



**UK National
Screening Committee**

Newborn screening for propionic acidaemia

External review against programme appraisal criteria
for the UK National Screening Committee (UK NSC)

Version: 1

Bazian Ltd.

May 2015

This analysis has been produced by Bazian Ltd for the UK National Screening Committee. Bazian Ltd has taken care in the preparation of this report, but makes no warranty as to its accuracy and will not be liable to any person relying on or using it for any purpose.

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. Current policies can be found in the policy database at <http://www.screening.nhs.uk/policies> and the policy review process is described in detail at <http://www.screening.nhs.uk/policyreview>

Template v1.2, June 2010

Abbreviations List

Abbreviation	Meaning
C0	Free carnitine
C2	Acetylcarnitine (Number indicates the number of carbon atoms in acyl chain attached to carnitine)
C3	Propionylcarnitine
C4	Butyrylcarnitine
C16	Hexadecanoylcarnitine
C4DC	Methylmalonyl/succinyl-carnitine
CblA	Cobalamin deficiency type A. Form of MMA caused by mutation in the MMAA gene
CblB	Cobalamin deficiency type B. Form of MMA caused by mutation in the MMAB gene
CblC	Cobalamin deficiency type C. Form of MMA caused by mutation in the MMACHC gene
CblD	Cobalamin deficiency type D. Form of MMA caused by mutation in the MMADHC gene
Gly	Glycine
HCSD	Holocarboxylase synthetase deficiency
Met	Methionine
MCD	Multiple carboxylase deficiency (includes HCSD and biotinidase deficiency)
MMA	Methylmalonic acidaemia/aciduria
MMAA	Methylmalonic aciduria (cobalamin deficiency) cblA type gene
MMAB	Methylmalonic aciduria (cobalamin deficiency) cblB type gene
MMACHC	Methylmalonic aciduria (cobalamin deficiency) cblC type with homocysteinuria gene
MMADHC	Methylmalonic aciduria (cobalamin deficiency) cblD type with homocysteinuria gene
MUT	Methylmalonyl CoA mutase gene
PA	Propionic acidaemia/aciduria
R4S	Region 4 Stork
RC	Reviewer calculated figure (i.e. not presented in original publication)
SD	Standard deviation
MS/MS	Tandem mass spectrometry

Plain English Summary

Propionic acidaemia (PA) is an organic-acid oxidation disorder. It is a rare condition passed on to those affected through altered genes from both mother and father. The condition is caused by a lack of or faulty chemical that is needed to help break down larger substances needed by the body. These include certain amino acid (which are the building blocks of protein) and types of fatty substances. These disorders lead to a build-up of harmful chemicals. The build-up of chemicals can cause damage to vital parts of the body, such as the brain, heart, liver and nervous system. The severity of symptoms and the age that they develop varies.

Newborns affected by forms of the condition early in life often show symptoms that include poor feeding, vomiting, loss of appetite, weak muscle tone (hypotonia), and lack of energy (lethargy). In more severely affected newborns PA can cause coma and death. Intellectual disability and delayed development can often develop in people who are affected by forms of the condition developing later in life.

Day to day treatment of PA is mainly through changes to the diet. These changes are through eating less of certain types of proteins and by adding more of another substance (L-carnitine). For some people treatment helps prevent the harmful effects of the condition. Others can still develop serious problems.

It has been suggested that newborn screening for PA would be beneficial. This is due to early treatment potentially preventing death and serious damage to vital parts of the body. The most recent review of PA in 2009 recommended against screening due to many uncertainties. This review searched for evidence between 2001 and January 2015. The focus was on some of the areas in the 2009 review that required further evidence or were unmet.

Separate reports have been prepared for MMA and PA. This report covers PA.

The review found that:

- there is no clear evidence of a correlation between genotype and phenotype or enzyme activity and phenotype in PA. Current evidence cannot help accurately predict the likely course of the condition following screening. This would help decide on treatment for affected individuals who may not have presented with clinical symptoms.

- the current screening test for PA picks up other conditions that are not PA. This would mean screening for PA would not be possible without also screening for other conditions. These other conditions may not be suitable for screening.
- the quality of evidence from studies is not high enough to understand whether screening is of long term clinical benefit. This is a view supported by the current European guidelines.
- it is not clear from guidelines how to treat newborns with no medical symptoms that are picked up at screen. This could result in overtreatment of mild cases of the disease that are not the target of screening.
- there is a lack of evidence exploring the impact of the wider ethical, legal and social issues associated with screening. PA is passed on through both parent's genes. Screening would also provide information relevant to the parents but the impact upon them has not been explored in literature.

Recommendation

The evidence considered suggests that the existing recommendation not to screen for Methylmalonic acidaemia and Propionic acidaemia in newborns should be retained.

Executive Summary

The Condition

Propionic acidaemia or aciduria (PA, also known as ketotic hyperglycinemia)¹ is an autosomal recessive disorder caused by mutations in the genes which encode the two subunits the propionyl-CoA carboxylase (PCC) enzyme. The PCC enzyme converts propionyl-coenzyme A (CoA) to methylmalonyl-CoA. This reaction is one step in the breakdown of certain amino acids, as well as lipids and cholesterol. Mutations in either subunit of the gene reduce enzyme function and lead to build up of propionyl-CoA and other harmful compounds.

The age of onset, type and severity of symptoms is variable. Newborns with early onset forms of the condition often present with symptoms including: poor feeding, vomiting, loss of appetite, weak muscle tone (hypotonia), and lack of energy (lethargy). These are signs of metabolic decompensation, which can cause coma and death. Later onset forms of the condition can lead to intellectual disability and delayed development. The condition can lead to damage to vital organs, such as the brain, heart, liver and nervous system. Even with treatment outcome in some cases is poor.

Treatment

Ongoing treatment of PA is mainly through dietary restriction and supplementation. A low protein diet with the avoidance of certain types of amino acids, such as methionine, threonine, valine, and isoleucine and odd chain fatty-acids is required, in addition to supplementation with L-carnitine. Rapid treatment of acute metabolic crises when they do occur also forms an important part of management.

Screening

It has been suggested that screening all newborns for PA would be beneficial due to early treatment potentially preventing death and reducing or preventing some of the serious complications like brain, heart, liver and nervous system damage. The initial screening test involves identifying raised propionylcarnitine (C3) on newborn dried blood spots using tandem mass spectrometry (MS/MS).

Previous/ Current UK NSC Review

The most recent UKNSC review of the policy on newborn screening for the organic acid disorders PA and a similar condition, methylmalonic acidaemia (MMA), was conducted in 2009. This used evidence based on HTA reviews from 1997 and 2004 on screening for inborn errors of metabolism and recommended against screening due to a number of uncertainties. Bazian Ltd were commissioned to undertake this current review of PA and MMA, which takes in literature between 2001 and January 2015 and focusses on some of the uncertainties raised in the last review and other recent UK NSC reviews of inborn errors of metabolism.

Separate reports have been prepared for PA and MMA. This report covers PA.

Findings

The review found that: -

- There is no clear evidence of a correlation between genotype and phenotype or enzyme activity and phenotype in PA. Current evidence is insufficient to accurately predict prognosis from screening and therefore decide on treatment for affected individuals who may not have presented with clinical symptoms.
- The current screening test has a poor predictive value (11% and lower) and the initial screen cannot distinguish between PA and MMA due to them utilising the same markers (mainly C3 and ratios involving C3).
- No large, robust studies allowing for comparison of treatment outcomes from screened and unscreened populations were identified. Current European guidelines support the conclusion that available data has not yet determined whether newborn screening for PA or MMA are of long term clinical benefit.
- The identified treatment guidelines did not give explicit recommendations about management of asymptomatic individuals identified through screening, or use specific genotype or level of PCC enzyme activity to guide management. This, and a lack of an accurate prediction of prognosis, could potentially result in treatment of asymptomatic, mild cases of the disease following screening that is aimed at more severe cases.
- Despite parents of the affected newborns being, by default, carriers of the mutation, no studies were identified in the update search which explored the implications of the carrier state identified as a result of screening. Additionally, no direct evidence was identified that explored the impact of newborn bloodspot screening for PA on wider ethical, legal, or social issues.

Given the rarity of PA, it is likely a prospectively constructed international study or registry would be needed to gain sufficient evidence to compare the impact of treatment following screening versus treatment following clinical detection.

Recommendation

The existing recommendation not to screen for Propionic acidaemia and Methylmalonic acidaemia in newborns should be retained.

Introduction

Propionic acidaemia

Propionic acidaemia or aciduria (PA, also known as ketotic hyperglycinemia)¹ is an autosomal recessive disorder caused by mutations in the genes which encode the two subunits the propionyl-CoA carboxylase (PCC) enzyme. These subunits are encoded by the *PCCA* and *PCCB* genes.

The PCC enzyme converts propionyl-coenzyme A (CoA) to methylmalonyl-CoA. This reaction is one step in the breakdown of certain amino acids, as well as lipids and cholesterol. Mutations in either subunit of the gene reduce enzyme function and lead to build up of propionyl-CoA and other compounds. Build-up of these metabolites leads to damage, for example, of the brain and nervous system, leading to the symptoms of PA.

Newborn screening for PA is through tandem mass spectrometry (MS/MS) on newborn dried blood spots. PA causes abnormal profiles of acylcarnitines in the blood, mainly raised propionylcarnitine (C3). This marker is also raised in the related organic acid disorder methylmalonic acidaemia, as well as the neonatal form of multiple carboxylase deficiency (MCD), and maternal vitamin B deficiency. Therefore there is the need for differential diagnosis between these conditions.

Confirmatory tests for PA include assays of urine organic acids, plasma acylcarnitines and amino acids.^{2,3} The presence of high levels of propionic acid in the urine, as well as raised plasma C3 levels and normal plasma homocysteine levels indicate the presence of PA. The activity of the PCC enzyme can also be assayed in fibroblasts to confirm reduced activity. MS/MS and urinary markers used in the screening and confirmation of PA are summarised in **Table 1**.

Incidence of PA identified in screening programmes has varied internationally, ranging from 1/27,000 in Saudi Arabia to 1/660,000 in Taiwan. Eighty-one mutations are reported to have been identified in the *PCCA* gene and 86 mutations in *PCCB*.⁴

Propionic acidaemia commonly presents with frequent vomiting, refusal to eat, lethargy, hypotonia, and failure to thrive.⁵ The condition can lead to metabolic acidosis (decompensation), coma and death.⁶ Decompensation is usually prompted by stressors such as infection, injury, surgery, hormonal changes, or significant dietary changes that involve increased protein intake.

Complications of the condition can include intellectual difficulties, neurological complications such as stroke-like episodes and seizures, as well as cardiac and gastrointestinal complications. Most patients have neonatal onset, but some have later onset, potentially in adolescence or adulthood. The latter tend to have milder symptoms and a higher survival rate.⁴

Ongoing treatment for PA focuses on limiting the amount of substrate flowing through the PCC biochemical pathway, by reducing protein intake. Plans to avoid potential precipitation of metabolic decompensation are also important, and acute treatment is required when metabolic decompensation occurs.

Basis for current recommendation

The most recent UKNSC review of the policy on newborn screening for the organic acid disorders propionic acidaemia and methylmalonic acidaemia was conducted in 2009. It used

evidence based on HTA reviews from 1997 and 2004 on screening for inborn errors of metabolism.

It concluded that universal screening for these disorders should not be offered due to uncertainties over:

- UK incidence
- The timing of the screening test and subsequent effectiveness of intervention (particularly those with neonatal onset)
- Sensitivity and specificity of the screening test using MS/MS
- The ability of the screening test to pick-up milder forms of MMA

Key uncertainties highlighted by the NSC in commissioning this external review included:

- Is the epidemiology and natural history of PA and MMA understood?
- Have any studies explored the implications of the carrier state identified as a result of screening?
- Do we have a precise and validated means of testing for PA or MMA at screen?
- Does treatment following screening add additional benefit over treatment in non-screened populations?
- Are guidelines available for screening for PA and MMA?
- Have any studies explored the wider ethical, legal and social impacts of screening for PA and MMA?

Current update review

The current review considers whether the volume and direction of the evidence produced since 2001 (the search date for the most recent HTA) indicates that the previous recommendation should be reconsidered. Six main criteria will be considered, with particular focus given to areas the NSC have identified as uncertain, or supported by insufficient evidence.

Separate reports have been prepared for PA and MMA. This report covers PA.

Table 1: Markers of the presence PA and MMA in newborn MS/MS screening⁷⁻¹⁰

Disorder	Primary marker for screening	Secondary markers for screening	Urine markers (confirmatory testing)
PA	↑C3	↑C3/C2 ↑C3/C16 ↑C3/C0 ↑C3/Met ↑Gly ↑C3/C4 ↑Lysine	↑ propionic acid † ↑3-hydroxypropionate ↑ methylcitrate ↑propionylglycine* ↑ N-tiglylglycine ↑ 2-methyl-3-oxovaleric acid ↑ 3-hydroxy-2-methylbutyric acid ↑ 2-methyl-3-oxobutyric acid ↑ 3-hydroxy-n-valeric acid ↑ 3-oxo-n-valeric acid
MMA	↑C3	↑C3/C2 ↑C3/C16 ↑C3/Met ↑ C4DC (methylmalonyl carnitine) ↑C3/C0 ↑Gly ↑C3/C4 ↑Lysine	↑ methylmalonate ↑ methylcitrate ↑3-hydroxypropionate ↑ propionylglycine*

Markers in bold were not reported to be elevated in the other condition. †Propionic acid reported as the confirmatory test for PA in ACMG guidance¹¹, but hydroxypropionate reported as raised in both MMA and PA. *Propionylglycine was reported as raised in both PA and MMA in one paper¹⁰ but reported as raised in only PA in a second.⁹

The main criteria and key questions reviewed are summarised in **Table 2**.

Table 2. Key questions for current PA update review

Criterion	Key Questions (KQ)	# Studies Included
2	<p>1) Is the epidemiology and natural history of PPA and MMA understood?</p> <p>1a) Are there known and understood genotype/ phenotype or enzyme activity/ phenotype correlations?</p> <p>1b) When do clinical presentations manifest and how does this relate to the timing of the screen in the UK (and elsewhere)</p>	19
4	2) Have any studies explored the implications of the carrier state identified as a result of screening?	0
5	<p>3) Do we have a precise and validated means of testing for PPA or MMA at screen?</p> <p>3a) Does the current screening marker give sufficient distinction between PA, MMA and other metabolic conditions tested for using acylcarnitine analysis?</p> <p>3b) Have cut off values for the screen marker been agreed and established in screened populations for each condition and are the test values acceptable?</p>	10
10	<p>4) Does treatment following screening add additional benefit over treatment in non-screened populations?</p> <p>4a) Is there any evidence that long-term treatment in screened populations adds benefit (early mortality, neurological symptoms, quality of life) over non-screened populations?</p> <p>4b) Is there any evidence through outcome data that early treatment through screen adds additional benefit over those detected clinically?</p>	4
11	<p>5) Are guidelines available for treatment of individuals identified through screening for PPA and MMA?</p> <p>5a) Is there an agreed set of criteria for who should receive treatment among asymptomatic cases (egg by genotype analysis or % enzyme activity) and what treatment they should receive?</p>	4
14	<p>6) Ethical, Legal and Social Issues (ELSI)</p> <p>Have any studies explored the wider impacts of screening for MMA and PA (egg reproductive decision making, diagnostic odyssey, implications for 'informed' consent when extending to additional NBS conditions etc.)?</p>	2

A systematic literature search of studies published between 2001 and January 2015 yielded 1811 references (a single search was performed for PA and MMA). Of these, 246 were assessed as being potentially relevant to the key questions outlined in Table 1. These studies were further filtered at title and abstract level, and 79 were selected for appraisal at full text. Additionally, the full texts appraised for the MMA review were also considered for inclusion (there was overlap in the full texts assessed). Relevant references identified in the preparation of reviews on fatty acid and amino acid disorders were also included. Broadly, systematic reviews, RCTs, other prospective studies and screening programme evaluations were prioritised. Each section below provides additional information on the evidence selection process for the given criterion.

Appraisal against UK NSC Criteria

These criteria are available online at <http://www.screening.nhs.uk/criteria>.

- 2. The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be a detectable risk factor, disease marker, latent period or early symptomatic stage**

Current UKNSC key question

1 Is the epidemiology and natural history of PA understood?

1a) Are there known and understood genotype/ phenotype or enzyme activity/ phenotype correlations?

1b) When do clinical presentations manifest and how does this relate to the timing of the screen in the UK (and elsewhere)?

For question (1a) any study design reporting on genotype- (or PCC enzyme activity-) phenotype correlations would have been included. Studies where the abstract (or full text) only reported on mutation spectrum without indicating that relationship with phenotype was assessed were excluded. Only studies looking at the correlation between genotype or enzyme activity and clinical phenotype (i.e. age at onset, symptoms, clinical outcome) were included. Studies assessing only non-clinical (intermediate) phenotypes e.g. mRNA levels, or metabolite levels were excluded. No study size criteria were set for inclusion, due to the small sizes of the studies identified.

For question (1b) studies reporting on the age at onset of PA symptoms in non-screened populations were included. Only studies which reported the median age at onset, or indicated how age at onset was spread across the individual weeks in the month after birth and afterwards were included. Studies presenting less detailed information about early onset cases e.g. just as 'within 1 month of birth' or 'within 1 year of birth' were excluded. This is to allow better comparison of age at onset with the age at which newborn screening and resulting diagnosis would place, as this would be within the first few days and weeks of life. Studies of screening included under criterion 5 which indicated the percentage of screen-identified cases that were symptomatic at time of diagnosis following screening were also included.

Description of the evidence

Overall, 23 studies were assessed as potentially relevant during title and abstract sifting for key question 1 and further assessed at full text. In addition, 2 studies extracted for criterion 5 were also relevant and described here.^{12, 13}

In total 19 studies were included and assessed by this review for this key question.

