and Challenge of False Positive Newborn Screening Test Results. *Arch Pediatr Adolesc Med.* 2000; 154:714-718
3. Burlina AB, Sherwood WG, Marchioro MV, Dalla Bernardina B, Gaburro D. 1988. Neonatal screening for biotinidase deficiency in north eastern Italy. *Eur J Pediatr.* 1988; 147:317-318.
4. Wiltink RC, Kruijshaar ME, Minkelen RV, Onkenhout W, Verheijen FW, Kemper EA, Spronsen FJ, Ploeg AT, Niezen-Koning KE, Saris JJ, Williams M. 2016. Neonatal screening for profound biotinidase deficiency in the Netherlands: consequences and considerations. *European Journal of Human Genetics.* 2016; 24:1424–1429.
5. Heard G, Wolf B, Jefferson LG, Weissbecker KA, Nance WE, Secor McVoy JR, Napolitano A, Mitchell PL, Lambert FW, Linyear AS. 1986. Neonatal Screening for BD: Results of a 1 year pilot study. *J PEDIATR.* 1986; 108:40-46.
6. Weber P, Scholl S, ER Baumgartner. 2004. Outcome in patients with profound biotinidase deficiency: relevance of newborn screening. *Developmental Medicine & Child Neurology.* 2004; 46: 481–484.

FLOWCHART



APPENDIX B. GRADE Evidence Profile (Outcome)

Certainty assessment							Nº of patients		Effect		Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Newborn Screening	No Screening	Relative (95% CI)	Absolute (95% CI)		
Visual Impairment												
1	observational studies	serious ^a	not serious ^b	not serious ^c	serious ^d	none	0/25 (0.0%)	4/12 (33.3%)	OR 0.04 (0.00 to 0.76)	314 fewer per 1,000 (from 58 fewer to -)	⊕○○○ VERY LOW	
Hearing Impairment												
1	observational studies	serious ^a	not serious ^b	not serious ^c	serious ^d	none	0/25 (0.0%)	4/12 (33.3%)	OR 0.04 (0.00 to 0.76)	314 fewer per 1,000 (from 58 fewer to -)	⊕○○○ VERY LOW	
Delayed Onset of Sitting												
1	observational studies	serious ^a	not serious ^b	not serious ^c	serious ^{d,e}	none	1/25 (4.0%)	3/12 (25.0%)	OR 0.13 (0.01 to 1.36)	208 fewer per 1,000 (from 247 fewer to 62 more)	⊕○○○ VERY LOW	
Delayed Onset of Walking												
1	observational studies	serious ^a	not serious ^b	not serious ^c	serious ^d	none	0/25 (0.0%)	5/12 (41.7%)	OR 0.03 (0.00 to 0.54)	396 fewer per 1,000 (from 138 fewer to -)	⊕○○○ VERY LOW	
Delayed Speech												
1	observational studies	serious ^a	not serious ^b	not serious ^c	serious ^d	none	4/25 (16.0%)	6/12 (50.0%)	OR 0.19 (0.04 to 0.90)	340 fewer per 1,000 (from 462 fewer to 26 fewer)	⊕○○○ VERY LOW	

CI: Confidence interval; OR: Odds ratio

Explanations

- a. High risk of bias since the study is only an observational study.
- b. only 1 study
- c. The outcome measured is only the surrogate outcome since there are no available studies on the direct outcome. Eventhough the outcome measured is a surrogate outcome, it is strongly associated with the direct outcome.
- d. small sample size (less than 30)
- e. wide confidence interval

APPENDIX C. GRADE Evidence Profile (Diagnostic Accuracy)

Question: Should newborn screening be used to screen for biotinidase deficiency in asymptomatic newborns?

Sensitivity	0.93 (95% CI: 0.67 to 0.99)
Specificity	1.00 (95% CI: 1.00 to 1.00)

Prevalences 0.00089%

Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested pre-test probability of 0.00089%	Test accuracy CoE	Importance
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias			
True positives (patients with biotinidase deficiency)	5 studies 3686573 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	serious ^b	serious ^c	none	0 (0 to 0)	⊕○○○ VERY LOW	
								0 (0 to 0)		
False negatives (patients incorrectly classified as not having biotinidase deficiency)	5 studies 3686573 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	serious ^b	serious ^c	none	1000 (1000 to 1000)	⊕○○○ VERY LOW	
								0 (0 to 0)		
True negatives (patients without biotinidase deficiency)	5 studies 3686573 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	serious ^b	serious ^c	none	1000 (1000 to 1000)	⊕○○○ VERY LOW	
								0 (0 to 0)		
Inconclusive	0 studies patients	-	-	-	-	-	-	-	-	-
Complications	0 studies patients								-	

Explanations

a. It is an observational study and the risk of bias is unclear with regards to patient sampling and the reference standard used (there is knowledge of the result of the index test before using the reference test and the reference standard used may have been done using the same test specimen but not exactly the same process or method of enzyme assay determination).

b. There is high heterogeneity.

c. There is still uncertainty since the true negative values were just calculated based on the assumption that there were no false negatives detected as reported by the studies but there were no confirmatory testing done or follow-up done on those with negative results.

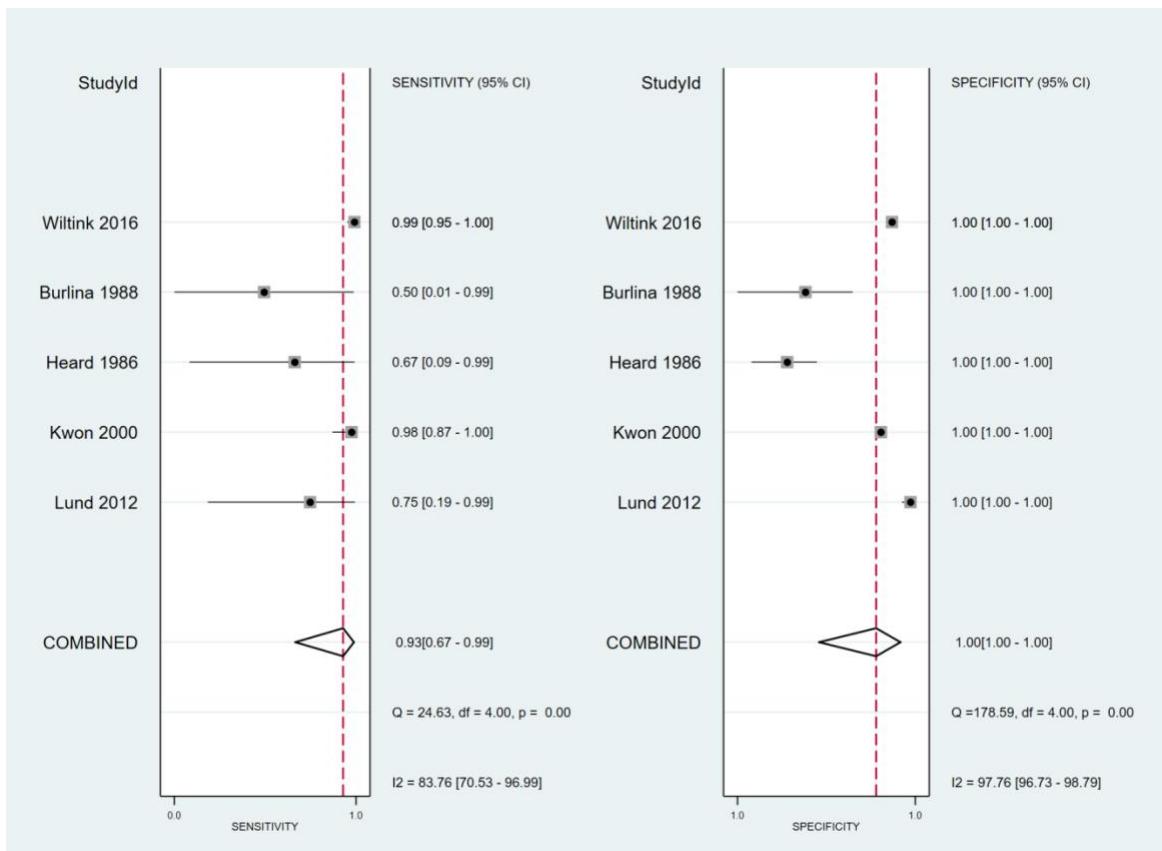
APPENDIX D. Characteristics of Included Studies (Diagnostic Accuracy)

STUDY ID	SETTING	STUDY DESIGN	POPULATION	SAMPLE SIZE	INDEX TEST (SCREENING)	INDEX TEST SPECIMEN	REFERENCE STANDARD (CONFIRMATORY)	REFERENCE STANDARD SPECIMEN	TP	FP	TN	FN	TEST PERFORMANCE DATA	NOTES
Lund 2012	Denmark, Faroe Islands and Greenland	Observational	Newborn (2009 to 2011)	140, 565	DBS	Serum	Quantitative Enzyme Assay	Serum	3	7	140,555	0	True Positive: 3; False Positive:7; False Negative: 0 PPV:30%; False Positive Rate: 0.00498%	*False negative was assumed to be zero since there was no false negative reported as stated in the study and true negative was then computed.
Kwon 2000	US	Observational	Newborn	4, 902, 516 (1991-1994)	DBS	Serum	Quantitative Enzyme Assay	Serum	40	589	2,525, 871	0	1993: PPV: 6.8%; FPR: 0.0229% 1994: PPV: 5.9%; FPR: 0.0237% PPV:1991-3.7%, 1992-9.1%, 1993-6.8%, 1994-5.9%; AVE PPV-6.4% 1993 & 1994: PPV: 6.35%; FPR: 0.0233%	* The data gave no indication of false-negative test result as stated in the study, therefore, false negative was assumed to be zero and true negative was then computed.

Heard 1986	Virginia	Observatio nal	Asymptomatic newborns for 1 year starting Jan 24, 1984	81,243	simple colorimetri c test in DBS	serum	quantitative enzyme assay	Serum	2	66	81, 175	0	False Positive Rate: 0.08% or 66/81,241 (based on the initial card) Positive Predictive Value: 2.94% (based on initial card) False Positive Rate: 0.09% (based on obtaining a 2nd card); Positive Predictive Value: 100% (based on 2 cards)	*False Negative was assumed to be zero since authors mentioned that they are not aware of any false negative test result and true negative was then computed.
Burlina 1988	North Eastern Italy	Observatio nal (Prospecti ve)	Newborns(6 month period)	24,300	semiquanti tative colorimetri c assay on DBS	Serum	quantitative colourimetric assay in serum	Serum	1	17	24,282	0	Positive Test: 18 Confirmed BD: 1 False Positive: 17 PPV: 5.56%; FPR: 0.06996%	*The author was unaware of any false negative results to date thus it can be assumed that False Negative is zero and true negative was then computed.

Wiltink 2016	South Western Netherlan ds	Observatio nal	Newborns (between 2007 and 2012)	Southwest ern Netherlan ds - 304, 982; Netherlan ds - 913, 965	DBS	Serum	quantitative colourimetric assay in serum	Serum	111	150	913,70 4	0	<p><u>South Western Netherlands</u> Positive Test: 92; Confirmed BD: 50 (Profound-6, Partial-44); False Positive: 42; PPV: 54.35%; <u>FPR: 0.0138%</u></p> <p><u>Netherlands</u> Positive Test: 261; Confirmed BD: 111; PPV: 42.53%; FPR: 0.0164%</p>	*No False Negative identified as stated in the study, thus it can be assumed that False Negative is zero and true negative was then computed.

APPENDIX E. STATA Analysis of Diagnostic Accuracy Studies



10. Newborn Screening for Beta-Ketothiolase Deficiency

APPENDIX A. Search Strategies

- a) NICE (National Institute for Health and Care Excellence)

#	QUERY	RESULTS
1	Organic acid disorders	0
2	Newborn screening	25
	Yield for: ▪ Beta-ketothiolase deficiency	0

- b) U.S. Preventive Services Task Force (USPSTF)

#	QUERY	RESULTS
1	Organic acid disorders	0
2	Newborn screening	131
	Filter: Pediatric	33
	Yield for: ▪ Beta-ketothiolase deficiency	0

- c) Canadian Task Force for Preventive Health Care (CTFPHC)

#	QUERY	RESULTS
1	Organic acid disorders	0
2	Newborn screening	6
	Yield for: ▪ Beta-ketothiolase deficiency	0

- d) PubMed

Search number	Query	Results	Time
32	# 2 OR #28 OR #31 AND #27	16	21:08:11
28	diagnostic accuracy	253,278	21:06:28
31	Sensitivity and specificity	748,387	21:06:08
17	(((((Newborn screening[MeSH Terms]) AND (beta ketothiolase deficiency)) OR (3-ketothiolase deficiency)) OR (3-oxothiolase deficiency)) OR (Alpha methylacetoacetic aciduria)) OR (Mitochondrial acetoacetyl-coenzyme A thiolase deficiency)) OR (T2 deficiency)	163	20:13:36
30	#2 AND #27 Filter: From 2015-2021	41	19:29:32
29	#27 AND #28	2	19:26:43
27	((((beta ketothiolase deficiency) OR (3-ketothiolase deficiency)) OR (3-oxothiolase deficiency)) OR (Alpha methylacetoacetic aciduria)) OR (Mitochondrial acetoacetyl-coenzyme A thiolase deficiency)) OR (T2 deficiency)	163	19:26:00
26	#2 AND #4 OR #10 OR # 11 OR #11 OR #12 OR #13 OR #14	Filter: Clinical Trial	2
25	#2 AND #4 OR #10 OR # 11 OR #11 OR #12 OR #13 OR #14	Filter: Meta-Analysis	0
24	#2 AND #4 OR #10 OR # 11 OR #11 OR #12 OR #13 OR #14	Filter: Randomized Controlled Trial	0
23	#2 AND #4 OR #10 OR # 11 OR #11 OR #12 OR #13 OR #14	Filter: Systematic Review	0
18	#4 AND #2		19:09:35
22	#4 AND #2	Filter: Clinical Trial	0
21	#4 AND #2	Filter: Meta-Analysis	0
20	#4 AND #2	Filter: Randomized Controlled Trial	0
19	#4 AND #2	Filter: Systematic Review	0
16	#4 OR #10 OR # 11 OR #11 OR #12 OR #13 OR #14 AND #3		19:02:19
14	T2 deficiency	118	18:47:37
13	Mitochondrial acetoacetyl-coenzyme A thiolase deficiency	84	18:46:56

12	Alpha methylacetoacetic aciduria	94	18:45:27
11	3-oxothiolase deficiency	96	18:44:59
10	3-ketothiolase deficiency	114	18:44:32
9	#1 OR #4 OR #3	735	18:42:38
8	#4 AND #3	0	18:41:27
7	#4 AND #3 Schema: All	0	18:41:27
6	(#1) AND (#2)	2,152	18:39:43
4	beta ketothiolase deficiency	93	18:38:43
5	beta ketothiolase deficiency	Filter: Systematic Review	0
3	(guideline [Publication Type] OR practice guideline [Publication Type] or recommendation*[Title] OR standard* [Title] OR standard* [Title] OR guideline* [Title])	238,276	18:35:48
2	Newborn screening[MeSH Terms]	10,793	18:35:03
1	"Metabolism, Inborn Errors"[Mesh]	159,795	18:33:41

e) Search from Guidelines International Network

#	QUERY	RESULTS
1	Organic acid disorders	0
2	Beta-ketothiolase deficiency	0
3	3-ketothiolase deficiency Alpha-methyl-acetoacetyl-CoA thiolase deficiency Mitochondrial acetoacetyl-coenzyme A thiolase deficiency T2 deficiency	0
4	3-oxothiolase deficiency	0
5	Alpha methylacetoacetic aciduria	0
6	Mitochondrial acetoacetyl-coenzyme A thiolase deficiency	0
7	T2 deficiency	0
8	Newborn Screening	2

f) Search from Cochrane Database

The screenshot shows a search interface with the following details:

- Search history:
 - #1 beta ketothiolase deficiency (Limits: 1)
 - #2 clinical practice guidelines (Limits: 17643)
 - #3 #1 and #2 (Limits: 1)
 - #4 systematic review (Limits: 32880)
 - #5 meta-analysis (Limits: 25771)
 - #6 #4 AND #5 AND #1 (Limits: 0)
 - #7 organic acid disorders (Limits: 173)
 - #8 Any MeSH descriptor in all MeSH products (MeSH: 0)
 - #9 Type a search term or use the S or MeSH buttons to compose (S: , MeSH: , Limits: N/A)
- Buttons: Save this search, View saved searches, Search help, View fewer lines, Print.
- Checkboxes: Clear all, Highlight orphan lines.

g) Search from Google Scholar

#	QUERY	RESULTS
1	Beta-ketothiolase deficiency screening	1910
2	"Beta-ketothiolase deficiency" screening AND clinical practice guidelines	310
3	"Beta-ketothiolase deficiency" screening AND clinical practice guidelines recommendations evidence; filter: 2015-2021	163

		Yield	

APPENDIX B. GRADE Evidence Profile

Certainty assessment							Summary of findings		
Participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event rates (%)		Impact
							With Selective Screening	With Screening for Beta Ketothiolase Deficiency	
Mortality									
24 (1 observational study)	serious ^a	not serious	serious ^b	serious ^c	publication bias strongly suspected ^d	⊕○○○ Very low	3/10 (30.0%)	0/14 (0.0%)	
Metabolic Decompensation									
29 (1 observational study)	serious ^a	not serious	serious ^b	serious ^c	publication bias strongly suspected ^d	⊕○○○ Very low	11/15 (73.3%)	3/14 (21.4%)	
Neurologic Impairment									
29 (1 observational study)	serious ^a	not serious	serious ^b	serious ^c	publication bias strongly suspected ^d	⊕○○○ Very low	7/15 (46.7%)	0/14 (0.0%)	

CI: confidence interval

Explanations

- a. Data was based on a single retrospective study from only 18 regions of China. Only available BKTD clinical reports that included the Chinese population were analyzed
- b. Reported outcomes were based on the regions NBS and selective metabolic screening database; no data on unscreened newborns reported
- c. Based on a single study with small size of subjects
- d. Not all outcomes from subjects identified through selective metabolic screening were retrieved

APPENDIX C. GRADE Evidence Profile for Diagnostic Accuracy

Sensitivity	0.52 (95% CI 0.35-0.68).					Prevalences	0.0004%		
Specificity	NA								
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested	Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 0.0004%	
True positives (patients with Beta Ketothiolase Deficiency)	2 studies 17588190 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	serious ^b	serious ^{a,c}	none	0 to 0	 VERY LOW
False negatives (patients incorrectly classified as not having Beta Ketothiolase Deficiency)								0 to 0	
True negatives (patients without Beta Ketothiolase Deficiency)	2 studies 17588190 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	serious ^b	serious ^{a,c}	none	999 to 1000	 VERY LOW
False positives (patients incorrectly classified as having Beta Ketothiolase Deficiency)								0 to 1	

Explanations

a. Only screen-positive patients underwent confirmatory testing with interpreters aware of positive result. Tests done on screen-negative patients were not stated

b. Both studies differ in reference standard or confirmatory test

c. Specificity cannot be derived since true disease state of screen-negative patients is unclear

11. Newborn Screening for Holocarboxylase Synthetase Deficiency

APPENDIX A. Search Strategies

5. Search from CPG Database

h) NICE (National Institute for Health and Care Excellence)

#	QUERY	RESULTS
1	Organic acid disorders	453
	Yield for:	
	<ul style="list-style-type: none"> ▪ Holocarboxylase synthetase deficiency ▪ Biotin-(propionyl-CoA-carboxylase) ligase deficiency ▪ Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency ▪ Early-onset biotin-responsive multiple carboxylase deficiency ▪ Early-onset combined carboxylase deficiency ▪ HLCS deficiency ▪ Infantile multiple carboxylase deficiency 	0 0 0 0 0 0 0 0
2	Newborn screening	2043
	Filter: evidence summary	538
	Yield for:	
	<ul style="list-style-type: none"> ▪ Holocarboxylase synthetase deficiency ▪ Biotin-(propionyl-CoA-carboxylase) ligase deficiency ▪ Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency ▪ Early-onset biotin-responsive multiple carboxylase deficiency ▪ Early-onset combined carboxylase deficiency ▪ HLCS deficiency ▪ Infantile multiple carboxylase deficiency 	0 0 0 0 0 0 0 0

i) U.S. Preventive Services Task Force (USPSTF)

#	QUERY	RESULTS
1	Organic acid disorders	0
	Yield for:	
	<ul style="list-style-type: none"> ▪ Holocarboxylase synthetase deficiency ▪ Biotin-(propionyl-CoA-carboxylase) ligase deficiency ▪ Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency ▪ Early-onset biotin-responsive multiple carboxylase deficiency ▪ Early-onset combined carboxylase deficiency ▪ HLCS deficiency ▪ Infantile multiple carboxylase deficiency 	0 0 0 0 0 0 0 0
2	Newborn screening	129
	Filter: Pediatric	8
	Yield for:	
	<ul style="list-style-type: none"> ▪ Holocarboxylase synthetase deficiency ▪ Biotin-(propionyl-CoA-carboxylase) ligase deficiency ▪ Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency ▪ Early-onset biotin-responsive multiple carboxylase deficiency ▪ Early-onset combined carboxylase deficiency ▪ HLCS deficiency ▪ Infantile multiple carboxylase deficiency 	0 0 0 0 0 0 0 0

j) Canadian Task Force for Preventive Health Care (CTFPHC)

#	QUERY	RESULTS
1	Organic acid disorders	0
2	Newborn screening	6
	Yield for: <ul style="list-style-type: none"> ▪ Holocarboxylase synthetase deficiency ▪ Biotin-(propionyl-CoA-carboxylase) ligase deficiency ▪ Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency ▪ Early-onset biotin-responsive multiple carboxylase deficiency ▪ Early-onset combined carboxylase deficiency ▪ HLCs deficiency ▪ Infantile multiple carboxylase deficiency 	0 0 0 0 0 0 0

6. Search from PubMed

Search	Actions	Details	Query	Results	Time
#11	...	!	> Search: #1 OR #2 OR #5 OR #6 OR #7 OR #8	37	11:29:25
#17	...	!	> Search: #11 AND #16	0	11:01:02
#16	...		> Search: Diagnostic accuracy	253,181	11:00:34
#15	...	!	> Search: #11 AND #14	2	10:59:47
#14	...		> Search: newborn screening[MeSH Terms]	10,793	10:59:36
#13	...	!	> Search: #9 AND #11	0	10:59:12
#12	...	!	> Search: #9 OR #11	238,306	10:58:59
#10	...		> Search: #1 and #9	0	10:57:21
#9	...		> Search: (guideline [Publication Type] OR practice guideline [Publication Type] or recommendation*[Title] OR standard* [Title] OR standard* [Title] OR guideline* [Title])	238,269	10:56:59
#8	...		> Search: Infantile multiple carboxylase deficiency[MeSH Terms]	37	10:55:23
#7	...		> Search: HLCs deficiency[MeSH Terms]	37	10:55:04
#6	...		> Search: Early-onset combined carboxylase deficiency[MeSH Terms]	37	10:54:49
#5	...		> Search: Early-onset biotin-responsive multiple carboxylase deficiency[MeSH Terms]	37	10:54:32
#4	...	!	> Search: Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency[MeSH Terms]	0	10:54:01
#3	...	!	> Search: Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency[MeSH Terms] - Schema: all	0	10:54:01
#2	...		> Search: Biotin-(propionyl-CoA-carboxylase) ligase deficiency[MeSH Terms]	37	10:53:37
#1	...		> Search: holocarboxylase synthetase deficiency[MeSH Terms]	37	10:49:29

Articles were of review type, case reports and diagnostic algorithm for patients with seizures and dermatologic concerns.

