frontiers in Pediatrics

SYSTEMATIC REVIEW
published: 19 March 2021
doi: 10.3389/fped.2021.606194

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

Chris Stinton[1], Hannah Fraser[1], Julia Geppert[1], Rebecca Johnson[2], Martin Connock[1], Samantha Johnson[3], Aileen Clarke[1] and Sian Taylor-Phillips[1]*

[1] Warwick Medical School, University of Warwick, Coventry, United Kingdom, [2] School of Nursing, Midwifery and Health, Coventry University, Coventry, United Kingdom, [3] Warwick Library, University of Warwick, Coventry, United Kingdom

OPEN ACCESS

Edited by:
Lambertus Petrus Van Den Heuvel, Radboud University Nijmegen Medical Centre, Netherlands

Reviewed by:
Kazumichi Fujioka, Kobe University, Japan
Jiang-Qin Liu, Shanghai First Maternity and Infant Hospital, China

***Correspondence:**
Chris Stinton
C.Stinton@warwick.ac.uk

Specialty section:
This article was submitted to Neonatology, a section of the journal Frontiers in Pediatrics

Received: 14 September 2020
Accepted: 03 February 2021
Published: 19 March 2021

Citation:
Stinton C, Fraser H, Geppert J, Johnson R, Connock M, Johnson S, Clarke A and Taylor-Phillips S (2021) 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. 9:606194. doi: 10.3389/fped.2021.606194

Background: Long-chain 3-hydroxyacyl-CoA dehydrogenase (LCHAD) and mitochondrial trifunctional protein (MTP) deficiencies are rare autosomal recessive fatty acid β-oxidation disorders. Their clinical presentations are variable, and premature death is common. They are included in newborn blood spot screening programs in many countries around the world. The current process of screening, through the measurement of acylcarnitines (a metabolic by-product) in dried blood spots with tandem mass spectrometry, is subject to uncertainty regarding test accuracy.

Methods: We conducted a systematic review of literature published up to 19th June 2018. We included studies that investigated newborn screening for LCHAD or MTP deficiencies by tandem mass spectrometry of acylcarnitines in dried blood spots. The reference standards were urine organic acids, blood acylcarnitine profiles, enzyme analysis in cultured fibroblasts or lymphocytes, mutation analysis, or at least 10-year follow-up. The outcomes of interest were sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Assessment of titles, abstracts, and full-text papers and quality appraisal were carried out independently by two reviewers. One reviewer extracted study data. This was checked by a second reviewer.

Results: Ten studies provided data on test accuracy. LCHAD or MTP deficiencies were identified in 23 babies. No cases of LCHAD/MTP deficiencies were identified in four studies. PPV ranged from 0% (zero true positives and 28 false positives from 276,565 babies screened) to 100% (13 true positives and zero false positives from 2,037,824 babies screened). Sensitivity, specificity, and NPV could not be calculated as there was no systematic follow-up of babies who screened negative.

Conclusions: Test accuracy estimates of screening for LCHAD and MTP deficiencies with tandem mass spectrometry measurement of acylcarnitines in dried blood were variable in terms of PPVs. Screening methods (including markers and thresholds) varied between studies, and sensitivity, specificity, and NPVs are unknown.

Keywords: systematic review, test accuracy, LCHAD deficiency, MTP deficiency, newborn blood spot screening

INTRODUCTION

Long-chain 3-hydroxyacyl-CoA dehydrogenase (LCHAD) and mitochondrial trifunctional protein (MTP) deficiencies (Enzyme Commission Number 1.1.1.211) are recessive autosomal fatty acid β-oxidation disorders. They are caused by mutations in the genes coding for MTP. LCHAD deficiency arises as a result of mutations in the HADHA gene; MTP deficiency arises from mutations in HADHA and HADHB genes (1). The conditions are characterized by lethargy, hypoglycemia, hypotonia, cardiomyopathy, and acute metabolic crisis (2, 3). Long-term complications include liver disease, peripheral neuropathy, and retinopathy (3, 4). Signs and symptoms may present immediately after birth or later in life (5). Three main forms of LCHAD/MTP deficiencies have been reported: an early-onset form, which is associated with cardiomyopathy, hypoglycemia, and sudden infant death; an infant-onset form, which is characterized by recurrent hypoketotic hypoglycemia and lethargy during illness or fasting; and a milder, late-onset form that is triggered by exercise, fasting, or infections and is associated with progressive peripheral neuropathy and recurrent rhabdomyolysis (6, 7). There is no cure for LCHAD or MTP deficiencies, and premature death is common. Approximately 38% of infants die before, or within 3 months of, diagnosis (5). A number of management strategies are available, namely a high-carbohydrate and fat-modified/decreased diet that is low in long-chain fatty acids, supplements (L-carnitine, docosahexaenoic acid, and medium-chain triglyceride oil, such as triheptanoin), and avoidance of fasting (1). There is some evidence that these treatments are associated with improved clinical outcomes (e.g., reduced mortality, delayed visual complications), but the effects are variable, study sample sizes are small, and few data are available from long-term follow-up studies (1, 8–10). The incidence of LCHAD and MTP deficiencies varies widely around the world. A recent estimate from the USA gives an incidence of 1:363,738 for LCHAD deficiency and 1:1,240,467 for MTP deficiency (11).