1a) Are there known and understood genotype/ phenotype or enzyme activity/ phenotype correlations?

Results

Seven case series were identified which assessed genotype and phenotype correlation; ^{4, 14-19} one guideline⁹ and a consensus conference document⁶ were also of relevance and included.

One European guideline based on evidence available up to the end of 2011 reported that no clear-cut correlation between genotype and phenotype has been found in PA.⁹ The lack of a clear genotype-phenotype correlation was also supported by an international consensus conference from 2011.⁶

The seven case series were small (range 4 to 54 participants), and several focused more on the effect of the mutations on gene transcription and protein stability, with genotype - phenotype correlation only touched on in the discussion.

Some papers described some relationship between genotype and phenotype or PCC enzyme activity and phenotype. For example:

- One paper in 14 Latin American PA patients suggested that most patients with early onset symptoms had frameshift mutations.¹⁹ However, not all early onset cases carried frameshift mutations, and they were found in a heterozygous state in both early and late onset patients.
- One paper in 10 Taiwanese PA patients identified two mutations which seemed to be associated with the classic early onset of PA.⁴ These cases had between <1% and 9.3% residual PCC enzyme activity.
- One paper in 8 Korean PA patients found 0% PCC activity in two early onset patients in whom it was measured, and they also mostly (4/5 cases) homozygous for one mutation.¹⁶ The 3 late onset cases had 1.6% to 9.6% PCC activity and 2/3 were heterozygous for the mutation found in the homozygous state in the early onset cases.
- One paper in 26 European and American patients described various mutations associated with different phenotypes, and suggested that patients bearing null mutations in both alleles are associated with the most severe presentation of the disease.¹⁷ However, they did note some exceptions to this.
- One paper in 10 patients with a mild form of PA in Japan, mostly identified by newborn screening, found a common mutation in these cases.²⁰ This was described as being different from that seen in severe patients, but mutation analysis for severe patients was not shown.

In general, the studies did not provide conclusive evidence of clear cut genotype and phenotype or PCC enzyme activity and phenotype correlations. This was acknowledged by the largest study (n=54),¹⁸ and a study in 10 people with PCCA mutations.¹⁴ Another study assessed mutations in 4 patients affecting splice sites, and despite these mutations being predicted to cause major disruption to the protein, the patients had mild phenotypes.¹⁵

The difficulty in identifying correlations is in part due to the small number of cases in the studies, as well as the wide range of mutations identified and the fact that most patients are compound heterozygotes. Given the wide range of mutations, correlation of enzyme activity with phenotype might provide a more universal measure than genotype. However, in the studies which assessed this, the relationship between this and phenotype also seemed variable.

This could in part be due to the use of different protocols for measuring it (e.g. some studies used fibroblasts, some leukocytes, and some lymphoblasts). It was often unclear in the studies how the cases had been selected and the clinical data seemed to have been obtained by retrospective assessment of medical records rather than prospective follow up of cases for the purpose of genotype-phenotype correlation.

1b) When do clinical presentations manifest and how does this relate to the timing of the screen in the UK (and elsewhere)?

Nine studies assessed age at onset of symptoms in clinically detected cases and were included for this key question.^{4, 16, 19, 21-26} Four studies reported on the proportion of screen detected cases with symptoms at the time of screening results and were included in this key question.^{12, 13, 27, 28}

The three largest studies in clinically presenting cases (n=55, n=49 and n=26) reported that between 64% and 84.6% of cases present in around the first week of life.²¹⁻²³ Two of these studies were in European populations, and one in Saudi Arabia.

The six smaller studies (between 4 and 14 cases) reported percentages between 12.5% and 100%.^{4, 16, 19, 24-26}

Their smaller size may make them less representative. The latter also tended to include more Asian countries, Latin American and middle Eastern countries which may be less representative of what might be seen in the UK than the European studies. One study also only included those cases presenting up to age 28 days, which may skew results.²⁵

The two largest studies in screen detected cases (n=7 and n=19) reported that 57% and 68% of cases were symptomatic at or before the time the screening results were known.^{27, 28} There was some overlap of the population in these papers. The other screening studies only included 1 and 2 cases of PA, and they found 100% and 50% were symptomatic before screening results were available.^{12, 13} Only the paper with a single screened case made it clear that a diagnosis had been made by the time screening results were available. Screen detected cases may have their symptoms noted in retrospect, because of their later diagnosis with PA.

Summary: Criterion 2 not met.

There is no clear cut evidence of a correlation between genotype and phenotype or PCC enzyme activity and phenotype. Existing case series tend to be small (most with fewer than 20 participants) and appear not to have been based on prospective assessment of patients for the purpose of genotype-phenotype correlation. A wide range of mutations have been identified, and most patients are compound heterozygotes, which complicates ability to detect correlations should they exist. Current evidence would not be a sufficient basis on which to predict prognosis, and therefore decide on treatment for affected individuals.

Studies suggest a high proportion of PA cases become symptomatic within the first week of life, potentially between two thirds and 85%. As in the UK the blood spot is only collected on day 5, these cases are likely to present before the screening results become available.

Assessments of screen detected cases from other countries supports that up to about two thirds of cases are likely to already be symptomatic and present by the time newborn screening results are available. Whether these symptoms would always result in diagnosis of PA in the absence of screening results is unclear.

4. If the carriers of a mutation are identified as a result of screening the natural history of people with this status should be understood, including the psychological implications.

Current UKNSC key question

2) Have any studies explored the implications of the carrier state identified as a result of screening?

For this question any studies assessing the psychological implications of carrier detection as part of or due to the screening programme (for example, through cascade testing) would have been included. In addition, studies which reported whether carriers of mutations which cause MMA were being identified by screening would also have been included.

Description of the evidence

One study was identified as potentially relevant to this question during title and abstract sifting and further assessed at full text.²⁹ It was excluded as while it looked at stress in parents of babies identified through newborn screening it did not look at the effect of carrier status specifically, or provide separate results for PA. Parents of babies with PA would be carriers, and other family members may also be carriers who could be identified by cascade testing.

No studies were identified which assessed levels of disease markers in blood spot screening for newborns heterozygous for mutation(s) in a gene(s) which causes PA. None of the studies on screening programmes assessed for Criterion 5 reported that newborn carriers of PA causing mutations were being identified by newborn blood spot screening. However, as mutation analysis may not routinely be carried out on those who turn out to be false positives, it is not clear whether this group might include carriers.

Results

No relevant studies were identified.

Summary: Criterion 4 not met.

No studies were identified in the update search which explored the implications of the carrier state identified as a result of screening. Newborn carriers have not been reported to be picked up by the screening programmes. However, parents of the affected newborns are by default carriers, and cascade testing could identify other carriers in the family. Separating the psychological impact of having an affected child (or affected relative) from that of being a carrier is likely to be challenging.

5. There should be a simple, safe, precise and validated screening test.

Current UKNSC key question

3) Do we have a precise and validated means of testing for MMA at screen?

a) Does the current screening marker give sufficient distinction between MMA, PA, and other metabolic conditions tested for using acylcarnitine analysis?

b) Have cut off values for the screen marker been agreed and established in screened populations for each condition and are the test values acceptable?

Description of the evidence

Overall, 50 studies were identified as potentially relevant during title and abstract sifting and further assessed at full text. Prospective cohort studies and screening programme evaluations assessing the clinical validity of the tandem mass spectrometry screening of newborn dried blood spots test for identifying newborns with PA were prioritised. In addition, guidelines were included where they provided background on screening recommendations and practice.

Studies were included if they provided figures to indicate true and false positive results specifically for PA in a newborn blood spot screening programme. Studies presenting only true positives without a false positive rate were excluded, as were studies which did not present figures which allowed estimation of test performance for detection of PA specifically, or which did not identify any cases of PA. Cases of MMA identified through C3 screening were treated as false positives for this report, and specificity calculated accordingly.

Studies solely reporting performance of tests on individuals presenting symptomatically, or analytical validity of tests on selected individuals known to have or not have the disorders were excluded, as were studies which did not present sufficient data to calculate sensitivity and/or specificity. Due to the rarity of PA, studies including fewer than 30,000 screened infants were excluded, as they are unlikely to yield any cases.

Of the 50 studies assessed at full text, 5 studies of newborn screening programmes were included in the final analysis. The main reasons for exclusion were not providing the required figures to assess clinical validity of PA screening, or assessing analytical rather than clinical validity. Only one study was excluded on the basis of size.

In addition, 3 guidelines addressing screening identified in the search were included.⁹⁻¹¹ One additional paper is cited for background information about markers used in screening.⁷ Two relevant papers on screening programme analyte cut-offs and diagnostic guidelines for newborns who screen positive from the reviews on newborn screening for the fatty acid and amino acid disorders was also included.^{8, 30}

Results

Background – guidelines addressing screening

Three US guidelines on newborn screening and subsequent diagnostic follow up for conditions including PA were identified.^{10, 11, 30} Two of these guidelines recommended screening for PA,^{10, 11} while the third only recommended follow up procedures after such screening based on expert consensus.³⁰ A more recent European guideline from 2014 did not recommend for or against newborn screening.⁹

In 2006 the American College for Medical Genetics published a practice guideline on newborn screening based on expert opinion and a systematic review of research evidence.¹¹ It recommended PA as a core condition which should be included in a newborn screening panel.

On the basis of its guidance the ACMG generated and keeps updated ACTION (ACT) sheets that describe the short term actions for a health professional if an infant screens positive and algorithms on the screening and diagnosis of the conditions.^{2, 3} **Table 3** summarises their

recommended diagnostic evaluation of infants screening positive for raised C3 levels, and guidance on the interpretation of results.

Table 3: ACMG recommendations for newborns with raised C3 levels in newborn screening

<ul style="list-style-type: none"> Diagnostic evaluation of newborns with elevated C3 includes assay of urine organic acids (UOA), plasma acylcarnitines (PACYL), plasma homocysteine 		
Assay results	Condition indicated	Optional confirmatory tests
Plasma C3 high Propionic acid in urine Plasma homocysteine normal	Propionyl CoA carboxylase deficiency (propionic academia, PA)	Propionyl-CoA carboxylase assay in fibroblasts
Plasma C3 high Methylmalonic acid in urine Plasma homocysteine normal	Methylmalonyl-CoA mutase (Mut ⁻ , Mut ⁰) deficiency, or CblA deficiency, or CblB deficiency	Mut assay Cbl complement studies in fibroblasts
Plasma C3 high Methylmalonic acid in urine Plasma homocysteine high	CblC deficiency, or CblD deficiency, or CblF deficiency, or Transcobalamin II, or Vitamin B12 deficiency	Cbl complement studies in fibroblasts
Plasma C3 and C4DC high Methylmalonic acid in urine Plasma homocysteine normal	Succinyl-CoA synthetase (SUCLA2) deficiency	SUCLA2 DNA sequencing
All values normal	False positive, consider maternal vitamin B12 deficiency	NA

The ACMG guidelines were developed in the US and relied to an extent on expert opinion as to the value of screening for the individual programmes. They were also considered in the context that newborn screening for various conditions is already being provided in some states, with exact conditions being screened for differing. This guideline may not be applicable to a UK setting, as views of experts in the UK may differ, and the context also differs, as screening is decided on nationally and is currently limited to a small panel of conditions.

The evidence available at that time was rated by two experts as mostly being from well-designed studies in relevant populations with at most only minor limitations, or consistent evidence from observational studies. The evidence itself was not presented.

The ACMG noted the potential for bias in the assessment of the severity of some conditions: without screening, the more severe forms are noticed first, biasing what is known about the effects of the condition. They also noted that until a large general population has been studied, understanding of the performance of the screening tests (in terms of the range of manifestations it identifies) is limited.

In 2009, the National Academy of Clinical Biochemistry published guidelines on the follow up testing of newborns identified through MS/MS screening.¹⁰ These guidelines were prepared by experts based on an assessment of the peer-reviewed literature. It strongly recommended

newborn screening for PA as it judged that there was good evidence that it improves important health outcomes, and that benefits substantially outweighed the harms. It provided follow-up recommendations for individuals screening positive.

The New York Mid-Atlantic Consortium for Genetic and Newborn Screening Services similarly published consensus guidelines on the follow up testing of newborns identified through MS/MS screening in 2010,³⁰ with the latest update of these published in 2014.³¹ These guidelines were based on expert consensus only and not explicitly on the research evidence.

These US guidelines may not be applicable to the UK context, particularly as they incorporated expert opinion and the UK viewpoint may differ. Other countries may also have similar national guidelines as to whether PA screening is offered and subsequent diagnostic and treatment approaches, but these were not identified in the search.

The European guideline published in 2014 noted that newborn screening for PA (and MMA) was feasible, but did not recommend for or against, and therefore did not provide guidelines as to how it should be carried out.⁹ It did provide recommendations relating to diagnosis of individuals presenting clinically, and some of these are also relevant for those identified through screening.

As this is a recent European guideline, using Scottish guideline methodology and with some members of the guideline development group coming from the UK, it is likely to be more applicable to the UK setting than the US guidelines. The evidence available was reported to be not of a high quality, and recommendations were mainly based on case series and case reports, and expert opinion.

Evidence from newborn screening programmes

The 5 studies on newborn screening programmes are summarised in **Table 4**, with additional information in the Appendix. Studies ranged in size from 239,415 (USA)³² to 1,321,123 (Taiwan).³³ The results of these studies are discussed in sections 3a and 3b below.

DNA testing was not always reported as being carried out to confirm or rule out diagnosis in all of the studies. Therefore mutational status of those described as true or false positives, or true or false negatives was not always known. There was a lack of systematic follow up in the studies to identify false negatives, and it is likely that some will have been missed. This may particularly be an issue for any cases of PA with later onset or milder phenotype, who might not present with metabolic decompensation in the newborn period.

Results from the studies may be applicable to a UK setting, as the UK has a newborn MS/MS screening programme. However, there was variation in the protocols used in the different studies (for example, in the age at which the dried blood spots were taken, range of markers and cut-offs used), which may reduce applicability to the UK of some studies. The variability of the screening protocols is also likely to impact on the results and test performance, rendering direct comparisons between different studies and generalisability difficult.

Table 4: Performance of newborn blood spot screening programmes for PA (other conditions identified considered as false positives)

Study	Country	n	Analytes assessed	True positives† (PA)	False positives*†	True negatives	False negatives	Specificity for PA†	Sensitivity for PA†	PPV for PA†
Niu et al. 2010 ³³	Taiwan	1,321,123	C3 C3/C2	2	73 (includes 14 MMA)	1,321,048	0	99.994%	100%	2.67%
Weisfeld-Adams et al. 2010 ³⁴	USA	1,006,298	C3 C3/C2 Met C4DC	5	150 (includes 26 MMA)	1,006,143	0	99.985%	100%	3.23%
Lund et al. 2012 ¹²	Denmark, Faroe Islands, Greenland	504,049	C3 C3/C2 C4DC	2	52 (includes 3 MMA)	503,992	0	99.990%	100%	3.70%
Schulze et al. 2003 ¹³	Germany	250,000	C3 C3/C0 C3/C2 C4DC	1	209 (includes 4 MMA)	249,790	0	99.916%	100%	0.48%
Frazier et al. 2006 ³²	USA	239,415	C3 C3/C2	1	8 (includes 5 MMA)	239,406	0	99.997%	100%	11.11%

*Cases of MMA were counted as false positives for this report; † reviewer calculated data; PPV positive predictive value

Variations in the incidence of PA in the UK and other countries may also impact on test performance, and so studies from countries with more genetically similar populations (e.g. European populations) will be more applicable to the UK and the purposes of this review.

3a) Distinction between PA, MMA, and other conditions tested for using acylcarnitine analysis

The primary marker for PA in newborn blood spot screening is an elevated level of the short chain acylcarnitine propionylcarnitine (referred to as C3, indicating the number of carbon atoms in the acyl part of the molecule). All of the included studies used C3 as the main screening marker for both PA and MMA. None of the newborn screening programme studies identified used any screening markers specific for PA.

Other conditions can also lead to elevated C3 levels, and data on distribution of C3 values on newborn screening from the international Region 4 Stork (R4S) data collection project shows that there is overlap between the C3 distributions seen with PA, MMA, Multiple Carboxylase Deficiency (MCD), and maternal B12 deficiency (see **Table 5**).⁸ C3 is only a secondary marker for neonatal onset Multiple Carboxylase Deficiency (MCD), also known as holocarboxylase synthetase deficiency, and its primary marker (C5-OH) is not shared with MMA and PA.⁷ The included studies did not report any cases of MCD being identified through PA screening specifically.

Table 4: Distribution of C3 values in PA, normal and other populations in newborn screening from the R4S project⁸

Population	N	Percentile (μM)		
		1 st	50th	99th
All newborns	93 members*	0.57	1.75	4.74
PA	201 cases	2.3	14	51
MMA: MUT/Cbl A, B	328 cases	1.1	9.5	40
MMA: Cbl C, D	124 cases	2.8	8.4	20
MCD	15 cases	1.6	4.6	16
Maternal B12 deficiency	47 cases	*2.55	7.8	22

1 μM = 1 $\mu\text{mol/L}$; *R4S members/centres who have submitted percentile values, overall estimated to be data from 25-30 million newborns in the database. Information from McHugh et al. 2011⁸

Secondary markers for PA (and MMA) include increased ratios of C3 to free carnitine (C0) and to certain other acylcarnitines (C2, C16, C4).^{7, 8} The R4S data also shows overlap in the distribution of these analytes for PA, MMA and maternal B12 deficiency (and for MCD for some analytes).⁸

Only two of the six newborn screening programme studies explicitly reported using a slightly different screening markers or cut-offs for PA and MMA.^{13, 34-37} The first used one first tier marker aimed specifically at MMA (reported as methylmalonylcarnitine), but also three shared first tier markers with the same cut-offs for both conditions.¹³ The second used a secondary marker (methionine) to identify CblC, D, and F diseases, but also three shared markers with the same cut-offs reported for both conditions (C3 and C3/C2 first tier and C4DC second tier).³⁴

3b) Cut offs

Cut-offs in MS/MS are set with reference to the distributions of analyte values in unaffected individuals. The R4S database project has collected and made available international data on analyte distributions from screening around 25-30 million newborns, and its main aim is to define evidence-based cut-off target ranges for all analytes.⁸ They recommend setting cut-offs for elevated markers in the interval between the cumulative 99th percentile of the normal population and the 5th percentile of the disorder with elevated values for that analyte (or the disorder with the lowest values for that analyte if it is elevated in multiple disorders). There was variation between different sites in analyte values, with mean variability of median values across sites of 27% for acylcarnitines.