7. Search from Guidelines International Network

#	QUERY	RESULTS
1	Organic acid disorders	0
	<ul style="list-style-type: none"> ▪ Holocarboxylase synthetase deficiency ▪ Biotin-(propionyl-CoA-carboxylase) ligase deficiency ▪ Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency 	0 0 0

	<ul style="list-style-type: none"> ▪ Early-onset biotin-responsive multiple carboxylase deficiency ▪ Early-onset combined carboxylase deficiency ▪ HLCS deficiency ▪ Infantile multiple carboxylase deficiency 	0 0 0 0
2	Newborn screening	2
3	Inborn Error of Metabolism	1

8. Search from Google Scholar

A. Holocarboxylase deficiency

#	QUERY	RESULTS
1	holocarboxylase synthetase deficiency screening newborns clinical practice guidelines recommendations evidence	482
2	"holocarboxylase synthetase deficiency" screening newborns clinical practice guidelines recommendations evidence	235
	Yield Lund, A. M., Hougaard, D. M., Simonsen, H., Andresen, B. S., Christensen, M., Dunø, M., Skogstrand, K., Olsen, R. K., Jensen, U. G., Cohen, A., Larsen, N., Saugmann-Jensen, P., Gregersen, N., Brandt, N. J., Christensen, E., Skovby, F., & Nørgaard-Pedersen, B. (2012). Biochemical screening of 504,049 newborns in Denmark, the Faroe Islands and Greenland — Experience and development of a routine program for expanded newborn screening. <i>Molecular Genetics and Metabolism</i> , 107(3), 281–293. https://doi.org/10.1016/j.ymgme.2012.06.006 Tangeraas, T., Sæves, I., Klingenberg, C., Jørgensen, J., Kristensen, E., Gunnarsdottir, G., Hansen, E. V., Strand, J., Lundman, E., Ferdinandusse, S., Salvador, C. L., Woldseth, B., Blidsrud, Y. T., Sagredo, C., Olsen, Y. E., Berge, M. C., Trømborg, A. K., Ziegler, A., Zhang, J. H., . . . Pettersen, R. D. (2020). Performance of Expanded Newborn Screening in Norway Supported by Post-Analytical Bioinformatics Tools and Rapid Second-Tier DNA Analyses. <i>International Journal of Neonatal Screening</i> , 6(3), 51. https://doi.org/10.3390/ijns6030051	2
3	"Biotin-(propionyl-CoA-carboxylase) ligase deficiency" screening newborns clinical practice guidelines recommendations evidence	0
4	Biotin-(propionyl-CoA-carboxylase) ligase deficiency screening newborns clinical practice guidelines recommendations evidence	125
	Yield Dupuis, L. (1996). Clustering of mutations in the biotin-binding region of holocarboxylase synthetase in biotin-responsive multiple carboxylase deficiency. <i>Human Molecular Genetics</i> , 5(7), 1011–1016. https://doi.org/10.1093/hmg/5.7.1011	
5	"Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency" screening newborns clinical practice guidelines recommendations evidence	0
6	Biotin-(propionyl-coenzyme A-carboxylase) ligase deficiency screening newborns clinical practice guidelines recommendations evidence	31
	Yield 7	0
7	Early-onset biotin-responsive multiple carboxylase deficiency screening newborns clinical practice guidelines recommendations evidence	255 000
8	"Early-onset biotin-responsive multiple carboxylase deficiency" screening newborns clinical practice guidelines recommendations evidence	2
	Yield 9	0
9	Early-onset combined carboxylase deficiency screening newborns clinical practice guidelines recommendations evidence	1010
10	"Early-onset combined carboxylase deficiency" screening newborns clinical practice guidelines recommendations evidence	0
	Yield	

11	HLCS deficiency screening newborns clinical practice guidelines recommendations evidence	278
12	HLCS deficiency screening newborns clinical practice guidelines recommendations evidence	8
	<p>Yield:</p> <p>Quinonez, S. C., Seeley, A. H., Lam, C., Glover, T. W., Barshop, B. A., & Keegan, C. E. (2016). Paracentric Inversion of Chromosome 21 Leading to Disruption of the HLCS Gene in a Family with Holocarboxylase Synthetase Deficiency. <i>JIMD Reports</i>, 55–61. https://doi.org/10.1007/8904_2016_9</p> <p>Slavin, T. P., Zaidi, S. J., Neal, C., Nishikawa, B., & Seaver, L. H. (2013). Clinical presentation and positive outcome of two siblings with holocarboxylase synthetase deficiency caused by a homozygous I216R mutation. <i>JIMD Reports</i>, 109–114. https://doi.org/10.1007/8904_2013_252</p>	2
13	Infantile multiple carboxylase deficiency screening newborns clinical practice guidelines recommendations evidence	1100
14	Infantile multiple carboxylase deficiency screening newborns clinical practice guidelines recommendations evidence	2
	Yield	0

APPENDIX B. GRADE Evidence Profile

Question: Should newborn screening be used to screen for holocarboxylase synthetase deficiency in asymptomatic newborns?

Sensitivity	1.00 (95% CI: N/A)				Prevalence	0.000093%			
Specificity	N/A								
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence						
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias		
True positives (patients with holocarboxylase synthetase deficiency)	1 study 504049 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	serious ^b	serious ^b	none	1 (0 to 0)	⊕○○○ VERY LOW
False negatives (patients incorrectly classified as not having holocarboxylase synthetase deficiency)								0 (1 to 1)	
True negatives (patients without holocarboxylase synthetase deficiency)	1 study 504049 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	serious ^b	serious ^b	none	0 (0 to 0)	⊕○○○ VERY LOW
False positives (patients incorrectly classified as having holocarboxylase synthetase deficiency)								999999 (999999 to 999999)	

Explanations

a. Interpreters were already aware of a positive newborn screening test during the conduct of the mutational analysis.

b. No other studies for comparison

APPENDIX C. Summary of Case Reports and Case Series

Study	Author	Country	# of Cases	Age	Sex	Screening	Confirmed?	Early or Late?	Onset of symptoms?	Manifestations	Treatment	Outcome
1	Xiong 2020	China	2	8 months	F	Urine organic acid analysis	Exome sequencing	Late	5 months	Erythematous dermatitis, motor delay	Oral biotin 30 mg/day --> 10 mg/day	Improved skin lesions
					M	Urine organic acid analysis	Exome sequencing	Late	1 month	Skin lesions	Oral biotin 30 mg/day --> 5 mg/day	Weight gain and improvement of skin lesions
2	Zheng 2020	China	1	1 year old	M	Urine organic acid analysis and plasma acylcarnitine	DNA sequencing analysis	Late	1 year	Skin lesions, hearing loss	Oral biotin 20 mg/day	Improved skin rash, normal acid-base, reversed hearing loss
3	De Castro 2015	Ghana	2	Newborn	M	Urine organic acid analysis and plasma acylcarnitine	DNA sequencing analysis	Early	2nd DOL	Dyspnea, metabolic acidosis	Oral biotin 20 mg/day	Expired
					Newborn	Urine organic acid analysis and plasma acylcarnitine	DNA sequencing analysis	Early	1st DOL	Tachypnea, metabolic acidosis	Oral biotin 20 mg/day --> 40 mg/day	Weaned from oxygen support, decreased lactate levels
4	Hui 2011	Vietnam	1	6 years old	M	Urine organic acid analysis and plasma acylcarnitine	DNA sequencing analysis	Late	18 months	Recurrent URTI, vomiting	Oral biotin 10 mg/day	Normal acid-base, resolved skin rash
5	Wilson 2005	Samoa	7	18 hours 24 hours 6 hours 12 hours 24 hours 24 hours 24 hours	N	Urine organic acid analysis and plasma acylcarnitine	Enzyme determination and DNA Sequencing	Early	Within 24 hours	Severe lactic acidosis	None	Expired after 3 days
					D						Oral biotin 20 mg	Slow improvement
											Oral biotin 10 mg	Expired after 3 months
											Oral biotin 20 mg	Moderate improvement
											None	Expired after 3 days
											Oral biotin 40 mg	Good recovery after 48 hours
											Oral biotin 20 mg --> 40 mg	Good recovery after 72 hours

6	Yokoi 2009	Japan	2	5 years	F	Urine organic acid analysis	DNA sequencing analysis	Early	Within 24 hours	Severe metabolic acidosis	Oral biotin	Mild psychomotor retardation
		Newborn	N	ND		Urine organic acid analysis	DNA sequencing analysis	Early	Within 24 hours	Severe metabolic acidosis	Oral biotin 60 mg/day	Clinically stable
7	Van Hove 2009	USA	1	Newborn	F	Urine organic acid analysis	DNA sequencing analysis	Early	Immediately after birth	Skin lesions, metabolic acidosis and respiratory distress	Oral biotin 10 mg/day --> 100 mg/day --> biotin-containing formula	Improved rash and no developmental delay
8	Dion 2021	USA	3	9 years old	M	Plasma acylcarnitine	DNA sequencing analysis	Late	5 months	Skin lesions	Oral biotin	Expired
			6 years old	M	NBS	Plasma acylcarnitine and DNA sequencing		Late	12 months	Skin lesions	Oral biotin 40 mg --> 30 mg	Dramatic resolution of dermatitis
			12 years old	M	Biochemical testing		DNA sequencing analysis	Late	12 months	Skin lesions, metabolic acidosis, seizures	Oral biotin 50 mg	Rapid skin improvement and normal developmental milestone
9	Santer 2003	Germany	1	8 year old	F	Urine organic acid analysis	Enzyme determination and DNA Sequencing	Late	10 months	Delayed psychomotor development, tachypnea and metabolic acidosis	Oral biotin 10 mg/day --> 200 mg/day	Mild muscular hypotonia but improved skin lesions and normal acid-base; noted with signs of attention deficit and visuomotoric dysfunction
10	Fatima 2020	Pakistan	1	2 months	F	Urine organic acid analysis	Enzyme determination	Early	12th day of life	Skin rash and infantile seizures	Oral biotin	Poor response to supplementation
11	Morrone 2002	Italy	4	5 years	M	Urine organic acid analysis	Enzyme determination and DNA Sequencing	Late	5 months	Erythematous dermatitis	Oral biotin 10 mg/day	Well and normal psychomotor development
			23 months	M	Urine organic acid analysis	Enzyme determination and DNA Sequencing	Late	1 month	Mild skin lesions	Oral biotin 10 mg/day	Normal psychomotor development and complete disappearance of skin lesions	
			6 months	M	Urine organic acid analysis	DNA sequencing analysis	Late	6 months	Lethargy, respiratory distress and severe metabolic acidosis	None	Expired	

2 years	M	Urine organic acid analysis	DNA sequencing analysis	Early	After birth	Lethargy, respiratory distress and severe metabolic acidosis	Oral biotin 20 mg/day --> 100 mg/day	Global developmental delay
---------	---	-----------------------------	-------------------------	-------	-------------	--	--------------------------------------	----------------------------

12. Newborn Screening for Isovaleric Acidemia

APPENDIX A. Search Strategy

1. Search from CPG Database

- a) NICE (National Institute for Health and Care Excellence)

#	QUERY	RESULTS
1	Organic acid disorders	453
	Yield for: <ul style="list-style-type: none">▪ Isovaleric acidemia	0
2	Newborn screening	2043
	Filter: evidence summary	538
	Yield for: <ul style="list-style-type: none">▪ Isovaleric acidemia	0

- b) U.S. Preventive Services Task Force (USPSTF)

#	QUERY	RESULTS
1	Organic acid disorders	0
	Yield for: <ul style="list-style-type: none">▪ Isovaleric acidemia	0
2	Newborn screening	129
	Filter: Pediatric	8
	Yield for: <ul style="list-style-type: none">▪ Isovaleric acidemia	0

- c) Canadian Task Force for Preventive Health Care (CTFPHC)

#	QUERY	RESULTS
1	Organic acid disorders	0
2	Newborn screening	6
	Yield for: <ul style="list-style-type: none">▪ Isovaleric acidemia	0

2. Search from CPG Database

Search	Actions	Details	Query	Results	Time
#33	...	>	Search: #24 and #25 Filters: in the last 5 years, Newborn: birth-1 month	63	08:55:09
#32	...	>	Search: #24 and #25 Filters: in the last 10 years, Newborn: birth-1 month	142	08:54:51
#31	...	>	Search: #24 and #25 Filters: in the last 10 years	1,010	08:54:47
#26	...	>	Search: #24 and #25	1,914	08:54:35
#30	...	>	Search: #26 and #27	29	08:48:50
#27	...	>	Search: diagnostic accuracy	251,561	08:48:02
#25	...	>	Search: organic acid disorders	6,621	08:47:14
#24	...	>	Search: screening	5,290,774	08:46:56
#23	...	>	Search: #22 and #18	8	06:43:40
#22	...	>	Search: #21 and #15	51	06:39:40
#21	...	>	Search: #4 and #20	733	06:39:21
#20	...	>	Search: inborn error of metabolism[MeSH Terms]	159,474	06:39:03
#19	...	>	Search: #3 and #18	1	06:37:29
#18	...	>	Search: tandem mass spectrometry	86,988	06:37:14
#7	...	>	Search: #4 and #6	1	06:36:43
#17	...	>	Search: #16 and #3	0	06:26:45
#16	...	>	Search: #4 and #15	309	06:26:25
#15	...	>	Search: newborn screening[MeSH Terms]	10,750	06:26:10
#14	...	>	Search: #4 and #12	0	06:25:18
#13	...	>	Search: beta-keto thialase deficiency - Spellcheck off	0	06:24:35
#12	...	>	Search: beta ketothiolase deficiency	93	06:24:35
#11	...	>	Search: #4 and #10	0	06:24:12
#10	...	>	Search: holocarboxylase deficiency	182	06:23:57
#9	...	>	Search: #4 and #8	3	06:23:32
#8	...	>	Search: biotinidase deficiency	632	06:23:24
#6	...	>	Search: isovaleric acidemia	235	06:22:42
#5	...	>	Search: #3 and #4	0	06:22:35
#4	...	>	Search: (guideline [Publication Type] OR practice guideline [Publication Type] or recommendation*[Title] OR standard* [Title] OR standard* [Title] OR guideline* [Title])	237,392	06:22:16
#3	...	>	Search: organic acid disorders[MeSH Terms]	83	06:21:42
#2	...	>	Search: organic acid disorders[MeSH Terms] Filters: Newborn: birth-1 month	10	06:21:14
#1	...	>	Search: organic acid disorder Filters: Newborn: birth-1 month	490	06:11:43

3. Search from Guidelines International Network

#	QUERY	RESULTS
1	Isovaleric acidemia	0
2	Newborn screening	2
3	Inborn Error of Metabolism	1

4. Search from PubMed

History and Search Details				 Download	 Delete
Search	Actions	Details	Query	Results	Time
#1	...	>	Search: "newborn screening" "isovaleric acid** Filters: from 2015 - 2021	30	19:03:33
#5	...	>	Search: "newborn screening" "isovaleric acid**"	54	19:03:27
#8	...	>	Search: "newborn screening" "isovaleric acid** Filters: Systematic Review	1	19:03:22
#7	...	>	Search: "newborn screening" "isovaleric acid** Filters: Meta-Analysis, Systematic Review	1	18:54:12
#6	...	>	Search: "newborn screening" "isovaleric acid** Filters: Systematic Review, from 2015 - 2021	0	18:54:03
#4	...	>	Search: "newborn screening" "isovaleric acid** Filters: Clinical Trial, from 2015 - 2021	0	18:53:39
#2	...	>	Search: "isovaleric acid**"	890	18:51:01

Showing 1 to 7 of 7 entries

5. Search from Google Scholar

A. Isovaleric acidemia / isovaleric aciduria

#	QUERY	RESULTS
1	isovaleric acidemia screening newborns clinical practice guidelines recommendations evidence	1110
2	isovaleric acidemia screening newborns clinical practice guidelines recommendations evidence; filter: 2015-2021	497
3	"isovaleric acidemia" screening newborns clinical practice guidelines recommendations evidence; filter: 2015-2021	304
4	"isovaleric aciduria" screening newborns clinical practice guidelines recommendations evidence; filter: 2015-2021	103
	Yield	

APPENDIX B. Characteristics of Included Studies

Table 1. Summary of study characteristics for Benefit of Screening

Author	Study design	Country	Number of patients	Intervention Group(s)	Control	Outcomes (Screened vs No Screening)
Mutze 2021	Prospective cohort Follow up range (years): Classic, screened = 12 Classic, Unscreened =11.2 Mild = 7.1	Germany	N=94	Screened=74 Mild=70 Classic=24	Unscreened, Classic=21	<p>Mortality: -Total: 1/26 vs. 29/155 -Classic (3.8%, 1/26 vs. 18.7%; 29/155; P = .08042) -Mild (0/70)</p> <p>With metabolic decompensation: -Classic (71% 17/24) -Mild (0/67)</p> <p>Total # of metabolic decompensations: Classic (28% 1.2 vs. 69% 3.3)</p> <p>Neurocognitive impairment: -Total: 21/91 vs. 41/108 -Classic (79.2%, 19/24) -Mild (3%)</p>
Szymanska 2020	Retrospective cohort Follow up range (years): NBS and Family screening = 2.5 (1.5-9) Unscreened= 17 (5-20)	Poland	N=10	Screened = 3 (NBS = 1; Family screening = 2)	Unscreened = 7	<p>Neurocognitive impairment: 0/3 vs. 4/7</p> <p>Metabolic decompensations: 3/3 vs. 6/7</p>
Couce 2016	Retrospective cohort Follow up range: 9 years, 10 months (6 months – 22 years)	Spain	N=16	Screened = 10 (NBS = 8, Family history = 2)	Unscreened = 6	<p>Mortality: 0/10 vs. 1/6</p> <p>Metabolic decompensation: 1/10 vs 2/6 (all with severe phenotype)</p> <p>PDI/IQ: 92.6 ± 8.9 vs. 87.4 ± 2.9 (but all were within normal)</p>

Table 2. Summary of study characteristics for Test Accuracy

Study ID	Setting/ Study design	Index Test/ Specimen	Population	Sample Size	Reference standard	Reference Standard Specimen	Outcomes
Sorensen 2020	Sweden Nov 2010 – July 2019	First-tier: Tandem mass	newborn	1,00,316 IVA=6	urine organic acids,	Urine, blood	Results: All recalls 18 TP=6 FP=12

	Prospective cohort	spectrometry (MS/MS) Second-tier: Ultraperformance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) Dried Blood Spot			plasma acylcarnitine profiles, DNA analysis		FN=0 TN=1,000,316 SN=100% SP=99.99% PPV=33.33% NPV=100% Before 2 nd -tier: All recalls=11 TP=0 FP=11 PPV=0% After 2 nd -tier: All recalls=7 TP=6 FP=1 PPV=86% Incidence after screening 1:170,000 Incidence before screening 1:530,000
Tangeraas 2020	Norway March 2012 – Feb 2020 Prospective cohort	First-tier: Tandem mass spectrometry (MS/MS) Second-tier: Ultraperformance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) Dried Blood Spot	newborn	461,369 IVA=2	DNA analysis Supplementary: urine organic acids, plasma acylcarnitine profiles	Urine, blood	Results: TP=2 FP=0 FN=0 TN=461,367 SN=100% SP=100% PPV=100% NPV=100% Incidence after screening 1:230,684 Incidence before screening 1:149,147
Lin 2020	Quanzhou, China Jan 2014 - Sept 2019 Retrospective cohort	Tandem mass spectrometry (MS/MS) NBS	newborn	422,346 IVA=5	Urine organic acids DNA analysis	Urine, blood	Results: TP=5 FP=401 Positivity rate=0.1% (406/422,346) PPV= 1.2% (5/406) Incidence: 1:84,469 SN,SP, and NPV could not be calculated as there was no systematic follow-up of

							babies who screened negative
Yunus 2016	Malaysia June 2006- Dec 2008 Prospective cohort	Tandem mass spectrometry (MS/MS) Dried blood spot	newborn	29,859 IVA=1	Urine organic acids	Urine	Results: TP=1 SN=100% Incidence: 1:29,859 SP, PPV, and NPV could not be calculated as there was specific data on the FP rate for IVA
Lim 2014	Singapore July 2006- April 2014 Prospective cohort	Tandem mass spectrometry (MS/MS) Dried blood spot	newborn	177,267 IVA=1	Biochemical study and DNA analysis	Urine, blood	Results: TP=1 FP=12 SP=99.99 PPV=8% Incidence: 1:177,000 SN and NPV could not be calculated as there was no systematic follow-up of babies who screened negative
Ensenauer 2011	Germany Retrospective cohort	Tandem mass spectrometry (MS/MS) Dried blood spot	newborn	1,612,105	Urine organic acids and DNA analysis	Urine, blood	Results: Recall rate = 0.024% TP=24 FN=0 SN=100% SP=99.98% PPV=7% Incidence: 1:67,000 NPV could not be calculated since FP and TN rates were not specified

APPENDIX C. GRADE Evidence Profile

NBS for IVA compared to no NBS for healthy asymptomatic newborns

Certainty assessment							Summary of findings				
Participants (studies) Follow-up	Risk of bias	Inconsiste- ncy	Indirectne- ss	Imprecisio- n	Publication bias	Overall certainty of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With no NBS	With NBS for IVA		Risk with no NBS	Risk difference with NBS for IVA

Mortality

197 (2 observational studies)	serious ^a	serious ^b	serious ^c	serious ^d	none	 Very low	30/161 (18.6%)	1/36 (2.8%)	OR 0.120 (0.016 to 0.950)	186 per 1,000	160 fewer per 1,000 (from 183 fewer to 8 fewer)
--	----------------------	----------------------	----------------------	----------------------	------	--	-------------------	-------------	---	---------------	---

Metabolic decompensation

26 (2 observational studies)	serious ^a	serious ^b	serious ^c	serious ^d	none	 Very low	8/13 (61.5%)	4/13 (30.8%)	OR 0.280 (0.055 to 1.410)	615 per 1,000	306 fewer per 1,000 (from 535 fewer to 77 more)
---------------------------------------	----------------------	----------------------	----------------------	----------------------	------	--	-----------------	-----------------	---	---------------	---

Neurocognitive Impairment

209 (2 observational studies)	serious ^a	serious ^b	serious ^c	serious ^d	none	 Very low	45/115 (39.1%)	21/94 (22.3%)	OR 0.48 (0.24 to 0.83)	391 per 1,000	155 fewer per 1,000 (from 258 fewer to 43 fewer)
--	----------------------	----------------------	----------------------	----------------------	------	--	-------------------	------------------	----------------------------------	---------------	--

Quality-Adjusted Life Year

NBS for IVA compared to no NBS for healthy asymptomatic newborns

Certainty assessment							Summary of findings
0 (1 observational study)	serious ^a	serious ^b	not serious ^c	serious ^d	none	 Very low	IVA patients noticeably benefit from early detection which extends their QALY by 3.69.