It has been proposed that earlier recognition and treatment of LCHAD/MTP deficiencies may be critical for improving health outcomes (5), and the two conditions are included in newborn screening programs in many countries. Screening is conducted through the measurement of acylcarnitines (primarily C16OH, C16:1 OH, C18OH) in dried blood spots using tandem mass spectrometry (TMS). These markers are not specific to LCHAD/MTP deficiencies, however, and levels can be raised in other conditions (e.g., carnitine palmitoyltransferase II deficiency, very long-chain acyl-CoA dehydrogenase deficiency) and in babies of very low birth weight/being treated in neonatal intensive care (12). The results of a recent report from the UK suggest that screening for LCHAD/MTP deficiencies would not lead to the identification of additional cases as compared with the current practice of clinical detection (13). In contrast, data from other countries have suggested that screening does lead to earlier detection of LCHAD/MTP deficiencies (14, 15).

To date, there has been one systematic review examining test accuracy of screening for LCHAD/MTP deficiencies (16). Searching up to 2012, Einoder-Moreno et al. (16) identified six studies and concluded that sensitivity, specificity, and negative predictive value (NPV) of acylcarnitine measurement in dried blood spots are close to 100%, and that the positive predictive value (PPV) ranges from 9 to 100%. However, three relevant papers were missed by their search (12, 17, 18), and the calculation of sensitivity, specificity, and negative predictive value were based on an assumption about the disease status of babies who screened negative, as no follow-up of these babies was conducted in the included studies. This approach can lead to overestimation of sensitivity and underestimation of specificity (19). The aim of the current paper, therefore, is to conduct a systematic review of test accuracy metrics (sensitivity, specificity, positive and negative predictive values) of acylcarnitine measurement in newborn screening dried blood spots (DBS) for LCHAD/MTP deficiencies using tandem mass spectrometry using a broader search than the previous review and taking into consideration whether or not babies who screen negative received follow-up assessment.

MATERIALS AND METHODS

The review protocol is registered at PROSPERO (registration number CRD42018094356).

Search Strategy

We conducted a search of the following electronic databases: MEDLINE, MEDLINE In-Process, MEDLINE Daily, MEDLINE ePub Ahead of Print, the Cochrane Library, Web of Science, and Embase. Search terms (free text and subject headings)

Abbreviations: C14OH, 3-hydroxytetradecanoylcarnitine; C14:1, tetradecenoylcarnitine; C16OH, 3-hydroxypalmitoylcarnitine; C16OH/C16, 3-hydroxypalmitoylcarnitine/palmitoylcarnitine; C16:1OH, 3-hydroxypalmi toleylcarnitine; C18OH, 3-hydroxystearoylcarnitine; C18:1OH, 3-hydroxyoleoylcarnitine; C18:1, oleoylcarnitine; DBS, dried blood spot; FP, false positive; LCHAD, long-chain 3-hydroxyacyl-CoA dehydrogenase; MTP, mitochondrial trifunctional protein; NPV, negative predictive value; PPV, positive predictive value; QUADAS-2, Quality Assessment Tool for Diagnostic Accuracy Studies 2; TMS, tandem mass spectrometry; TP, true positive.

related to the disease area (e.g., "mitochondrial trifunctional protein," "long-chain-3-hydroxyacyl-CoA dehydrogenase," "fatty acid oxidation disorder") and screening (e.g., "newborn screening," "dried blood spot," "tandem mass spectrometry"). Full details of the search are provided in **Supplement 1**. The reference lists of included articles and relevant systematic reviews were also examined. The search was conducted on 19th June 2018, with no restrictions on the publication date or language of articles.

Eligibility Criteria

We included journal articles and reports that investigated newborn screening for LCHAD or MTP deficiencies by TMS analysis of acylcarnitines in dried blood spots. The reference standards were urine organic acids, blood acylcarnitine profiles, enzyme analysis in cultured fibroblasts or lymphocytes, mutation analysis, or at least 10-year follow-up. These could be on their own or in any combination. Appropriate study designs were cross-sectional test accuracy studies, case-control studies, and cohort studies. The outcomes of interest were sensitivity, specificity, PPV, and NPV (or sufficient data to allow us to calculate these). We excluded non-human studies, letters, editorials, communications, conference abstracts, gray literature, studies of fatty acid β-oxidation disorders where data for LCHAD/MTP deficiencies could not be separated from data for other conditions, studies with no extractable data, and studies where more than 10% of the study sample did not meet our inclusion criteria.

Screening and Data Extraction

Titles, abstracts, and full-text papers were independently screened by two reviewers. Data extraction was conducted by a single reviewer and checked by a second reviewer. At each stage of the review, disagreements were resolved through discussion between the reviewers, with the involvement of a third reviewer if consensus could not be achieved.

Quality Appraisal

Two reviewers independently assessed risk of study bias and applicability concerns using the Quality Assessment Tool for Diagnostic Accuracy Studies (QUADAS-2) (20), which was tailored to the research question. Tailoring comprised defining cut-offs for exclusions, identifying appropriate reference standards, selecting a suitable interval between index tests and reference standards, and producing guidance notes. Disagreements were resolved through discussion by the two reviewers, leading to a consensus on assessment of risk of bias and applicability concerns for all studies. The QUADAS-2 tool is presented in **Supplement 2** and the guidance notes in **Supplement 3**.