There was variability in how cut-offs were set in the included studies, with several basing these on the mean value in an unaffected or general population sample plus 4-8 standard deviations³²⁻³⁴ and others basing this on percentiles in the unaffected or general population.^{13, 36, 38} This variability may be due to the programs being set up before the start of the R4S project in 2005, or when it was relatively new.

After initially setting the cut-offs at the start of their programs, several reported modifying these over time to optimise sensitivity and false positive rate.^{12, 32, 33, 35, 36, 38}

All of the studies reported using C3 as a marker for PA, and cut-offs ranged from at or above 4.74 μ M to 7 μ M.³³ They also all used the C3/C2 ratio, and cut-offs ranged from over 0.15³² to over 0.39.¹³

Those with lower cutoff values often designated these “borderline” values, and required a repeat test to confirm the value before referral. In addition, studies varied in the markers they used, so variations in cut-offs for individual markers may reflect different combinations and approaches being used to optimise test performance in the local population.

Test performance

Overall, screening for PA showed a similarly high specificity across all studies, ranging from to 99.916%¹³ to 99.997%.³²

Studies requiring testing of a second sample for those with ‘borderline’ values also markedly reduced the number of referrals compared to if they had used just a single borderline value for a referral (where figures were provided).³²⁻³⁴ For example, in one study of 1,321,123 newborns, over 3,114 had a borderline result, and only 39 of these still had a value above the borderline in the second test (1 PA case was identified among these 39). Age of the babies at collection of the second sample was not reported in the studies. This need for a second test does involve additional resource and potential psychological impact on the family as they wait for the second result.

PPV in the studies ranged from 0.48%¹³ to 11.11%.³² The study with the highest PPV was a US study, and required repeat testing of borderline C3 or C3/C2 values, which reduced false positives. The lowest value in this range was obtained in Germany,¹³ where screening for MMA and PA was stopped in 2005 due to the high recall and false positive rate for C3.³⁹

No false negative PA cases were reported in any of the studies, implying a sensitivity of 100%. However, there is no systematic follow up to identify false negatives, and it is likely that some will have been missed. This may particularly be the case for any cases of PA with later onset or milder phenotype, which might be less likely to present and be identified.

While some studies reported algorithms for how primary and secondary markers were combined to determine which babies were counted as screen positives,^{33, 34, 36, 37, 40} others did not state how this was reconciled (e.g. what happened if the primary marker was not raised but a secondary marker was, and vice versa).

Summary: Criterion 5 not met.

Screening for PA and MMA utilises the same markers (mainly C3 and ratios involving C3), and it is not possible to definitively distinguish between them at the screening stage alone.

The combinations of markers and cut-offs of screening programmes for PA and MMA vary internationally. There is now an international collaboration set up to provide evidence-based cut-offs for MS/MS analytes in newborn screening and this may lead to greater standardisation.

Specificity of screening for PA is high (over 99.90%) but positive predictive value is low (11% and lower). The highest PPV came from a US study which utilised repeat testing of borderline values to reduce false positives, however, another programme in the US, and one in Taiwan used similar measures and did not achieve as high PPVs. The range of different protocols makes comparison between programmes difficult.

No false negatives have been reported (hence implied sensitivity is 100%), but the lack of systematic follow up or monitoring for false negatives means that some are likely to be missed. This is particularly likely to be the case for any cases with milder symptoms or later onset.

10. There should be an effective treatment or intervention for patients identified through early detection, with evidence of early treatment leading to better outcomes than late treatment.

Current UKNSC key question

4) Does treatment following screening add additional benefit over treatment in non-screened populations?

4a) Is there any evidence that long-term treatment in screened populations adds benefit (early mortality, neurological symptoms, quality of life) over non-screened populations?

4b) Is there any evidence through outcome data that early treatment through screen adds additional benefit over those detected clinically?

Description of the evidence

Overall, 11 studies were identified as potentially relevant during title and abstract sifting and further assessed at full text. Systematic reviews and randomised controlled trials (RCTs) were prioritised. As no studies of this type were identified non-randomised comparisons of outcomes of screen-detected and non-screen detected or untreated cases were included. Studies were excluded if they reported outcomes of PA pooled with other conditions, with no data available for PA alone.

Of the 11 studies assessed at full text, 3 were included in the final analysis,²⁷⁻²⁹ plus one guideline for context.⁹ The main reasons for exclusion were that the studies did not provide a comparison of screened and clinically detected or untreated PA cases, or did not provide separate results for PA alone.

Results

4) Does treatment following screening add additional benefit over treatment in non-screened populations?

Due to the small number of relevant studies, identified results have not been split into short term and longer term outcomes.

One European guideline developed using Scottish Intercollegiate Guideline Network (SIGN) methodology (utilising systematic review of the evidence supplemented by expert opinion as needed) noted in its recommendations that while newborn screening for PA and MMA was technically feasible, available data had not determined whether it was of long term clinical benefit.⁹ It also noted that prospective and retrospective studies on PA and MMA have suggested that intellectual disability and developmental delay may not be prevented by early treatment.

Three studies were identified which provided some comparison of screen detected and clinically detected PA cases.²⁷⁻²⁹ None specifically reported what treatment was received by the two groups, and there was overlap in the patients reported in two of the studies.

One comparative case series of 55 individuals with PA diagnosed through either newborn screening or clinically found no difference in groups in the number of metabolic decompensations, hospital admissions, or IQ. More of the clinically detected cases than screen detected cases attended a kindergarten for children with special needs once they reached 3 years and over (55% vs. 21%) but the significance of this finding was not reported.

There was a non-significant trend for reduced mortality in the screen detected group (0%) versus the clinically detected group (12%; p value not reported). However, the screen detected group were generally younger than the clinically detected group and this may affect results. The study groups were relatively small and may have lacked power to detect clinically important differences.

The second study was a poorly reported comparative case series of 20 individuals with PA diagnosed through either newborn screening or clinically.²⁷ No direct statistical comparisons between the groups were made in this study. Mortality was slightly higher in the clinically detected group with neonatal onset PA (33%) than in the screen detected cases (22%), but this only represented 3 and 2 deaths respectively. Metabolic problems appeared to be common in both groups, but it was unclear if this outcome was consistently assessed and described in the study. Cognitive outcome was also not consistently described for the two groups.

The third was a small comparative case series of 9 individuals with PA diagnosed through either newborn screening or clinically.²⁹ In the clinically detected PA group it reported one death (1/6, 17%), two cases needing NICU care before diagnosis (2/6, 33%), and at least one child up to the age of 3 and one over the age of 3 (unclear if the same individuals) with mental retardation. None of the small group of 3 children identified by screening had these outcomes. This small study did not provide statistical analysis for this comparison.

Summary: Criterion 10 not met.

This criterion has not been met as no large, robust studies were identified which allowed comparison of treatment outcomes from screened and unscreened populations. One relatively large study in 55 individuals suggested that mortality was higher in clinically detected cases, but this difference was not statistically significant, and treatments received were not clear.

This lack of robust evidence was also noted by a recent European guideline, which also concluded that available data has not yet determined whether newborn screening for PA or MMA are of long term clinical benefit.

Given the rarity of PA, it is likely a prospectively constructed international study or registry would be needed to gain sufficient evidence to compare the impact of treatment following screening versus treatment following clinical detection.

11. There should be agreed evidence based policies covering which individuals should be offered treatment and the appropriate treatment to be offered.

Current UKNSC key question

5) Are guidelines available for treatment of individuals identified through screening for PA?

5a) Is there an agreed set of criteria for who should receive treatment among asymptomatic cases (e.g. by genotype analysis or % enzyme activity) and what treatment they should receive?

Description of the evidence

For this question, clinical guidelines from national, state or professional organisations relating to newborns screening for or treatment of PA were included. Of the 7 studies assessed for this question, 4 met inclusion criteria.^{9, 41-43} The excluded studies were guidelines which did not cover treatment and a primary study which assessed potential markers of impending metabolic decompensation but did not make recommendations for treatment.

Results

5a) Is there an agreed set of criteria for who should receive treatment among asymptomatic cases (e.g. by genotype analysis or % enzyme activity) and what treatment they should receive?

One European guideline published in 2014⁹ and a set of guidelines developed at an international consensus conference in 2011 were included.⁴¹⁻⁴³ These guidelines did not explicitly cover the treatment of asymptomatic individuals with PA identified through screening.

These guidelines focused on acute management on symptomatic presentation, and long term management. The European guideline aimed to provide a European consensus on the diagnosis and management of PA and MMA, and was informed by a systematic review of the evidence.⁹ The consensus conference discussions were reported to be informed by literature reviews, but these were not explicitly systematic reviews.¹

The long-term management described in this guideline includes various prophylactic interventions attempting to reduce the build-up of potentially toxic metabolites. This included a low protein diet, with amino acid supplementation as needed, plus L-carnitine supplementation, and oral antibiotics continuously or intermittently to control intestinal propionic acid producing bacteria. In addition, careful regular monitoring is recommended.

While, for example, the low protein diet is recommended to be titrated individually, none of the recommendations are explicitly stated as applying to only symptomatic individuals.

The section on acute management in the international consensus guideline recommends that, in a symptomatic individual in whom PA is known or highly suspected as a result of newborn screening, treatment should be initiated immediately, but does not explicitly state what treatment should be offered to asymptomatic individuals.⁴¹ None of the recommendations on chronic treatment of PA in either guideline were explicitly further guided by genotype or % enzyme activity. The lack of explicit recommendations otherwise, and the seriousness of, for example, metabolic decompensation should it arise, suggests that all individuals identified as having PA would be considered for ongoing prophylactic interventions.

Summary: Criterion 11 not met.

The two treatment guidelines identified did not give explicit recommendations about management of asymptomatic individuals identified through screening, or use specific genotype or level of PCC enzyme activity to guide management. The lack of explicit recommendations

otherwise, and the seriousness of metabolic decompensation should it arise, suggests that all individuals identified as having PA would be considered for ongoing management aimed at reducing build-up of toxic metabolites.

Only one of the guidelines explicitly reported performing a systematic review of the evidence. It noted that the available evidence was generally not of high quality. The evidence base on which the other guidelines were developed may not have been complete, and there may have been bias in the evidence selection.

14. There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/ intervention) is clinically, socially and ethically acceptable to health professionals and the public.

Current UKNSC key question

6) Ethical, Legal and Social Issues (ELSI)

Have any studies explored the wider impacts of screening for PA (e.g. reproductive decision making, diagnostic odyssey, implications for 'informed' consent when extending to additional NBS conditions etc.)?

For this question any studies assessing the wider ethical, legal and social impact of screening for PA specifically were included. Studies assessing the financial impact or cost-effectiveness of screening for PA were not included. Studies which looked at the impact of newborn screening more broadly (i.e. not for PA specifically) were also excluded.

Description of the evidence

Six studies were identified as potentially relevant to this question during title and abstract sifting and further assessed at full text. One study assessing age at diagnosis was identified as relevant to this key question and included.²⁸ One guideline was also included for context on reproductive decision making.⁹ The other four were excluded. Reasons for exclusion were not looking at relevant outcomes, not reporting on the impact of a population newborn screening programme, or providing data on newborn screening as a whole rather than by disorder.

No studies were identified which looked at the impact of PA screening specifically on other ethical, legal, or social issues.

Results

Reproductive decision making

A European guideline reports that prenatal testing for PA and MMA is feasible.⁹ This suggests that once one affected child is identified (whether by screening or clinical detection) prenatal diagnosis could be an option available for parents to inform reproductive decision making. However, no studies were identified in the search which assessed or compared reproductive decision making after screen or clinical detection of PA in a family.

Age at diagnosis

One study (n=55 individuals with PA) comparing screening with clinical diagnosis found that infants diagnosed through newborn screening had a significantly reduced age at diagnosis for PA (median 6 days, range 4 to 13 days) compared with clinical diagnosis (median 12 days, range 3 days to 8 years; p=0.018).²⁸ This suggests that the diagnostic odyssey may be reduced with newborn screening, although how long individuals were symptomatic before diagnosis in the

clinically diagnosed group was not discussed. The study did not assess the impact of this difference on the families. There may have been some selection bias in the sample, particularly in the clinically diagnosed group due to lack of follow up, and inclusion of only those under the care of a metabolic centre, which may be more severely affected cases.

Summary: Criterion 14 not met.

No direct evidence was identified on the wider impact of newborn bloodspot screening for PA on wider ethical, legal, or social issues. One comparative case series suggested that diagnosis as a result of screening occurs earlier than clinical diagnosis and this may mean a reduced diagnostic odyssey. One guideline noted that prenatal testing in PA is feasible, but no studies were identified which assessed or compared reproductive decision making after screen or clinical detection of PA in a family.

Conclusions

This report assesses newborn bloodspot screening for propionic acidaemia (PA) against select UK National Screening Committee (UK NSC) criteria for appraising the viability, effectiveness and appropriateness of a screening programme. It assessed key questions to determine if evidence published since 2004 supports a recommendation for newborn bloodspot screening for PA in the UK.

There was no clear cut evidence of a correlation between genotype (or enzyme activity) and phenotype, or of the benefit of early treatment following screening compared to treatment following clinical detection. Evidence on performance of the existing newborn bloodspot screening programmes for detecting PA was identified, as was evidence regarding natural history of symptom onset in PA.

The volume, quality and direction of evidence published since 2001 does not indicate that the evidence has changed sufficiently to warrant recommending newborn bloodspot screening for PA in the UK. Uncertainties remain across key criteria, including:

- Lack of a screening marker which can distinguish between PA, MMA, and other conditions. The same primary marker (C3) is used in PA and MMA screening, and is also raised in the neonatal form of multiple carboxylase deficiency and maternal vitamin B12 deficiency. In countries which screen for PA, there is not an agreed combination of markers and cut-offs for newborn bloodspot screening programmes. These programmes have thus far shown low positive predictive values for PA detection.
- Lack of studies on the implications of the carrier state identified as a result of screening. While newborn screening does not pick up heterozygous carriers of PA-causing mutations, parents of affected newborns will themselves be carriers, as may other members of the families who could be identified by cascade testing.
- Lack of clear evidence on the genotype-phenotype or PCC enzyme activity-phenotype relationship. This limits ability to predict prognosis and identify individuals who will have later onset or less severe manifestations of the condition, and on which to decide appropriate treatment.
- Lack of agreed criteria to decide who should receive treatment among asymptomatic cases and what treatment they should receive. All cases identified by screening are therefore likely to be treated similarly, despite differing clinical courses.
- Lack of robust evidence of a benefit of screen detection as opposed to clinical detection of PA. Studies suggest a high proportion of PA cases become symptomatic within the first week of life, and may therefore present before the screening results become available. No large, robust studies were identified which allowed comparison of treatment outcomes from screened and clinically detected populations. In order to establish the additional benefit of early treatment opportunities presented by screen detection, sufficiently large studies that assess variation in outcomes according to age of treatment initiation are necessary.
- Lack of direct evidence on the wider impact of newborn bloodspot screening for PA on wider ethical, legal, or social issues.

Methodology

The draft update report was prepared by Bazian Ltd., and then adapted in line with comments from the National Screening Committee.

Search strategies

Combined searches for both MMA and PA were performed, and the search strategies are given below.