CI: confidence interval; **OR:** odds ratio

Explanations

- a. Outcome assessors knew whether the participants had been screened or clinically detected
- b. Clinical heterogeneity: data includes both mild and classic/severe phenotypes
- c. Non-contemporaneous controls were used
- d. Event rate very low

Should newborn screening be used to screen for isovaleric acidemia in asymptomatic newborns?

Patient or population: asymptomatic newborns:

Pooled sensitivity:1.00 (95% CI: -- to --)|Pooled specificity:1.00 (95% CI: -- to --)

Test result	Number of results per 100,000 patients tested (95% CI)	Number of participants (studies)	Certainty of the Evidence (GRADE)	Comments
True positives	Prevalence0.0001% Typically seen in	1461703 (2)	 Low ^{a,b,c}	Sorensen, 2020: Nov 2010 - July 2019; screened 1,000,334; IVA: All recalls 18, TP 6, FP 12, FN 0, TN 1,000,316, Incidence after screening 1:170,000; Incidence before screening 1:530,000
False negatives	0 (0 to 0)			Tangeraas, 2020: March 2012 - Feb 2020; screened 461,369; IVA: TP 2, FP 0, FN 0, TN 461,367, Incidence after screening 1:230,684; Incidence before screening 1:149,147

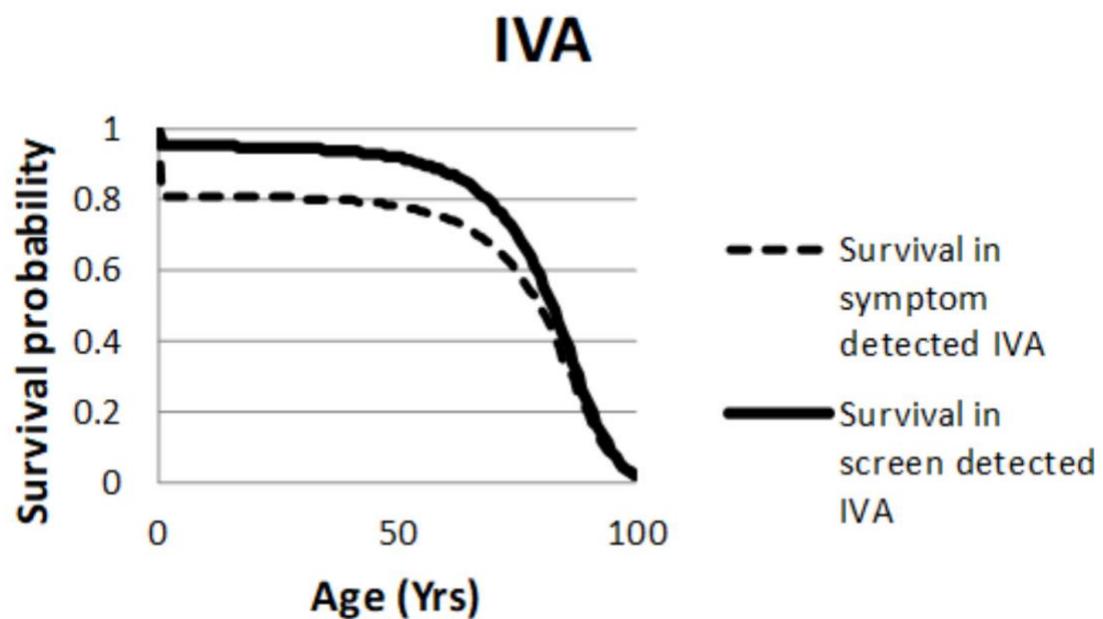
Test result	Number of results per 100,000 patients tested (95% CI)	Number of participants (studies)	Certainty of the Evidence (GRADE)	Comments
	Prevalence 0.0001% Typically seen in			
True negatives	100000 (0 to 0)	1461703 (2)	-	
False positives	0 (100000 to 100000)			

CI: confidence interval

Explanations

- a. Outcome assessors knew whether the participants had been screened or clinically detected
- b. Methodological heterogeneity: with 2nd-tier testing and DNA analysis as confirmatory test
- c. Event rate is low

APPENDIX D. Modelled Survival Estimates in the Screened and Non-Screened Cohorts (16)



13. Screening for Developmental Delay

APPENDIX A. Search Strategy

Last Search Done: September 2, 2021

C. Clinical Practice Guidelines Databases

6. National Institute for Health and Care Excellence (NICE)

QUERY	RESULTS	YIELD
Developmental screening	38	1

7. US Preventive Services Task Force (USPSTF)

QUERY	RESULTS	YIELD
Developmental screening Filter: Development and Behavior, Pediatric	9	0

8. UK National Screening Committee

QUERY	RESULTS	YIELD
Developmental screening	0	0
Developmental delay	0	0

9. Canadian Task Force for Preventive Health Care

QUERY	RESULTS	YIELD
Developmental screening	0	0
Developmental delay	1	1

D. PubMed

Maternal risk factors in general yield: 39

(("developmental"[All Fields] OR "developmentally"[All Fields]) AND ("delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]) AND ("loatrrfree full text"[Filter] AND "humans"[MeSH Terms]) AND ((("maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]) AND ("risk factors"[MeSH Terms] OR ("risk"[All Fields] AND "factors"[All Fields]) OR "risk factors"[All Fields]) AND ("loatrrfree full text"[Filter] AND "humans"[MeSH Terms]))) AND ((ffrft[Filter] AND (humans[Filter]))

Translations

developmental: "developmental"[All Fields] OR "developmentally"[All Fields]

delay: "delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]

ffrft[Filter]: loatrrfree full text[subset]

humans[Filter]: humans[MH]

maternal: "maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]

risk factors: "risk factors"[MeSH Terms] OR ("risk"[All Fields] AND "factors"[All Fields]) OR "risk factors"[All Fields]

ffrft[Filter]: loatrrfree full text[subset]

humans[Filter]: humans[MH]

Search specific to SMOKING yield: 159

(("maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]) AND ("smoke"[MeSH Terms] OR "smoke"[All Fields] OR "smoke s"[All Fields] OR "smoked"[All Fields] OR "smokes"[All Fields] OR "smoking"[MeSH

Terms] OR "smoking"[All Fields] OR "smokings"[All Fields] OR "smoking s"[All Fields]) AND ((("developmental"[All Fields] OR "developmentally"[All Fields]) AND ("delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]))) AND ((ffrft[Filter]) AND (humans[Filter]))

Translations

maternal: "maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]

smoking: "smoke"[MeSH Terms] OR "smoke"[All Fields] OR "smoke's"[All Fields] OR "smoked"[All Fields] OR "smokes"[All Fields] OR "smoking"[MeSH Terms] OR "smoking"[All Fields] OR "smokings"[All Fields] OR "smoking's"[All Fields]

developmental: "developmental"[All Fields] OR "developmentally"[All Fields]

delay: "delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]

Search specific to ALCOHOL yield: 93

((("maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]) AND ("alcohol drinking"[MeSH Terms] OR ("alcohol"[All Fields] AND "drinking"[All Fields]) OR "alcohol drinking"[All Fields] OR "alcohol"[All Fields] OR "alcohol use"[All Fields]) AND ("loatrtfree full text"[Filter] AND "humans"[MeSH Terms]) AND ((("developmental"[All Fields] OR "developmentally"[All Fields]) AND ("delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]))) AND ((ffrft[Filter]) AND (humans[Filter]))

Translations

maternal: "maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]

alcohol use: "alcohol drinking"[MeSH Terms] OR ("alcohol"[All Fields] AND "drinking"[All Fields]) OR "alcohol drinking"[All Fields] OR ("alcohol"[All Fields]) OR "alcohol use"[All Fields]

ffrft[Filter]: loatrtfree full text[subset]

humans[Filter]: humans[MH]

developmental: "developmental"[All Fields] OR "developmentally"[All Fields]

delay: "delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]

Search specific to ANEMIA yield: 25

((("maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]) AND ("anaemia"[All Fields] OR "anemia"[MeSH Terms] OR "anemia"[All Fields] OR "anaemias"[All Fields] OR "anemias"[All Fields]) AND ("loatrtfree full text"[Filter] AND "humans"[MeSH Terms]) AND ((("developmental"[All Fields] OR "developmentally"[All Fields]) AND ("delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]))) AND ((ffrft[Filter]) AND (humans[Filter]))

Translations

maternal: "maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]

anemia: "anaemia"[All Fields] OR "anemia"[MeSH Terms] OR "anemia"[All Fields] OR "anaemias"[All Fields] OR "anemias"[All Fields]

ffrft[Filter]: loatrtfree full text[subset]

humans[Filter]: humans[MH]

developmental: "developmental"[All Fields] OR "developmentally"[All Fields]

delay: "delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]

Search specific to DIABETES yield: 137

((("diabetes, gestational"[MeSH Terms] OR ("diabetes"[All Fields] AND "gestational"[All Fields]) OR "gestational diabetes"[All Fields] OR ("maternal"[All Fields] AND "diabetes"[All Fields]) OR "maternal diabetes"[All Fields]) AND ("loatrtfree full text"[Filter] AND "humans"[MeSH Terms]) AND ((("developmental"[All Fields] OR "developmentally"[All Fields]) AND ("delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]))) AND ((ffrft[Filter]) AND (humans[Filter]))

Translations

maternal diabetes: "diabetes, gestational"[MeSH Terms] OR ("diabetes"[All Fields] AND "gestational"[All Fields]) OR "gestational diabetes"[All Fields] OR ("maternal"[All Fields] AND "diabetes"[All Fields]) OR "maternal diabetes"[All Fields]

ffrft[Filter]: loatrfree full text[subset]

humans[Filter]: humans[MH]

developmental: "developmental"[All Fields] OR "developmentally"[All Fields]

delay: "delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]

Search specific to OBESITY yield: 153

((("developmental"[All Fields] OR "developmentally"[All Fields]) AND ("delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields])) AND ("loatrfree full text"[Filter] AND "humans"[MeSH Terms]) AND ((("obesity, maternal"[MeSH Terms] OR ("obesity"[All Fields] AND "maternal"[All Fields]) OR "maternal obesity"[All Fields] OR ("maternal"[All Fields] AND "obesity"[All Fields])) AND ("loatrfree full text"[Filter] AND "humans"[MeSH Terms]))) AND ((ffrft[Filter]) AND (humans[Filter])))

Translations

developmental: "developmental"[All Fields] OR "developmentally"[All Fields]

delay: "delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]

ffrft[Filter]: loatrfree full text[subset]

humans[Filter]: humans[MH]

maternal obesity: "obesity, maternal"[MeSH Terms] OR ("obesity"[All Fields] AND "maternal"[All Fields]) OR "maternal obesity"[All Fields] OR ("maternal"[All Fields] AND "obesity"[All Fields])

ffrft[Filter]: loatrfree full text[subset]

humans[Filter]: humans[MH]

Search specific to HYPERTENSION yield: 47

((("developmental"[All Fields] OR "developmentally"[All Fields]) AND ("delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields])) AND ("loatrfree full text"[Filter] AND "humans"[MeSH Terms]) AND ((("maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]) AND ("hypertense"[All Fields] OR "hypertension"[MeSH Terms] OR "hypertension"[All Fields] OR "hypertension s"[All Fields] OR "hypertensions"[All Fields] OR "hypertensive"[All Fields] OR "hypertensive s"[All Fields] OR "hypertensives"[All Fields]) AND ("loatrfree full text"[Filter] AND "humans"[MeSH Terms]))) AND ((ffrft[Filter]) AND (humans[Filter])))

Translations

developmental: "developmental"[All Fields] OR "developmentally"[All Fields]

delay: "delay"[All Fields] OR "delayed"[All Fields] OR "delaying"[All Fields] OR "delays"[All Fields]

ffrft[Filter]: loatrfree full text[subset]

humans[Filter]: humans[MH]

maternal: "maternally"[All Fields] OR "maternities"[All Fields] OR "maternity"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields]

hypertension: "hypertense"[All Fields] OR "hypertension"[MeSH Terms] OR "hypertension"[All Fields] OR "hypertension's"[All Fields] OR "hypertensions"[All Fields] OR "hypertensive"[All Fields] OR "hypertensive's"[All Fields] OR "hypertensives"[All Fields]

ffrft[Filter]: loatrfree full text[subset]

humans[Filter]: humans[MH]

APPENDIX B. AGREE II Score Sheet on NICE Guidelines on Developmental follow-up of children and young people born preterm

Domain	Score (1-7)	
Domain 1. Scope and Purpose		
24. The overall objective/s of the guidelines is/are specifically described.	6	7
25. The health question/s covered by the guideline is/are specifically described.	5	7
26. The population to whom the guideline is meant to apply is specifically described	6	7
TOTAL	88.89%	
Domain 2. Stakeholder Involvement		
27. The guideline development group includes individuals from all relevant professional groups	7	7
28. The views and preferences of the target population have been sought.	6	7
29. The target users of the guideline are clearly defined.	6	7
TOTAL	94.44%	
Domain 3. Rigour of Development		
30. Systematic methods were used to search for evidence.	7	7
31. The criteria for selecting the evidence are clearly described.	6	7
32. The strength and limitations of the body of evidence are clearly described.	7	7
33. The methods for formulating the recommendations are clearly described.	6	7
34. The health benefits, side effects, and risks have been considered in formulating the recommendations.	7	7
35. There is explicit link between the recommendations and the supporting evidence.	7	7
36. The guideline has been externally reviewed by experts prior to its publication.	6	7
37. A procedure for updating the guideline is provided.	5	7
TOTAL	94.79%	
Domain 4. Clarity of Presentation		
38. The recommendations are specific and unambiguous.	7	7
39. The different options for management of the condition or health issue are clearly presented.	6	7
40. Key recommendations are easily identifiable.	6	6
TOTAL	91.67%	
Domain 5. Applicability		
41. The guideline describes facilitators and barriers to its application.	5	7
42. The guideline provides advice and/or tools on how the recommendations can be put into practice.	5	7
43. The potential resource implications of applying the recommendations have been considered.	7	7
44. The guideline presents monitoring and/or auditing criteria.	6	7
TOTAL	89.58%	
Domain 6. Editorial Independence		
45. The views of the funding body have not influenced the content of the guideline.	7	7
46. Competing interests of guideline development group members have been recorded and addressed.	7	7
TOTAL	100%	
OVERALL GUIDELINE ASSESSMENT		
Rate the overall quality of this guideline.	6	7
I would recommend this guideline for use	Yes	Yes

APPENDIX C. Summary of studies on the association between gestational age and cerebral palsy

Study	Data Source	Sample and Population studied	Measures of Outcome	Adjustment	Prognostic Outcomes	Study Quality
Odd 2013	Prospective regional cohort study	N=13,843 Analysis compares moderate to late preterm infants (32-36 weeks) to full term (37-42 weeks)	Infants with cerebral palsy were identified from hospital and community health service records, and the diagnosis confirmed at age 4 using the Standard Recording of Motor Deficit	Ethnicity, housing, crowding, maternal education, socio-economic group, car ownership, age, gender, parity, weight, length, head circumference at birth, mode of delivery, maternal hypertension, pyrexia, need for resuscitation at birth	Cerebral palsy at 7 years age Term: reference 32-36 weeks: OR 6.38 (2.28- 17.76)	Moderate
Hirvonen 2014	Population based retrospective cohort using national registry data	Overall sample: n=1039263 Sample size after exclusions: n=1018302 (included for comparisons of cerebral palsy risk at different gestational ages) n=53078	All inpatient and outpatient visits due to a CP diagnosis in public hospitals were registered. The diagnosis of CP in Finland is based on medical history, ultrasound and MRI data, and multidisciplinary evaluations in the Paediatric neurology units of 20 secondary level central hospitals and 5 tertiary level university hospitals. The diagnosis is included in the database as soon as it has been established	Maternal age, maternal smoking status, primiparous, previous C-section, maternal diabetes, multiple pregnancy, order of fetuses, assisted reproductive technology, cervical cerclage, chorionic villus sampling, PROM, preeclampsia, time of birth, antenatal steroid use, place of birth, mode of delivery, gender, gestational weight, birth weight <1500g, Apgar score, umbilical artery pH, admission to neonatal unit, ventilator, resuscitation at birth, phototherapy, antibiotic therapy, RDS, sepsis, intracranial haemorrhage, convulsions and hyperbilirubinaemia.	By the age of 7 years Cerebral palsy Gestational age Term: Reference <32 weeks: OR 9.37 (7.34- 11.96) ^[11] 32+0 to 33+6 weeks: OR 5.12 (4.13-6.34) 34+0 to 36+6 weeks: OR 2.35 (1.99 to 2.77)	Low
Petrini 2009	Regional retrospective cohort study	n=141321 Analysis compares preterm infants to full term (37-41 weeks)	ICD 9 codes of patient diagnoses in electronic medical records were used to identify cases of cerebral palsy and developmental delay/mental retardation.	Maternal ethnicity, sex, multiple pregnancy and size for gestational age.	During follow-up time of up to 5.5 years ^[11] Cerebral palsy Term: Reference 34-36 weeks: HR 3.39 (2.54- 4.52) 30-33 weeks: HR	Moderate

					7.87 (5.38- 11.51)	
Sukhov 2012	Retrospective cohort study using population registry data	n=6,145,357 Analysis compares different groups of preterm infants to term (≥ 37 weeks)	Cerebral palsy was identified through an administrative database from 21 non- profit regional centres which provide therapy services to people with developmental disabilities including CP.	Maternal age, parity, maternal education, payer-source, ethnicity, timing of initiation of prenatal care, number of prenatal visits, gestational age, birthweight, multiple pregnancy, gender, placental abruption, fetal distress, mild to severe birth asphyxia, birth defects, birth trauma, meningitis and cord prolapse.	At between 5 and 15 years Cerebral palsy ^[1] Term: Reference ^[1] 32-36 weeks: OR 2.20 (2.05- 2.36) 28-31 weeks: OR 8.83 (8.04- 9.70) ^[SEP] < 28 weeks: OR 18.21 (16.70-19.86)	Moderate

APPENDIX D. Summary of studies on the association between gestational age and intellectual disability

Study	Data Source	Sample and Population studied	Measures of Outcome	Adjustment	Prognostic Outcomes	Study Quality
Woythaler 2011	Population based prospective cohort study	n=1200 preterm infants (34-36+6 weeks) n=6300 term infants (≥ 37 weeks)	The mental development index (MDI) of the Bayley Short Form Research edition (BSF-R) were used to identify developmental delay and psychomotor developmental delay. Abnormal scores were identified as mild abnormality (between 1SD and 2SD below the mean score) and severe abnormality (<2SD below the mean score).	Gestational age, plurality, maternal race, education, marital status, depression, prenatal care, primary language, infant gender, poverty level, delivery type, fetal growth and any breast milk feeding.	At 2 years chronological age Severe developmental delay Term: Reference 34-36+6 weeks: OR 1.51 (1.26-1.82) Mild developmental delay Term: Reference 34-36+6 weeks: OR 1.43 (1.22-1.67)	Moderate
Serenius 2013	Population based prospective cohort study (EXPRESS)	n=456 preterm infants (<27 weeks) ^[1] n=701 full term controls (37-41 weeks)	Cognitive, language and motor development were all assessed with the Bayley- Scales of Infant and Toddler Development (Bayley- III). Cognitive, language and motor development was considered normal if the	Maternal country of birth (Nordic/non-Nordic), maternal and paternal educational level	At 2.5 years corrected age Mild cognitive impairment Term: Reference ^[1] <27 weeks: OR 4.3 (2.3-7.9) Mild mental developmental delay Term: Reference ^[1] <27 weeks: OR 3.0 (1.8-5.0) Moderate mental developmental delay ^[SEP]	Moderate

			composite score on the respective Bayley-III scale was within 1 SD of the norm, mildly impaired if the score was between 1 and 2SD below the norm, moderately impaired if the score was between 2 and 3 SD below the norm, and severely impaired if the score was < 3SD below the norm.		Term: Reference ^[1] <27 weeks: OR 6.4 (2.4-17.1)	
Larroque 2008	Population based prospective cohort study (EPIPAGE)	n=1534 preterm children born at 22 to 32 completed weeks gestation n=320 term controls born at 39-40 weeks	Mental Processing Composite (MPC) of the Kaufmann Assessment Battery for Children (K-ABC) was used to assess intellectual disability. Scores of <2SD below the mean were taken as abnormal.	Maternal age, parity, maternal education, maternal birthplace and socioeconomic status.	At age 5 years Intellectual disability (MPC score 55-69) Term: Reference 22-32 weeks: OR 3.4 (1.8- 6.4)	High
Petrini 2009	Regional retrospective cohort study	n=141321 Analysis compares preterm infants to full term (37-41 weeks)	ICD 9 codes of patient diagnoses in electronic medical records were used to identify cases of cerebral palsy and developmental delay/mental retardation.	Maternal ethnicity, sex, multiple pregnancy and size for gestational age.	During follow-up time of up to 5.5 years For the outcome of Developmental delay/ mental retardation ^[1] Term: Reference ^[1] 34-36 weeks: HR 1.25 (1.01- 1.54) ^[2] 30-33 weeks: HR 1.90 (1.34- 2.71)	Moderate
Singh 2013	Cross sectional survey	n=85,535 Separated into premature children (born at <37 weeks) and term children (\geq 37 weeks)	Parents were asked to self-report whether their child had been diagnosed with one of the disorders by a doctor or health care provider.	Household composition, place of residence and highest household/ parental education.	During follow-up period of between 2 and 17 years Intellectual disability/ mental retardation Term: Reference <37 weeks: OR 2.74 (2.02- 3.73)	Low
Helderman 2012	Multicentre Prospective cohort study	Sample recruited: n=1506 ^[3] Sample eligible for assessment: n=1200	The assessment of developmental delays (determined by cognitive impairment Mental	Single mother, BMI>30, vaginal/cervical infection,	Intellectual disability (developmental delay - Mental Developmental Index [MDI]) ^[4] Gestational age 23–24 week - (RR	Moderate (the study was downgraded for risk of bias