Data Summary and Synthesis

Due to incomplete 2 × 2 tables and heterogeneity between study designs, a narrative summary of the evidence is provided. We calculated confidence intervals for test accuracy metrics using the Wilson score method with continuity correction (21).

RESULTS

Searching, Sifting, and Sorting

Full details of the flow of studies through the review are outlined in **Figure 1**. One thousand one hundred and ninety-four unique records were identified through searching electronic databases. After examination of titles and abstracts, 39 papers were retained for full-text assessment. Eleven of these papers met the review's inclusion criteria (12–15, 17, 18, 22–26). Two papers included overlapping cohorts (17, 22). Only the data from the larger, more recent paper by Lindner et al. (17) [which included all of the data from Schulze et al. (22)] are reported here. A list of excluded studies [with reasons for exclusion] is provided in **Supplement 4**.

Quality Appraisal

A summary of the risks of bias and applicability concerns of the included papers is provided in **Figure 2**. Ratings of risks of biases and applicability concerns for each individual study are provided in **Supplement 5**.

Risk of bias was considered to be high in two or more domains for nine of the 10 studies (90%) (12–14, 17, 18, 23–26) and in one domain for the remaining study (10%) (15). In the patient selection domain, risk of bias was rated as unclear in seven (70%) studies due to unclear/incomplete reporting (12, 15, 17, 23–26), and low in three (30%) studies (13, 14, 18). There were considerable concerns regarding applicability to the UK context in seven (70%) studies as blood samples were taken before day 5 in six studies (12, 15, 17, 18, 23, 25), and the incidence of MTPD was lower than expected in one study (1:300,000 compared with 1:149,254 in the UK) (24). Applicability concerns were unclear in two studies (14, 26) and low in one study (13).

In the index test domain, five studies (50%) were rated as having high risk of bias as the cut-off for "screen positive" was altered during the study period (13) or was not pre-specified (12, 23, 24, 26). Of the remaining studies, two were at unclear risk of bias (18, 25) and three were at low risk of bias (14, 15, 17). Two (20%) studies had high applicability concerns as one included additional markers (C14:1, C14-OH) (15) and one included both blood and urine samples (18). Applicability concerns were unclear in two studies (17, 25) and low in six studies (12–14, 23, 24, 26).

In the reference standard domain, risk of bias was rated as unclear in all 10 studies because it was not possible to tell the reference standard results were interpreted without knowledge of the results of the index test (12–15, 17, 18, 23–26). Applicability concerns were unclear in six (60%) studies (12, 15, 23–26) and low in the remaining four (40%) (13, 14, 17, 18).

Finally, all studies were judged to be at high risk of bias in the flow and timing domain (12–15, 17, 18, 23–26). The reasons for this were that the reference standards used to confirm disease status for screen positives and screen negatives were not the same, follow-up of those children who screened negative was not defined or not conducted, and losses to follow-up were not reported.

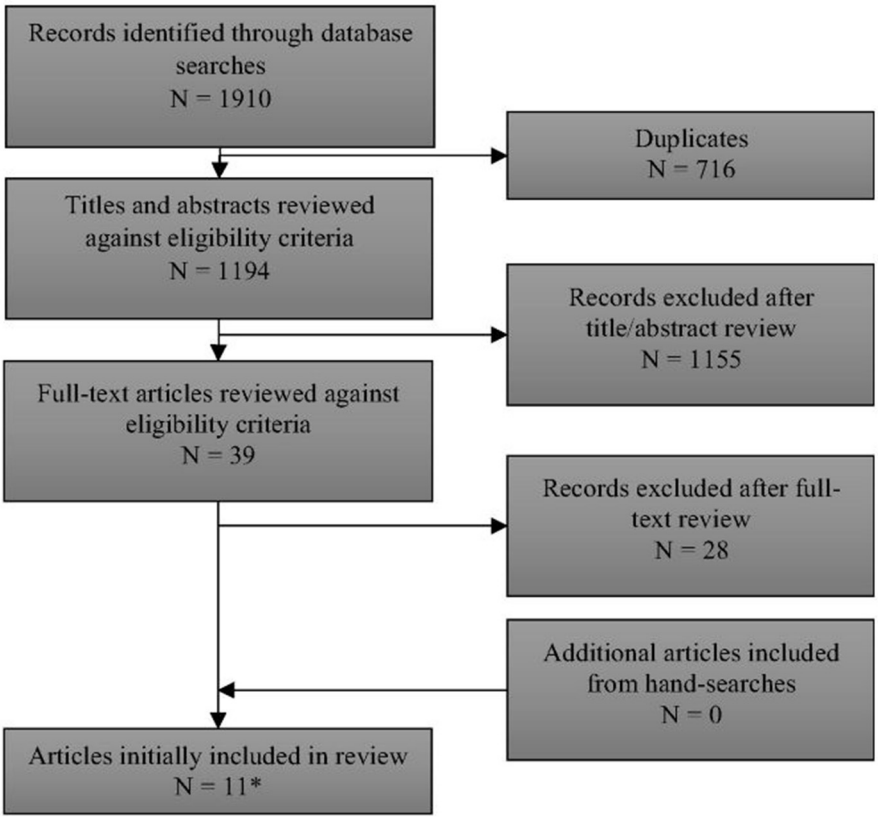

FIGURE 1 | PRISMA flow diagram of records through the systematic review. *See **Supplement 4** for list of excluded studies with reasons.