MEDLINE

- 1 methylmalonic acidemia.ti,ab. (392)
- 2 methylmalonic acidaemia.ti,ab. (105)
- 3 methylmalonic aciduria.ti,ab. (449)
- 4 methylmalonicaciduria.ti,ab. (14)
- 5 mcm deficien\$.ti,ab. (11)
- 6 Amino Acid Metabolism, Inborn Errors/ (5059)
- 7 or/1-6 (5616)
- 8 Propionic Acidemia/ (89)
- 9 Methylmalonyl-CoA Decarboxylase/ (300)
- 10 propionic acidemia.ti,ab. (359)
- 11 propionic acidaemia.ti,ab. (129)
- 12 "hyperglycinemia with ketoacidosis and leukopenia".ti,ab. (1)
- 13 ketotic hyperglycinemia.ti,ab. (94)
- 14 pcc deficien\$.ti,ab. (15)
- 15 pa deficien\$.ti,ab. (46)
- 16 propionic aciduria.ti,ab. (38)
- 17 propionicacidemia.ti,ab. (12)
- 18 "propionyl-coa carboxylase deficien\$.ti,ab. (44)
- 19 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 (899)
- 20 7 or 19 (6158)
- 21 limit 20 to (english language and yr="2002 - 2015") (1337)
- 22 (neonat\$ adj2 screen\$.ti,ab. (3407)
- 23 (newborn\$ adj2 screen\$.ti,ab. (5051)
- 24 neonatal screening/ (7379)
- 25 mass screening/ (82342)
- 26 exp Infant, Newborn/ (498630)
- 27 25 and 26 (4768)
- 28 22 or 23 or 24 or 27 (15015)
- 29 exp Metabolism, Inborn Errors/ (142531)
- 30 (inborn adj2 error\$.ti,ab. (5145)
- 31 (metaboli\$ adj error\$.ti,ab. (259)
- 32 29 or 30 or 31 (144285)
- 33 28 and 32 (3036)
- 34 Tandem Mass Spectrometry/ (22480)
- 35 Spectrum Analysis, Mass/ (73663)
- 36 (mass adj2 spect\$.ti,ab. (172907)
- 37 (tandem adj2 mass).ti,ab. (30548)
- 38 (tandem adj2 ms).ti,ab. (2118)
- 39 34 or 35 or 36 or 37 or 38 (209941)
- 40 33 and 39 (541)
- 41 limit 40 to (english language and yr="2002 -Current") (444)
- 42 21 or 41 (1705)

EMBASE

Database: Embase <1996 to 2015 January 15>

Search Strategy:

-
- 1 methylmalonic acidemia.ti,ab. (403)
 - 2 methylmalonic acidaemia.ti,ab. (69)
 - 3 methylmalonic aciduria.ti,ab. (433)
 - 4 methylmalonicaciduria.ti,ab. (2)
 - 5 mcm deficien\$.ti,ab. (6)
 - 6 methylmalonic acidemia/ (581)
 - 7 or/1-6 (1089)
 - 8 propionic acidemia/ (547)
 - 9 methylmalonyl coenzyme A decarboxylase/ (47)
 - 10 propionic acidemia.ti,ab. (340)
 - 11 propionic acidaemia.ti,ab. (90)
 - 12 "hyperglycinemia with ketoacidosis and leukopenia".ti,ab. (0)
 - 13 ketotic hyperglycinemia.ti,ab. (76)
 - 14 pcc deficien\$.ti,ab. (6)
 - 15 pa deficien\$.ti,ab. (39)
 - 16 propionic aciduria.ti,ab. (37)
 - 17 propionicacidemia.ti,ab. (0)
 - 18 "propionyl-coa carboxylase deficien\$.ti,ab. (12)
 - 19 or/8-18 (749)
 - 20 7 or 19 (1620)
 - 21 limit 20 to (english language and yr="2002 -Current") (1296)
 - 22 newborn screening/ (10999)
 - 23 (neonat\$ adj2 screen\$.ti,ab. (3367)
 - 24 (newborn\$ adj2 screen\$.ti,ab. (6364)
 - 25 mass screening/ (28637)
 - 26 newborn/ (221179)
 - 27 25 and 26 (892)
 - 28 22 or 23 or 24 or 27 (13625)
 - 29 exp "inborn error of metabolism"/ (128452)
 - 30 (inborn adj2 error\$.ti,ab. (4710)
 - 31 (metaboli\$ adj error\$.ti,ab. (132)
 - 32 29 or 30 or 31 (130184)
 - 33 28 and 32 (3636)
 - 34 tandem mass spectrometry/ (38100)
 - 35 mass spectrometry/ (145782)
 - 36 (mass adj2 spect\$.ti,ab. (172908)
 - 37 (tandem adj2 mass).ti,ab. (33597)
 - 38 (tandem adj2 ms).ti,ab. (2336)
 - 39 34 or 35 or 36 or 37 or 38 (235512)
 - 40 33 and 39 (729)
 - 41 limit 40 to (english language and yr="2002 -Current") (627)
 - 42 21 or 41 (1852)
 - 43 limit 42 to exclude medline journals (113)

Cochrane library

UK NSC External Review

#1	"methylmalonic acidemia":ti,ab,kw	1	
#2	"methylmalonic acidaemia":ti,ab,kw	2	
#3	"methylmalonic aciduria":ti,ab,kw	1	
#4	methylmalonicaciduria:ti,ab,kw	0	
#5	MeSH descriptor: [Amino Acid Metabolism, Inborn Errors] this term only		12
#6	mcm next deficien*:ti,ab,kw	0	
#7	#1 or #2 or #3 or #4 or #5 or #6	12	
#8	MeSH descriptor: [Propionic Acidemia] this term only	0	
#9	MeSH descriptor: [Methylmalonyl-CoA Decarboxylase] this term only	1	
#10	"propionic academia":ti,ab,kw	0	
#11	"propionic acidaemia":ti,ab,kw	1	
#12	"hyperglycinemia with ketoacidosis and leukopenia":ti,ab,kw	0	
#13	"ketotic hyperglycinemia":ti,ab,kw	0	
#14	pcc next deficien*:ti,ab,kw	0	
#15	pa next deficien*:ti,ab,kw	0	
#16	"propionic aciduria":ti,ab,kw	0	
#17	propionicacidemia:ti,ab,kw	0	
#18	"propionyl-coa carboxylase deficien*":ti,ab,kw	0	
#19	#8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18		2
#20	#7 or #19	14	
#21	(neonat* near/2 screen*):ti,ab,kw	369	
#22	(newborn* near/2 screen*):ti,ab,kw	338	
#23	MeSH descriptor: [Neonatal Screening] this term only		271
#24	MeSH descriptor: [Mass Screening] this term only	4296	
#25	MeSH descriptor: [Infant, Newborn] explode all trees	13259	
#26	#24 and #25	119	
#27	#21 or #22 or #23 or #26	565	
#28	MeSH descriptor: [Metabolism, Inborn Errors] explode all trees		1795
#29	(inborn near/2 error*):ti,ab,kw	173	
#30	(metaboli* near error*):ti,ab,kw	172	
#31	#28 or #29 or #30	1844	
#32	#27 and #31	65	
#33	#32 or #20	58	

Quality

Several factors were assessed to determine the quality of the identified evidence, including study design and methodology, risk of bias, directness and generalisability of the evidence. Factors that were determined to be pertinent to the body of evidence identified for each criterion are outlined in the results section as well as the comment section of the Appendix tables. The overall level of evidence was assessed by considering the quantity, quality and consistency of evidence across the body of studies for each criterion reviewed.

Appendices

Appendix number	1
Relevant criteria	KQ1
Publication details	(1) Kraus JP, Spector E, Venezia S, et al. Mutation analysis in 54 propionic acidemia patients. <i>J Inherit Metab Dis.</i> 2012;35(1):51-63. ¹⁸ (2) Grunert SC, Mullerleile S, De Silva L, et al. Propionic acidemia: clinical course and outcome in 55 pediatric and adolescent patients. <i>Orphanet J Rare Dis.</i> 2013;8:6. ²¹
Study details	Case series, Europe. (Same case series as reported in Grunert et al. 2012 ²⁸)
Study objectives	(1) To describe the mutations identified in European PA patients (2) To describe the clinical course and outcome in paediatric and adolescent patients with PA
Inclusions	Individuals aged ≤20 years with PA seen at 16 metabolic centres in Germany, Austria, and Switzerland between September 2007 and March 2008.
Exclusions	Patients aged >20 years (as treatment more than 20 years ago may have differed).
Population	n=54 (Kraus et al. 2012) ¹⁸ n=55 (Grunert et al. 2013) ²¹ 35 identified by clinical presentation or family history (1 excluded from the mutation paper as mutation analysis was not carried out), and 20 identified by newborn screening
Intervention/test	NA
Comparator	NA
Results/outcomes	KQ1a (Kraus et al. 2012): Mutations identified in the 54 patients were reported, as were their PCC enzyme activities (range 0% to 2% of control activity where reported), and symptoms. 23 patients had mutations in PCCA, 30 patients had mutations in PCCB, 3 patients had mutations in both PCC genes, no mutations were identified in 4 patients. The authors reported that no genotype-phenotype correlations had been found and most of the mutations were private. KQ1b (Grunert et al. 2013): <ul style="list-style-type: none"> • Age at diagnosis ranged between 1 day and 8 years (median 7 days). • Median age at onset was 4 days • Only 4 patients (7%) have remained without acute symptoms by last follow up (age 6 months to 6.5 years) • 73% of patients were symptomatic in the first 5 days of life

	<ul style="list-style-type: none"> • 3 patients (5%) presented after day 90 (late onset) • Remaining patients (about 15%) presumably presented between 5 and 90 days [RC]
Comments	See comments for Grunert et al. 2012 ²⁸ below

Appendix number	
Relevant criteria	KQ1
Publication details	Perez B, Angaroni C, Sanchez-Alcudia R, et al. The molecular landscape of propionic acidemia and methylmalonic aciduria in Latin America. J Inherit Metab Dis. 2010;33(Suppl 2):S307-14. ¹⁹
Study details	Mutation analysis of Latin American PA patients.
Study objectives	To bring together clinical and genetic data for PA in Latin American patients and provide insight into the genotype-phenotype correlation.
Inclusions	Patients referred to Madrid from Argentina, Brazil, Chile, and Venezuela for genetic analysis.
Exclusions	None reported
Population	n=14 PA patients diagnosed after presenting clinically
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1a:</p> <ul style="list-style-type: none"> • 11 patients had mutations in PCCB, 3 in PCCA • Among the patients had early onset symptoms, “most” of them carriers of out-of-frame mutations: 4 carried two frameshift mutations, 1 carried one frameshift mutation and one missense mutation, and 3 carried no frameshift mutations (mostly missense, also one insertion which might have led to frameshift) • 5 late-onset PCCB patients were reported to be alive, and all were homozygous or hemizygous for the missense mutations p.E168K or p.R165W, 4/5 carried one frameshift mutation. One late onset PCCA case was alive and carried on missense and once frameshift mutation • 3 early onset cases were homozygous for the G407fs mutation in PCCB, and one carried one copy of this mutation. 3 late onset cases carried one copy of this mutation. <p>KQ1b:</p> <ul style="list-style-type: none"> • 8 cases (57%; 2 PCCA and 6 PCCB) had early onset symptoms: 7 (50%) in the first 5 days of life (includes 2 reported as “first days”) and 1 (7%) had onset at 2 weeks

	<ul style="list-style-type: none"> 6 cases (43%; 1 PCCA and 5 PCCB) had late onset symptoms: from 3 months to 9 months
Comments	<p>The paper did not provide much assessment of genotype-phenotype correlations. The small sample is likely to make identification of correlations difficult.</p> <p>The information on clinical presentation appeared to have been collected retrospectively.</p>

Appendix number	2
Relevant criteria	KQ1
Publication details	Chiu YH, Liu YN, Liao WL, et al. Two frequent mutations associated with the classic form of propionic acidemia in Taiwan. <i>Biochem Genet.</i> 2014;52(9-10):415-29. ⁴
Study details	Mutation analysis of Taiwanese PA patients.
Study objectives	To bring together clinical and genetic data for PA in Latin American patients and provide insight into the genotype-phenotype correlation.
Inclusions	Individuals from Taiwan with PA.
Exclusions	None reported
Population	n=10 PA patients including two sibling pairs (1 diagnosed prenatally, 3 detected by newborn screening, 6 presented clinically)
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1a:</p> <ul style="list-style-type: none"> 1 patient had PCCA mutations and 9 PCCB mutations All PCCB-deficient patients who carried the c.1301C→T mutation (6 patients) and/or the c.-4156_183+3713del mutation (3 patients, 2 also carrying c.1301C→T) showed low PCC enzyme activity (<1% to 9.3%) and had symptoms fitting the classic phenotype (poor feeding, hyperglycinemia, hyperammonemia, metabolic acidosis, and recurrent infection in their early age). Three patients with these mutations were reported to have mild to moderate developmental delay. <p>KQ1b:</p> <ul style="list-style-type: none"> 2/5 (40%) of those presenting clinically presented in the first 5 days of life, one at 2 weeks, one at 40 days, and one at 1 year and 8 months
Comments	<p>How the cases were selected for inclusion was not reported.</p> <p>The small sample is likely to make identification of correlations difficult, for example, making it difficult to be sure that mutations are consistently associated with “classic” PA. Mutation information was not available for 2 patients, but this information was available</p>

	for their younger siblings. The information on clinical presentation appeared to have been collected retrospectively.
--	--

Appendix number	3
Relevant criteria	KQ1
Publication details	Clavero S, Martinez MA, Perez B, et al. Functional characterization of PCCA mutations causing propionic acidemia. <i>Biochim Biophys Acta</i> . 2002;1588(2):119-25. ¹⁴
Study details	Expression analysis of 11 PCCA mutations with discussion of genotype-phenotype correlations
Study objectives	To establish the pathogenicity of 11 PCCA mutations (10 missense and an in-frame deletion) by expression studies.
Inclusions	PCCA mutations identified in individuals with PA.
Exclusions	None reported
Population	n=10 PA patients with PCCA mutations (possibly 2 additional patients who had the same mutation as one of the ten – reporting unclear)
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1a:</p> <ul style="list-style-type: none"> Overall, 4 cases with mild phenotypes had between 0.6% and 4% PCC activity in their fibroblasts or leukocytes; one person with a moderate phenotype had 1.3% PCC activity while the other had no PCC activity data, the 3 severe cases had between ≤1% and 2% activity (1 had no data). M373K and G631R changes are present in homozygous patients presenting a severe phenotype (these mutations led to null activity of PCC in expression studies, the patients showed 2% and ≤1% activity in their fibroblasts respectively) Patients with A138T (homozygous) and I164T (hemizygous with a frameshift mutation on the second allele) had a mild phenotype (these mutations led to partial residual activity in expression studies, 9% and 16%, respectively; and 4% activity in patients' fibroblasts). Steady-state protein was reported to be undetectable in these patients' fibroblasts. A75P was present in a heterozygous patient, with a moderate psychomotor delay (this mutation led to 26.7% residual activity in expression studies, but the protein appeared the least stable in cell-free studies; there was no PCC activity data from the patient) M229K was present with an as yet unidentified mutation in patient AG with a mild phenotype, very low PCC activity was detected in the patient's leukocytes (0.6%)

	<ul style="list-style-type: none"> • Patient MSH, homozygous for delEx21, had a moderate phenotype, despite expression studies suggesting that the resulting protein has no enzyme activity and is rapidly degraded; the patient’s fibroblasts showed 1.3% PCC activity • The authors concluded that “The high genetic heterogeneity in the PCCA gene defies the drawing of simplistic conclusions on the genotype–phenotype correlations.”
Comments	<p>The authors note that the establishment of genotype–phenotype correlations is hampered by genetic heterogeneity and the fact that most patients are compound heterozygotes.</p> <p>The paper focused mainly on expression studies and only discussed genotype–phenotype correlations in the discussion.</p> <p>The cases were selected for inclusion as they carried the mutations being studied, but it was not clear whether these were all the patients with these mutations whose data was available to the authors. The patients had been reported in previous publications, and it was unclear whether the authors only utilised the published data on these patients or had access to their records.</p> <p>The small sample is likely to make identification of correlations difficult.</p>

Appendix number	4
Relevant criteria	KQ1
Publication details	Clavero S, Perez B, Rincon A, et al. Qualitative and quantitative analysis of the effect of splicing mutations in propionic acidemia underlying non-severe phenotypes. Hum Genet. 2004;115(3):239-47. ¹⁵
Study details	The effects of splicing mutations in the PCCA and PCCB genes were assessed in vitro.
Study objectives	To analyse splicing mutations identified in PA patients to clarify their functional effects and involvement in the disease phenotype.
Inclusions	PCCA and PCCB splice site mutations identified in individuals with PA.
Exclusions	None reported
Population	n=4 PA patients with PCCA mutations (possibly 2 additional patients who had the same mutation as one of the ten – reporting unclear)
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1a:</p> <ul style="list-style-type: none"> • One patient (5626) homozygous for the IVS21+3del4 PCCA exon skipping mutation had 1.3% of normal PCC activity in fibroblasts (this mutation had been described in Clavero et al 2002, and shown to have no enzyme activity in vitro). Despite this, their condition was mild, they were alive and had adequate

	<p>psychomotor development. Analysis suggested some normal transcript was being made</p> <ul style="list-style-type: none"> • One patient (17475) is homozygous for the PCCA IVS22-2A>G exon skipping mutation, they had moderate phenotype with slight psychomotor delay , and residual PCC activity in fibroblasts of between 4% and 14% (suggesting that some normal transcript was made, confirmed by RT-PCR) • One patient (9132) carried PCCB mutations c.653A>G which affects splicing, and R165Q which was functionally null, and had 1.3% of normal PCC activity in fibroblasts. Despite this their condition was mild, they were alive and had adequate psychomotor development. The authors suggested that there might be other undiscovered effects of the splice site mutation. No normal transcript could be detected in RT-PCR or hybridization experiments. • One patient (13225) had PCCB mutations IVS10-11del6 and R410W, they had 11% residual PCC activity in their fibroblasts, and remained asymptomatic until age 5 when a sudden and fatal crisis occurred
Comments	<p>The authors suggest that results from other studies “consistently point to certain missense mutations retaining partial activity and which are associated with a less severe phenotype, while other missense mutations with no activity, along with nonsense and splicing mutations, out of frame insertions and deletions predictably resulting in truncated proteins, are associated with the most severe phenotypes.” However, only 4 studies were cited to support this, some were papers reported in this review¹⁴ or assessed and excluded.⁴⁴ Despite this statement, they also note that they observed remarkably mild phenotypes in the 4 patients assessed despite carrying splicing mutations which should cause major disruption of the protein. The presence of some correctly spliced transcript was identified in some cases which might contribute to the results seen.</p> <p>The paper focused mainly on expression studies and not specifically on genotype–phenotype correlation.</p> <p>The cases were selected for inclusion as they carried the types of mutations being studied, but it was not clear whether these were all the patients with these mutations whose data was available to the authors. Some of the patients had been reported in previous publications and there is overlap with at least one other paper reported here.¹⁴</p> <p>The small sample is likely to make identification of correlations difficult.</p>

Appendix number	5
Relevant criteria	KQ1
Publication details	Kim SN, Ryu KH, Lee EH, et al. Molecular analysis of PCCB gene in Korean patients with propionic acidemia.[Erratum appears in Mol Genet Metab. 2003 Mar;78(3):236-7]. Mol Genet Metab. 2002;77(3):209-16. ¹⁶
Study details	Mutation and protein analysis in Korean patients.
Study objectives	To identify whether there was any correlation between the biochemical, molecular and

	immunological features of individuals with PA and their clinical features.
Inclusions	Korean patients with PA
Exclusions	None reported
Population	n=8 PA patients
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1a:</p> <ul style="list-style-type: none"> • All patients had PCCB mutations • 5 patients had neonatal onset PA, the two who had PCC activity measured had 0% activity in lymphoblasts • 4/5 neonatal onset patients were homozygous for the Thr428Ile mutation, 2/3 of the late onset patients were heterozygous for this mutation • The other neonatal onset patient was homozygous for the 1527del3 mutation • The 3 patients with late onset PA had 1.6% to 9.6% PCC activity lymphoblasts, all three were compound heterozygotes <p>KQ1b:</p> <ul style="list-style-type: none"> • 1 case (12.5%) presented at 1 week, and the other 4 neonatal cases (50%) presented at between 1 week and 3 weeks • The 3 late onset cases (37.5%) presented at between 2 and 12 months
Comments	<p>It was unclear whether the cases described were the only Korean cases, or how they were selected for inclusion.</p> <p>The paper focused mainly on sequencing and expression studies and not as much on genotype–phenotype correlation.</p> <p>The small sample is likely to make identification of correlations difficult.</p>

Appendix number	6
Relevant criteria	KQ1
Publication details	Perez B, Desviat LR, Rodriguez-Pombo P, et al. Propionic acidemia: identification of twenty-four novel mutations in Europe and North America. <i>Mol Genet Metab.</i> 2003;78(1):59-67. ¹⁷
Study details	Mutation analysis in European and American PA patients.
Study objectives	To identify the mutations in European and American PA patients.
Inclusions	Patients with PA from Spain, Italy, Croatia, Belgium, Austria, USA and Brazil whose samples had been referred to the authors' laboratory for analysis.