		Sample analysed after exclusions: n=921	Development Index [MDI]) at 24- months adjusted age at 24-months included the Bayley Scales of Infant Development- 2nd Edition (BSID-II). Cognitive impairment was defined as an MDI <70. An MDI <55 was considered severe cognitive impairment.	caesarean delivery, BWZ <2, mother's education <12 years or >16 years, Hospital cluster	[95% CIs]) Referent group is infants with MDI <70 MDI < 55: 1.9 (0.97, 3.6) MDI = 55–69: 1.0 (0.5, 1.9) Gestational age 25–26 week - (RR [95% CIs]) Referent group is infants with MDI <70 MDI < 55: 1.2 (0.7, 2.1) MDI = 55–69: 0.8 (0.5, 1.3)	because the confounders for adjustment were not reported clearly)
Hillemeier 2011	National longitudinal cohort study	n=7,200	Cognitive delay was assessed at 24 and 48 months age using the Bayley Short Form- Research Edition (BSF-R). Children scoring the lowest 10% of the scale were considered to have cognitive delay. At 48 months, Bayley assessment was not possible due to age, therefore a standardised assessment developed for other large studies of child development. Children scoring lowest 10% were considered to have cognitive delay	Adjustment for sex, age, race/ethnicity, socioeconomic variables, characteristics of gestation and infant status at birth	At 24 months: Cognitive delay Gestational age Term Ref Moderately preterm (33-36 weeks) OR 1.07 (NS, 95% CI not presented) Very preterm (<=32 weeks) 1.52 (NS) The model adjusted for sex, age, race/ethnicity, socioeconomic variables, characteristics of gestation and infant status at birth. At 48 months: Cognitive delay Gestational age Term Ref Moderately preterm (33-36 weeks) 1.10 (NS) Very preterm (<=32 weeks) 1.86 (NS) The model adjusted for sex, age, race/ethnicity, socioeconomic variables, characteristics of gestation and infant status at birth.	Low

APPENDIX E. Summary of studies on the association between gestational age and speech and/or language disorder

Study	Data Source	Sample and Population studied	Measures of Outcome	Adjustment	Prognostic Outcomes	Study Quality
Serenius 2013	Population based prospective cohort study	n=456 preterm infants (<27 weeks) n=701 full term controls (37-41)	Cognitive, language and motor development were all assessed with the Bayley-Scales of Infant and Toddler	Maternal country of birth (Nordic/non-Nordic), maternal and paternal educational level	Mild language impairment at 2.5 years corrected age Term: Reference <27 weeks: OR 3.5 (1.9-6.4)	Moderate

	(EXPRESS)	weeks)	Development (Bayley- III). Cognitive, language and motor development was considered normal if the composite score on the respective Bayley-III scale was within 1 SD of the norm, mildly impaired if the score was between 1 and 2SD below the norm, moderately impaired if the score was between 2 and 3 SD below the norm, and severely impaired if the score was < 3SD below the norm.		Moderate language impairment ^[1] Term: Reference ^[1] <27 weeks: OR 5.1 (1.9-13.8)	
Rabie 2015	Retrospective cohort study using population registry data	n=38802 Analysis compares late preterm infants to full term (39-41+6 weeks)	ICD-9 codes from Medicaid files were used to identify children with ADHD and developmental speech and/or language delay.	Birth weight, SGA and LGA, gender, ethnicity, hospital characteristics and maternal medical comorbidities (diabetes, hypertension, anaemia, chronic lung disease, herpes, neurologic disorder, coagulation disorder, obesity, depression).	At age 3-5 years. Developmental speech and/or language delay Term: Reference ^[1] 34-36+6: HR 1.36 (1.23-1.50)	Low
Wolke 2008	National cohort study	n=308 children born <=25 gestational weeks n=241 children survived to follow-up ^[1] n=160 full-term born children as comparison group, matched by age and sex	Serious impairment in receptive and expressive language ability, evaluated using the Preschool Language Scale-3 (UK) (PLS-3) which comprises Auditory Comprehension and Expressive Communication scales. ◦Total score ◦Auditory comprehension ◦Expressive communication ◦Articulation screener Outcome were dichotomized	Adjusting for MPC score (cognitive ability)	Outcomes assessed at median age of 6 years and 4 months: Serious impairment in language abilities ^[1] Total score ^[1] Full-term Extremely preterm Ref 1.3 (0.3-5.3) Auditory comprehension: Full-term Extremely preterm Ref 1.6 (0.3-9.8) Expressive communication: Full-term Extremely preterm Ref 1.2 (0.2-6.5) Articulation screener:	Low

			a priori using a cut-off of 2 SD or the 10th/90th percentiles as appropriate (not specified which one was used for this outcome).		Full-term Extremely preterm Ref 1.1 (0.3-4) Model adjusted for cognitive impairment score (MPC score).	
--	--	--	---	--	--	--

APPENDIX F. Summary of studies on the association between gestational age and attention deficit hyperactivity disorder

Study	Data Source	Sample and Population studied	Measures of Outcome	Adjustment	Prognostic Outcomes	Study Quality
Burnett 2014	Prospective geographical cohort study	n=215 early preterm/extremely low birth weight infants n=157 normal birth weight (>2499 g) controls n=372 in total	Standardized face-to- face clinical interview and questionnaires were used to assess the mental health status in late adolescence ADHD, any type (All ADHD types assessed with the ADHD module of the Children's Interview for Psychiatric Syndromes (ChIPS)) ADHD, combined type ADHD, inattentive type ADHD, hyperactive/impulsive type	Adjusting for sex, parental education and childhood SES.	At age 18 years: ADHD, any type Normal BW EP/ELBW Reference 2.67 (1.08-6.58) ADHD, combined type Normal BW EP/ELBW Reference 4.9 (0.56-43.24) ADHD, hyperactive/impulsive type Normal BW EP/ELBW Reference NR (0 cases in the control group)	Low
Rogers 2013	Cross sectional survey	n=39 preterm (34-36 weeks) n=154 full term (40-41 weeks)	The Preschool Age Psychiatric Assessment (PAPA) was used to establish DSM-IV Axis 1 diagnoses. It was administered by bachelor's or master's level clinicians and final diagnoses were derived using computerised algorithms.	Sex, family income, IQ and ethnicity.	At age 3-6 years Risk of ADHD Term: Reference 34-36 weeks: OR 0.81 (0.29-2.29) ADHD-inattentive Term: Reference 34-36 weeks: OR 1.21 (0.11-13.22)	Low
Rabie 2015	Retrospective cohort study using population registry data	n=38802 Analysis compares late preterm infants to full term (39-41+6 weeks)	ICD-9 codes from Medicaid files were used to identify children with ADHD and developmental speech and/or language	Birth weight, SGA and LGA, gender, ethnicity, hospital characteristics and maternal medical comorbidities (diabetes,	At age 3-5 years. ADHD Term: Reference 34-36+6 weeks: HR 1.21 (0.98-1.49)	Low

			delay.	hypertension, anaemia, chronic lung disease, herpes, neurologic disorder, coagulation disorder, obesity, depression). OR are unadjusted, as adjustment for sex and socioeconomic status did not affect the results significantly.		
Johnson 2010	Population based prospective cohort study (EPICure)	n=219 preterm children born at <26 weeks n=152 term controls (exact gestation not described)	The Development and Wellbeing Assessment was administered via a telephone interview with parents. Potential cases were identified using computer based scoring algorithms, and final DSM-IV diagnoses were assigned by two child and adolescent psychiatrists on review of summary sheets and clinical transcripts	OR are unadjusted, as adjustment for sex and socioeconomic status did not affect the results significantly.	At age 11 years ^[1] ADHD ^[1] Term: Reference ^[1] <26 weeks: OR 4.3 (1.5-13.0) ADHD inattentive subtype Term: Reference <26 weeks: OR 10.5 (1.4-81.1) ^[1] ADHD combined type Term: Reference <26 weeks: OR 2.1 (0.5-7.9)	Moderate
Singh 2013	Cross sectional survey	n=85,535 Separated into premature children (born at <37 weeks) and term children (≥ 37 weeks)	Parents were asked to self-report whether their child had been diagnosed with one of the disorders by a doctor or health care provider.	Household composition, place of residence and highest household/ parental education.	During follow-up period of between 2 and 17 years ADHD ^[1] Term: Reference <37 weeks: OR 1.49 (1.29-1.73)	Low

APPENDIX G. Summary of studies on the association between gestational age and autism spectrum disorder

Study	Data Source	Sample and Population studied	Measures of Outcome	Adjustment	Prognostic Outcomes	Study Quality
Kuzniewicz 2014	Regional prospective cohort study	n=195021 Analysis compares preterm infants to term (≥ 37 weeks)	Cases of autistic spectrum disorder identified through a regional autism registry. Cases were defined as children with at least one	Gestational age, sex, maternal age, maternal education and small for gestational age.	During follow-up period of age 2-11 ^[1] Autism spectrum disorder Term: Reference 34-36 weeks: HR 1.3 (1.1-1.4) ^[1]	High

			diagnosis of ASD made at an ASD evaluation centre, or by a clinical specialist, or by a general paediatrician.		27-33 weeks: HR 1.4 (1.1-1.8) ^[1] 24-26 weeks: HR 2.7 (1.5- 5.0)	
Singh 2013	Cross sectional survey	n=85,535 Separated into premature children (born at <37 weeks) and term children (≥ 37 weeks)	Parents were asked to self-report whether their child had been diagnosed with one of the disorders by a doctor or health care provider	Household composition, place of residence and highest household/ parental education.	During follow-up period of between 2 and 17 years Autism spectrum disorder Term: Reference <37 weeks: OR 2.26 (1.69- 3.03)	Low

APPENDIX H. Summary of studies on the association between gestational age and specific learning difficulty

Study	Data Source	Sample and Population studied	Measures of Outcome	Adjustment	Prognostic Outcomes	Study Quality
Johnson 2011	Population based prospective cohort study (EPICure)	n=219 preterm children born at <26 weeks n=153 term controls ^[1] (exact gestation not described)	Wechsler Individual Achievement Test to measure mathematics and reading ability. Scores of <2SD below the mean were taken as abnormal.	OR are unadjusted, as adjustment for maternal education and socioeconomic status did not affect the results significantly.	At age 11 years Reading impairment Term: Reference < 26 weeks: OR 21.6 (6.6-70.4) ^[1] Mathematics impairment Term: Reference < 26 weeks: OR 58.7 (14.2-242.9)	Moderate

APPENDIX I. Summary of studies on the association between gestational age and composite outcomes

Study	Data Source	Sample and Population studied	Measures of Outcome	Adjustment	Prognostic Outcomes	Study Quality
Kent 2012	Population based longitudinal cohort study	Sample size N=2701 Followed up at 2- 3 years: n=1473	Assessment of outcome involved examination of 4 domains: developmental, neurologic, vision, and hearing Developmental assessment used the Griffiths Mental Developmental Scales or Bayley Scales of Infant Development II Neurologic assessment included evaluation of muscle tone, primitive reflexes, automatic reactions, and volitional movement Cerebral palsy was diagnosed if the child had non-progressive motor impairment	Multiple regression analysis adjusted for male versus female, gestational age, birth weight percentiles, antepartum haemorrhage, pregnancy- induced hypertension, foetal stress, emergency caesarean delivery, Apgar score < 7 at 5 min, outborn versus inborn	At 2-3 years corrected age Gestational age: ^[1] 27-28 weeks GA: reference 22-26 weeks GA: OR 2.444 (1.831-3.263)	High

			characterised by abnormal muscle tone and a decreased range or decreased control of movements, accompanied by neurologic signs Moderate to severe functional disability was defined as one or more of the following: developmental delay (<2SD below the mean for adjusted age determined by the Griffiths Mental Developmental Scales or BSID-II), cerebral palsy (unable to walk without aids), bilateral blindness (visual acuity <6/60 in better eye), or bilateral deafness (requiring bilateral hearing aids or cochlear implants)			
Toome 2013	Population based prospective cohort study	n=155 preterm infants (<32 weeks) n=153 full term controls (\geq 37 weeks)	Cerebral palsy was defined according to the guidelines of the Surveillance of Cerebral Palsy in Europe collaborative group. The Bayley Scales of Infant and Toddler Development were used to generate composite scores for cognitive, language and motor skills. A composite outcome measure of neurodevelopmental impairment was used. This includes any one (or more) of the following criteria: CP with GMFCS level 2,3,4 or 5; cognitive and/or language composite scores of \leq 2SD below the norm; hearing loss corrected with hearing aids or deafness; vision moderately reduced or blindness.	Antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or PVL grade 2-4, BPD, ROP stage 3-5 with laser therapy, positive blood culture sepsis, NEC stage 2-3, weight $<$ 10th percentile at discharge, maternal age, maternal higher education, single mother, paternal age, paternal higher education and low income of the family	At corrected age of 2 years Moderate/Severe neurodevelopmental disability (CP with GMFCS level 2,3,4 or 5; cognitive and/or language composite scores of \leq 2SD below the norm; hearing loss corrected with hearing aids or deafness; vision moderately reduced or blindness.) OR 0.7 (0.6-0.9) per additional week of gestational age	High

APPENDIX J. Summary of studies on the association. Between developmental delay and maternal smoking during pregnancy

Study	Data Source	Sample and population studied	Measures of outcomes	Adjustment	Prognostic outcomes	Study quality
Moore 2020	Healthy Start mother-child pairs followed through 6 years of age	246 mother-child pairs, maternal urinary cotinine at ~27 weeks age of gestation	Ages and Stages Questionnaire (ASQ)-3 domains dichotomized as fail/monitor and pass	Sex, maternal age, maternal education, daily caloric intake during pregnancy,	Fine motor OR 0.9 (0.2-3.8) Gross motor OR 1.4 (0.1-20.9)	Low

				race/ethnicity, household income, and maternal psychiatric disorders	Communication OR 0.7 (0.2-2.9) Problem solving OR 1.0 (0.1-4.6)	
De Moura 2010	2004 Pelotas birth cohort	3,869 (out of 4231) children born in Pelotas, Brazil in 2004, assessed at 2 years of age	Battelle Screening Developmental Inventory (BDSI) to determine suspected developmental delay	Maternal sociodemographic, reproductive and gestational characteristics, and child and environmental characteristics	Suspected developmental delay PR 1.29 (0.86-1.92)	Low

APPENDIX K. Summary of studies on the association between developmental delay and maternal alcohol use during pregnancy

Study	Data Source	Sample and population studied	Measures of outcomes	Adjustment	Prognostic outcomes	Study quality
O' Leary 2013	Western Australian Health, mental health, and drug and alcohol data sets from Australian Data Linkage Unit; Western Australian Midwives Notification System (1983-2007); and Intellectual Disability Among Exploring Answers database (1983-2001)	64,842 children born to mothers between 1983 to 2007; 1,660 identified to have with intellectual disability	Intellectual disability as defined by Disability Services Commission and severity according to the DSM-IV recommendations	Exposed cohort matched on maternal age within maternal race who had never had an alcohol-related diagnosis recorded on the administrative data sets	Intellectual disability OR 1.81 (1.53-2.14)	Low
Singer 2017	Study to Explore Early Development (SEED), a multi-site case-control study	Children born from September 2003 to August 2006 in the United States; 684 children with ASD, 869 other developmental disabilities, and 962 general population	Autism spectrum disorder (ASD) using Social Communication Questionnaire (SCQ) and general developmental assessment,	Matched with children with non-ASD developmental disorders and general population, with potential	Use during trimester 3: ASD OR 0.4 (0.3-0.7) Other developmental disabilities OR 0.7 (0.5-0.9)	Low

			followed by Autism Diagnostic Observation Schedule (ADOS)	confounders: maternal race/ethnicity, household income, maternal education, pregnancy history, psychiatric conditions, pre-pregnancy body mass index, and child sex		
--	--	--	---	---	--	--

APPENDIX L. Summary of studies on the association between developmental delay and maternal anemia during pregnancy

Study	Data Source	Sample and population studied	Measures of outcomes	Adjustment	Prognostic outcomes	Study quality
De Moura 2010	2004 Pelotas birth cohort	3,869 (out of 4231) children born in Pelotas, Brazil in 2004, assessed at 2 years of age	Battelle Screening Developmental Inventory to determine suspected developmental delay	Maternal sociodemographic, reproductive and gestational characteristics, and child and environmental characteristics	Suspected developmental delay PR 1.48 (0.95-2.29)	Low

APPENDIX M. Summary of studies on the association between developmental delay and maternal diabetes during pregnancy

Study	Data Source	Sample and population studied	Measures of outcomes	Adjustment	Prognostic outcomes	Study quality
Li 2016	Children from Boston Birth Cohort who completed at least 1 postnatal study visit at Boston Medical Center between 1998 and 2014; maternal diabetes diagnosis	2,734 mother-child pairs; 102 Autism Spectrum Disorder cases	Previous diagnosis of Autism Spectrum Disorder, Attention-deficit hyperactivity disorder, and other developmental disorders	Year of birth, maternal age, gender of the child, and maternal parity	Autism Spectrum Disorder HR 1.86, (0.92-3.76) Attention-deficit hyperactivity disorder HR	Low

	based on medical records				0.99 (0.50-1.94)	
Cordero 2019	Study to Explore Early Development (SEED), a multi-site case-control study	Children born in 2003 to 2006 enrolled at 2 to 5 years of age; 2,564 mothers identified (246 with diabetes); Autism Spectrum Disorder 698, non-ASD developmental delay 887, and general population 979	Autism Spectrum Disorder using Social Communication Questionnaire (SCQ) or a previous diagnosis of Autism Spectrum Disorder	Maternal age at conception, maternal race/ethnicity, maternal education, and study site	ASD OR 1.10, (0.77-1.56)	Low
Girchenko 2018	Prediction and Prevention of Pre-eclampsia and Intrauterine Growth Restriction (PREDO study with data from the Finnish Medical Birth Register	2,504 mother-child dyads	Developmental milestones using the Ages and Stages Questionnaire (ASQ)	Maternal age at childbirth, parity, delivery mode, maternal smoking during pregnancy, child's gestational age at delivery, child's birthweight, child's sex, maternal alcohol use during pregnancy, maternal education level, child's age at follow up	Communication OR 2.17 (1.28-3.66) Fine motor OR 1.11 (0.63-1.95) Gross motor OR 1.40 (0.83-2.35) Problem solving OR 1.47 (0.85-2.53) Personal social OR 0.94 (0.47-1.85)	Low
Krakowiak 2012	Childhood Autism Risks from Genetics and the Environment Study (CHARGE) study, population based case-control investigation between January 2003 and June 2010	Children born in California, had parents who spoke English or Spanish, and were living with a biological parent in selected regions of California; 517 Autism Spectrum Disorder, 172 other	Children previously diagnosed with autism or autism spectrum disorder were reevaluated using the Autism Diagnostic Observation Schedule (ADOS); for screening, the Social	Mother's age at delivery, race/ethnicity, education level, delivery payer, calendar time, child's gender and age, and catchment area	ASD OR 1.52 (0.82-2.83) Developmental delay OR 2.33 (1.08-5.05)	Low

		developmental delays, 315 controls	Communication Questionnaire (SCQ) was used; all children underwent Mullen Scales of Early Learning (MSEL) and Vineland Adaptive Behavior Scales (VABS)			
De Moura 2010	2004 Pelotas birth cohort	3,869 (out of 4231) children born in Pelotas, Brazil in 2004, assessed at 2 years of age	Battelle Screening Developmental Inventory to determine suspected developmental delay	Maternal sociodemographic, reproductive and gestational characteristics, and child and environmental characteristics	Suspected developmental delay PR 2.77 (1.34-5.75)	Low

APPENDIX N. Summary of studies on the association between developmental delay and maternal obesity during pregnancy