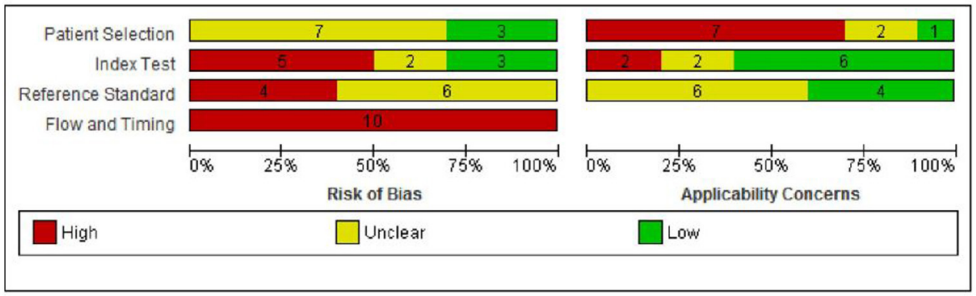

FIGURE 2 | Risk of bias and applicability concern graph: review authors' judgments about each domain presented as percentages across included studies.

Characteristics of Included Studies

Details of the included studies are provided in **Table 1** and **Supplement 6**. Ten studies were included, four pilot programs (12, 13, 23, 25), three mixed pilot and national/regional screening programs (14, 17, 24), and three national/regional screening programs (15, 18, 26). Six studies took place in Europe [Denmark, Faroe Islands, and Greenland (14), Germany (15, 17), Slovenia (23), Spain (18), UK (13)], two in Asia (China (26), Hong Kong (25)), and two in North America [USA (12, 24)]. Sample sizes ranged from 2,440 (25) to 1,200,000 (15). In total, the 10 studies screened 3,951,358 newborns. Twenty-three true positives (TP) and 40 false positives (FP) were identified. The 23 true positives comprised 11 babies with LCHAD deficiency, two with MTP deficiency, and 10 babies for whom LCHAD and MTP deficiencies were not differentiated.

Description of Screening and Diagnostic Tests

Details of the screening methodology and diagnostic tests used in each study are provided in **Supplement 6**.

TABLE 1 | Accuracy of newborn screening tests for LCHADD/MTPD using TMS measurement of acylcarnitines.

References	Number screened	Marker(s)	Cut-off(s) (μmol/L)	2 × 2 table				Test accuracy[a]			
				TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)*	NPV (95% CI)
Couce et al. (18)	210,165	C16OH, C18:1OH, C18OH	NR	2[b]	0	NR	NR	NA	NA	1 (0.20, 1)	NA
Frazier et al. (24)	239,415	C16OH, C18:1, C18:1OH	>0.18, >4.08, >0.14	2[b]	0	NR	NR	NA	NA	1 (0.20, 1)	NA
Lindner et al. (17)	1,084,195	C14OH (and/or): C16:1OH; C16OH; C18:1OH; C18OH	>0.12, >0.22, >0.20, >0.12, >0.11	6 (5 NBS, 1 cascade testing)[b]	0	NR	NR	NA	NA	1 (0.52, 1)	NA
Lund et al. (14)	504,049	Primary: C16OH; secondary: C18:1OH	>0.12U, >0.1U	3[c]	0	NR	NR	NA	NA	1 (0.31, 1)	NA
Mak et al. (25)	2,440	Unclear	Unclear	0	2	NR	NR	NA	NA	0 (0, 0.80)	NA
Sander et al. (15)	1,200,000	C16OH, C18:1OH, C14:1, C14OH	>0.08, >0.06, >0.35, >0.2	9 (7 LCHADD, 2 MTPD)	10	NR	NR	NA	NA	0.47 (0.25, 0.71)	NA
Smon et al. (23)	10,048	C16:1OH, C16OH, C18:1OH, C18 OH, C16OH/C16	0.042[d], 0.028[d], 0.014[d], 0.009[d], 0.015[d]	0	8	NR	NR	NA	NA	0 (0, 0.40)	NA
UK NSC (13)	436,969	Primary: C16OH; secondary: C16-OH; C16:1-OH; C18-OH	>0.12, >0.15–lowered from 0.2	1[c]	2	NR	NR	NA	NA	0.33 (0.02, 0.87)	NA
Yang et al. (26)	100,077	LCHADD: C16OH; MTPD: C18OH, C18:1OH	>0.04, >0.03, >0.05	0	13 (4 LCHADD, 9 MTPD)	NR	NR	NA	NA	0 (0, 0.28)	NA
Zytkovicz et al. (12)	164,000	C16OH/d-C16	0.1[e]	0	5	NR	NR	NA	NA	0 (0, 0.54)	NA

*[a]Confidence intervals calculated using Wilson score method with continuity correction. [b]LCHAD deficiency. [c]LCHAD/MTP deficiencies. [d]Data provided by study authors. [e]Determined during the study to ensure that no more than 0.02% of population would be flagged. CI, confidence interval; FN, false negative; FP, false positive; LCHADD, long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency; MTPD, mitochondrial trifunctional protein deficiency; NA, not applicable; NBS, newborn bloodspot; NPV, negative predictive value; NR, not reported; PPV, positive predictive value; TN, true negative; TP, true positive; UK NSC, UK National Screening Committee. *The values were calculated by the review authors.*