Exclusions	None reported
Population	n=26 PA patients (15 European, 11 American)
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1a:</p> <ul style="list-style-type: none"> • 10 patients had mutations in PCCA, and 16 in PCCB • 30% of the patients with PCCA mutations were heterozygous, and 69% of the patients with PCCB mutations • The authors reported in their discussion that patients bearing null mutations in both alleles are associated with the most severe presentation of the disease, as previously described • However, they note that they have found exceptions, such as one patient with an in frame deletion of exon 23, who lacks an essential biotinylation domain, but had a mild phenotype. They had detected some normal transcript which may explain the discrepancy • The alleles G246V (homozygous), R165Q (hetero- and homo-zygous and, and E168K (heterozygous) were reported to be associated with a mild phenotype – although those carrying R165Q varied from early onset and mild outcome to late onset and moderate outcome • The alleles G112D (heterozygous), V107M (homozygous), and G188R (heterozygous) in PCCB were found in patients with early onset presentation • The authors suggest that genetic modifiers and environmental factors may significantly affect clinical presentation of PA
Comments	<p>The paper focused mainly on sequencing and expression studies and not as much on genotype–phenotype correlation.</p> <p>The small sample is likely to make identification of correlations difficult.</p>

Appendix number	7
Relevant criteria	KQ1
Publication details	Yorifuji T, Kawai M, Muroi J, et al. Unexpectedly high prevalence of the mild form of propionic acidemia in Japan: presence of a common mutation and possible clinical implications.[Erratum appears in Hum Genet. 2003 Jan;112(1):100.]. Hum Genet. 2002;111(2):161-5. ²⁰
Study details	Mutation analysis in Japanese PA patients with a mild form of the condition.
Study objectives	To assess the mutations associated with a mild form of PA in Japanese patients.

Inclusions	Patients with mild PA from Japan.
Exclusions	None reported
Population	n=10 mild PA patients, 8 of whom were identified by newborn screening. The 8 identified by screening were aged 1-3 years of age at the time of reporting and were still asymptomatic (natural protein intake was not restricted in two patients, and 0.8 to 2.0 g/kg body weight in the others). One of the non-screen detected patients presented with convulsion and left hemiparesis at age 10 months, had an episode of basal ganglia crisis at age 13, and was aged 14 years at the time of reporting (natural protein intake 1 g/kg body weight). The other was diagnosed in adulthood with mild mental retardation and abnormal cardiac scintigram (whether this was the first presentation, or just the presentation leading to diagnosis was not reported), the patient was 22 years old at the time of reporting and clinically stable (natural protein intake 1 g/kg body weight).
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1a:</p> <ul style="list-style-type: none"> • The patients had PCC activities in leukocytes between 2.4% and 12.3% of normal where available, this was reported to be higher than severe phenotype patients, but no figures were presented for comparison • Mutation spectrum was reported to be quite different from severe phenotype patients (authors' unpublished data) • 5 patients were homozygous for the Y435C allele, and present in the heterozygous state in 3 patients • Carrier frequency for the Y435C allele was 1/86.5 in a sample of 173 normal Japanese volunteers • In the other two patients, one had one mutation (A1288C) was identified which led to multiple exon skipping, and one which led to absence of mRNA
Comments	<p>The authors noted that their screening programme in Japan found an incidence of PA was more than 10 times that previously reported, and that the Y435C mutation accounted for many of the mutations in these patients. The exact number of affected individuals and their mutations was not reported.</p> <p>While the authors reported that mutation spectrum was different from that of severe (early onset) cases, the data for these was not reported. The lack of data for a late onset group for comparison means it is not possible to rule out that the Y435C occurs in this group as well.</p> <p>The authors suggest that even the currently asymptomatic cases may go on to show symptoms if not treated, as they have low residual PCC activity.</p> <p>The small sample is likely to make identification of correlations difficult.</p>

Appendix number	8
Relevant criteria	KQ1
Publication details	Rafique M. Clinical spectrum of propionic acidaemia. J Nutr Metab. 2013;2013:975964. ²³
Study details	Retrospective case series, Saudi Arabia
Study objectives	To describe natural history of PA patients in Saudi Arabia.
Inclusions	All patients admitted with PA at tertiary care hospital in Saudi Arabia between 2001 and 2012.
Exclusions	None reported
Population	n=26 PA patients. No newborn screening programme was reported to be in place, therefore presentation was likely to be clinical.
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1b:</p> <p>Age at clinical onset/diagnosis was:</p> <ul style="list-style-type: none"> • 84.6% in the first week of life • 92% presented in the first 2 weeks of life • 7.7% from the second week of life to 3 months • 7.7% after 3 months of age (late onset) <p>Average age at diagnosis was 0.13 year (about 7 weeks; range 2 days to 1 year)</p>
Comments	<p>Units of the average age at diagnosis were not reported, but this seems likely to be years given the range presented.</p> <p>Whether any of the patients were identified by screening or due to having an affected sibling was not reported. The authors call for more extensive centralised newborn screening, suggesting that such a programme is not in place.</p>

Appendix number	9
Relevant criteria	KQ1
Publication details	Vatanavicharn N, Liammongkolkul S, Sakamoto O, et al. Clinical characteristics and mutation analysis of propionic acidemia in Thailand. World J Pediatr. 2014;10(1):64-8. ²⁴
Study details	Retrospective case series, Thailand.
Study objectives	To describe natural history of PA patients in Thailand.
Inclusions	Patients with PA diagnosed at two hospitals in Thailand between 2000 and 2010, after clinical presentation.

Exclusions	None reported
Population	n=4 PA patients presenting clinically.
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1b:</p> <ul style="list-style-type: none"> All cases presented in the first (75%) or second (25%) weeks of life, although in some cases initial symptoms were attributed to sepsis or not attributed, and the diagnosis not made until later. Age at diagnosis was between 11 days and 11 months.
Comments	It was unclear whether this represented all cases seen in the study period, but this was implied.

Appendix number	10
Relevant criteria	KQ1
Publication details	Joshi SN, Venugopalan P. Clinical characteristics of neonates with inborn errors of metabolism detected by Tandem MS analysis in Oman. Brain Dev. 2007;29(9):543-6. ²⁵
Study details	Case series, Oman.
Study objectives	To describe natural history of PA patients in Oman.
Inclusions	Patients with PA diagnosed at one tertiary hospital in Oman between June 1998 and May 2005. The hospital was the only referral centre for inborn errors of metabolism (IEM) in Oman. Patients were referred due to having at least one major risk factor, either clinical, biochemical, or a family history.
Exclusions	None reported
Population	n=8 PA patients referred to the centre up to the age of 28 days (6 presenting with acute neonatal encephalopathy with or without seizures, hypoglycaemia, or ketosis), and 2 due to family history of known or suspected IEM).
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1b:</p> <ul style="list-style-type: none"> Clinically presenting cases all presented in the first week of life (median 2 days, range 1 to 4 days)
Comments	Very limited information was provided on the cases. Only neonates were included in the study (up to age 28 days) so this would exclude late presenting cases.

Appendix number	11
Relevant criteria	KQ1
Publication details	Sass JO, Hofmann M, Skladal D, et al. Propionic acidemia revisited: a workshop report. Clin Pediatr (Phila). 2004;43(9):837-43. ²²
Study details	Case series, central Europe.
Study objectives	To describe natural history of PA patients in central Europe.
Inclusions	Patients with PA attending 18 metabolic centres in central Europe between June 1998 and April 1999.
Exclusions	Two cases diagnosed due to having an affected older sibling were excluded from analyses of age at onset.
Population	n=49 PA patients
Intervention/test	NA
Comparator	NA
Results/outcomes	<p>KQ1b:</p> <p>Early onset cases:</p> <ul style="list-style-type: none"> • 86% (42/49) presented in the first 90 days of life (early onset)(possibly 1 case excluded due to older sibling) <ul style="list-style-type: none"> ○ 74% of these presented in the first 8 days of life ○ 26% of these presented between day 11 and day 90 • Median age at onset was 4 days (range 1 to 90 days; based on 38 patients for which this data available) • Median age at diagnosis was 13.5 days, range 4 to 195 days <p>Late onset cases:</p> <ul style="list-style-type: none"> • 12% (6/49) had late onset (1 with an older affected sibling) • Median age of onset was 16 months (range 11 to 69 months)
Comments	Data was from a period when newborn screening was not available.

Appendix number	12
Relevant criteria	KQ3
Publication details	Niu DM, Chien YH, Chiang CC, et al. Nationwide survey of extended newborn screening by

	tandem mass spectrometry in Taiwan. J Inherit Metab Dis. 2010;33(Suppl 2):S295-305.
Study details	Screening programme evaluation from 3 centres in Taiwan
Study objectives	To report the incidences and outcomes of inborn errors of metabolism in nationwide newborn screening in Taiwan
Inclusions	Newborns having screening between March 2000 and June 2009
Exclusions	None reported
Population	n=1,321,123 newborns
Intervention/test	<p>Newborn screening for PA, MMA and other conditions using tandem mass spectrometry (MS/MS) on dried blood spot collected by heel prick at 1 day after 1st feeding or 2-3 days of life. Preterm infants had a 2nd DBS taken and tested at 1 month of age.</p> <p><u>Cutoff values:</u></p> <p>The system used included two cut-offs: a 'positive' cutoff which warranted immediate referral, and a borderline cutoff – values between this and the positive cutoff resulted in repeat DBS sampling and testing.</p> <p>Borderline values initially set at mean + 4 standard deviations (SD) and positive cut-offs at twice borderline values. These cut-offs were modified over time to minimise false positives and negatives.</p> <p>C3 borderline: $\geq 7 \mu\text{M}$ (2 centres), $\geq 4.74 \mu\text{M}$ (1 centre)</p> <p>C3 positive: $\geq 12 \mu\text{M}$ (2 centres) $> 8.8 \mu\text{M}$ (1 centre)</p> <p>Or (in 2 centres)</p> <p>C3/C2 borderline: ≥ 0.2 (2 centres)</p> <p>C3/C2 positive: ≥ 0.3 (1 centre), ≥ 0.25 (1 centre)</p> <p>Borderline values needed to be confirmed in a repeat blood sample before an infant was referred. The timing of collection of the repeat sample was not reported.</p> <p>Confirmatory tests are reported below as the 'Comparator' for the screening test results as they formed the gold standard against which false and true positives and negatives were judged.</p>
Comparator	Confirmatory tests included UOA, acyl carnitine analysis before and after 3 day carnitine loading (100 mg/kg per day; used because acylcarnitine results tended to improve on repeated samples in many OA cases), serum B12 and plasma homocysteine analysis.
Results/outcomes	<p>KQ3:</p> <p>Requiring confirmation of borderline values in a repeat sample cut the number of cases requiring referrals dramatically (from 3,114 to 39).</p> <p>Results below are for the process using confirmed borderline values and positive values to indicate referral.</p> <p>True positives: 2 PA</p> <p>False positives: 73 (these numbers include 14 cases of MMA)</p> <p>True negatives: 1,321,048</p>

	<p>False negatives: 0 (none reported)</p> <p>Sensitivity: 100% [RC]</p> <p>Specificity: 99.994% [RC]</p> <p>PPV: 2.67% [RC]</p> <p><u>Incidence:</u></p> <p>PA: 1 per 660,561 [RC]</p>
Comments	No systematic approaches to identifying false negatives was described, therefore some may be missed, particularly among those with a single borderline C3 or C3/C2 value.

Appendix number	13
Relevant criteria	KQ3
Publication details	Weisfeld-Adams JD, Morrissey MA, Kirmse BM, et al. Newborn screening and early biochemical follow-up in combined methylmalonic aciduria and homocystinuria, cblC type, and utility of methionine as a secondary screening analyte. Mol Genet Metab. 2010;99(2):116-23.
Study details	Screening programme evaluation from the USA
Study objectives	To report on newborn screening data from 10 patients with CblC born in New York State since 2005, including description of the impact of incorporating a new secondary marker for CblC/D/F
Inclusions	Newborns screened in New York State between January 2005 and December 2008
Exclusions	None reported
Population	N=1,006,325 newborns
Intervention/test	<p>Newborn DBS screening for PA, MMA (including CblC), and other conditions using MS/MS.</p> <p><u>Cutoff values:</u></p> <p>The C3 cutoff value was mean of a normal population sample plus about 8 SD (7 μM)</p> <p>Category 1: C3>7 μM and C3/C2 >0.2 (immediate referral)</p> <p>Category 2: C3>7 μM and C3/C2 <0.2</p> <p>Category 3: C3 in range of 5 to7 μM and C3/C2 >0.2</p> <p>For Categories 2&3, a repeat sample is requested and tested and the infant referred if C3 >7 μM , or C3 5-7 μM and C3/C2 >0.2. The timing of collection of the second sample was not reported.</p> <p>In 2005 C4DC was added as a secondary marker for MMA/PA, a value of C4DC >1 μM on initial and repeat sample leads to referral.</p> <p>From late 2008 (variably reported as September or November), methionine (Met) was added as a secondary marker to identify CblC, D, and F diseases. Levels of Met were considered in those falling into Category 3 before repeat sample was requested. The following approach was then taken:</p>

	<p>Category 3a: Met >13.4 µM, repeat sample requested and tested as for Category 3 above</p> <p>Category 3b: Met <13.4 µM, immediate referral (suspicion of CbIC, D, or F)</p> <p>This aimed to reduce time to referral in those with a higher suspicion of being CbIC, D, or F cases.</p> <p>Met was already being assessed in the MS/MS screen as a marker for cystathione beta synthase deficiency when elevated above a specified cutoff.</p>
Comparator	Confirmatory tests were not explicitly reported, but DNA analysis was used, as was complementation testing.
Results/outcomes	<p>True positives: 5 [RC]</p> <p>False positives: 150 (includes 26 MMA cases) [RC]</p> <p>False negatives: 0 (None reported)</p> <p>True negatives: 1,006,143</p> <p>There were 27 screen positives with screening results unresolved, including 2 who had died, 9 lost to follow up, and 16 'remaining open' at the time of publication. Figures have been calculated excluding these from the analyses</p> <p>Sensitivity: 100% [RC]</p> <p>Specificity: 99.997% [RC]</p> <p>PPV: 3.23% [RC]</p>
Comments	<p>No systematic follow up of screen negatives, or monitoring for false negatives was reported, so these may have been missed.</p> <p>The inability to include the 26 unresolved screen positives in the analyses may mean the accuracy results are not truly representative of test performance</p>

Appendix number	14
Relevant criteria	KQ3
Publication details	Frazier DM, Millington DS, McCandless SE, et al. The tandem mass spectrometry newborn screening experience in North Carolina: 1997-2005. J Inher Metab Dis. 2006;29(1):76-85.
Study details	Screening programme evaluation from the USA
Study objectives	To assess the performance of the newborn screening programme in North Carolina over 8 years
Inclusions	Newborns undergoing screening from July 1997 to July 2005
Exclusions	None reported
Population	n=944,078 newborns
Intervention/test	Newborn screening for PA, MMA and other conditions through first and second tier tests using tandem mass spectrometry (MS/MS) on dried blood spot.