Study	Data Source	Sample and population studied	Measures of outcomes	Adjustment	Prognostic outcomes	Study quality
Sanchez 2018	Systematic review of 41 studies, meta-analysis of 32 of 41 studies (6 case-control and 26 cohort studies) (total of 36 cohorts)	Mothers with obesity prior and during pregnancy and children with neurodevelopmental disorders	Previous diagnosis and parent report of Attention-deficit hyperactivity or using child behavior checklist and DSM-IV		PRE-PREGNANCY Compromised neurodevelopmental outcomes OR 1.51 (1.35-1.69) Attention-deficit hyperactivity disorder OR 1.62 (1.23-2.14) Autism spectrum disorder OR 1.36 (1.08-1.70) Developmental delay OR 1.58 (1.39-1.79)	Moderate
Girchenko 2018	Prediction and Prevention of Pre-	2,504 mother-child dyads	Developmental milestones using	Maternal age at childbirth,	Communication OR 1.12 (0.63-2.0)	Low

	eclampsia and Intrauterine Growth Restriction (PREDO study with data from the Finnish Medical Birth Register		the Ages and Stages Questionnaire (ASQ)	parity, delivery mode, maternal smoking during pregnancy, child's gestational age at delivery, child's birthweight, child's sex, maternal alcohol use during pregnancy, maternal education level, child's age at follow up	Fine motor OR 1.66 (1.08-2.57) Gross motor OR 1.13 (0.69-1.86) Problem solving OR 1.01 (0.60-1.70) Personal-social OR 0.85 (0.54-1.35)	
Duffany 2016	Retrospective cohort study using a population-based New York City data warehouse with linked birth and Early Intervention data; data from the Longitudinal Study of Early Development data warehouse	Birth cohort of children born in New York City (NYC) of resident mothers from 1994 to 2001 (N=541,816) with the study focusing on those receiving a referral to the NYC Early Intervention Program through 2004 (N=59,589) and a subset of those referred and found eligible by delay for Early Intervention services (N=45,709)	Referral based on assessment of delays across five functional domains: communication, cognitive, physical, adaptive, and socio-emotional domains	Maternal and infant characteristics, maternal race/ethnicity, age, education level, insurance payer, parity, and drug use during pregnancy	Communication RR 1.0 (0.99-1.02) Cognitive RR 1.04 (1.02-1.07) Physical RR 1.04 (1.01-1.08) Social-emotional RR 1.02 (0.98-1.08) Adaptive RR 1.02 (0.97-1.07) Global developmental delay RR 1.05 (1.01-1.08)	Moderate
Windham 2019	Study to Explore Early Development (SEED), a multi-	Children born in 2003 to 2006; Autism Spectrum Disorder 540, other	All caregivers were administered the Social Communication	Maternal age, race/ethnicity, education, and child sex,	ASD OR 1.37 (0.98-1.92)	Low

	site case-control study	developmental delays or disorders 720, control 776	Questionnaire (SCQ) or previous diagnosis of Autism Spectrum Disorder or presence of symptoms noted by research clinician and evaluated using Autism Diagnostic Interview-revised	maternal smoking and hypertension		
--	-------------------------	--	---	-----------------------------------	--	--

APPENDIX O. Summary of studies on the association between developmental delay and maternal hypertension during pregnancy

Study	Data Source	Sample and population studied	Measures of outcomes	Adjustment	Prognostic outcomes	Study quality
Wang 2021	Medical birth registries Psychiatric Central Research Registry of Denmark and Sweden, Swedish National Patient Register and Danish National Patient Register	Study population of 4,489,044 live-born singletons in Denmark during 1978 to 2012 and Sweden during 1987-2010, sample of 2,402,666 children in Denmark and 2,398,429 children in Sweden	Previous diagnosis of Autism Spectrum Disorder, Attention-deficit hyperactivity disorder, and intellectual disability based on registries	Child sex, calendar period of birth, parity, maternal age at birth, maternal country of origin, maternal education level, maternal cohabitation status at birth, and maternal psychiatric disorder before childbirth	Attention-deficit hyperactivity disorder HR 1.24, 1.20-1.28 Autism Spectrum Disorder HR 1.29, 1.24-1.34	Moderate
Girchenko 2018	Prediction and Prevention of Pre-eclampsia and Intrauterine Growth Restriction (PREDO study with data from the	2,504 mother-child dyads	Developmental milestones using the Ages and Stages Questionnaire (ASQ)	Maternal age at childbirth, parity, delivery mode, maternal smoking during pregnancy, child's gestational age at delivery, child's birthweight,	Pre-eclampsia: Communication OR 1.81 (0.83-3.93) Fine motor 0.52 (0.16-1.68) Gross motor 2.33 (1.17-4.26)	Low

	Finnish Medical Birth Register			child's sex, maternal alcohol use during pregnancy, maternal education level, child's age at follow up	Problem solving 1.75 (0.85-3.62) Personal social 0.95 (0.34-2.67)	
Cordero 2019	Study to Explore Early Development (SEED), a multi-site case-control study	Children born in 2003 to 2006 enrolled at 2 to 5 years of age; 2,564 mothers identified (386 with hypertension); Autism Spectrum Disorder 698, non-ASD developmental delay 887, and general population 979	Autism Spectrum Disorder using Social Communication Questionnaire (SCQ) or a previous diagnosis of Autism Spectrum Disorder	Maternal age at conception, maternal race/ethnicity, maternal education, and study site	ASD OR 1.71, (1.30-2.25)	Low
Krakowiak 2012	Childhood Autism Risks from Genetics and the Environment Study (CHARGE) study, population based case-control investigation between January 2003 and June 2010	Children born in California, had parents who spoke English or Spanish, and were living with a biological parent in selected regions of California; 517 Autism Spectrum Disorder, 172 other developmental delays, 315 controls	Children previously diagnosed with autism or autism spectrum disorder were reevaluated using the Autism Diagnostic Observation Schedule (ADOS); for screening, the Social Communication Questionnaire (SCQ) was used; all children underwent Mullen Scales of Early Learning (MSEL) and Vineland	ASD OR 1.52 (0.82-2.83) Developmental delay OR 2.33 (1.08-5.05)	ASD OR 2.84 (0.94-8.56) Developmental delay OR 3.58 (0.93-13.78)	Low

			Adaptive Behavior Scales (VABS)			
De Moura 2010	2004 Pelotas birth cohort	3,869 (out of 4231) children born in Pelotas, Brazil in 2004, assessed at 2 years of age	Battelle Screening Developmental Inventory to determine suspected developmental delay	Maternal sociodemographic, reproductive and gestational characteristics, and child and environmental characteristics	Suspected developmental delay PR 1.28 (0.85-1.93)	Low

APPENDIX P. GRADE Evidence Profile

Maternal smoking

de Moura DR, Costa JC, Santos IS, et al. Risk factors for suspected developmental delay at age 2 years in a Brazilian birth cohort. *Paediatr Perinat Epidemiol.* 2010;24(3):211-221. doi:10.1111/j.1365-3016.2010.01115.x

Moore BF, Shapiro AL, Wilkening G, Magzamen S, Starling AP, Allshouse WB, Adgate JL, Dabelea D. Prenatal Exposure to Tobacco and Offspring Neurocognitive Development in the Healthy Start Study. *J Pediatr.* 2020 Mar;218:28-34.e2. doi: 10.1016/j.jpeds.2019.10.056. Epub 2019 Nov 20. PMID: 31759580; PMCID: PMC7042047.

Question: What is the risk of developmental delay among apparently healthy children whose mothers smoked and drank alcoholic beverages and were anemic, diabetic, obese, and hypertensive during pregnancy?

Nº of studies	Certainty assessment						Effect			Certainty	Importance
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nº of events	Nº of individuals	Rate (95% CI)		
Smoking (follow-up: mean 2 years)											
1	observational studies	very serious ^a	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect	128	3869	event rate 3.3 per 100 (2.7 to 3.8)	⊕⊕○ ○ Low	IMPORTANT
Smoking (follow-up: mean 6)											
1	observational studies	very serious ^b	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕○ ○ Low	IMPORTANT

Explanations

a. Attrition rate approximately 10%, possible confounding by postnatal smoking

b. Large attrition rate (246 out of 1,410), possible confounding by postnatal smoking

Maternal alcohol use

O'Leary C, Leonard H, Bourke J, D'Antoine H, Bartu A, Bower C. Intellectual disability: population-based estimates of the proportion attributable to maternal alcohol use disorder during pregnancy. *Dev Med Child Neurol.* 2013 Mar;55(3):271-7. doi: 10.1111/dmcn.12029. Epub 2012 Dec 14. PMID: 23241019.

Singer AB, Aylsworth AS, Cordero C, Croen LA, DiGuiseppi C, Fallin MD, Herring AH, Hooper SR, Pretzel RE, Schieve LA, Windham GC, Daniels JL. Prenatal Alcohol Exposure in Relation to Autism Spectrum Disorder: Findings from the Study to Explore Early Development (SEED). *Paediatr Perinat Epidemiol.* 2017 Nov;31(6):573-582. doi: 10.1111/ppe.12404. Epub 2017 Sep 7. PMID: 28881390; PMCID: PMC5690833.

Question: What is the risk of developmental delay among apparently healthy children whose mothers smoked and drank alcoholic beverages and were anemic, diabetic, obese, and hypertensive during pregnancy?

Nº of studies	Certainty assessment						Effect			Certainty	Importance
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nº of events	Nº of individuals	Rate (95% CI)		
Alcohol											
1	observational studies	very serious ^a	not serious	serious ^b	not serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕○ ○ Low	IMPORTANT
Alcohol											
1	observational studies	very serious	not serious	not serious	serious	all plausible residual confounding would suggest spurious effect, while no effect was observed	684 cases 962 controls		- 0.0%	⊕⊕○ ○ Low	IMPORTANT

Explanations

a. Majority of alcohol diagnoses were for acute intoxication and we could not examine the exact timing of exposure and whether there were extended periods of sobriety which could have reduced risk to the fetus.

b. Alcohol intake may mean presence of maternal alcohol use disorder or fetal alcohol syndrome.

Maternal anemia

de Moura DR, Costa JC, Santos IS, et al. Risk factors for suspected developmental delay at age 2 years in a Brazilian birth cohort. *Paediatr Perinat Epidemiol.* 2010;24(3):211-221. doi:10.1111/j.1365-3016.2010.01115.x

Question: What is the risk of developmental delay among apparently healthy children whose mothers smoked and drank alcoholic beverages and were anemic, diabetic, obese, and hypertensive during pregnancy?

Nº of studies	Certainty assessment						Effect			Certainty	Importance
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nº of events	Nº of individuals	Rate (95% CI)		
Anemia											
1	observational studies	very serious ^a	not serious	not serious	serious	all plausible residual confounding would suggest spurious effect, while no effect was observed				⊕⊕○ ○ Low	IMPORTANT

Explanations

a. Attrition rate approximately 10%, possible confounding from causes of anemia

Gestational diabetes

- Li M, Fallin MD, Riley A, Landa R, Walker SO, Silverstein M, Caruso D, Pearson C, Kiang S, Dahm JL, Hong X, Wang G, Wang MC, Zuckerman B, Wang X. The Association of Maternal Obesity and Diabetes With Autism and Other Developmental Disabilities. *Pediatrics*. 2016 Feb;137(2):e20152206. doi: 10.1542/peds.2015-2206. Epub 2016 Jan 29. PMID: 26826214; PMCID: PMC4732357.
- Cordero C, Windham GC, Schieve LA, Fallin MD, Croen LA, Siega-Riz AM, Engel SM, Herring AH, Stuebe AM, Vladutiu CJ, Daniels JL. Maternal diabetes and hypertensive disorders in association with autism spectrum disorder. *Autism Res*. 2019 Jun;12(6):967-975. doi: 10.1002/aur.2105. Epub 2019 Apr 10. PMID: 30969030; PMCID: PMC6546522.
- Girchenko P, Tuovinen S, Lahti-Pulkkinen M, Lahti J, Savolainen K, Heinonen K, Pyhälä R, Reynolds RM, Hämaläinen E, Villa PM, Kajantie E, Pesonen AK, Laivuori H, Räikkönen K. Maternal early pregnancy obesity and related pregnancy and pre-pregnancy disorders: associations with child developmental milestones in the prospective PREDO Study. *Int J Obes (Lond)*. 2018 Jun;42(5):995-1007. doi: 10.1038/s41366-018-0061-x. Epub 2018 Apr 23. PMID: 29686379.
- Krakowiak P, Walker CK, Bremer AA, Baker AS, Ozonoff S, Hansen RL, Hertz-Pannier I. Maternal metabolic conditions and risk for autism and other neurodevelopmental disorders. *Pediatrics*. 2012 May;129(5):e1121-8. doi: 10.1542/peds.2011-2583. Epub 2012 Apr 9. PMID: 22492772; PMCID: PMC3340592.
- de Moura DR, Costa JC, Santos IS, et al. Risk factors for suspected developmental delay at age 2 years in a Brazilian birth cohort. *Paediatr Perinat Epidemiol*. 2010;24(3):211-221. doi:10.1111/j.1365-3016.2010.01115.x

Question: What is the risk of developmental delay among apparently healthy children whose mothers smoked and drank alcoholic beverages and were anemic, diabetic, obese, and hypertensive during pregnancy?

Nº of studies	Certainty assessment						Effect			Certainty	Importance
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nº of events	Nº of individuals	Rate (95% CI)		
Diabetes											
1	observational studies	very serious ^a	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕○ ○ Low	IMPORTANT
Diabetes											
1	observational studies	very serious ^b	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕○ ○ Low	IMPORTANT

Nº of studies	Certainty assessment						Effect			Certainty	Importance	
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nº of events	Nº of individuals	Rate (95% CI)			
Diabetes												
1	observational studies	very serious ^c	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕○ ○ Low	IMPORTANT	
Diabetes												
1	observational studies	very serious ^d	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect	517 cases 315 controls		- 0.0%	⊕⊕○ ○ Low	IMPORTANT	
Diabetes												
1	observational studies	very serious ^e	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕○ ○ Low	IMPORTANT	

Explanations

- a. High likelihood of misclassification error because diagnosis is tentative. Selection bias still possible.
- b. Selection bias likely
- c. Outcomes parent reported and prone to bias, some children with mild symptoms may be undetected.
- d. Self-reported medical conditions and lack of biological measurements or diagnostic tests.
- e. Attrition rate approximately 10%, possible confounding by postnatal smoking

Maternal obesity

Sanchez CE, Barry C, Sabhlok A, Russell K, Majors A, Kollins SH, Fuemmeler BF. Maternal pre-pregnancy obesity and child neurodevelopmental outcomes: a meta-analysis. *Obes Rev*. 2018 Apr;19(4):464-484. doi: 10.1111/obr.12643. Epub 2017 Nov 22. PMID: 29164765; PMCID: PMC6059608.

Girchenko P, Tuovinen S, Lahti-Pulkkinen M, Lahti J, Savolainen K, Heinonen K, Pyhälä R, Reynolds RM, Hämäläinen E, Villa PM, Kajantie E, Pesonen AK, Laivuori H, Räikkönen K. Maternal early pregnancy obesity and related pregnancy and pre-pregnancy disorders: associations with child developmental milestones in the prospective PREDO Study. *Int J Obes (Lond)*. 2018 Jun;42(5):995-1007. doi: 10.1038/s41366-018-0061-x. Epub 2018 Apr 23. PMID: 29686379.

Duffany KO, McVeigh KH, Kershaw TS, Lipkind HS, Ickovics JR. Maternal Obesity: Risks for Developmental Delays in Early Childhood. *Matern Child Health J*. 2016 Feb;20(2):219-30. doi: 10.1007/s10995-015-1821-z. PMID: 26694046.

Windham GC, Anderson M, Lyall K, Daniels JL, Kral TVE, Croen LA, Levy SE, Bradley CB, Cordero C, Young L, Schieve LA. Maternal Pre-pregnancy Body Mass Index and Gestational Weight Gain in Relation to Autism Spectrum Disorder and other Developmental Disorders in Offspring. *Autism Res.* 2019 Feb;12(2):316-327. doi: 10.1002/aur.2057. Epub 2018 Dec 21. PMID: 30575327; PMCID: PMC7778460.

Question: What is the risk of developmental delay among apparently healthy children whose mothers smoked and drank alcoholic beverages and were anemic, diabetic, obese, and hypertensive during pregnancy?

Nº of studies	Certainty assessment						Effect			Certainty	Importance
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nº of events	Nº of individuals	Rate (95% CI)		
1	observational studies	very serious ^c	not serious	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕⊕ ○ Moderate	IMPORTANT
Obesity											
1	observational studies	very serious ^d	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect	0 cases -	0 controls 0.0%		⊕⊕○ ○ Low	IMPORTANT

Explanations

- a. Assessments for delay were not the same across studies and methods used to address confounding also differed across studies.
- b. Outcomes parent reported and prone to bias, some children with mild symptoms may be undetected.
- c. Potential selection bias because source of data (Early Intervention) may be unknown to a particular population.
- d. Self-report of weight, authors claim some necessary data are missing. Selection bias due to inability to contact eligible participants.

Gestational hypertension

Wang H, László KD, Gissler M, Li F, Zhang J, Yu Y, Li J. Maternal hypertensive disorders and neurodevelopmental disorders in offspring: a population-based cohort in two Nordic countries. *Eur J Epidemiol.* 2021 May;36(5):519-530. doi: 10.1007/s10654-021-00756-2. Epub 2021 May 4. PMID: 33948753; PMCID: PMC8159819.

Girchenko P, Tuovinen S, Lahti-Pulkkinen M, Lahti J, Savolainen K, Heinonen K, Pyhälä R, Reynolds RM, Hämäläinen E, Villa PM, Kajantie E, Pesonen AK, Laivuori H, Räikkönen K. Maternal early pregnancy obesity and related pregnancy and pre-pregnancy disorders: associations with child developmental milestones in the prospective PREDO Study. *Int J Obes (Lond).* 2018 Jun;42(5):995-1007. doi: 10.1038/s41366-018-0061-x. Epub 2018 Apr 23. PMID: 29686379.

Cordero C, Windham GC, Schieve LA, Fallin MD, Croen LA, Siega-Riz AM, Engel SM, Herring AH, Stuebe AM, Vladutiu CJ, Daniels JL. Maternal diabetes and hypertensive disorders in association with autism spectrum disorder. *Autism Res.* 2019 Jun;12(6):967-975. doi: 10.1002/aur.2105. Epub 2019 Apr 10. PMID: 30969030; PMCID: PMC6546522.

Krakowiak P, Walker CK, Bremer AA, Baker AS, Ozonoff S, Hansen RL, Hertz-Pannier I. Maternal metabolic conditions and risk for autism and other neurodevelopmental disorders. *Pediatrics.* 2012 May;129(5):e1121-8. doi: 10.1542/peds.2011-2583. Epub 2012 Apr 9. PMID: 22492772; PMCID: PMC3340592.

de Moura DR, Costa JC, Santos IS, et al. Risk factors for suspected developmental delay at age 2 years in a Brazilian birth cohort. *Paediatr Perinat Epidemiol.* 2010;24(3):211-221. doi:10.1111/j.1365-3016.2010.01115.x

Question: What is the risk of developmental delay among apparently healthy children whose mothers smoked and drank alcoholic beverages and were anemic, diabetic, obese, and hypertensive during pregnancy?

Nº of studies	Certainty assessment						Effect			Certainty	Importance
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nº of events	Nº of individuals	Rate (95% CI)		
Hypertension											
1	observational studies	very serious	not serious	not serious	not serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕○ Moderate	IMPORTANT
Hypertension											
1	observational studies	very serious	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕○ Low	IMPORTANT

Nº of studies	Certainty assessment						Effect			Certainty	Importance
	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Nº of events	Nº of individuals	Rate (95% CI)		
Hypertension											
1	observational studies	very serious ^b	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect				⊕⊕○ ○ Low	IMPORTANT
Hypertension											
1	observational studies	very serious ^c	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect	517 cases 315 controls		- 0.0%	⊕⊕○ ○ Low	IMPORTANT
Hypertension											
1	observational studies	very serious ^a	not serious	not serious	serious	all plausible residual confounding would reduce the demonstrated effect	128	3869	event rate 3.3 per 100 (2.7 to 3.8)	⊕⊕○ ○ Low	IMPORTANT

Explanations

- a. Attrition rate approximately 10%, possible confounding by postnatal smoking
- b. Selection bias likely
- c. Outcomes parent reported and prone to bias, some children with mild symptoms may be undetected.
- d. Self-reported medical conditions and lack of biological measurements or diagnostic tests.
- e. Misclassification bias possible if cases of mild hypertension are classified as controls.