In brief, eight different analytes or ratios were used as markers to screen for LCHAD/MTP deficiencies: 3-hydroxytetradecanoylcarnitine (C14OH), tetradecenoylcarnitine (C14:1), 3-hydroxypalmitoylcarnitine (C16OH), 3-hydroxypalmitoleylcarnitine (C16:1OH), 3-hydroxystearoylcarnitine (C18OH), 3-hydroxyoleoylcarnitine (C18:1OH), oleoylcarnitine (C18:1), and 3-hydroxypalmitoylcarnitine/palmitoylcarnitine (C16OH/C16), with each study employing a unique combination. The cut-offs used varied between studies (e.g., cut-offs ranged from >0.04 to >0.20 μmol/L for C16OH, >0.03 to >0.15 μmol/L for C18OH, and >0.05 to >0.14 μmol/L for C18:1OH), with no two studies using the same combination of markers and thresholds. Screening samples were collected between 24 h and day 37 of life (25, 26).

The reference standards used varied between and within studies. For screen-positive babies, reference standards were blood acylcarnitines, urinary organic acids, and DNA analysis

(13, 24, 25); enzyme and/or molecular studies (18); enzyme activity in fibroblasts/lymphocytes and mutation analysis (15); acylcarnitine profile in plasma/DBS and/or genotype and/or enzyme activity (17); urine organic acids, plasma acylcarnitines, and molecular-genetic analyses (14); organic acid in urine, next-generation sequencing, and an additional acylcarnitine profile in DBS (23); and urinary organic acids or DNA analysis (26). Lastly, one study used "standard metabolic criteria" (12). No systematic follow-up of babies who screened negative was conducted in any of the studies.

Accuracy of Screening Tests

The cut-offs used to classify a positive case of LCHAD/MTP deficiencies and the diagnostic tests used to confirm this varied between studies. Therefore, we report positive screening results as those that met/exceeded the cut-off and were diagnostically confirmed as presented in the individual study. Test accuracy data are show in **Table 1**.

Positive and Negative Predictive Values

PPV varied considerably between studies. It was 0% in four studies, with zero true positives and 28 false positives from 276,565 babies screened (12, 23, 25, 26), 33% in one study, with one true positive and two false positives from 436,969 babies screened (13), 47% in one study, with 9 true positives and 10 false positives from 1,200,000 babies screened (15), and 100% in four studies, with 13 true positives from 2,037,824 babies screened (14, 17, 18, 24). In the UK study, the single case reported as a true positive was being treated for LCHADD at the point of screening, as they had already been detected clinically (13). Confidence intervals were wide due to the small number of cases of LCHAD/MTP deficiencies detected (23 in total, zero to nine per study). It was not possible to calculate NPV as newborns who screened negative were not systematically followed up.

Sensitivity and Specificity

We were not able to determine sensitivity or specificity due to a lack of information on babies who screened negative.

DISCUSSION

We assessed the test accuracy of acylcarnitine measurement in newborn DBS using TMS for LCHAD/MTP deficiencies. Ten relevant studies were identified. All studies had a high risk of bias in at least one domain, and 9/10 (90%) studies had a high risk of bias in at least two domains. Across the 10 studies, ∼4,000,000 babies were screened and 23 cases of LCHAD/MTP deficiencies were identified; 11 babies had LCHAD deficiency, two had MTP deficiency, and 10 had undifferentiated LCHAD/MTP deficiencies. One of the cases reported as a true positive had already been detected clinically at the point at which screening took place (13). Arguably, the PPV for this study should be 0% rather than 33%, as reported in the study. Forty additional babies screened positive but were subsequently found not to have LCHAD or MTP deficiency. In four studies, no cases of LCHAD/MTP deficiencies were identified (12, 23, 25, 26). However, in three of these studies, the sample sizes were too small to be likely to detect such rare diseases [screening population sizes were 2,440 (25), 10,048 (23), and 100,077 (26)]. The fourth study included a larger sample ($n = 164{,}000$) but only included one marker (C16OH), which might have made the screening process less accurate (12).

The only measure of test accuracy that was consistently reported (or where sufficient data were present to allow us to calculate it) was PPV. PPV in the 10 studies ranged from 0% (zero 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 NPV as there was no systematic follow-up of babies who had screened negative. In a pilot or national screening program for a rare disease using a "promising" test, negative tests will inevitably represent the vast majority of test results. While some studies provided very high PPV, PPV is not intrinsic to the test itself, and at any particular values of sensitivity and specificity, the estimates of PPV (and also NPV) are strongly dependent on disease prevalence. This relationship is illustrated in **Supplement 7** over a range of prevalence values similar to those in the included studies and over a range of specificity values. In order to provide a complete assessment of test accuracy, all four metrics (sensitivity, specificity, NPV, PPV) are required.

Whether newborn screening for LCHAD/MTP deficiencies with acylcarnitines measurement in dried blood spots using TMS is appropriate is currently unclear due to a lack of data on babies who screen negative and a lack of consistency between screening test methods. Partial verification bias is a key issue in the included studies; from nearly 4,000,000 babies screened, only 63 (those who screened positive) received a reference standard. Therefore, we cannot know the true disease state of the babies who screened negative. Partial verification bias is common in studies of test accuracy because it is often impractical, unethical, and not cost-effective to follow-up every participant. Alternative approaches to whole population follow-up include statistical methods to attempt to correct for the bias, follow-up of samples of participants who screen negative, and searching disease registers to find false negatives. Statistical methods may introduce other forms of bias (27, 28).