	<p><u>Cut-off values:</u></p> <p>Initially when the pilot started in 2002 the cut-offs were set at about mean + 4SD, with mean based on 2,000 samples. These were modified over time as more samples were analysed and more clinical data was available.</p> <p>For 2003-2004 the cutoff values were as below:</p> <p>‘Diagnostic’ cutoff values:</p> <p>C3 >9.0 µM</p> <p>And</p> <p>C3/C2 >0.15</p> <p>Borderline cutoff values:</p> <p>C3 >4.82 µM</p> <p>And</p> <p>C3/C2 >0.15</p> <p>(Elevated C3/C2 ratios alone are not considered significant).</p> <p>If an elevated result was obtained, the test was repeated on a second disc from the same sample card.</p> <p>For confirmed borderline results a second sample was requested for testing, the exact timing of this sample was not reported. If this was also borderline the same procedure as for diagnostic cut-offs was followed.</p> <p>For confirmed results above the diagnostic cutoff the infant’s primary care physician was contacted with results and recommendations so that the infant could be referred to a metabolic centre as needed, ideally within 24-48 hours.</p>
<p>Comparator</p>	<p>Confirmatory tests included a repeat of the MS/MS screen, UOA, and PACYLC, serum B12, and plasma homocysteine. Enzyme and mutation analyses were carried out where available and approved by reimbursers. Results of the latter tests were not needed to start treatment, as the aim was to start treatment within 10-14 days of birth.</p>
<p>Results/outcomes</p>	<p><u>Screening test performance:</u></p> <p><u>For 2003-2004:</u></p> <p>True positives (TP): 1</p> <p>False positives (FP): 8 (includes 5 MMA) [RC]</p> <p>True negatives (TN): 239,406 [RC]</p> <p>False negatives (FN): 0</p> <p>Sensitivity: 100% [RC]</p> <p>Specificity: 99.997% [RC]</p> <p>Positive predictive value (PPV): 3.70% [RC]</p> <p><u>Incidence:</u></p>

	PA: 1 in 300,000
Comments	<p>Limited information on false positives (for 2003-2004 only) was provided. This limited the analyses of test performance which were possible.</p> <p>No systematic follow up of screen negatives or alerting system for false negatives was reported, so they may be under ascertained. However, some false negatives of other disorders (not PA) were identified.</p>

Appendix number	15
Relevant criteria	KQ1, KQ3
Publication details	Lund AM, Hougaard DM, Simonsen H, 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):281-93.
Study details	Screening programme evaluation from Denmark, the Faroe Islands and Greenland
Study objectives	To present results of expanded newborn screening programme data from a 9 year period
Inclusions	Newborns born in Denmark, the Faroe Islands and Greenland between 1 st February 2002 and 31 st March 2011 whose parents accepted NBS
Exclusions	Non-participation in newborn screening (n=82,930)
Population	n=504,049 newborns
Intervention/test	<p>Newborn screening for PA, MMA and other conditions using tandem mass spectrometry (MS/MS) on dried blood spot collected by heel prick at 4-9 days of age (trial period 2002-2009; median age 5 days) or 2-3 days of age (routine screening period 2009-2011; median age 2.5 days). Results were available 2-7 days (trial period) or 2-3 days (routine period) after samples received by the screening lab.</p> <p><u>Cutoff values:</u> These were initially set based on literature, in-house experience on stores DBS samples from affected individuals and statistical assessment of reference cohorts. Cut-offs were adjusted over time to optimise screening performance and because of changing analytical method (different MS/MS machines and analysis kits used over time). The main change was after the trial period where samples were collected earlier and an underivatized screening method was used rather than derivatised. Performance of the screening lab was monitored by the CDC Newborn Screening Quality Assurance Program.</p> <p><u>Trial period:</u></p> <p>Primary: C3 >5.1 µM</p> <p>Secondary: C3/C2 >0.35 or C4DC >0.4U</p> <p><u>Routine period:</u></p> <p>Primary: C3 >6.0 µM</p>

	<p>Secondary: C3/C2 >0.25</p> <p>Confirmatory tests are reported below as the 'Comparator' for the screening test results as they formed the gold standard against which false and true positives and negatives were judged.</p>
<p>Comparator</p>	<p>Confirmatory tests (diagnostic 'gold standard'): DBS with values above cut-offs were reanalysed in duplicate, and if abnormality was confirmed the diagnosis was confirmed or ruled out through testing of urine organic acids (UOA), plasma acylcarnitines (PACYLC), plasma amino acids (PAA), DNA, and propionate incorporation into fibroblasts. If the disease could not be confirmed in these tests the screening result was considered a false positive.</p> <p>False negatives were also recorded where a child had had screened negative, but was later identified as being affected by population screening of through affected families.</p> <p>Historical cohort: Infants who had clinical diagnoses of the disorders during the study period, but who were born before the expanded NBS panel was introduced, between January 1st 1992 and December 31st 2001, were used as a comparator in terms of presentation. The screening centre also carried out diagnosis and treatment of all children from the participating areas, so the majority or all of affected children diagnosed in this period were believed to have been identified.</p>
<p>Results/outcomes</p>	<p>KQ1:</p> <p><u>Presentation and outcome of infants identified or missed by screening:</u></p> <ul style="list-style-type: none"> • 1/2 (50%) presented clinically before screening results were available at 4 days of age. The other case presented at 6 months of age with metabolic decompensation which left him developmentally delayed – his initial screening sample had not been properly processed due to a technical error, but the stored sample was positive on re-testing. • 1/2 (50%) was well with normal psychomotor development and developmental age at last follow up; timing of last follow up was not reported and no formal cognitive testing was carried out • There were no deaths <p><u>Historical cohort:</u></p> <p>1 case of PA was identified, who presented at age 4 days with metabolic decompensation. There were 674,754 births in the period.</p> <p><u>Incidence of MMA/PA:</u></p> <p>Screening cohort: 1 in 100,809 in those receiving screening, and 1 in 65,219 if the case in the screening refusal group included</p> <p>Historical cohort: 1 in 61,341</p> <p>KQ3:</p> <p><u>Screening test performance:</u></p>

	<p>True positives (TP): 2</p> <p>False positives (FP): 52 (includes 3 MMA cases)</p> <p>True negatives (TN): 503,992</p> <p>False negatives (FN): 0</p> <p>Sensitivity: 50% [RC]</p> <p>Specificity: 99.99% [RC]</p> <p>PPV: 5.56% [RC]</p>
Comments	<p>Although there was not systematic follow up of the screen negative infants, the fact that diagnostic tests and management of metabolic disorders for all three regions seemed to be carried out at the centre conducting the study, increases likelihood that missed cases would be detected and also known cases in the historical period assessed. In both periods individuals who died from or with the conditions without having ever been screened or diagnosed would have been missed.</p> <p>Exact treatment received by affected individuals was not reported.</p> <p>The authors noted that the high false positive rate could be reduced by a second tier MS/MS assessment of methylmalonic acid, 2-methylcitric acid, and total homocysteine in dried blood spots, but this would further delay availability of screening results. They suggest that before introducing such an assessment the basis for inclusion of MMA and PA in the screening panel should be reviewed.</p> <p>The incidence of the conditions did not differ between the screening and non-screening time periods, suggesting that additional cases might not be being detected by screening. However, given the rarity of these individual conditions, even larger numbers of births might be needed to be able to detect differences.</p>

Appendix number	16
Relevant criteria	KQ1, KQ3
Publication details	Schulze A, Lindner M, Kohlmuller D, et al. Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. <i>Pediatrics</i> . 2003;111(6 Pt 1):1399-406.
Study details	Screening programme evaluation from Germany
Study objectives	To assess the impact of an expanded newborn screening programme on detection of IEM in one region in Germany over a 42 month period, and to assess outcome of those diagnosed with IEM
Inclusions	Newborns born in Baden-Wurtemberg, Germany between April 1998 and September 2001 whose parents accepted NBS
Exclusions	Non-participation in newborn screening (n=16,200)
Population	n=250,000 newborns

Intervention/test	<p>Newborn screening for PA, MMA and other conditions using tandem mass spectrometry (MS/MS) on dried blood spot collected by heel prick at 3-7 days of age (median 5 days; 0.88% carried out before day 1, and 1.65% after day 7). Preterm infants had a second sample taken after 14 days of age.</p> <p><u>Cutoff values:</u></p> <p>Cutoff set at 99.5th percentile of values from MS/MS on DBS from 10,000 healthy neonates. These were reported as:</p> <p>C3 >6.8 μM (MMA and PA)</p> <p>C3/C0 >0.19 (MMA and PA)</p> <p>C3/C2 >0.39 (MMA and PA)</p> <p>Methylmalonylcarnitine >1 (MMA only) (analyte likely to be C4DC, containing methylmalonyl and succinylcarnitine)</p> <p>Repeat analysis was carried out on samples with values above the cutoff for one or more analytes.</p> <p>If the 1st test was >30% over the cutoff and the 2nd test was normal, a 3rd test was carried out and the mean of the three values used.</p> <p>Samples where both tests were positive for any analyte were taken to the next step: assessment by an experienced metabolic disease specialist.</p> <p>The specialist used a rating system based on level of deviation from normal and overall analyte profile to decide if the sample was a screen positive. The system was optimised based on measurements of infants with known disorders, reports in the literature, previous screening test performance (sensitivity and specificity), and assessor experience.</p> <p>A second DBS was requested from screen positives, or the infant was referred for hospital admission if avoidance of delay was felt to be essential for their wellbeing (conditions falling into this latter category not reported).</p> <p>Confirmatory tests are reported below as the 'Comparator' for the screening test results as they formed the gold standard against which false and true positives and negatives were judged.</p>
Comparator	<p>Confirmatory tests (diagnostic 'gold standard'): Diagnosis was confirmed or ruled out through testing of UOA (MMA: methylmalonic acid, PA: tiglylglycine, 3-hydroxypropionic acid, methylcitrate), plasma testing (MMA: ammonia, homocysteine; PA: ammonia), and enzyme activity in fibroblasts.</p> <p>Monthly questionnaires were sent to all pediatric hospitals and metabolic centres in Germany to identify any children with IEM missed by the screening programme (false negatives).</p>
Results/outcomes	<p>KQ1b:</p> <p><u>Presentation and outcome of infants identified by screening:</u></p> <ul style="list-style-type: none"> The 1 PA case was symptomatic and diagnosed before the result of screening was

	<p>available</p> <p>Details of symptoms were not provided</p> <p>KQ3:</p> <p><u>Screening test performance:</u></p> <p>True positives (TP): 1</p> <p>False positives (FP): 209 (includes 3 confirmed and 1 suspected MMA) [RC]</p> <p>True negatives (TN): 249,790 [RC]</p> <p>False negatives (FN): 0</p> <p>Positive predictive value (PPV): 0.48% [RC]</p> <p>Sensitivity: 100% [RC]</p> <p>Specificity: 99.916% [RC]</p>
Comments	<p>Although there was not systematic follow up of all of the screen negative infants, the fact that monthly questionnaires were sent to centres where the infants might be treated (paediatric hospitals and metabolic centres), increases likelihood that missed cases would be detected.</p> <p>This pilot explicitly used a rating system based on various criteria (including cut-offs) and assessed by a metabolic specialist in order to assign screen positives, rather than based on the cutoff value alone. To what extent results differed from what would have been obtained just using set cut-offs alone was not investigated.</p> <p>A later paper³⁹ reported that due to an unfavourable balance of sensitivity and specificity, disorders of propionate metabolism assessed by elevated C3 levels on NBS were excluded from the conditions recommended to be screened for nationally in German guidelines from 2005. The paper aimed to develop a statistical approach to new parameter combinations to give improved specificity and 100% sensitivity for disorders of propionate metabolism (not reported here as it was not an actual screening programme).</p>

Appendix number	17
Relevant criteria	KQ1
Publication details	Hoffmann GF, von Kries R, Klose D, et al. Frequencies of inherited organic acidurias and disorders of mitochondrial fatty acid transport and oxidation in Germany. <i>Eur J Pediatr.</i> 2004;163(2):76-80. ²⁶
Study details	Case series with historical control group, Germany
Study objectives	To compare numbers of clinically diagnosed cases with those detected by newborn screening, to give an estimate of the maximum and minimum frequency of clinical disease in candidate conditions for expanded newborn screening.
Inclusions	Infants identified as having diagnoses of OA or fatty acid disorders in 1999 to 2000 in Germany through new clinical presentation, and detected by active surveillance (monthly

	<p>inquiries to all paediatric departments, 3-monthly inquiries to all metabolic laboratories). Information on diagnoses and short term outcomes were obtained from physicians.</p> <p>Screened cases were identified through newborn MS/MS screening in two regions in Germany in the same period.</p>
Exclusions	Asymptomatic siblings of clinically diagnosed cases were excluded.
Population	N=3 cases of PA detected by screening and 5 PA cases detected clinically.
Intervention/test	Newborn MS/MS screening.
Comparator	Clinical detection of cases.
Results/outcomes	<p>KQ1:</p> <ul style="list-style-type: none"> • 3/5 (60%) clinically detected cases of PA presented within 7 days of birth with severe metabolic decompensation • 2/5 (40%) clinically detected cases of PA presented more than 7 days after birth (15 and 31 days at diagnosis), and both were reported to have had metabolic crisis • No clinically symptomatic case were reported to be detected in the screened cohort – it was unclear whether this meant that no false negatives were identified, or that all of the screen detected cases (including the 2 PA cases) were asymptomatic at the time of detection or afterwards.
Comments	<p>Whether both groups had similar assessment and monitoring of outcomes was unclear. Outcome data was very limited and appeared to only represent information about presentation for clinically detected MMA cases, and particularly unclear for the screened cases.</p> <p>It was unclear whether the asymptomatic siblings of clinically diagnosed cases were tested for MMA. If they had the condition and were excluded, this could bias towards poorer outcome for the unscreened group, as those included were all symptomatic.</p> <p>Any treatment received was not reported.</p> <p>The small number of cases and the limitations above make it difficult to draw firm conclusions about differences in the outcome of screen detected and clinically detected cases, particularly relating to the impact of treatment.</p>

Appendix number	18
Relevant criteria	KQ1, KQ4
Publication details	Dionisi-Vici C, Deodato F, Roschinger W, et al. 'Classical' organic acidurias, propionic aciduria, methylmalonic aciduria and isovaleric aciduria: long-term outcome and effects of expanded newborn screening using tandem mass spectrometry. <i>J Inherit Metab Dis.</i> 2006;29(2-3):383-9. ²⁷
Study details	Comparative case series

Study objectives	To compare the natural history of patients with classical organic acidurias (including PA) diagnosed on a clinical basis with those diagnosed through newborn MS/MS screening.
Inclusions	Clinically detected cases of PA (and other organic acidurias) from a children's hospital in Italy from 1983, and cases detected by newborn screening in Australia and Germany (time period not stated).
Exclusions	NR
Population	N=13 clinically diagnosed PA cases and 7 cases of PA detected by screening. (The latter may include infants diagnosed due to affected older sibling).
Intervention/test	Newborn MS/MS screening.
Comparator	Clinical detection of cases.
Results/outcomes	<p>KQ1b:</p> <ul style="list-style-type: none"> • 4/7 (57%) screen detected cases were symptomatic before the result of screening was known • Age at onset of clinical cases with PA was not reported <p>KQ4:</p> <p><u>Clinically detected cases:</u></p> <p>Among the 9 PA cases with neonatal onset for which this information was presented:</p> <ul style="list-style-type: none"> • 3/9 died (33%, one at about age 6, one at age 9, and one at about age 15 years), the 6 survivors were aged between 4 and 14 years. • 8/9 (89%) had at least one coma, 8/9 (89%) had at least one metabolic decompensation (the one who was not indicated as having a decompensation had two comas and died, which may indicate decompensation), and 7/9 (78%) had multiple metabolic decompensations and the 2 who did not die at around ages 6 and 9 <p>No information on the metabolic outcomes or mortality of the remaining 4 PA cases was reported, including any late onset PA cases.</p> <p>3 PA cases (3/9, 33%) were described as having cardiomyopathy. One child with normal cognitive level died suddenly at age 9, they had had only one metabolic decompensation (a severe episode neonatally).</p> <p>1/9 PA case (11%) had acute pancreatitis</p> <p><u>Screen detected cases:</u></p> <p>The 7 cases appeared to be aged from neonates to around 4-5 years of age, outcomes included:</p> <ul style="list-style-type: none"> • 2/7 (22%) died, at ages 5 days and 13 days • 5/5 (100%) survivors had metabolic instability: 2/5 (40%) had mild metabolic instability, 3/5 (60%) had severe metabolic instability (neither term defined, nor

	<p>how this related to metabolic decompensation)</p> <ul style="list-style-type: none"> • 1/5 (20%) survivors had mild neurocognitive impairment, 1/5 (20%) had severe neurocognitive impairment (terms not defined), and 3/5 (60%) did not have neurocognitive impairment
Comments	<p>It was unclear whether the cases described were all of the cases detected, or a subset.</p> <p>Whether both groups had similar assessment and monitoring of outcomes was unclear. Different outcomes were reported for the groups (e.g. metabolic decompensations were reported for the clinically detected group, while “metabolic instability” was reported for the screen detected group).</p> <p>The time period in which the cases were detected was unclear, as the clinically detected group were diagnosed from 1983, when MS/MS would not have been available until later, it is likely the screened and clinically diagnosed groups are not contemporaneous. As a result, the duration of follow up is likely to have been shorter for the screen detected cases.</p> <p>It was also unclear whether the treatment regimens were the same across the countries or groups.</p> <p>These limitations and the small number of cases make it difficult to draw firm conclusions about differences in the outcome of screen detected and clinically detected cases, particularly relating to the impact of treatment.</p>