Diagnostic accuracy of Ages and Stages Questionnaire in correctly identifying intellectual disability in preterm children

Question: ASQ compared to no screening for developmental delay among preterm infants

Quality assessment							Summary of findings: diagnostic accuracy				Quality	Importance		
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Sens (95% CI)	Spec (95% CI)	LR+ (95% CI)	LR- (95% CI)				
Screening: ASQ-3 Psychometric values (< -2 SD) among children born preterm (GA 32-36wks) assessed at 8 months, 18 months, or 30 months corrected age														
Diagnosis tool: Bayley III < -1 SD														
1 (Schonhaut 2013)	Cross sectional	serious ^a	N/A	not serious	serious ^b	none	0.80 (0.61-0.91)	0.73 (0.63-0.81)	2.9 (2.0-4.3)	0.27 (0.1-0.6)	 Low	CRITICAL		
Screening: ASQ-3 Psychometric values (< 2 SD) among children born preterm (GA < 32wks) assessed at 8 months, 18 months, or 30 months corrected age														
Diagnosis tool: Bayley III < -1 SD														
1 (Schonhaut 2013)	Cross sectional	serious ^a	N/A	not serious	serious ^b	none	0.86 (0.60-0.96)	0.86 (0.73-0.93)	6.0 (2.9-12.3)	0.17 (0.05-0.6)	 Low	CRITICAL		
Screening: ASQ < -1 SD among children born at 25.4 weeks' GA (range: 23.0-31.0 weeks) assessed at 18-22 months corrected age														
Diagnosis: BSID-II < -2 SD either MDI or PDI														
1 (Woodward 2012)	Follow-up of RCT, cross sectional study	serious ^c	N/A	not serious	not serious	none	0.94 (0.89-1.00)	0.32 (0.23-0.40)	1.39 (1.21-1.60)	0.16 (0.05-0.49)	 Moderate	CRITICAL		
Screening: ASQ < -2 SD among children born at 25.4 weeks' GA (range: 23.0-31.0 weeks) assessed at 18-22 months corrected age														
Diagnosis: BSID-II < -2 SD either MDI or PDI														
1 (Woodward 2012)	Follow-up of RCT, cross sectional study	serious ^c	N/A	not serious	not serious	none	0.73 (0.60-0.84)	0.65 (0.55-0.73)	2.05 (1.58-2.76)	0.42 (0.27-0.65)	 Moderate	CRITICAL		
Screening: ASQ < -2 SD among children born at 25.4 weeks' GA (range: 23.0-31.0 weeks) assessed at 18-22 months corrected age														
Diagnosis: BSID-II < -1 SD either MDI or PDI														
1 (Woodward 2012)	Follow-up of RCT, cross sectional study	serious ^c	N/A	not serious	not serious	none	0.63 (0.53-0.72)	0.76 (0.64-0.85)	2.47 (1.58-3.86)	0.50 (0.38-0.67)	 Moderate	CRITICAL		
Screening: ASQ < -2 SD among children born at < 31 weeks' GA assessed at 18 months corrected age														
Diagnosis: Bayley MDI < -1 SD														
1 (Skellern 2001)	Cross sectional study	serious ^d	N/A	not serious	serious ^b	none	0.50 (0.01-0.99)	0.91 (0.71-0.99)	5.5 (0.81-37.2)	0.55 (0.14-2.2)	 Low	CRITICAL		

Quality assessment							Summary of findings: diagnostic accuracy				Quality	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Sens (95% CI)	Spec (95% CI)	LR+ (95% CI)	LR- (95% CI)		

Screening: ASQ < -1 SD among children born at 29-36 weeks' GA assessed at 12 months corrected age

Diagnosis: BSID-II MDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	serious ^b	none	0.60 (0.39-0.81)	0.68 (0.59-0.77)	1.83 (1.17-2.87)	0.60 (0.36-1.01)	 Low	CRITICAL
-----------------	---	-------------	-----	-------------	----------------------	------	------------------	------------------	------------------	------------------	---	----------

Screening: ASQ < -1.5 SD among children born at 29-36 weeks' GA assessed at 12 months corrected age

Diagnosis: BSID-II MDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	serious ^b	none	0.45 (0.23-0.67)	0.78 (0.71-0.87)	2.25 (1.23-4.11)	0.68 (0.46-1.01)	 Moderate	CRITICAL
-----------------	---	-------------	-----	-------------	----------------------	------	------------------	------------------	------------------	------------------	--	----------

Screening: ASQ < -2 SD among children born at 29-36 weeks' GA assessed at 12 months corrected age

Diagnosis: BSID-II MDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	serious ^b	none	0.20 (0.02-0.38)	0.88 (0.82-0.95)	1.50 (0.53-4.21)	0.93 (0.75-1.15)	 Moderate	CRITICAL
-----------------	---	-------------	-----	-------------	----------------------	------	------------------	------------------	------------------	------------------	--	----------

Screening: ASQ < -1 SD among children born at 29-36 weeks' GA assessed at 12 months corrected age

Diagnosis: BSID-II PDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	very serious ^e	none	0.52 (0.38-0.67)	0.90 (0.83-0.96)	5.04 (2.46-10.3)	0.53 (0.38-0.74)	 Low	CRITICAL
-----------------	---	-------------	-----	-------------	---------------------------	------	------------------	------------------	------------------	------------------	---	----------

Screening: ASQ < -1.5 SD among children born at 29-36 weeks' GA assessed at 12 months corrected age

Diagnosis: BSID-II PDI < 85

Quality assessment							Summary of findings: diagnostic accuracy				Quality	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Sens (95% CI)	Spec (95% CI)	LR+ (95% CI)	LR- (95% CI)		
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	very serious ^a	none	0.39 (0.24-0.53)	0.96 (0.92-1.00)	7.33 (2.62-20.5)	0.65 (0.51-0.83)		CRITICAL

Screening: ASQ < -2 SD among children born at 29-36 weeks' GA assessed at 12 months corrected age

Diagnosis: BSID-II PDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	very serious ^a	none	0.25 (0.12-0.38)	0.97 (0.94-1.00)	9.85 (2.29-42.4)	0.76 (0.64-0.91)		CRITICAL
-----------------	---	-------------	-----	-------------	---------------------------	------	------------------	------------------	------------------	------------------	---	----------

Screening: ASQ < -1 SD among children born at 29-36 weeks' GA assessed at 24 months corrected age

Diagnosis: BSID-II MDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	serious ^b	none	0.92 (0.81-1.00)	0.558 (0.45-2.69)	2.07 (1.59-2.69)	0.14 (0.04-0.53)		CRITICAL
-----------------	---	-------------	-----	-------------	----------------------	------	------------------	-------------------	------------------	------------------	---	----------

Screening: ASQ < -1.5 SD among children born at 29-36 weeks' GA assessed at 24 months corrected age

Diagnosis: BSID-II MDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	serious ^b	none	0.88 (0.74-1.00)	0.72 (0.63-0.82)	3.34 (2.27-4.90)	0.16 (0.05-0.46)		CRITICAL
-----------------	---	-------------	-----	-------------	----------------------	------	------------------	------------------	------------------	------------------	---	----------

Screening: ASQ < -2 SD among children born at 29-36 weeks' GA assessed at 24 months corrected age

Diagnosis: BSID-II MDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	serious ^b	none	0.75 (0.58-0.92)	0.78 (0.69-0.87)	3.46 (2.17-5.51)	0.33 (0.17-0.63)		CRITICAL
-----------------	---	-------------	-----	-------------	----------------------	------	------------------	------------------	------------------	------------------	---	----------

Quality assessment							Summary of findings: diagnostic accuracy				Quality	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Sens (95% CI)	Spec (95% CI)	LR+ (95% CI)	LR- (95% CI)		

Screening: ASQ < -1 SD among children born at 29-36 weeks' GA assessed at 24 months corrected age

Diagnosis: BSID-II PDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	serious ^b	none	0.50 (0.31-0.69)	0.73 (0.64-0.83)	1.82 (1.09-3.03)	0.69 (0.47-1.02)	Moderate	CRITICAL
-----------------	---	-------------	-----	-------------	----------------------	------	------------------	------------------	------------------	------------------	----------	----------

Screening: ASQ < -1.5 SD among children born at 29-36 weeks' GA assessed at 24 months corrected age

Diagnosis: BSID-II PDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	serious ^b	none	0.50 (0.31-0.69)	0.73 (0.64-0.83)	1.82 (1.09-3.03)	0.69 (0.47-1.02)	Moderate	CRITICAL
-----------------	---	-------------	-----	-------------	----------------------	------	------------------	------------------	------------------	------------------	----------	----------

Screening: ASQ < -2 SD among children born at 29-36 weeks' GA assessed at 24 months corrected age

Diagnosis: BSID-II PDI < 85

1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	not serious	N/A	not serious	serious ^b	none	0.31 (0.13-0.49)	0.92 (0.86-0.98)	3.95 (1.51-10.36)	0.77 (0.59-0.97)	Moderate	CRITICAL
-----------------	---	-------------	-----	-------------	----------------------	------	------------------	------------------	-------------------	------------------	----------	----------

Screening: ASQ score < 270 among children born at ≤ 35 weeks' GA assessed at 5 years

Diagnosis: IQ score < 70 on WPPSI-III

1 (Halbwachs 2013)	Cross sectional study	not serious	N/A	not serious	serious ^b	none	0.85 (0.68-0.94)	0.81 (0.77-0.85)	4.46 (3.47-5.7)	0.18 (0.07-0.45)	Moderate	CRITICAL
--------------------	-----------------------	-------------	-----	-------------	----------------------	------	------------------	------------------	-----------------	------------------	----------	----------

Screening: ASQ score < 280 among children born at ≤ 35 weeks' GA assessed at 5 years

Diagnosis: IQ score < 85 on WPPSI-III

1 (Halbwachs 2013)	Cross sectional study	not serious	N/A	not serious	not serious	none	0.80 (0.71-0.87)	0.54 (0.48-0.60)	1.74 (1.50-2.02)	0.37 (0.24-0.56)	High	CRITICAL
--------------------	-----------------------	-------------	-----	-------------	-------------	------	------------------	------------------	------------------	------------------	------	----------

ASQ: Ages and Stages Questionnaire; BSID: Bayley Scales of Infant Toddler Development; GA: gestational age; IQ: intelligence quotient; LR+: positive likelihood ratio; LR-: negative likelihood ratio; MDI: mental developmental index; N/A: not applicable; PDI: psychomotor developmental index; SD: standard deviation; Sens: sensitivity; Spec: specificity; WPPSI: Wechsler Preschool and Primary Scale of Intelligence

Explanations

a. Evidence was downgraded by 1 level because of the selection bias in the sample recruited

b. Evidence was downgraded by 1 level due to the wide confidence intervals on some accuracy estimates

- c. Evidence was downgraded by 1 level because of the selection bias of the sample recruited (follow-up study of an earlier RCT)
- d. Evidence was downgraded by 1 level because the study did not clearly report whether diagnosis outcome assessors were blinded to the screening results
- e. Evidence was downgraded by 1 level due to the very wide confidence intervals on some accuracy estimates

14. Screening for Autism Spectrum Disorder

APPENDIX A. Search strategies

1. Cochrane Central Register of Controlled Trials (CENTRAL), in the Cochrane Library

(Search date: August 2, 2021)

#1MeSH descriptor Early Intervention (Education), this term only

#2MeSH descriptor Behavior Therapy, this term only

#3Lovaas*

#4(intens* NEAR/3 (intervent* or therap* or treat* or program*))

#5(IFI or EIFI)

#6applied NEXT behavio* NEXT analy* or ABA

#7(#1 OR #2 OR #3 OR #4 OR #5 OR #6)

#8MeSH descriptor Child Development Disorders, Pervasive explode all trees

#9(pervasive development* disorder* or PDD or PDDs)

#10Rett*

#11Asperger*

#12autis* or ASD or ASDs

#13Kanner*

#14childhood schizophren*

#15MeSH descriptor Communication Disorders, this term only

#16MeSH descriptor Speech Disorders, this term only

#17MeSH descriptor Language Development Disorders, this term only

#18MeSH descriptor Child Behavior Disorders, this term only

#19communicat* NEAR/3 disorder*

#20speech NEAR/3 (delay* or disorder*)

#21(child* NEAR/3 behavio* NEAR/3 disorder*)

#22(language NEAR/3 (delay* or disorder*))

#23(#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)

#24(baby or babies or infant* or toddler* or child* or pre-school* or preschool* or boy* or girl*)

#25MeSH descriptor Child explode all trees

#26MeSH descriptor Infant, this term only

#27(#24 OR #25 OR #26)

#28(#7 AND #23 AND #27)

FILTER: Aug 2017 - August 2, 2021

2. Conference Proceedings Citation Index - Social Sciences & Humanities Web of Science
(Search date: August 2, 2021)

#7 AND #6

DocType=All document types; Language=All languages;

#7 TS=(baby or babies or infant* or toddler* or child* or pre-school* or preschool* or boy* or girl*)

DocType=All document types; Language=All languages;

#6 #5 AND #4

DocType=All document types; Language=All languages;

#5 TS=(autis* or asperger* or ASD or ASDs or Pervasive development* disorder* or PDD or PDDs or Rett* or Kanner* or childhood schizophren*)

DocType=All document types; Language=All languages;

#4 #3 OR #2 OR #1

DocType=All document types; Language=All languages;

#3 TS=(“applied behav* analy*” or ABA)

DocType=All document types; Language=All languages;

#2 TS=(lovaas OR IBI or EIBI)

DocType=All document types; Language=All languages;

#1 TS=(intens* NEAR/3 (interven* or therap* or treat* or program*))

DocType=All document types; Language=All languages;

FILTER: Aug 2017 - August 2, 2021

3. Epistemonikos (Search date: August 2, 2021)

(title:(title:(asd OR autis* OR asperger* OR pervasive OR rett) AND (title:(early intervention* OR intensive behav* OR applied behav*) OR abstract:(early intervention* OR intensive behav* OR applied behav*))) OR abstract:(title:(asd OR autis* OR asperger* OR pervasive OR rett) AND (title:(early intervention* OR intensive behav* OR applied behav*) OR abstract:(early intervention* OR intensive behav* OR applied behav*))))

FILTER: Aug 2017 - August 2, 2021

4. ClinicalTrials.gov (Search date: August 2, 2021)

Autism OR ASD OR Asperger OR PDD OR " pervasive developmental " | EIBI OR IBI OR ABA OR "early behavioural" OR "early behavioral" OR "applied behavioral" OR "applied behavioural" OR "Intensive behavioral" OR "Intensive behavioural" | Child

FILTER: Aug 2017 - August 2, 2021

5. WorldCat OCLC (Search date: August 2, 2021)

'kw:(“intens* behav*” OR EIBI OR IBI OR ABA OR “applied behav*”) AND kw:(autis* OR asd* OR asperg* OR PDD* OR “pervasive development* disorder*”)) AND kw:(child* OR infant* OR baby OR babies OR toddler* OR preschool* OR pre-school*)' > 'Thesis/dissertation'

FILTER: Aug 2017 - August 2, 2021

PubMed search strategy to update 2016 USPSTF Systematic Review on ASD screening

Search Terms	Results
1 autistic[tiab] OR autism[tiab] OR autistic disorder[mh] OR asperger syndrome[mh] OR child development disorders, pervasive[mh:noexp] OR asperger[tiab] OR asperger's[tiab] OR aspergers[tiab] OR pervasive development[tiab] OR pervasive developmental[tiab]	59 439
2 mass screening[mh] OR screening[tiab] OR screened[tiab] OR screen[tiab] OR screener[tiab] OR screeners[tiab] OR early diagnosis[mh] OR identify[tiab] OR identification[tiab]	2 416 139
3 #1 AND #2 AND eng[la] AND humans[mh]	6 612
4 newspaper article[pt] OR letter[pt] OR comment[pt] OR case reports[pt] OR review[pt] OR practice guideline[pt] OR news[pt] OR editorial[pt] OR historical article[pt] OR meta-analysis[pt] OR legal cases[pt] OR jsubsetk	7 319 828
5 #3 NOT #4 AND 2015:2021[dp] Filter: Clinical Trial, Meta-Analysis, Randomized Controlled Trial, Review, Systematic Review, from 2015 - 2021	162

APPENDIX B. PRISMA Diagram

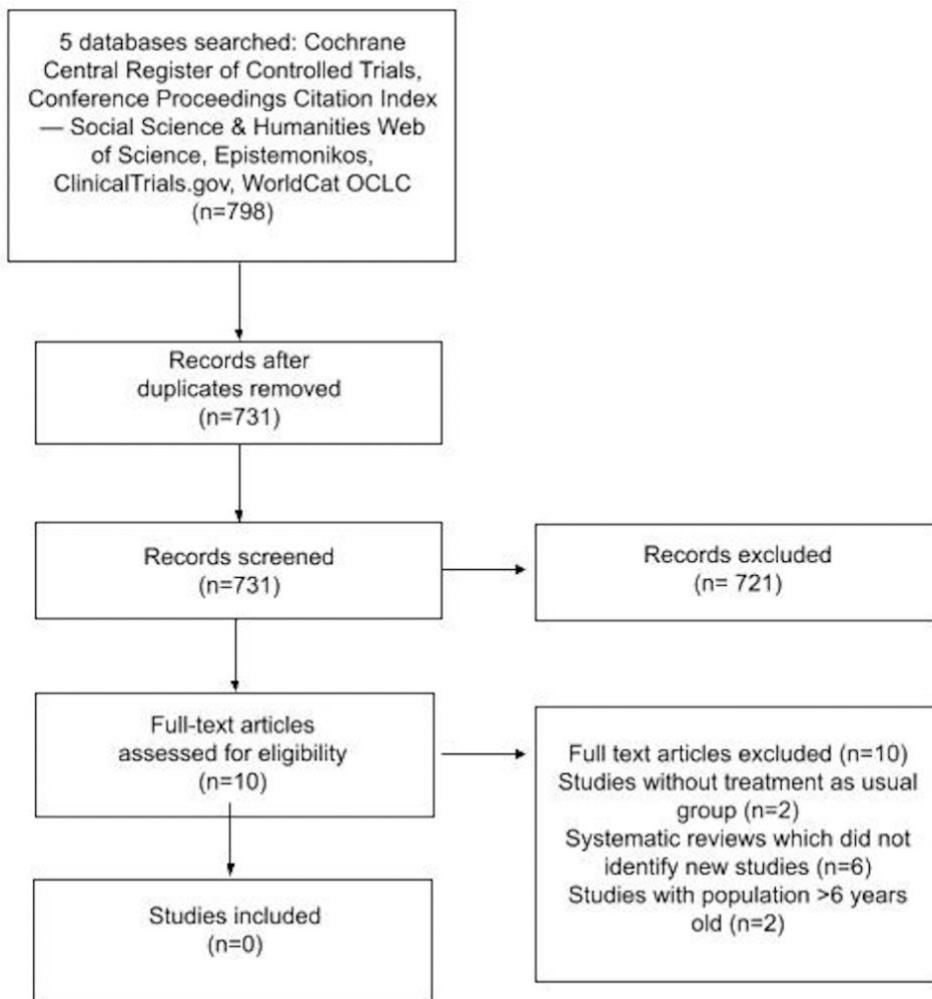


Figure 1. PRISMA diagram in updating the Cochrane 2018 Systematic Review on the effects of EIBI on ASD

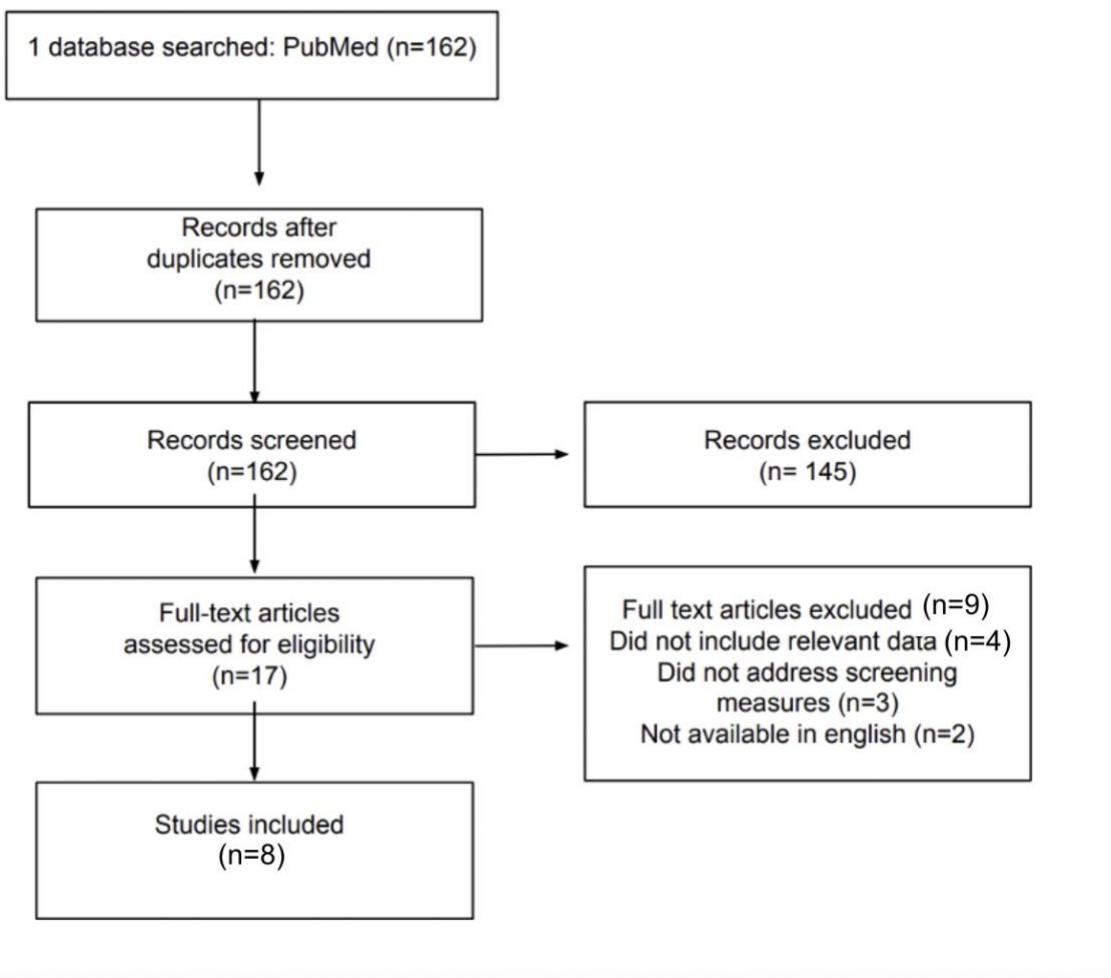


Figure 2. PRISMA diagram in updating the 2016 USPSTF systematic review on ASD screening

APPENDIX C. Characteristics of included studies

Table 1. Characteristics of included studies in update of 2016 USPSTF systematic review on ASD screening

Study ID	Index test	Reference Standard	Population (age in months)	Total participants
Samango-Sprouse 2016	PDDST-II	MSEL, PLS-4, ELM-2, DSM-IV	4-9	1024
Zahorodny 2018	PDQ-1, SCQ	ADI-R, DSM-IV	18-36	1959
Janvier 2016	M-CHAT F	ADOS, DSM-IV	16-76	361
Sturmer 2016	M-CHAT F	ADOS-2, MSEL, DSM-V	14.7-40.8	5071
Baduel 2017	M-CHAT F	ADOS-G, PEP-R, VABS, DSM-IV	24	1227
Pandey 2008	M-CHAT F	ADI-R, ADOS, CARS, DSM-IV	16-30	6050
Khawaja 2015	M-CHAT R/F	MSEL, Vineland, BASC, ADI-R, ADOS, CARS, DSM-IV-TR	16-30	11 845
Hardy 2015	M-CHAT R/F	Vineland II, ASI, ADOS, MSEL	16-31	2848

APPENDIX D. GRADE Evidence Profiles

Table 1
GRADE Evidence Profile for EIBI in ASD

Author(s): Corinna M. Puyat, Kathryn Baltazar-Braganza, MD

Question: EIBI compared to TAU for treatment of ASD in young children?

Setting: Outpatient

Bibliography: Smith T, Groen AD, Wynn JW. Randomized trial of intensive early intervention for children with pervasive developmental disorder. Am J Ment Retard. 2000;105(4):269–85. Cohen H, Amerine-Dickens M, Smith T. Early intensive behavioral treatment: replication of the UCLA model in a community setting. J Dev Behav Pediatr. 2006;27(2 Suppl):S145–55. Howard JS, Stanislaw H, Green G, Sparkman CR, Cohen HG. Comparison of behavior analytic and eclectic early interventions for young children with autism after three years. Res Dev Disabil. 2014;35(12):3326–44. Magiati I, Charman T, Howlin P. A two-year prospective follow-up study of community-based early intensive behavioural intervention and specialist nursery provision for children with autism spectrum disorders. J Child Psychol Psychiatry. 2007;48(8):803–12. Remington B, Hastings RP, Kovshoff H, degli Espinosa F, Jahr E, Brown T, et al. Early intensive behavioral intervention: outcomes for children with autism and their parents after two years. Am J Ment Retard. 2007;112(6):418–38.

Certainty assessment							Effect		Certainty
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Relative (95% CI)	Absolute (95% CI)	
5	1 RCT, 4 CCTs	very serious ^{a,b}	not serious	not serious	not serious	none	-	MD 9.58 points higher (5.57 higher to 13.6 higher)	⊕⊕○○ LOW

CI: Confidence interval; MD: Mean difference

Explanations

a. One study was conducted via RCT design, and the other studies were conducted via CCT design.

b. All studies were downgraded for performance bias and detection bias.