There were substantial differences between screening test methods. For example, there were differences in the extraction and calibration methods (in-house or commercially available test kits); analysis as acylcarnitine butyl esters or free acids (underivatised); screening markers used (e.g., C14OH, C14:1, C16OH, C16:1OH, C18OH, C18:1, C18:1OH, C16OH/C16 ratio); whether markers were employed in isolation or in combination with each other [two studies did not report which marker(s) were used (18, 25)] and variability in the cut-offs between studies, with cut-offs not specified in three studies (18, 23, 25). The majority of FP were found in studies which used lower thresholds for C16OH (>0.03 to >0.10 μmol/L), C16:1OH (>0.04 to 0.15 μmol/L), and C18:1OH (>0.01 to 0.06 μmol/L).

While screening for LCHADD/MTP is conducted in the newborn bloodspot programs of a number of countries, there is little published data on the benefits and harms of these. Taylor-Phillips et al. (29) reviewed the evidence on national

policy recommendations on screening newborn babies for rare diseases. They highlight three elements that might determine the balance of benefits and harms from screening programs: test accuracy, the benefit of early detection and treatment, and overdiagnosis (the detection and subsequent treatment of disease that would never have caused symptoms within a person's lifetime). Many of the national policy recommendations (including for LCHADD) did not assess all of these three elements. In relation to screening for LCHADD/MTP, the current review suggests that the evidence on test accuracy is uncertain. A recent systematic review has examined the potential benefit of early detection and treatment, comparing the health outcomes of people with LCHADD/MTP who were treated with pre-symptomatic dietary management following screen detection of the conditions compared with people detected following symptomatic presentation (30). There was some evidence of an association between timing of intervention and outcomes, such as mortality, heart problems, liver problems, visual problems, motor/muscular problems, and hypoglycemia. However, the majority of included studies found no statistically significant differences in outcomes between the two groups. Furthermore, the review identified few studies from which to draw conclusions and high risks of bias in included studies. There is no published evidence on overdiagnosis. Overall, the paucity of data and variability between studies lead to considerable uncertainty regarding the benefits and harms of screening for LCHADD/MTP.

Our review has a number of limitations. First, we were not able to synthesize (meta-analyze) our results numerically due to a lack of data on FN and TN and because of variability between screening test methods. Second, we tailored the QUADAS-2 to reflect newborn screening in the UK; this resulted in high concerns regarding applicability in the patient selection domain, as screening is often conducted sooner after birth in other countries. The definition of a high applicability concern is likely to differ in other countries.

There is currently insufficient evidence to clearly judge test accuracy. This is driven, in part, by a wide range of markers and thresholds being used in the included studies: PPV estimates differed greatly by study, with some suggesting good PPV, albeit on small numbers of cases. It was not possible for us to combine data from different studies or determine which combination of markers and thresholds may yield good accuracy as results were not presented by marker. Future research could involve collaboration between researchers to report scores on a range of relevant markers for cases of LCHAD, cases of MTP, and in the general population using consistent units. There is a precedent for this approach in the form of the Region 4 Stork (R4S) project and accompanying multivariate pattern recognition software (subsequently developed into the interactive web tool Collaborative Laboratory Integrated Reports (CLIR), https://clir.mayo.edu/). The R4S project aimed "(a) to achieve uniformity of testing panels by MS/MS to maximize detection of affected newborns within the region; (b) to improve overall analytical performance; and (c) to set and sustain the lowest achievable rates of false positive and false negative results" (31). Reference and disease ranges for LCHAD/MTP markers were reported for the R4S project (31). To date, the CLIR tool has been used in a small number of research projects (32–35). An additional piece of work should aim to clarify the disease states of babies who screen negative. This could be achieved in a number of ways, such as searching hospital/primary care records or disease registers, or following up samples of babies who have screened negative.

CONCLUSIONS

Measurement of acylcarnitines in newborn dried blood spots using TMS may prove to be a useful way to screen for LCHAD/MTP deficiencies, but currently, there are significant concerns regarding the high number of false positives in some of the studies, risks of bias in the studies, heterogeneity in the methods used, and a lack of data on sensitivity, specificity, or negative predictive values. Clinicians interested in the identification of LCHAD/MTP may consider partnership development across clinical and research networks to address the knowledge gaps identified from this study, including data available for long-term follow-up studies and alignment of diagnostic methodologies.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

CS undertook project planning and research design, coordinated the review process, conducted all aspects of the review, and co-wrote the paper with HF. HF undertook project planning, undertook all sifting, sorting, data extraction, and quality appraisal, and co-wrote the paper with CS. JG contributed to sifting and sorting and commented on draft and final versions of the paper. RJ contributed to sifting and sorting and quality appraisal and commented on first and final drafts of the paper. MC undertook project planning, and commented on first and final drafts of the paper. SJ developed and conducted the literature searches, managed references, and helped obtain full-text references. AC undertook project planning, oversight of search strategies and methods, and commented on first and final drafts of the paper. ST-P undertook project planning and research design and commented on first and final drafts of the paper. All members of the team contributed to the development of the protocol. All authors contributed to the article and approved the submitted version.