Appendix number	19
Relevant criteria	KQ4
Publication details	Waisbren SE, Albers S, Amato S, et al. Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. <i>Jama</i> . 2003;290(19):2564-72. ²⁹
Study details	Prospective comparative case series (comparing inception cohorts), USA
Study objectives	To compare newborn identification by expanded screening with clinical identification of biochemical genetic disorders.
Inclusions	Children identified with biochemical genetic disorders (including MMA and PA) through expanded newborn screening in 3 states (2 in New England plus Pennsylvania), and children identified clinically in any of the 6 New England states between January 1999 and June 2002, and assessed by December 2002. Parents of diagnosed children were invited to participate between 5 and 30 months after diagnosis. 82% of the newborn screening group (those not enrolled reported to have the same diagnoses as those enrolled, except for 2 disorders [not MMA or PA]) and 97% of the clinically identified group agreed to participate. Children were given standard medical examination as well as having their medical records obtained for assessment.
Exclusions	Infants who died before enrolment (5 from screened group – none with MMA or PA), children from centres which failed to obtain internal approval for the study (43 families of

	clinically identified children, diagnoses NR) and those who could not be contacted (10 children in the newborn screening group, diagnoses NR).
Population	N=3 PA cases in screened group and 6 PA cases in clinically diagnosed group
Intervention/test	Newborn MS/MS screening.
Comparator	Clinical detection of cases.
Results/outcomes	<p>KQ4:</p> <ul style="list-style-type: none"> • 1 child (1/6, 16.7%) with clinically detected PA died during the study, while none of those with screen detected PA (0/3, 0%) were reported to have died (no statistical comparisons provided) • 2 children (2/6, 33.3%) with clinically detected PA required NICU care before diagnosis compared to none of those with screen detected PA (0/3, 0%; no statistical comparisons provided) • At least one child with PA up to the age of 3 and at least one over the age of 3 (exact number not specified, unclear if the same individuals) had developmental status in the range of mental retardation. None of the children with PA in the screen detected group (0/3, 0%; no statistical comparisons provided) were in this range
Comments	<p>The study did report other outcomes, but for children with metabolic disorders as a whole and not for PA separately. There was no statistical analysis for the PA outcomes separately.</p> <p>Age at diagnosis and evaluation was lower for the newborn screened group than the clinically detected group as a whole (diagnosis: median 5 days [range 1 to 180] screened vs. 4 months [range 0.1 months to 5.9 years] clinically detected; evaluation: median 9 months [range 5 to 91] screened vs. 34 months [range 4 to 101] clinically detected, $p < 0.001$ for both). These figures were not reported for PA specifically; therefore it is difficult to determine whether the children being compared with these diagnoses are comparable in terms of age.</p> <p>Follow up was reported to be short (exact duration not specified).</p> <p>Details of treatments received were not provided.</p>

Appendix number	20
Relevant criteria	KQ1, KQ4, KQ6
Publication details	Grunert SC, Mullerleile S, de Silva L, et al. Propionic acidemia: neonatal versus selective metabolic screening. <i>J Inherit Metab Dis.</i> 2012;35(1):41-9. ²⁸
Study details	Comparative case series, Europe. A questionnaire was used to extract retrospective clinical information from medical records. Patients (n=52) were examined clinically during a routine outpatient visit and 3 were hospitalised during the study visit. Psychological testing was carried out using the Snijders-Omen test for those aged 2.5 to 8.5 years old,

	<p>and the culture-free test in those aged older than this. Whether these clinical and psychological assessments were carried out prospectively was not clearly stated.</p> <p>Comparisons were made between screen detected and non-screen detected groups, and between individuals with different ages at diagnosis (regardless of presentation method leading to diagnosis). Deaths from PA at participating centres in the last 20 years were assessed to see whether there was bias due to only following patients remaining alive and under the care of a centre at the time of the study.</p>
Study objectives	To assess the potential benefits of expanded newborn screening for propionic acidemia.
Inclusions	Individuals aged ≤ 20 years with PA seen at 16 metabolic centres in Germany, Austria, and Switzerland between September 2007 and March 2008.
Exclusions	Patients aged >20 years (as treatment more than 20 years ago may have differed).
Population	<p>55 individuals (35 male, 25 female) with PA from 50 families in Germany, Austria, and Switzerland seen at 16 centres between September 2007 and March 2008.</p> <p><i>Newborn screening (NBS) group:</i></p> <p>20 were diagnosed as a result of NBS. Median age 3 years 9 months. At this point in most cases screening was being performed by individual centres and not as part of an official programme.</p> <p><i>Comparator group:</i></p> <p>35 were diagnosed as a result of clinical findings, family history, or routine laboratory tests (referred to as “clinically detected” here for simplicity).</p> <p>Seven of these were identified immediately after birth due to an older sibling with the disease, and one at age 5 years as a result of positive NBS result for a younger sibling – these individuals were excluded from statistical comparisons of the groups.</p> <p>Median age was 8 years 10 months (unclear if age at investigation or follow up). Age at investigation was between 5 days and 18 years 7 months. It was unclear whether these figures included those diagnosed as a result of a sibling.</p>
Intervention/test	Newborn MS/MS screening.
Comparator	Clinical detection of cases.
Results/outcomes	<p>KQ1b:</p> <ul style="list-style-type: none"> 63% (12/19) of the screen detected group were symptomatic by the time of positive screening results, and 5% (1/19) developed symptoms on the day of the result. 21% (4/19) were diagnosed while asymptomatic but developed symptoms later, and 11% (2/19) remained asymptomatic The authors suggest that this indicates that NBS is likely to offer benefit to milder phenotypes with later clinical presentation, and that their findings could not verify a reduction in the effects of the initial metabolic decompensation with NBS Overall 93% of patients showed symptoms (51/55), while 7% (4/55) are asymptomatic at age 6 months to 6.5 years (2 in NBS group, 2 in ‘clinically

detected' group identified through siblings). Whether and how the asymptomatic individuals were being treated was not reported.

- In 73% of symptomatic patients, symptoms were noted in the first 5 days of life (unclear if this included symptomatic patients on both groups)
- There was no significant difference in residual PCC activity in the groups for the individuals with this data available (0.9% in NBS group [n=7] vs. 1% in clinical group [n=14], p=0.58)

	Clinically detected	NBS†
Symptomatic*	94% (33/35)	90% (17/19)
<i>Early onset</i>	91% (31/34)	79% (15/19)
<i>Late onset</i>	3% (1/34)	11% (2/19)
Asymptomatic	6% (2/35)	10% (2/20)

* One individual in the clinically detected group reported as symptomatic, but age at onset not known. Early onset refers to diagnosis within first 90 days of life. †Unclear why NBS group denominator varied by 1 in different categories.

KQ4:

Effect of age at diagnosis:

- There was no significant difference in number of metabolic decompensations between those diagnosed before 5 days and those diagnosed after (numbers and p value not reported), with the same findings for groups diagnosed before and after 6 days or 7 days.

Effect of NBS vs. clinical detection:

- The number of **metabolic decompensations** was reported to be similar in both groups (overall numbers and p value not reported). The authors suggested that this indicated that earlier diagnosis through NBS could not prevent metabolic crises in infancy. Median number of crises in 1st year of life was 2 in both groups (range NR), and decreased in 2nd and 3rd years of life.
- The number of **hospital admissions** was also reported to be similar in both groups, although the range was larger in the clinically detected group (overall numbers, range and p value not reported).
- There was a trend towards lower **plasma ammonia concentration in the initial metabolic crisis** in those detected by NBS, but this was not significant (median 207 µmol/L [range 56 to 962] with NBS vs. 392 µmol/L [range 53 to 1,936] with clinical detection; p=0.065; number of people analysed not reported)
- **School attended:** More clinically detected cases attended a kindergarten for children with special needs once they reached 3 years and over (21% with NBS vs. 55% with clinical detection; p value not reported; normal kindergarten: 57% with

	<p>NBS vs. 42% with clinical detection; status of remaining 22% of NBS group and 3% of clinically detected group unclear)</p> <ul style="list-style-type: none"> • Most participants aged 6 years and over attended a special school (100% with NBS vs. 68% with clinical detection; p value not reported); only 2 NBS patients had reached this age. 18% of those identified through clinical detection attended regular school (status of remaining 14% unclear). • IQ: No differences in cognitive outcome were found between NBS and clinical detection groups (45 people assessed, numbers displayed graphically, p values not reported). Overall 76% were at least mildly mentally retarded (IQ<69), 16% had IQ 70 to 84, and 9% had an average IQ (85 to 114). • Number of metabolic crises showed negative correlation with IQ ($r=-0.422$, $p=0.007$), but age and residual PCC activity did not. • There was lower mortality in the NBS group but this did not reach significance (0% with NBS vs. 12% with clinical detection [6 reported to be known to have died]; difference reported as not significant, p value not reported) but the NBS group were noted to be usually younger than the clinically detected group. <p>They concluded that overall the reduced mortality supported screening for PA, but that no significant benefit for surviving patients could be shown</p> <p>KQ6:</p> <ul style="list-style-type: none"> • NBS significantly reduced age at diagnosis compared with clinical detection (median: 6 days [range 4 to 13 days] with NBS vs. 12 days [range 3 days to 8 years]; $p=0.018$)
<p>Comments</p>	<p>The study adjusted for multiple comparisons, which reduces the likelihood of differences being found by chance.</p> <p>Some of the participants had been reported in previous studies, such as that by Dionisi-Vici et al 2006²⁷</p> <p>9 clinically detected patients were either not asked to participate (2 due to loss to follow up) or refused to participate compared to 1 NBS detected patient who refused. The authors note that this and the inclusion of only those who were under the care of a metabolic centre may bias results, particularly for the clinically detected group.</p> <p>The difference in ages between the NBS and clinically detected groups may have influenced results. Age at death of those who died, and average age of the clinically detected group including those who died was not reported. Ideally a survival curve comparison would be used to take into account differences in age. The small numbers of cases is likely to limit the ability to detect differences in this and other outcomes. Not all individuals were included in all analyses, which could further reduce power.</p> <p>Often numbers and p values were not reported for the analyses, limiting ability to judge</p>

	whether lack of significance could be due to lack of power in the analyses. The treatment received by the groups was not reported, so it is unclear whether these were comparable.
--	--

Appendix number	21
Relevant criteria	KQ5
Publication details	American College of Medical Genetics (ACMG). Newborn screening: toward a uniform screening panel and system. 2006. ¹¹
Study details	<p>Practice guideline, USA, developed by a multidisciplinary expert group (n=292) working with a steering committee and expert working groups. The group provided their opinions on the extent to which individual met criteria relating to:</p> <ol style="list-style-type: none"> 1. The availability and characteristics of the screening test 2. The availability and complexity of diagnostic services 3. The availability and efficacy of treatments related to the conditions. <p>The conditions were then ranked based on results and then the evidence base was assessed in depth. The evidence was reported to be based on a systematic review of clinical evidence, cost/economic evidence and modelling, reference lists obtained from PubMed and Medline, books, health technology assessments commissioned by the UK NSC, internet searches, and professional guidelines.</p> <p>At least 2 experts assessed the data and level and quality of the evidence for each condition. The results of the evidence review and expert survey were considered by the ACMG along with over-riding principles, and other existing technology and condition specific recommendations to place the conditions into (a) core screening panel, (b) secondary targets (part of the differential diagnosis of a core panel condition, (c) not appropriate for newborn screening.</p> <p>The ACMG also assessed to what extent screening programmes met goals such as the availability of educational programmes, uptake of screening and follow up based on data from the National Newborn Screening and Genetics Resource Centre 2002 and experts. A brief cost-effectiveness analysis was also carried out.</p>
Study objectives	<ol style="list-style-type: none"> 1. To analyse the scientific literature on the effectiveness of newborn screening. 2. To gather expert opinion to outline the best evidence for screening for specified conditions and develop recommendations focused on newborn screening, including the development of a uniform condition panel. 3. To consider other parts of the newborn screening system that are critical to achieving the expected outcomes in those screened.
Inclusions	Specific inclusion criteria for the systematic review part of the process were not reported.
Exclusions	Specific exclusion criteria for the systematic review part of the process were not reported.

Population	Newborns
Intervention/test	Newborn screening for conditions including PA (and MMA)
Comparator	NA
Results/outcomes	<p>Final outcome was as follows:</p> <ul style="list-style-type: none"> • Core panel: 29 conditions, including PA (and also MMA caused by MUT mutations, and Cbl A,B; and MCD) • Secondary panel: 25 conditions, including MMA caused by Cbl C,D • Not appropriate for newborn screening: 27 conditions <p>They reported that they decided to subdivide PA, MUT, Cbl A,B, C, and D since they had quite different natural histories and treatment options. The MUT form of MMA was the higher scoring primary target (i.e. identified by experts as a better candidate for screening), while CblA, B and PA were lower scoring primary targets.</p> <p>For MMA there was “credible evidence of less than ideal sensitivity with the current testing technology (affected cases with normal concentration when tested at birth) and specificity (relatively high rate of false-positive results, including cases with relatively high levels that are followed up by perfectly normal plasma acylcarnitine and urine organic acid profiles). As the screening test for MMA and PA are the same, the issue with false positives will apply to screening for both conditions. The ACMG suggested that a second tier test to detect methylmalonic acid in bloodspots could improve test performance for MMA, however, a consideration of how this might impact detection of PA was not reported.</p> <p>On the basis of the guideline, fact sheets confirmatory algorithms were developed which are keep updated on an ongoing basis.</p>
Comments	<p>This guidance was developed in the US, and relied on expert views as well as research evidence. The UK viewpoint may differ.</p> <p>The ACMG note that some conditions have limited evidence, often because they are very rare. It noted that there was there potential for bias in the assessment of some conditions, as without screening, the more severe forms are noticed first, biasing what is known about the effects of the condition. Until a large general population has been studied understanding of the performance of the screening test (in terms of the range of manifestations it identifies) is limited.</p>

Appendix number	22
Relevant criteria	KQ3
Publication details	Dietzen DJ, Rinaldo P, Whitley RJ, et al. National academy of clinical biochemistry laboratory medicine practice guidelines: follow-up testing for metabolic disease identified by expanded newborn screening using tandem mass spectrometry; executive summary. Clin Chem. 2009;55(9):1615-26. ¹⁰

Study details	<p>A practice guideline for laboratory confirmation of positive newborn screening results issued by the National Academy of Clinical Biochemistry (NACB). The guideline was developed by experts in metabolic genetic diseases based on the available peer-reviewed evidence, guidelines by other groups such as the ACMG, and their experience and opinion.</p> <p>The panel graded the evidence and recommendations based on the US Preventive Services Task Force system. This was as follows:</p> <p><i>Recommendation gradings:</i></p> <p>A. The NACB strongly recommends adoption; there is good evidence that it improves important health outcomes and concludes that benefits substantially outweigh harms.</p> <p>B. The NACB recommends adoption; there is at least fair evidence that it improves important health outcomes and concludes that benefits outweigh harms.</p> <p>C. The NACB recommends against adoption; there is evidence that it is ineffective or that harms outweigh benefits.</p> <p>I. The NACB concludes that the evidence is insufficient to make recommendations; evidence that it is effective is lacking, of poor quality, or conflicting and the balance of benefits and harms cannot be determined.</p> <p><i>Evidence gradings:</i></p> <p>I. Evidence includes consistent results from well-designed, well-conducted studies in representative populations.</p> <p>II. Evidence is sufficient to determine effects, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies; generalizability to routine practice; or indirect nature of the evidence.</p> <p>III. Evidence is insufficient to assess the effects on health outcomes because of limited number or power of studies, important flaws in their design or conduct, gaps in the chain of evidence, or lack of information.</p>
Study objectives	To provide standard guidelines for follow up testing for metabolic diseases identified by expanded newborn screening using MS/MS.
Inclusions	NA
Exclusions	NA
Population	Newborns
Intervention/test	Newborn screening and follow up testing for conditions including PA (and MMA)
Comparator	NA
Results/outcomes	<p>KQ3:</p> <p><i>Expanded newborn screening:</i></p> <p>Newborn screening for PA (and MMA mutase, Cbl A, B, C and D) were rated as level A-II – i.e. strongly recommended based on evidence sufficient to determine effects but with limitations. This was reported to be based on the ACMG assessment, early screening</p>

	<p>programmes in Australia and the US, and the Region 4 international programme data and experience.</p> <p><i>Follow up in screen positives:</i></p> <p>Generic recommendations for follow up were given (not reported here), as well as disease specific recommendations.</p> <p>For PA the specific recommendations (grade A-I) were:</p> <ul style="list-style-type: none"> • Screening marker: C3-acylcarnitine • Follow up analyses: urine organic acids • Follow-up markers: 3-OH propionic acid, tiglyl-glycine, methylcitrate • Additional testing: B12 studies <p>They note that propionic acid is not routinely detected in urine organic acid analysis as it is volatile, but propionylglycine and 3-OH propionic acid are detected.</p> <p>For MMA the specific recommendations (grade A-I) were:</p> <ul style="list-style-type: none"> • Screening marker: C3-acylcarnitine • Follow up analyses: urine organic acids • Follow-up markers: methylmalonic, 3-OH propionic acid, tiglyl-glycine, methylcitrate • Additional testing: Complementation analysis, B12 studies <p><i>Patient outcomes from expanded newborn screening:</i></p> <p>The authors note the difficulty in meeting the traditional criteria for a screening program for all disorders detected through newborn MS/MS. For example, a number of the disorders, propionic acidemia, and several of the methylmalonic acidemias (not specified) can present with catastrophically with clinically severe metabolic decompensations at age 2 to 5 days, well before the 5 to 7 days required in many programs for completion of MS/MS newborn screening in the reference laboratory. They also note that the rarity of these disorders, and genetic and biochemical heterogeneity of the MMAs in particular, make assessment of patient outcomes and treatment effectiveness very difficult for these diseases.</p> <p>They rated the evidence for improved patient outcomes in neonatally detected patients with PA or MMA as grade A-II, based on the paper by Dionisi-Vici et al 2006.²⁷ They also say that cost effectiveness and health care use analyses confirm the utility of newborn screening for the 6 disorders they list as having evidence of improved outcomes, whether this applied to PA and MMA specifically was not reported.</p>
<p>Comments</p>	<p>This guidance was developed in the US, and relied on expert views as well as research evidence. The UK viewpoint may differ.</p> <p>The methods for searching for and selecting evidence for consideration was not described, not was the method of assessment of the quality of the individual studies. Therefore</p>

	whether the methods were subject to bias cannot be assessed.
Appendix number	23
Relevant criteria	KQ3
Publication details	Kronn D, Mofidi S, Braverman N, et al. Diagnostic guidelines for newborns who screen positive in newborn screening. Genet Med. 2010;12(12 Suppl):S251-5. ³⁰
Study details	<p>A consensus diagnostic guideline for newborns who screen positive in newborn screening results issued by the New York Mid-Atlantic Consortium for Genetic and Newborn Screening Services.</p> <p>The guideline was developed by a nationwide panel of metabolic experts. They met in groups to debate the issues and made recommendations based on this. These were then shared with additional specialists in the region and wider US for confirmation</p> <p>These guidelines were reported to be reviewed and revised regularly. The most recent version of the guideline was published in 2014.³¹</p>
Study objectives	To provide standard guidelines for diagnostic testing for diseases identified by expanded newborn screening using MS/MS, for use across the 8 screening programmes across the region.
Inclusions	NA
Exclusions	NA
Population	Newborns
Intervention/test	Newborn screening and follow up testing for conditions including PA (and MMA)
Comparator	NA
Results/outcomes	<p>KQ3:</p> <p>For PA specifically the most recent recommendations (2014) were for:</p> <ul style="list-style-type: none"> • Initial diagnostic tests at the referral centre: urine organic acids (UOA), plasma acylcarnitine profile, plasma methylmalonic acid, plasma amino acids (PAA), total homocysteine, and B12 levels in patient and mother • Additional testing to consider at the time of the initial consultation: electrolytes (especially if infant is ill), glucose, ammonia, carnitine (total and free), urine ketones • The abnormal metabolites expected are: <ul style="list-style-type: none"> ○ UOA: Elevated 3-OH-propionate, propionylglycine, tiglylglycine, propionate(volatile, so not always detected) and methylcitrate ○ Plasma: Elevated glycine; normal methylmalonic acid and homocysteine ○ Electrolytes abnormalities common in sick patients