Table 2
GRADE Evidence Profile for M-CHAT F in ASD Screening

Question: Should the M-CHAT-F be used to screen for ASD in asymptomatic, apparently healthy children?

Sensitivity	0.83 (95% CI: 0.55 to 0.95)	Prevalences	1.85 %
Specificity	0.95 (95% CI: 0.84 to 0.99)		

Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 100 patients tested	Test accuracy CoE
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias		
True positives (patients with ASD)	6 studies 13313 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	very serious ^b	very serious ^c	none	2 (1 to 2)	○ ○ ○ VERY LOW
False negatives (patients incorrectly classified as not having ASD)								0 (0 to 1)	
True negatives (patients without ASD)	6 studies 13313 patients	cross-sectional (cohort type accuracy study)	serious ^a	not serious	very serious ^d	not serious	none	94 (83 to 97)	○ ○ ○ VERY LOW
False positives (patients incorrectly classified as having ASD)								4 (1 to 15)	

Explanations

a. Additional behavioral tests differed between the 6 studies; 1 study used varying behavioral tests among their sample.

b. Heterogeneity statistics for sensitivity included $Q = 37.74$, $p = 0.00$, and $I^2 = 84.10$, indicating great heterogeneity between the studies.

c. 95% CI of the pooled sensitivity = 0.54 - 0.95.

d. Heterogeneity statistics for specificity included $Q = 291.87$, $p = 0.00$, and $I^2 = 97.94$, indicating great heterogeneity between the studies.

Table 3**GRADE Evidence Profile for M-CHAT R/F in ASD Screening**

Question: Should M-CHAT R/F be used to screen for ASD in asymptomatic, apparently healthy children?

Sensitivity	0.85 (95% CI: 0.79 to 0.92)				Prevalences	0.67%			
Specificity	0.99 (95% CI: 0.99 to 0.99)								Test accuracy CoE
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested	
True positives (patients with ASD)	1 study 15612 patients	cross-sectional (cohort type accuracy study)	serious ^{a,b}	not serious	not serious	not serious	none	6 (5 to 6)	 Moderate
False negatives (patients incorrectly classified as not having ASD)								1 (1 to 2)	
True negatives (patients without ASD)	1 study 15612 patients	cross-sectional (cohort type accuracy study)	serious ^{a,b}	not serious	not serious	not serious	none	983 (983 to 983)	 Moderate
False positives (patients incorrectly classified as having ASD)								10 (10 to 10)	

Explanations

a. Patient selection: It was unclear if a consecutive or random sample of patients was enrolled

b. Flow and timing: It was unclear if all patients received the same reference standard, and not all patients were included in the 2x2 table

Question: Should the PDQ be used to screen for ASD in asymptomatic, apparently healthy children?

Sensitivity	0.85 (95% CI: -- to --)				Prevalences	1.3%			
Specificity	0.99 (95% CI: -- to --)								
Outcome	Nº of studies (Nº of patients)	Study design	Factors that may decrease certainty of evidence					Effect per 1,000 patients tested	
			Risk of bias	Indirectness	Inconsistency	Imprecision	Publication bias	pre-test probability of 1.3%	
True positives (patients with ASD)	1 studies 1959 patients	cross-sectional (cohort type accuracy study)	not serious	not serious	not serious	very serious ^a	none	11 (0 to 0)	
False negatives (patients incorrectly classified as not having ASD)								2 (13 to 13)	
True negatives (patients without ASD)	1 studies 1959 patients	cross-sectional (cohort type accuracy study)	not serious	not serious	not serious	very serious ^a	none	977 (0 to 0)	
False positives (patients incorrectly classified as having ASD)								10 (987 to 987)	

Explanations

a. The study did not indicate a 95% confidence interval.

APPENDIX E. Forest plots

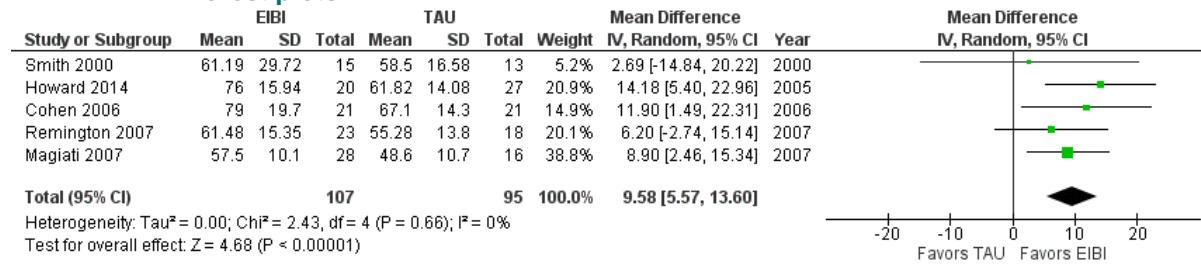


Figure 1. Forest plot of pooled effect of EBI vs TAU on adaptive behavior in children less than 6 years old with a diagnosis of ASD.

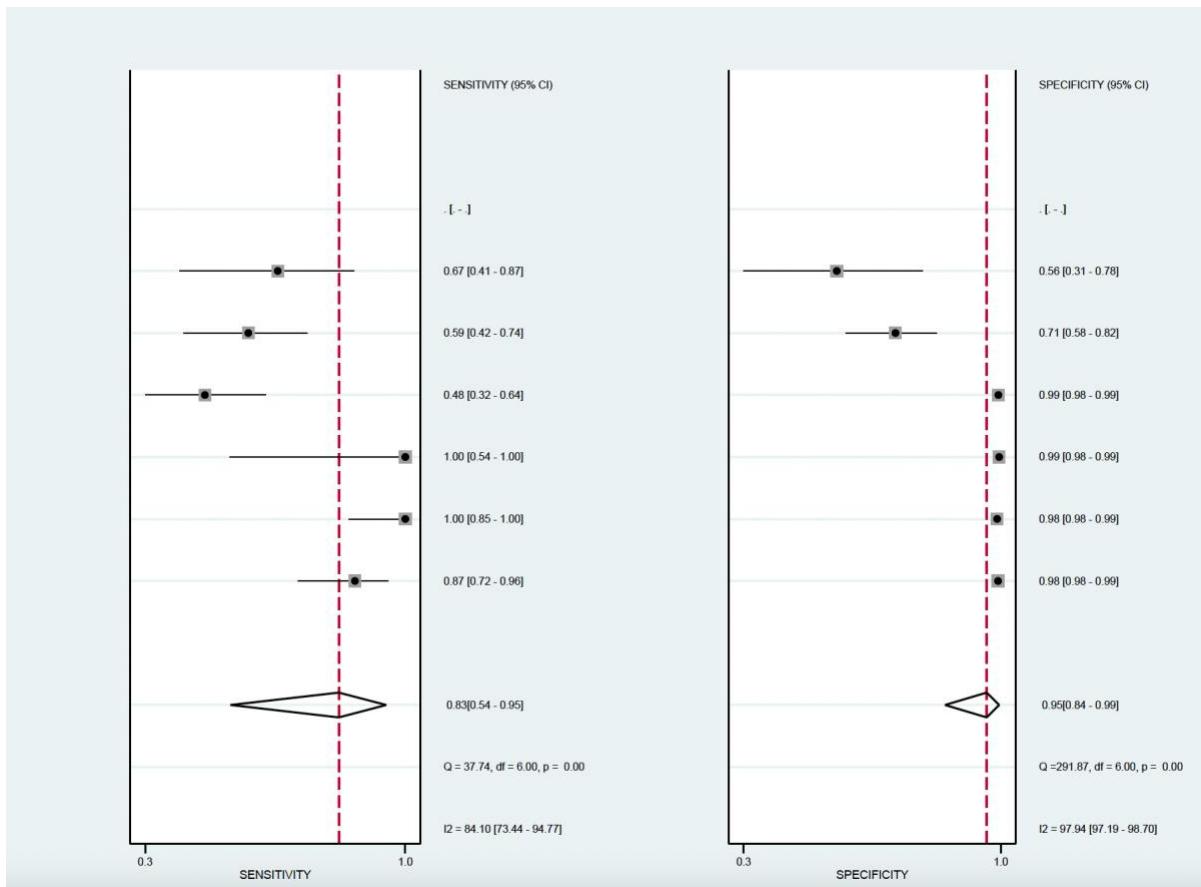


Figure 2. Forest plot of pooled sensitivity and specificity of M-CHAT F in asymptomatic, apparently healthy children <6 years old

15. Screening for Learning Disorders

APPENDIX A. Benefit and Harm of Educational intervention versus any control

Table O.10: Learning Experiences and Alternative Program for Pre-schoolers and Their Parents (LEAP) – full replication condition versus manual-only attention control

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With attention control	With educational intervention		Risk with attention control	Risk difference with educational intervention (95% CI)
Behaviour that challenges (severity) – post-treatment (measured with: Change score¹; Better indicated by lower values)											
294 (1 study)	serious ²	no serious inconsistency	serious ³	serious ⁴	undetected	⊕⊖⊖ VERY LOW ^{2,3,4} due to risk of bias, indirectness, imprecision	117	177	-	The mean behaviour that challenges (severity) – post-treatment in the intervention groups was 0.19 standard deviations lower (0.42 lower to 0.04 higher)	
Adaptive functioning (social) – post-treatment (Better indicated by lower values)											
294 (1 study)	serious ²	no serious inconsistency	serious ³	serious ⁴	undetected	⊕⊖⊖ VERY LOW ^{2,3,4} due to risk of bias,	117	177	-	The mean adaptive functioning (social) – post-treatment in the	
Quality assessment							Summary of findings				
							indirectness, imprecision				intervention groups was 0.76 standard deviations higher (0.52 to 1 higher)
Adaptive functioning (communication) – post-treatment (Better indicated by lower values)											
294 (1 study)	serious ²	no serious inconsistency	serious ³	serious ⁴	undetected	⊕⊖⊖ VERY LOW ^{2,3,4} due to risk of bias, indirectness, imprecision	117	177	-	The mean adaptive functioning (communication) – post-treatment in the intervention groups was 0.94 standard deviations higher (0.7 to 1.19 higher)	

¹ Due to significant baseline differences, standard deviation of change and estimates of mean change were derived using initial and final mean values and utilising $r = 0.5$. Sensitivity analyses were used to explore the impact of altering assumptions about the calculation of the effect size, but this resulted in no change to conclusions.

² Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect

Source: National Collaborating Centre for Mental Health (UK). (2015). Challenging behaviour and learning disabilities: prevention and interventions for people with learning disabilities whose behaviour challenges..Appendix O. Table O.10. Learning Experiences and Alternative Program for Pre-schoolers and Their Parents (LEAP) – full replication condition versus manual-only attention control. Page 18-19

APPENDIX B. Benefit and Harm of home based versus community based early behavioral intervention

Table O.11: Home-based Building Blocks programme versus centre-based Building Blocks programme

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)	Relative effect (95% CI)	Anticipated absolute effects	Risk with centre-based early behavioural intervention	Risk difference with home-based early behavioural intervention (95% CI)
Behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
44 (1 study)	serious ¹	no serious inconsistency	no serious indirectness	very serious ²	undetected	⊕⊕⊕⊕ VERY LOW ^{1,2} due to risk of bias, imprecision	22	22	-		The mean behaviour that challenges (severity) – post-treatment in the intervention groups was 0.11 standard deviations lower (0.7 lower to 0.48 higher)
Adaptive functioning (social) – post-treatment (Better indicated by lower values)											
56 (1 study)	serious ¹	no serious inconsistency	no serious indirectness	very serious ²	undetected	⊕⊕⊕⊕ VERY LOW ^{1,2} due to risk of bias, imprecision	29	27	-		The mean adaptive functioning (social) – post-treatment in the intervention
Challenging behaviour and learning disabilities							Summary of findings				
Quality assessment											
Adaptive functioning (communication) – post-treatment (Better indicated by lower values)											
55 (1 study)	serious ¹	no serious inconsistency	no serious indirectness	very serious ²	undetected	⊕⊕⊕⊕ VERY LOW ^{1,2} due to risk of bias, imprecision	29	26	-		The mean adaptive functioning (communication) – post-treatment in the intervention groups was 0.46 standard deviations lower (1 lower to 0.07 higher)

¹ Crucial limitation for one criterion or some limitations for multiple criteria sufficient to lower ones confidence in the estimate of effect

² Optimal information size not met; small, single study

Source: National Collaborating Centre for Mental Health (UK). (2015). Challenging behaviour and learning disabilities: prevention and interventions for people with learning disabilities whose behaviour challenges.. Appendix O. Table O.11. Home-based Building Blocks programme versus centre-based Building Blocks programme. Page 20-21

APPENDIX C. Benefit and Harm of parent education, support and skills training

Table O.14: Parent training (plus centre based EBI) versus treatment as usual (centre-based EBI)

Quality assessment							Summary of findings				
Participants (studies) Follow up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
Behaviour that challenges (severity) – post-treatment (Better indicated by lower values)											
57 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ¹	undetected	⊕⊕⊕ LOW ¹ due to imprecision	28	29	-		The mean behaviour that challenges (severity) – post-treatment in the intervention groups was 0.79 lower to 0.05 higher)
Adaptive functioning (global) – post-treatment (Better indicated by lower values)											
58 (1 study)	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ¹	undetected	⊕⊕⊕ LOW ¹ due to imprecision	28	30	-		The mean adaptive functioning (global) – post-treatment in the intervention groups was 0.25 standard deviations higher (0.27 lower to 0.77 higher)
Adaptive functioning (global) – follow-up (Better indicated by lower values)											
119 (2 studies) 26 to 52 weeks	serious ²	no serious inconsistency	no serious indirectness	serious ³	undetected	⊕⊕⊕ LOW ^{2,3} due to risk of bias, imprecision	56	63	-		The mean adaptive functioning (global) – follow-up in the intervention groups was 0.52 standard deviations higher
Quality assessment							Summary of findings				
Adaptive functioning (communication) – follow-up (Better indicated by lower values)											
68 (1 study) 26 weeks	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious ¹	undetected	⊕⊕⊕ LOW ¹ due to imprecision	33	35	-		The mean adaptive functioning (communication) – follow-up in the intervention groups was 0.75 standard deviations higher (0.26 to 1.25 higher)

¹ Optimal information size not met; small, single study

² Most information is from studies at moderate risk of bias

³ Optimal information size not met

Source: National Collaborating Centre for Mental Health (UK). (2015). Challenging behaviour and learning disabilities: prevention and interventions for people with learning disabilities whose behaviour challenges.. Appendix O. Table O.14: Parent training (plus centre based EBI) versus treatment as usual (centre-based EBI). Page 24-26

APPENDIX D. AMSTAR Tables

AMSTAR Checklist for Galuschka 2020

Galuschka 2020	
AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both	
<p>6. Did the review authors perform data extraction in duplicate?</p> <p>For Yes, either ONE of the following:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> at least two reviewers achieved consensus on what data to extract from included studies <input type="checkbox"/> one reviewer extracted data from a sample of eligible studies and achieved good agreement (at least 90 percent), with the remainder extracted by one reviewer. 	
<p>7. Did the review authors provide a list of excluded studies and justify the exclusion?</p> <p>For Partial Yes:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> provided a list of all potentially relevant studies that were read in full-text but excluded from the review <p>For Yes, must also have:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Justified the exclusion from the review of each potentially relevant study <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No</p>	
<p>8. Did the review authors describe the included studies in adequate detail?</p> <p>For Partial Yes (ALL the following):</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> described populations <input checked="" type="checkbox"/> described interventions <input checked="" type="checkbox"/> described comparators <input checked="" type="checkbox"/> described outcomes <input checked="" type="checkbox"/> described research designs <p>For Yes, should also have ALL the following:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> described population in detail <input checked="" type="checkbox"/> described intervention in detail (including doses where relevant) <input type="checkbox"/> described comparator in detail (including doses where relevant) <input checked="" type="checkbox"/> described study's setting <input checked="" type="checkbox"/> timeframe for follow-up <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No</p>	
<p>9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?</p> <p>RCTs</p> <p>For Partial Yes, must have assessed RoB from:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> un concealed allocation, and <input checked="" type="checkbox"/> lack of blinding of patients and assessors when assessing outcome(s) necessary for objective outcomes such as all-cause mortality) <p>For Yes, must also have assessed RoB from:</p> <ul style="list-style-type: none"> <input type="checkbox"/> allocation sequence that was not truly random, and <input type="checkbox"/> selection of the reported result from among multiple measurements or analyses of a specified outcome <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No <input type="checkbox"/> Includes only NRSI</p>	
<p>NRSI</p> <p>For Partial Yes, must have assessed RoB from:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> from confounding, and <input type="checkbox"/> from selection bias <p>For Yes, must also have assessed RoB:</p> <ul style="list-style-type: none"> <input type="checkbox"/> methods used to ascertain exposure and outcome, and <input type="checkbox"/> selection of the reported result from among multiple measurements or analyses of a specified outcome <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No <input type="checkbox"/> Includes only RCTs</p>	
<p>10. Did the review authors report on the sources of funding for the studies included in the review?</p> <p>For Yes:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p>	

Galuscha 2020	
AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both	
<p>11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?</p> <p>RCTs:</p> <p>For Yes:</p> <ul style="list-style-type: none"> <input type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present <input type="checkbox"/> AND investigated the cause(s) of any heterogeneity <p>For NRSI</p> <p>For Yes:</p> <ul style="list-style-type: none"> <input type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present <input type="checkbox"/> AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available <input type="checkbox"/> AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review <p>12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?</p> <p>For Yes:</p> <ul style="list-style-type: none"> <input type="checkbox"/> included only low risk of bias RCTs <input type="checkbox"/> OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect <p>13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?</p> <p>For Yes:</p> <ul style="list-style-type: none"> <input type="checkbox"/> included only low risk of bias RCTs <input type="checkbox"/> OR, if RCTs with moderate or high RoB, or NRSI were included the review performed a discussion of the likely impact of RoB on the results <p>14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?</p> <p>For Yes:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> There was no significant heterogeneity in the results <input type="checkbox"/> OR, if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review <p>15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?</p> <p>For Yes:</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias <input type="checkbox"/> No <input type="checkbox"/> OR, no meta-analysis conducted 	

AMSTAR Checklist for Toffalini 2021

Toffalini 2021
AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

1. Did the research questions and inclusion criteria for the review include the components of PICO?	
For Yes:	Optional (recommended)
<input checked="" type="checkbox"/> Population	<input type="checkbox"/> Timeframe for follow-up
<input checked="" type="checkbox"/> Intervention	<input checked="" type="checkbox"/> Yes
<input checked="" type="checkbox"/> Comparator group	<input type="checkbox"/> No
<input checked="" type="checkbox"/> Outcome	
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	
For Partial Yes: The authors state that they had a written protocol or guide that included ALL the following:	
<input checked="" type="checkbox"/> review question(s)	<input checked="" type="checkbox"/> Yes
<input checked="" type="checkbox"/> a search strategy	<input checked="" type="checkbox"/> Partial Yes
<input checked="" type="checkbox"/> inclusion/exclusion criteria	<input type="checkbox"/> No
<input type="checkbox"/> a risk of bias assessment	
3. Did the review authors explain their selection of the study designs for inclusion in the review?	
For Yes, the review should satisfy ONE of the following: <i>Explanation</i> for including only RCTs <input checked="" type="checkbox"/> Yes <input type="checkbox"/> OR <i>Explanation</i> for including only NRSI <input type="checkbox"/> No <input checked="" type="checkbox"/> OR <i>Explanation</i> for including both RCTs and NRSI	
4. Did the review authors use a comprehensive literature search strategy?	
For Partial Yes (all the following): For Yes, should also have (all the following):	
<input checked="" type="checkbox"/> searched at least 2 databases (relevant to research question)	<input type="checkbox"/> Yes
<input checked="" type="checkbox"/> provided key word and/or search strategy	<input checked="" type="checkbox"/> Partial Yes
<input checked="" type="checkbox"/> justified publication restrictions (e.g. language)	<input type="checkbox"/> No
5. Did the review authors perform study selection in duplicate?	
For Yes, either ONE of the following: <input checked="" type="checkbox"/> at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include <input checked="" type="checkbox"/> Yes <input type="checkbox"/> OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer	

Toffalini 2021
AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

6. Did the review authors perform data extraction in duplicate?	
For Yes, either ONE of the following: <input checked="" type="checkbox"/> at least two reviewers achieved consensus on which data to extract from included studies <input checked="" type="checkbox"/> Yes <input type="checkbox"/> OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer <input type="checkbox"/> No	
7. Did the review authors provide a list of excluded studies and justify the exclusions?	
For Partial Yes: provided a list of all potentially relevant studies that were refuted in full-text form but excluded from the review	
For Yes, must also have: <input checked="" type="checkbox"/> Justified the exclusion from the review of each potentially relevant study <input checked="" type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No	
8. Did the review authors describe the included studies in adequate detail?	
For Partial Yes (ALL the following): For Yes, should also have ALL the following:	
<input checked="" type="checkbox"/> described populations	<input type="checkbox"/> described population in detail
<input checked="" type="checkbox"/> described interventions	<input type="checkbox"/> described intervention in detail (including doses where relevant)
<input checked="" type="checkbox"/> described comparators	<input type="checkbox"/> described comparator in detail (including doses where relevant)
<input checked="" type="checkbox"/> described outcomes	<input type="checkbox"/> described outcomes
<input type="checkbox"/> described research designs	<input type="checkbox"/> described study's setting and timeframe for follow-up
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	
RCTs: For Partial Yes, must have assessed RoB from: <input type="checkbox"/> uncontrolled allocation, and <input type="checkbox"/> lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)	
For Yes, must also have assessed RoB from: <input type="checkbox"/> allocation sequence that was not truly random, and <input type="checkbox"/> selection of the reported result from among multiple measurements or analyses of a specified outcome <input type="checkbox"/> Yes <input checked="" type="checkbox"/> Partial Yes <input type="checkbox"/> No <input type="checkbox"/> Includes only NRSI	
NRSI: For Partial Yes, must have assessed RoB: <input type="checkbox"/> from confounding, and <input type="checkbox"/> from selection bias	
For Yes, must also have assessed RoB: <input type="checkbox"/> methods used to ascertain outcomes and outcomes, and <input type="checkbox"/> selection of the reported result from among multiple measurements or analyses of a specified outcome <input type="checkbox"/> Yes <input checked="" type="checkbox"/> Partial Yes <input type="checkbox"/> No <input type="checkbox"/> Includes only RCTs	
10. Did the review authors report on the sources of funding for the studies included in the review?	
For Yes: <input type="checkbox"/> Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No	

Toffalini 2021
AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	
RCTs: For Yes: <input type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present <input type="checkbox"/> AND investigated the cause(s) of any heterogeneity	
<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted	
NRSI: For Yes: <input type="checkbox"/> The authors justified combining the data in a meta-analysis <input checked="" type="checkbox"/> AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present <input type="checkbox"/> AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available <input type="checkbox"/> AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review	
<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted	
12. If a meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	
For Yes: <input type="checkbox"/> included only low risk of bias: RCTs <input type="checkbox"/> OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect	
<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted	
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	
For Yes: <input type="checkbox"/> included only low risk of bias: RCTs <input type="checkbox"/> OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results	
<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No	
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	
For Yes: <input type="checkbox"/> There was no significant heterogeneity in the results <input checked="" type="checkbox"/> OR, if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review	
<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No	
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	
For Yes: <input checked="" type="checkbox"/> performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias	
<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted	

Toffalini 2021
AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	
For Yes: <input type="checkbox"/> The authors reported no competing interests OR <input checked="" type="checkbox"/> The authors described their funding sources and how they managed potential conflicts of interest	
<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No	

To cite this tool: Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;356:j4008.