FUNDING

This research was commissioned by the UK National Screening Committee.

ACKNOWLEDGMENTS

This review was funded by the UK National Screening Committee. This research was supported by the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West Midlands (NIHR CLAHRC WM), now recommissioned as NIHR Applied Research Collaboration West Midlands. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fped.2021.606194/full#supplementary-material

REFERENCES

1. Spiekerkoetter U, Lindner M, Santer R, Grotzke M, Baumgartner M R, Boehles H, et al. Management and outcome in 75 individuals with long-chain fatty acid oxidation defects: results from a workshop. *J Inherit Metab Dis*. (2009) 32:488–97. doi: 10.1007/s10545-009-1125-9
2. Baruteau J, Sachs P, Broué P, Brivet M, Abdoul H, Vianey-Saban C, et al. Clinical and biological features at diagnosis in mitochondrial fatty acid beta-oxidation defects: a French pediatric study from 187 patients. Complementary data. *J Inherit Metab Dis*. (2014) 37:137–9. doi: 10.1007/s10545-013-9628-9
3. Immonen T, Turanlahti M, Paganus A, 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) 105:549–54. doi: 10.1111/apa.13313
4. Karall D, Brunner-Krainz M, Kogelnig K, Konstantopoulou V, Maier EM, Möslinger D, et al. 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) 10:21. doi: 10.1186/s13023-015-0236-7
5. den Boer ME, Wanders RJ, Morris AA, IJlst L, Heymans HS, Wijburg FA. Long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency: clinical presentation and follow-up of 50 patients. *Pediatrics*. (2002) 109:99–104. doi: 10.1542/peds.109.1.99
6. Spiekerkoetter U, Sun B, Khuchua Z, Bennett MJ, Strauss AW. Molecular and phenotypic heterogeneity in mitochondrial trifunctional protein deficiency due to beta-subunit mutations. *Hum Mutat*. (2003) 21:598–607. doi: 10.1002/humu.10211
7. Spiekerkoetter U, Khuchua Z, Yue Z, Bennett MJ, Strauss AW. General mitochondrial trifunctional protein (TFP) deficiency as a result of either alpha- or beta-subunit mutations exhibits similar phenotypes because mutations in either subunit alter TFP complex expression and subunit turnover. *Pediatr Res*. (2004) 55:190–6. doi: 10.1203/01.PDR.0000103931.80055.06
8. 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
9. Roe CR, Brunengraber H. Anaplerotic treatment of long-chain fat oxidation disorders with triheptanoin: review of 15 years experience. *Mol Genet Metab*. (2015) 116:260–8. doi: 10.1016/j.ymgme.2015.10.005
10. Fahnehjelm KT, Holmström G, Ying L, Haglind CB, Nordenström A, Halldin M, et al. Ocular characteristics in 10 children with long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency: a cross-sectional study with long-term follow-up. *Acta Ophthalmol*. (2008) 86:329–37. doi: 10.1111/j.1600-0420.2007.01121.x
11. Therrell BL Jr, Lloyd-Puryear MA, Camp KM, Mann MY. Inborn errors of metabolism identified via newborn screening: ten-year incidence data and costs of nutritional interventions for research agenda planning. *Mol Genet Metab*. (2014) 113:14–26. doi: 10.1016/j.ymgme.2014.07.009
12. Zytkovicz TH, Fitzgerald EF, Marsden D, Larson CA, Shih VE, Johnson DM, et al. Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two-year summary from the New England Newborn Screening Program. *Clin Chem*. (2001) 47:1945–55. doi: 10.1093/clinchem/47.11.1945
13. UK National Screening Committee. *Screening for Mitochondrial Trifunctional Protein Disorders, Including Long-Chain 3-hydroxyacyl-CoA Dehydrogenase Deficiency: External Review Against Programme Appraisal Criteria for the UK National Screening Committee*. (2019). Available online at: https://legacyscreening.phe.org.uk/lchadd
14. Lund AM, Hougaard DM, Simonsen H, Andresen BS, Christensen M, Dunø M, 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) 107:281–93. doi: 10.1016/j.ymgme.2012.06.006
15. Sander J, Sander S, Steuerwald U, Janzen N, Peter M, Wanders RJ, et al. Neonatal screening for defects of the mitochondrial trifunctional protein. *Mol Genet Metab*. (2005) 85:108–14. doi: 10.1016/j.ymgme.2005.02.002
16. Einöder-Moreno M, Atienza Merino G. *Efectividad clínica del cribado neonatal de errores congénitos del metabolismo mediante espectrometría de masas en tándem. Parte I: enfermedad de la orina con olor a jarabe de arce, homocistinuria, acidemia glutárica tipo I, acidemia isovalérica y deficiencia de 3-hidroxi-acil-CoA deshidrogenasa de cadena larga. [Clinical Effectiveness of Newborn Screening for Inborn Errors of Metabolism Using Mass Spectrometry: Maple Syrup Urine Disease, Homocystinuria, Glutaric Aciduria Type I, Isovaleric Acidaemia, Long-chain 3-Hydroxyacyl CoA Dehydrogenase Deficiency]*. Santiago de Compostela: Galician Agency for Health Technology Assessment (AVALIA-T) (2013).
17. Lindner M, Gramer G, Haege G, Fang-Hoffmann J, Schwab KO, Tacke U, et al. Efficacy and outcome of expanded newborn screening for metabolic diseases–report of 10 years from South-West Germany. *Orphanet J Rare Dis*. (2011) 6:44. doi: 10.1186/1750-1172-6-44
18. Couce ML, Castiñeiras DE, Bóveda MD, Baña A, Cocho JA, Iglesias AJ, et al. Evaluation and long-term follow-up of infants with inborn errors of metabolism identified in an expanded screening programme. *Mol Genet Metab*. (2011) 104:470–5. doi: 10.1016/j.ymgme.2011.09.021
19. Whiting PF, Rutjes AW, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. *J Clin Epidemiol*. (2013) 66:1093–104. doi: 10.1016/j.jclinepi.2013.05.014
20. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. (2011) 155:529–36. doi: 10.7326/0003-4819-155-8-201110180-00009
21. Wilson EB. Probable Inference, the law of succession, and statistical inference. *J Am Stat Assoc*. (1927) 22:209–12. doi: 10.1080/01621459.1927.10502953
22. Schulze A, Lindner M, Kohlmüller D, Olgemöller K, Mayatepek E, Hoffmann GF. Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatrics*. (2003) 111(6 Pt 1):1399–406. doi: 10.1542/peds.111.6.1399
23. Smon A, Repic Lampret B, Groselj U, Zerjav Tansek M, Kovac J, Perko D, et al. Next generation sequencing as a follow-up test in an expanded newborn screening programme. *Clin Biochem*. (2018) 52:48–55. doi: 10.1016/j.clinbiochem.2017.10.016
24. Frazier DM, Millington DS, McCandless SE, Koeberl DD, Weavil SD, Chaing SH, et al. The tandem mass spectrometry newborn screening experience in North Carolina: 1997-2005. *J Inherit Metab Dis*. (2006) 29:76–85. doi: 10.1007/s10545-006-0228-9