	<ul style="list-style-type: none"> ○ Blood glucose depends on fed status of patient ○ Ammonia can be elevated in sick patients ○ Normal or low carnitine levels ○ Elevated urine ketones especially in sick patients ● Elevated propionate and methylcitrate are generally accepted for diagnosis ● If initial testing is negative the disorder is considered ruled out ● Forms of MMA are differential diagnoses (MUT, Cbl A, B, C and D1)
Comments	<p>This guidance was developed in the US, and relied on expert views as well as research evidence. The UK viewpoint may differ.</p> <p>This was a consensus guideline with no searching for or assessment of evidence reported.</p>

Appendix number	24
Relevant criteria	KQ1, KQ3, KQ4, KQ6
Publication details	Baumgartner MR, Horster F, Dionisi-Vici C, et al. Proposed guidelines for the diagnosis and management of methylmalonic and propionic acidemia. Orphanet J Rare Dis. 2014;9:130. ⁹
Study details	<p>European guideline, USA, developed using Scottish Intercollegiate Guideline Network (SIGN) methodology. A multidisciplinary guideline development group (GDG) (including representation from the UK) performed the systematic review (search date December 2011), drafted the guideline, discussed it with other GDG members and revised it based on this and other discussions with external stakeholders.</p> <p>Statements (recommendations) were graded:</p> <ul style="list-style-type: none"> ● A: based on the Level 1 evidence (meta-analyses, systematic reviews of RCTs, RCTs) ● B: based on Level 2 evidence (case-control or cohort studies or systematic reviews of these) ● C: based on level 3 evidence (non-analytic studies such as case reports and case series) ● D: based on level 4 evidence (expert opinion)
Study objectives	To standardise the diagnosis, therapy and long-term management of PA (and MMA) in Europe based on the highest level of evidence available.
Inclusions	NA
Exclusions	NA
Population	Newborns
Intervention/test	Diagnosis and management of PA (and MMA)

Comparator	NA
Results/outcomes	<p>70 recommendations are given for the diagnosis and management of PA and MMA. Those most relevant to the key questions of the current review are reported here.</p> <p>KQ1a:</p> <ul style="list-style-type: none"> • Statement 9 (Grade B) No clear-cut genotype-phenotype correlations have been found in PA. <p>It noted that although there are some common mutations, most mutations are private.</p> <p>KQ3:</p> <ul style="list-style-type: none"> • Statement 6 (Grade B-C) Determination of organic acids in urine and the acylcarnitine profile in blood are the most commonly used investigations to detect MMA and PA. Determination of amino acid concentrations may help in diagnosis and treatment. In addition total plasma homocysteine allows differentiation between the various types of MMA. • Statement 7 (Grade B/D) Enzymatic studies and/or molecular genetic analyses should be performed to confirm diagnosis (B). This is ideally done in specialized laboratories (D). <p>KQ4:</p> <ul style="list-style-type: none"> • Statement 51 (Grade C-D) Despite medical treatment, MMA and PA are associated with a high frequency of intellectual disability. <p>The text to support this statement suggested that prospective and retrospective studies on PA and MMA have suggested that intellectual disability and developmental delay may not be prevented by early treatment. For PA this appeared to be based on two case series, one small (n=5) and one published before the search date of this current NSC review, as well as one comparative study included in the current NSC review.</p> <p>KQ4&5:</p> <ul style="list-style-type: none"> • Statement 11 (Grade C-D) Newborn screening for MMA and PA is technically feasible. So far available data about outcome has not answered the question as to whether newborn screening in MMA/PA is of long term clinical benefit. <p>KQ5:</p> <p>Additional recommendations covered long term management, which included:</p> <ul style="list-style-type: none"> • Dietary management, mainly low protein diet (exact levels reported to be guided by age, growth, metabolic stability and severity of condition) with PA precursor free amino acid supplements as needed to make up for any deficit. • L-carnitine supplementation to enhance propionyl group elimination and transform certain toxic metabolites into less toxic forms • Oral antibiotics continuously or intermittently to control intestinal propionic acid producing bacteria • Avoidance of certain drugs such as steroids and chemotherapy drugs

	<p>KQ6:</p> <ul style="list-style-type: none"> Statement 10 (Grade D) Prenatal testing in both diseases is feasible. Prior to testing, it is desirable that the index case has been confirmed biochemically and/or genetically, and the carrier status of the parents has been confirmed by mutation analysis.
Comments	The authors noted that the consensus recommendations, although based on best available evidence, often only represented expert opinion, and were meant to be followed flexibly, applying the practitioner's own experience, and taking into account the individual patient.

Appendix number	25
Relevant criteria	KQ1, KQ5
Publication details	<p>(1) Chapman KA, Summar ML. Propionic acidemia consensus conference summary. Mol Genet Metab. 2012;105(1):3-4.¹</p> <p>(2) Chapman KA, Gropman A, MacLeod E, et al. Acute management of propionic acidemia. Mol Genet Metab. 2012;105(1):16-25.⁴¹</p> <p>(3) Sutton VR, Chapman KA, Gropman AL, et al. Chronic management and health supervision of individuals with propionic acidemia. Mol Genet Metab. 2012;105(1):26-33.⁴²</p> <p>(4) Schreiber J, Chapman KA, Summar ML, et al. Neurologic considerations in propionic acidemia. Mol Genet Metab. 2012;105(1):10-5.⁴³</p> <p>(5) Pena L, Franks J, Chapman KA, et al. Natural history of propionic acidemia. Mol Genet Metab. 2012;105(1):5-9.⁶</p>
Study details	<p>Guidelines developed as part of an international multidisciplinary consensus conference in the US in 2011.</p> <p>The recommendations were reported to come from an exhaustive search of the English language literature (not further described) which "informed expert opinion and allowed for debate of controversial issues". This led to production of recommendations on acute and chronic treatment and monitoring plus reviews of the natural history and neurological complications of PA.</p>
Study objectives	To better describe treatment approaches, identify complications, and provide baseline treatment recommendations based on best literature available and expert opinion.
Inclusions	Specific inclusion criteria for literature were not reported.
Exclusions	Specific exclusion criteria for literature were not reported.
Population	Individuals with PA.
Intervention/test	Natural history and management of PA
Comparator	NA

<p>Results/outcomes</p>	<p>KQ1a:</p> <p>As of yet no clear genotype-phenotype correlations are known.</p> <p>Most affected individuals are compound heterozygotes, which makes establishing genotype-phenotype correlations difficult.</p> <p>Mutations in <i>PCCB</i> are most common, with the c.1218_1231del14ins12 mutation found in 32% of mutant alleles in a Caucasian population. Seven homozygotes for this mutation have been reported and they all had early onset of symptoms, developmental delay and hypotonia.</p> <p>The c.1304T>C (p.Y435C) mutation is common in the Japanese population (25% of mutant alleles) and was reported to be characterised as mild.</p> <p>Various mutations in <i>PCCA</i> have been described, with varying age of onset and symptoms.</p> <p>KQ5:</p> <p>The sections on acute and chronic treatment and monitoring, and neurological complications did not explicitly cover the treatment of asymptomatic individuals with PA identified through screening. None of the recommendations were guided by specific genotype or level of enzyme activity.</p> <p>The section on acute management recommended that in a symptomatic individual in whom PA is known or highly suspected as a result of newborn screening treatment should be initiated immediately, but did not explicitly state what approach should be taken for asymptomatic individuals.</p>
<p>Comments</p>	<p>Details of the literature searching, selection and appraisal procedures were not reported, therefore this cannot be considered a systematic review, and the risk of bias in the procedure cannot be assessed.</p>

References

1. Chapman KA, Summar ML. Propionic acidemia consensus conference summary. *Mol Genet Metab.* 2012;105(1):3-4.
2. American College of Medical Genetics. Newborn screening confirmatory algorithm: C3 elevated (isolated) [Internet]. [updated 2009; cited 21 May 2015]. Available from: <https://www.acmg.net/StaticContent/ACT/Algorithms/Visio-C3.pdf>.
3. American College of Medical Genetics. Newborn Screening ACT Sheet [Elevated C3 Acylcarnitine] Propionic Acidemia and Methylmalonic Acidemia [Internet]. [updated 2012; cited 21 May 2015]. Available from: <https://www.acmg.net/StaticContent/ACT/C3.pdf>.
4. Chiu YH, Liu YN, Liao WL, et al. Two frequent mutations associated with the classic form of propionic acidemia in Taiwan. *Biochem Genet.* 2014;52(9-10):415-29.
5. Desviat LR, Perez B, Perez-Cerda C, et al. Propionic acidemia: mutation update and functional and structural effects of the variant alleles. *Mol Genet Metab.* 2004;83(1-2):28-37.
6. Pena L, Franks J, Chapman KA, et al. Natural history of propionic acidemia. *Mol Genet Metab.* 2012;105(1):5-9.
7. Fernandez-Lainez C, Aguilar-Lemus JJ, Vela-Amieva M, et al. Tandem mass spectrometry newborn screening for inborn errors of intermediary metabolism: abnormal profile interpretation. *Curr Med Chem.* 2012;19(26):4511-22.
8. McHugh D, Cameron CA, Abdenur JE, 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(3):230-54.
9. Baumgartner MR, Horster F, Dionisi-Vici C, et al. Proposed guidelines for the diagnosis and management of methylmalonic and propionic acidemia. *Orphanet J Rare Dis.* 2014;9:130.
10. Dietzen DJ, Rinaldo P, Whitley RJ, et al. National academy of clinical biochemistry laboratory medicine practice guidelines: follow-up testing for metabolic disease identified by expanded newborn screening using tandem mass spectrometry; executive summary. *Clin Chem.* 2009;55(9):1615-26.
11. American College of Medical Genetics. Newborn screening: toward a uniform screening panel and system. 2006. Available from: https://www.acmg.net/ACMG/Publications/Practice_Guidelines_docs/NBS_report.aspx.
12. Lund AM, Hougaard DM, Simonsen H, 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):281-93.
13. Schulze A, Lindner M, Kohlmuller D, et al. 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.
14. Clavero S, Martinez MA, Perez B, et al. Functional characterization of PCCA mutations causing propionic acidemia. *Biochim Biophys Acta.* 2002;1588(2):119-25.
15. Clavero S, Perez B, Rincon A, et al. Qualitative and quantitative analysis of the effect of splicing mutations in propionic acidemia underlying non-severe phenotypes. *Hum Genet.* 2004;115(3):239-47.
16. Kim SN, Ryu KH, Lee EH, et al. Molecular analysis of PCCB gene in Korean patients with propionic acidemia. [Erratum appears in *Mol Genet Metab.* 2003 Mar;78(3):236-7]. *Mol Genet Metab.* 2002;77(3):209-16.

17. Perez B, Desviat LR, Rodriguez-Pombo P, et al. Propionic acidemia: identification of twenty-four novel mutations in Europe and North America. *Mol Genet Metab.* 2003;78(1):59-67.
18. Kraus JP, Spector E, Venezia S, et al. Mutation analysis in 54 propionic acidemia patients. *J Inherit Metab Dis.* 2012;35(1):51-63.
19. Perez B, Angaroni C, Sanchez-Alcudia R, et al. The molecular landscape of propionic acidemia and methylmalonic aciduria in Latin America. *J Inherit Metab Dis.* 2010;33(Suppl 2):S307-14.
20. Yorifuji T, Kawai M, Muroi J, et al. Unexpectedly high prevalence of the mild form of propionic acidemia in Japan: presence of a common mutation and possible clinical implications.[Erratum appears in *Hum Genet.* 2003 Jan;112(1):100.]. *Hum Genet.* 2002;111(2):161-5.
21. Grunert SC, Mullerleile S, De Silva L, et al. Propionic acidemia: clinical course and outcome in 55 pediatric and adolescent patients. *Orphanet J Rare Dis.* 2013;8:6.
22. Sass JO, Hofmann M, Skladal D, et al. Propionic acidemia revisited: a workshop report. *Clin Pediatr (Phila).* 2004;43(9):837-43.
23. Rafique M. Clinical spectrum of propionic acidemia. *J Nutr Metab.* 2013;2013:975964.
24. Vatanavicharn N, Liammongkolkul S, Sakamoto O, et al. Clinical characteristics and mutation analysis of propionic acidemia in Thailand. *World J Pediatr.* 2014;10(1):64-8.
25. Joshi SN, Venugopalan P. Clinical characteristics of neonates with inborn errors of metabolism detected by Tandem MS analysis in Oman. *Brain Dev.* 2007;29(9):543-6.
26. Hoffmann GF, von Kries R, Klose D, et al. Frequencies of inherited organic acidurias and disorders of mitochondrial fatty acid transport and oxidation in Germany. *Eur J Pediatr.* 2004;163(2):76-80.
27. Dionisi-Vici C, Deodato F, Roschinger W, et al. 'Classical' organic acidurias, propionic aciduria, methylmalonic aciduria and isovaleric aciduria: long-term outcome and effects of expanded newborn screening using tandem mass spectrometry. *J Inherit Metab Dis.* 2006;29(2-3):383-9.
28. Grunert SC, Mullerleile S, de Silva L, et al. Propionic acidemia: neonatal versus selective metabolic screening. *J Inherit Metab Dis.* 2012;35(1):41-9.
29. Waisbren SE, Albers S, Amato S, et al. Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *Jama.* 2003;290(19):2564-72.
30. Kronn D, Mofidi S, Braverman N, et al. Diagnostic guidelines for newborns who screen positive in newborn screening. *Genet Med.* 2010;12(12 Suppl):S251-5.
31. Kronn D, Mofidi S, Braverman N, et al. Diagnostic Guidelines for Confirmation of Screen-Positive Newborn Screening Results [Internet]. [cited 06/05/2015]. Available from: http://www.wadsworth.org/newborn/nymac/docs/DX_Guidelines_2014-10-01.pdf.
32. Frazier DM, Millington DS, McCandless SE, et al. The tandem mass spectrometry newborn screening experience in North Carolina: 1997-2005. *J Inherit Metab Dis.* 2006;29(1):76-85.
33. Niu DM, Chien YH, Chiang CC, et al. Nationwide survey of extended newborn screening by tandem mass spectrometry in Taiwan. *J Inherit Metab Dis.* 2010;33(Suppl 2):S295-305.
34. Weisfeld-Adams JD, Morrissey MA, Kirmse BM, et al. Newborn screening and early biochemical follow-up in combined methylmalonic aciduria and homocystinuria, cblC type, and utility of methionine as a secondary screening analyte. *Mol Genet Metab.* 2010;99(2):116-23.

35. la Marca G, Malvagia S, Casetta B, et al. Progress in expanded newborn screening for metabolic conditions by LC-MS/MS in Tuscany: update on methods to reduce false tests. *J Inherit Metab Dis.* 2008;31 Suppl 2:S395-404.
36. Lim JS, Tan ES, John CM, et al. Inborn Error of Metabolism (IEM) screening in Singapore by electrospray ionization-tandem mass spectrometry (ESI/MS/MS): An 8 year journey from pilot to current program. *Mol Genet Metab.* 2014;113(1-2):53-61.
37. Shigematsu Y, Hata I, Tajima G. Useful second-tier tests in expanded newborn screening of isovaleric acidemia and methylmalonic aciduria. *J Inherit Metab Dis.* 2010;33(Suppl 2):S283-8.
38. Loukas YL, Soumelas GS, Dotsikas Y, et al. Expanded newborn screening in Greece: 30 months of experience. *J Inherit Metab Dis.* 2010;33 Suppl 3:S341-8.
39. Lindner M, Ho S, Kolker S, et al. Newborn screening for methylmalonic acidurias-- optimization by statistical parameter combination. *J Inherit Metab Dis.* 2008;31(3):379-85.
40. Matern D, Tortorelli S, Oglesbee D, et al. Reduction of the false-positive rate in newborn screening by implementation of MS/MS-based second-tier tests: the Mayo Clinic experience (2004-2007). *J Inherit Metab Dis.* 2007;30(4):585-92.
41. Chapman KA, Gropman A, MacLeod E, et al. Acute management of propionic acidemia. *Mol Genet Metab.* 2012;105(1):16-25.
42. Sutton VR, Chapman KA, Gropman AL, et al. Chronic management and health supervision of individuals with propionic acidemia. *Mol Genet Metab.* 2012;105(1):26-33.
43. Schreiber J, Chapman KA, Summar ML, et al. Neurologic considerations in propionic acidemia. *Mol Genet Metab.* 2012;105(1):10-5.
44. Perez-Cerda C, Clavero S, Perez B, et al. Functional analysis of PCCB mutations causing propionic acidemia based on expression studies in deficient human skin fibroblasts. *Biochim Biophys Acta.* 2003;1638(1):43-9.