AMSTAR Checklist for Macarthur 2018

McArthur 2018

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

1. Did the research question and inclusion criteria for the review include the components of PICO?	
For Yes:	Optional (recommended)
<input checked="" type="checkbox"/> Population	<input type="checkbox"/> Timeframe for follow-up
<input checked="" type="checkbox"/> Intervention	<input type="checkbox"/> Yes
<input checked="" type="checkbox"/> Comparator group	<input type="checkbox"/> No
<input checked="" type="checkbox"/> Outcome	
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	
For Partial Yes: The authors state that they had a written protocol or guide that included ALL the following:	
<input checked="" type="checkbox"/> a review question(s)	<input type="checkbox"/> Yes
<input checked="" type="checkbox"/> a search strategy	<input type="checkbox"/> Partial Yes
<input checked="" type="checkbox"/> inclusion/exclusion criteria	<input type="checkbox"/> No
<input checked="" type="checkbox"/> a risk of bias assessment	<input type="checkbox"/> justification for deviations from the protocol
3. Did the review authors explain their selection of the study designs for inclusion in the review?	
For Yes, the review should satisfy ONE of the following: <input type="checkbox"/> Explanation for including only RCTs <input type="checkbox"/> OR Explanation for including only NRSI <input checked="" type="checkbox"/> OR Explanation for including both RCTs and NRSI	
4. Did the review authors use a comprehensive literature search strategy?	
For Partial Yes (all the following): <input checked="" type="checkbox"/> searched at least 2 databases (relevant to research question) <input checked="" type="checkbox"/> provided key word and/or search strategy <input checked="" type="checkbox"/> justified publication restrictions (e.g. language)	
For Yes, should also have (all the following): <input checked="" type="checkbox"/> searched the reference lists / bibliographies of included studies <input checked="" type="checkbox"/> searched trial/study registries <input checked="" type="checkbox"/> included/consulted content experts in the field <input checked="" type="checkbox"/> where applicable, searched for grey literature <input checked="" type="checkbox"/> conducted search within 24 months of completion of the review	
5. Did the review authors perform study selection in duplicate?	
For Yes, either ONE of the following: <input checked="" type="checkbox"/> at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include <input type="checkbox"/> OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer	

McArthur 2018

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

6. Did the review authors perform data extraction in duplicate?	
For Yes, either ONE of the following: <input checked="" type="checkbox"/> at least two reviewers achieved consensus on which data to extract from included studies <input type="checkbox"/> OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer	
7. Did the review authors provide a list of excluded studies and justify the exclusions?	
For Partial Yes: <input checked="" type="checkbox"/> provided a list of all potentially relevant studies that were read in full-text form but excluded from the review	
For Yes, must also have: <input checked="" type="checkbox"/> identified the exclusion from the review of each potentially relevant study	
8. Did the review authors describe the included studies in adequate detail?	
For Partial Yes (ALL the following): For Yes, should also have ALL the following: <input checked="" type="checkbox"/> described populations in detail <input checked="" type="checkbox"/> described interventions in detail (including doses where relevant) <input checked="" type="checkbox"/> described comparators in detail (including doses where relevant) <input checked="" type="checkbox"/> described outcomes in detail (including doses where relevant) <input checked="" type="checkbox"/> described research designs in detail (including doses where relevant)	
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	
RCTs For Partial Yes, must have assessed RoB from: <input checked="" type="checkbox"/> concealed allocation, and lack of blinding of patients and assessors when assessing outcome (unnecessary for objective outcomes such as all-cause mortality)	
For Yes, must also have assessed RoB from: <input checked="" type="checkbox"/> allocation sequence that was not only random, and selection of the reported result from among multiple measurements or analyses of a specified outcome	
NRSI For Partial Yes, must have assessed RoB from: <input checked="" type="checkbox"/> from confounding, and selection bias	
For Yes, must also have assessed RoB: methods used to ascertain exposures and outcomes, and selection of the reported result from among multiple measurements or analyses of a specified outcome	
10. Did the review authors report on the sources of funding for the studies included in the review?	
For Yes: <input checked="" type="checkbox"/> Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies	

McArthur 2018

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	
RCTs For Yes: <input checked="" type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present <input type="checkbox"/> AND investigated the causes of any heterogeneity	
For NRSI For Yes: <input checked="" type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present <input checked="" type="checkbox"/> AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available <input checked="" type="checkbox"/> AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review	
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	
For Yes: <input type="checkbox"/> included only low risk of bias RCTs <input checked="" type="checkbox"/> OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect	
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	
For Yes: <input type="checkbox"/> included only low risk of bias RCTs <input checked="" type="checkbox"/> OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results	
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	
For Yes: <input type="checkbox"/> There was no significant heterogeneity in the results <input checked="" type="checkbox"/> OR, if heterogeneity was present the author performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review	
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	
For Yes: <input checked="" type="checkbox"/> performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of publication bias	

McArthur 2018

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	
For Yes: <input type="checkbox"/> The authors reported no competing interests OR <input checked="" type="checkbox"/> The authors described their funding sources and how they managed potential conflicts of interest	

To cite this tool: Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008.

AMSTAR Checklist for Lee 2015

Lee 2015

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

1. Did the research questions and inclusion criteria for the review include the components of PICO?	
For Yes:	Optional (recommended) <input type="checkbox"/> Population <input checked="" type="checkbox"/> Intervention <input type="checkbox"/> Comparator group <input checked="" type="checkbox"/> Outcome
<input type="checkbox"/> Timeline for follow-up <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No	
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	
For Partial Yes:	The authors state that they had a written protocol or guide that included ALL the following: <input checked="" type="checkbox"/> review question(s) <input checked="" type="checkbox"/> a search strategy <input checked="" type="checkbox"/> inclusion/exclusion criteria <input type="checkbox"/> a risk of bias assessment
For Yes: As for partial yes, plus the protocol should be registered and should also have specified: <input checked="" type="checkbox"/> a meta-analysis/synthesis plan, if appropriate, and <input checked="" type="checkbox"/> a list for investigating causes of heterogeneity <input type="checkbox"/> justification for any deviations from the protocol	
3. Did the review authors explain their selection of the study designs for inclusion in the review?	
For Yes, the review should satisfy ONE of the following:	<input type="checkbox"/> Explanation for including only RCTs <input type="checkbox"/> OR Explanation for including only NRSI <input type="checkbox"/> OR Explanation for including both RCTs and NRSI
4. Did the review authors use a comprehensive literature search strategy? For Partial Yes (all the following): <input checked="" type="checkbox"/> searched at least 2 databases (relevant to research question) <input checked="" type="checkbox"/> provided key word and/or search strategy <input type="checkbox"/> justified publication restrictions (e.g. language)	
For Yes, should also have (all the following): <input type="checkbox"/> searched the reference lists / bibliographies of included studies <input checked="" type="checkbox"/> searched trial/study registries <input type="checkbox"/> included/consulted content experts in the field <input type="checkbox"/> where relevant, searched for grey literature <input type="checkbox"/> conducted search within 24 months of completion of the review	
5. Did the review authors perform study selection in duplicate?	
For Yes, either ONE of the following:	<input checked="" type="checkbox"/> at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include <input type="checkbox"/> OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.
<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No	

Lee 2015

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	
RCTs	For Yes: <input type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present. <input type="checkbox"/> AND investigated the causes of any heterogeneity
	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted
NRSI	For Yes: <input checked="" type="checkbox"/> The authors justified combining the data in a meta-analysis <input type="checkbox"/> AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present <input checked="" type="checkbox"/> AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, by averaging raw data when adjusted effect estimates were not available <input type="checkbox"/> AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review
	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	
For Yes:	<input type="checkbox"/> included only low risk of bias RCTs <input type="checkbox"/> OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.
	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted
13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?	
For Yes:	<input type="checkbox"/> included only low risk of bias RCTs <input type="checkbox"/> OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results
	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	
For Yes:	<input type="checkbox"/> There was no significant heterogeneity in the results <input checked="" type="checkbox"/> OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review
	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
15. If the performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	
For Yes:	<input type="checkbox"/> performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias
	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No meta-analysis conducted

Lee 2015

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

6. Did the review authors perform data extraction in duplicate?	
For Yes, either ONE of the following:	<input checked="" type="checkbox"/> at least two reviewers achieved consensus on which data to extract from included studies <input type="checkbox"/> OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer
7. Did the review authors provide a list of excluded studies and justify the exclusions?	
For Partial Yes:	For Yes, must also have: <input type="checkbox"/> provided a list of all potentially relevant studies that were read in full-text form and excluded from the review
	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partial Yes <input checked="" type="checkbox"/> No
8. Did the review authors describe the included studies in adequate detail?	
For Partial Yes (ALL the following):	For Yes, should also have ALL the following: <input checked="" type="checkbox"/> described populations <input type="checkbox"/> described interventions <input type="checkbox"/> described comparators <input checked="" type="checkbox"/> described outcomes <input checked="" type="checkbox"/> described research designs
	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No
9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?	
RCTs	For Partial Yes, must have assessed RoB from: <input type="checkbox"/> un concealed allocation, and <input type="checkbox"/> lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)
	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No <input type="checkbox"/> Includes only NRSI
NRSI	For Partial Yes, must have assessed RoB: <input type="checkbox"/> from confounding, and <input type="checkbox"/> from selection bias
	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No <input type="checkbox"/> Includes only RCTs
10. Did the review authors report on the sources of funding for the studies included in the review?	
For Yes:	<input type="checkbox"/> Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies
	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	
For Yes:	<input type="checkbox"/> The authors reported no competing interests OR <input checked="" type="checkbox"/> The authors described their funding sources and how they managed potential conflicts of interest
	<input checked="" type="checkbox"/> Yes <input type="checkbox"/> No

Lee 2015

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

To cite this tool: Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008.

APPENDIX E. Summary of findings for phonics training versus control for English-speaking poor readers

Summary of findings for the main comparison. Phonics training versus control (no training or alternative training) for English-speaking poor readers

Phonics training versus control (no training or alternative training) for English-speaking poor readers

Patient or population: English-speaking poor readers

Setting: English-speaking countries

Intervention: phonics

Comparison: control (no training or alternative training)

Outcomes	Illustrative comparative risks (SMD* 95% CI*)		Relative effect (95% CI)	N* of participants (studies)	Quality of the evidence (GRADE)*	Comments
	Assumed risk	Corresponding risk				
	Control (no training or alternative training)	Phonics training				
Mixed/regular word reading accuracy Assessed with: various scales Follow-up: immediate	— The mean score in the intervention groups was 0.51 standard deviations higher (0.13 higher to 0.90 higher)	—	—	701 (11 studies)	⊕⊕⊕ Low^a	A standard deviation of 0.51 represented a moderate effect between groups. Phonics training "may improve" outcome (Ryan 2016).
Non-word reading accuracy Assessed with: various scales Follow-up: immediate	— The mean score in the intervention groups was 0.67 standard deviations higher (0.26 higher to 1.07 higher)	—	—	682 (10 studies)	⊕⊕⊕ Low^a	A standard deviation of 0.67 presented a moderate effect between groups. Phonics training "may improve" outcome (Ryan 2016).
Irregular word reading accuracy Assessed with: various scales Follow-up: immediate	— The mean score in the intervention groups was 0.84 standard deviations higher (0.30 higher to 1.39 higher)	—	—	294 (4 studies)	⊕⊕⊕ Moderate^{a,c}	A standard deviation of 0.84 presented a large effect between groups. Phonics training "probably improves" outcome (Ryan 2016).
Mixed/regular word reading fluency Assessed with: various scales Follow-up: immediate	— The mean score in the intervention groups was 0.45 standard deviations higher (0.19 higher to 0.72 higher)	—	—	224 (4 studies)	⊕⊕⊕ Moderate^b	A standard deviation of 0.45 presented a moderate effect between groups. Phonics training "probably improves" outcome (Ryan 2016).
Non-word reading fluency Assessed with: various scales Follow-up: immediate	— The mean score in the intervention groups was 0.39 standard deviations higher (0.10 higher to 0.68 higher)	—	—	188 (3 studies)	⊕⊕⊕ Moderate^b	A standard deviation of 0.39 presented a moderate effect between groups. Phonics training "probably improves" outcome (Ryan 2016).
Reading comprehension Assessed with: various scales Follow-up: immediate	— The mean score in the intervention groups was 0.28 standard deviations higher (0.07 lower to 0.62 higher)	—	—	343 (5 studies)	⊕⊕⊕ Low^a	A standard deviation of 0.28 presented a small effect between groups. Phonics training "may improve" outcome (Ryan 2016).
Spelling Assessed with: various scales Follow-up: immediate	— The mean score in the intervention groups was 0.47 standard deviations higher (0.07 lower to 1.01 higher)	—	—	158 (3 studies)	⊕⊕⊕ Low^a	A standard deviation of 0.47 presented a moderate effect between groups. Phonics training "may improve" outcome (Ryan 2016).

SMD: standardised mean difference. Different studies used different continuous measures. Thus, effect sizes are reflected by size of phonics training effect as indexed using SMDs. The results are expressed as standard deviation (SD) units. As a general rule, 0.2 SMD represents a small effect size, 0.5 a moderate effect size, and 0.8 a large effect size.

CI: confidence interval.

GRADE: Working Group grades of evidence

High quality: we are very confident that the true effect lies close to the that of the estimate of the effect.

Moderate quality: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low quality: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of effect.

Very low quality: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

^aDowngraded two levels due to very serious imprecision: very wide confidence intervals (greater than 0.6; Schünemann 2011b).

^bDowngraded one level due to serious imprecision: wide confidence intervals (0.3 to 0.6; Schünemann 2011b).

APPENDIX F. Academic Screening Tools chart for reading

About DBI	Tools Charts	Implementation & Intervention	Training	Special Topics	Resource by Audience
-----------	--------------	-------------------------------	----------	----------------	----------------------

Academic Screening Tools Chart

Legend  Convincing evidence  Partially convincing evidence  Unconvincing evidence  Data unavailable  Disaggregated data available

Title	Area	Age/Grade	Classification Accuracy Fall	Classification Accuracy Winter	Classification Accuracy Spring	Reliability	Validity	Sample Representativeness	Bias Analysis Conducted
Acadience Reading (aka DIBELS Next)	Composite Score	Kindergarten						Regional without Cross-Validation	Yes
Acadience Reading (aka DIBELS Next)	Composite Score	Grade 1						Regional without Cross-Validation	Yes
Acadience Reading (aka DIBELS Next)	Composite Score	Grade 2						Regional without Cross-Validation	Yes
Acadience Reading (aka DIBELS Next)	Composite Score	Grade 3						Regional with Cross-Validation	Yes
Acadience Reading (aka DIBELS Next)	Composite Score	Grade 4						Regional with Cross-Validation	Yes
Title	Area	Age/Grade	Classification Accuracy Fall	Classification Accuracy Winter	Classification Accuracy Spring	Reliability	Validity	Sample Representativeness	Bias Analysis Conducted
Acadience Reading (aka DIBELS Next)	Composite Score	Kindergarten						Regional without Cross-Validation	Yes
Achieve3000's LevelSet	Reading	Grade 3						Local without Cross-Validation	Yes
Achieve3000's LevelSet	Reading	Grade 4						Local without Cross-Validation	Yes
Achieve3000's LevelSet	Reading	Grade 5						Local without Cross-Validation	Yes
Classworks Universal Screener	Reading	Grade 2						National without Cross-Validation	No
DIBELS 8th Edition	Composite	Kindergarten						National without Cross-Validation	No
DIBELS 8th Edition	Letter Naming Fluency	Kindergarten						National without Cross-Validation	No
DIBELS 8th Edition	Nonsense Word Fluency Correct Letter Sounds	Kindergarten						National without Cross-Validation	No
DIBELS 8th Edition	Oral Reading Fluency	Grade 1						Regional without Cross-Validation	No

PERIODIC HEALTH EXAMINATION TASK FORCE ON CONGENITAL AND DEVELOPMENTAL DISORDERS 2021

Task Force Steering Committee

Chair Mary Ann R. Abacan, MD, MSc, FPPS
Co-chair Kathryn R. Baltazar-Braganza, MD, MSc (cand.)
Members: Mary Anne D. Chiong, MD, MSc
Anna Guia O. Limpoco, MD, MSc, DFM, FPAFP
Maria Isabel O. Quilendrino, MD, DPPS, FPSDBP
Angelica Cecilia V. Tomas, MD

Technical Working Group

Technical Coordinators

Kathryn R. Baltazar-Braganza, MD, MSc (cand.)
Ian Theodore G. Cabaluna, RPh, MD, Gdip (Clin Epid), MSc (cand.)

Evidence Review Experts

Christine Mae S. Avila, MD, DPPS
Hazel S. Baconga, MD, DPPS
Reginald B. Balmeo, MD, DPPS, DPAPP
Liza Marie P. Bejemonio, MD, DPPS
Fides Roxanne M. Castor, MD, DPPS
Ma. Theresa M. Collante, MD, FPPS, FPRA, CCD
Melissa A. Dator, MD, MBA, DPPS, DPSN
Ma. Margarita S. de Rivera, MD, DPPS
Vaneza Leah A. Espino, MD, DPPS, DPAPP
Natasha Ann R. Esteban-Ipac, MD, FPPS, DPSAMS
Erena S. Kasahara, MD, DPPS
Mark Andrew O. Perez, MD, DPPS, DPSN, DPNSP
Corinna Victoria M. Puyat, MD
Marie Julianne C. Racoma, MD, DPPS
Charlotte Averill Y. Tan, MD
Grazielle S. Verzosa, MD, DPPS

Consensus Panel

Maria Melanie Liberty B. Alcausin, MD, FPPS
Abigael C. Andal, MD, FPAFP
Martin Graciano Raymundo S. Baquiran, MD,
MPH, FPPS, FAAP

Hairam R. Encendencia, MD, MCHM, DFM

Institute of Human Genetics
Philippine Academy of Family Medicine
Catholic Educational Association of the
Philippines

Association of Municipal Health Officers of the
Philippines

Cindy C. Llego, MD, FPPS, FPSDBP

Philippine Society for Developmental and Behavioral Pediatrics

Janet Marriane Go-Nierva, MD, FPPS, FPSNbM

Philippine Pediatric Society

Florence B. Nitafan

Homeschoolers Association of the Philippine Islands

Maria Socorro L. Romabiles, PhD

Social Scientist

Benjamin P. Sablan Jr., MD, MDM

Philippine Ambulatory Pediatrics Association

Maria Wilda T. Silva, MD, MBA in Health

Department of Health

Consensus Panel Meeting Facilitator

Carlo Irwin A. Panelo, MD, MA

Technical Writer

Isabella S. Ocampo, MD, DPPS

Administrative Officer

Claudette V. Silva

PERIODIC HEALTH EXAMINATION PHASE 2 TASK FORCE 2021

Central Steering Committee

Antonio L. Dans, MD, MSc
Dante D. Morales, MD
Marissa M. Alejandria, MD, MSc
Beverly Lorraine C. Ho, MD
Aileen R. Espina, MD
Maria Vanessa Villaruz-Sulit RN MSc

Oversight Committee

Chair: Dante D. Morales, MD
Co-Chair: Antonio L. Dans, MD, MSc
Members:
Angela Abanilla-Du, MD
Camilo G. Te, MD
Maria Vanessa V. Sulit, RN MSc

Project Leader

Leonila F. Dans, MD, MSc

Co-Project Leader

Marissa M. Alejandria, MD, MSc

Project Managers

Ian Theodore G. Cabaluna RPh, MD, GDip (Clin Epi), MSc (cand.)
Josephine T. Sanchez, RN

Administrative Staff

Lailanie Ann C. Tejuco
Maria Pamela Tagle

CONFLICT OF INTEREST DECLARATION

Panelist	COI based on Oversight Committee	Remarks
Maria Melanie Liberty B. Alcausin, MD, FPPS (Institute of Human Genetics)	Manageable A	Paper on advocacy for birth defects prevention and care, Director-Institute of Human Genetics, past director of Newborn Screening Center
Abigael C. Andal, MD, FPAFP (Philippine Academy of Family Medicine)	Manageable A	Section editor-The Filipino Family Physician
Martin Graciano Raymundo S. Baqirran, MD, MPH, FPPS, FAAP (Catholic Educational Association of the Philippines)	Manageable A	HealthDev Integrated Clinics-private practice
Hiram R. Encendencia, MD, MCHM, DFM (Association of Municipal Health Officers of the Philippines)	Manageable A	MHO
Cindy C. Llego, MD (Philippine Society for Developmental and Behavioral Pediatrics)	Manageable A	Board member PPS North Central Mindanao Chapter
Janet Marriane Go-Nierva, MD, FPPS, FPSNbM (Philippine Pediatric Society)	Acceptable	No COI
Florence B. Nitafan (Homeschoolers Association of the Philippine Islands)	Acceptable	No COI
Maria Socorro L. Romabiles, PhD (Social Scientist)	Acceptable	No COI
Benjamin P. Sablan Jr., MD, MDM (Philippine Ambulatory Pediatrics Association)	Manageable A	Asst. to the Associate Dean-UPCM; Head, Medical Affairs-Vaccine Sanofi Pasteur
Maria Wilda T. Silva, MD, MBA in Health (Department of Health)	Manageable B	Team Lead-Family Health Division of DOH