25. Mak CM, Law EC, Lee HH, Siu WK, Chow KM, Au Yeung SK, et al. The first pilot study of expanded newborn screening for inborn errors of metabolism and survey of related knowledge and opinions of health care professionals in Hong Kong. *Hong Kong Med J*. (2018) 24:226–37. doi: 10.12809/hkmj176939
26. Yang CJ, Wei N, Li M, Xie K, Li JQ, Huang CG, et al. Diagnosis and therapeutic monitoring of inborn errors of metabolism in 100,077 newborns from Jining city in China. *BMC Pediatr*. (2018) 18:110. doi: 10.1186/s12887-018-1090-2
27. de Groot JA, Janssen KJ, Zwinderman AH, Bossuyt PM, Reitsma JB, Moons KG. Correcting for partial verification bias: a comparison of methods. *Ann Epidemiol*. (2011) 21:139–48. doi: 10.1016/j.annepidem.2010.10.004
28. Irwig L, Glasziou PP, Berry G, Chock C, Mock P, Simpson JM. Efficient study designs to assess the accuracy of screening tests. *Am J Epidemiol*. (1994) 140:759–69. doi: 10.1093/oxfordjournals.aje.a117323
29. Taylor-Phillips S, Stinton C, Ferrante di Ruffano L, Seedat F, Clarke A, Deeks JJ. Association between use of systematic reviews and national policy recommendations on screening newborn babies for rare diseases: systematic review and meta-analysis. *BMJ*. (2018) 361:k1612. doi: 10.1136/bmj.k1612
30. Fraser H, Geppert J, Johnson R, Johnson S, Connock M, Clarke A, et al. 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) 14:258. doi: 10.1186/s13023-019-1226-y
31. McHugh D, Cameron CA, Abdenur JE, Abdulrahman M, Adair O, Al Nuaimi SA, et al. Clinical validation of cutoff target ranges in newborn screening of metabolic disorders by tandem mass spectrometry: a worldwide collaborative project. *Genet Med*. (2011) 13:230–54. doi: 10.1097/GIM.0b013e31820d5e67
32. Keller R, Chrastina P, Pavlíková M, Gouveia S, Ribes A, Kölker S, et al. Newborn screening for homocystinurias: Recent recommendations versus current practice. *J Inherit Metab Dis*. (2019) 42:128–39. doi: 10.1002/jimd.12034
33. Tortorelli S, Eckerman JS, Orsini JJ, Stevens C, Hart J, Hall PL, et al. Moonlighting newborn screening markers: the incidental discovery of a second-tier test for Pompe disease. *Genet Med*. (2018) 20:840–46. doi: 10.1038/gim.2017.190
34. Minter Baerg MM, Stoway SD, Hart J, Mott L, Peck DS, Nett SL, et al. Precision newborn screening for lysosomal disorders. *Genet Med*. (2018) 20:847–54.
35. Mørkrid L, Rowe AD, Elgstoen KB, Olesen J H, Ruijter G, Hall P L, et al. Continuous age- and sex-adjusted reference intervals of urinary markers for cerebral creatine deficiency syndromes: a novel approach to the definition of reference intervals. *Clin Chem*. (2015) 61:760–8. doi: 10.1373/clinchem.2014.235564

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Stinton, Fraser, Geppert, Johnson, Connock, Johnson, Clarke and Taylor-Phillips. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.